Kromatic Symmetric Functions and Cographs

Soham Samanta
soham2020sam@gmail.com

Under the direction of
Laura Pierson

September 2026

Kromatic Symmetric Functions and Cographs

Soham Samanta

Abstract

We study how much of a graph can be recovered from its Kromatic symmetric function (KSF). This function was introduced by Crew, Pechenik, and Spirkl as a K-theoretic refinement of Stanley's chromatic symmetric function (CSF). The smallest failures of reconstruction occur on eight vertices: there are four pairs of nonisomorphic graphs with the same KSF. We describe these pairs and constructions that produce larger examples. We also show how the KSF separates many pairs arising from known equal-CSF constructions. Our main reconstruction result concerns cographs, obtained from single vertices by successive disjoint unions and joins. Let

$$F = (P_3 \sqcup K_1) \odot 2K_1.$$

For an F-free cograph, we prove that its KSF determines the graph up to isomorphism. This includes claw-free cographs and cographs with canonical cotree height at most three. The proof first recovers the component KSFs, then uses Newton polytopes to identify the factors in a join. Computations through twelve vertices find no two nonisomorphic cographs with the same KSF. We conjecture that reconstruction holds for all cographs.

Keywords: Kromatic symmetric function, chromatic symmetric function, cographs, graph reconstruction, independence polynomials, Newton polytopes

Contents

1 Introduction **4**

2 Background and notation **5**

3 Graphs with equal Kromatic symmetric functions **10**

4 The Orellana–Scott construction **14**

5 The split-graph construction **17**

6 Information determined by the KSF **24**

7 Reconstructing connected components **29**

8 Reconstructing cographs **32**

9 The cograph conjecture **38**

Acknowledgments **44**

References **44**

1 Introduction

Stanley's chromatic symmetric function (CSF) records proper graph colorings by their color-class sizes [Sta95]. Each coloring contributes a monomial, and summing these monomials gives a symmetric function that refines the chromatic polynomial. A central question is how much this function tells us about the graph. It distinguishes several families of trees [Mor05, AZ14, MMW08], but different graphs can have equal CSFs. Infinite families of such pairs were constructed by Orellana and Scott [OS14] and by Aliste-Prieto, Crew, Spirkl, and Zamora [ACSZ21].

The Kromatic symmetric function (KSF) of Crew, Pechenik, and Spirkl [CPS26] allows a vertex to receive several colors. Each vertex receives a finite, nonempty color set, and no color may occur at both ends of an edge. Summing the corresponding monomials defines the KSF. We denote the two functions by X_G and $\overline{X}_G$, respectively. The terms of least degree in $\overline{X}_G$ are exactly X_G; the higher-degree terms may separate graphs whose CSFs agree.

We use Pierson's description of the information in $\overline{X}_G$ [Pie25]: it is equivalent to the multiset obtained by taking the independence polynomial of every induced subgraph of G. Theorem 2.3 recalls this result. An independence polynomial records the numbers of independent sets, with the coefficient of x^k counting those of size k. When only one graph realizes a given independence polynomial, its multiplicity in this multiset counts induced copies of that graph in G. These counts will provide several of our KSF comparisons.

Pierson asked whether the KSF distinguishes all graphs [Pie24, Conjecture 2]. Section 3 presents the smallest counterexamples: four pairs on eight vertices. It also describes operations that preserve equality of KSFs. Along with disjoint unions and joins, we consider two attachments: joining a new graph to every vertex outside a chosen clique, and joining it only to a chosen vertex.

We then study the Orellana–Scott and Aliste-Prieto–Crew–Spirkl–Zamora constructions in Sections 4 and 5. These give graphs with the same CSF, so we ask when the two graphs also have the same KSF. By counting induced claws in the first construction and certain seven-vertex induced subgraphs in the second, we obtain conditions under which their KSFs are different. Sections 3–5 build on joint work of Pierson and Samanta [PS26]. The eight-vertex examples and the attachment constructions in Section 3 come from Section 2 of that paper. Theorem 4.2 extends its claw comparison to the setwise automorphism hypotheses of Theorem 4.1. Theorem 5.2 gives an exact difference formula for the split-graph construction; Corollary 5.7 recovers the sufficient condition in [PS26, Theorem 4.2]. The cograph reconstruction results in Sections 6–9 are not treated in [PS26].

These counterexamples leave open the reconstruction question for particular graph families. We focus on *cographs*. Starting with single vertices, disjoint unions and joins produce exactly this class; equivalently, cographs are the graphs without an induced four-vertex path [CLS81]. Tsujie showed that the CSF does not distinguish all cographs [Tsu18]. The eight-vertex KSF counterexamples lie outside this class, and our computations through twelve vertices found no counterexample among cographs. This leads to the following conjecture.

Conjecture 1.1. *If G, H are cographs and $\overline{X}_G = \overline{X}_H$, then $G \cong H$.*

Our main theorem proves reconstruction when one further induced subgraph is excluded. Here P_3 is the three-vertex path, K_1 is a single vertex, and $\sqcup$ and $\odot$ denote disjoint union and join. We use F-free to mean that no vertex subset induces a copy of F.

Theorem (Theorem 8.8)**.** *Let $F = (P_3 \sqcup K_1) \odot 2K_1$. If G is an F-free cograph and H is any graph with $\overline{X}_G = \overline{X}_H$, then $G \cong H$.*

This includes all claw-free cographs and all cographs whose canonical cotrees have height at most three. There are also F-free cographs with arbitrarily large cotree height, since we can keep taking

disjoint unions and adding universal vertices. For connected claw-free cographs, we prove that even the polynomial

$$P_G(u,v) = \sum_{S \subseteq V(G)} u^{|S|} v^{e(G[S])}$$

determines the graph. This polynomial counts induced subgraphs according to their numbers of vertices and edges, and it can be computed from the KSF.

To prove these results, we first show that the KSF determines the KSF of each connected component, so it is enough to consider connected cographs. Every connected cograph with at least two vertices is a join of smaller cographs. We use a polynomial J_G with the same information as the KSF for which $J_{G \odot H} = J_G J_H$. When the join factors are disjoint unions of cliques, we show that different factors give polynomials with no common nonconstant factor. The proof uses the Newton polytopes of these polynomials. We can then determine the join factors from their product, and extend the result by taking disjoint unions and adding universal vertices.

After the background in Section 2, Sections 3–5 develop the constructions and comparisons described above. Section 6 recovers vertex data and subgraph counts from the KSF. Component reconstruction follows in Section 7 and the cograph theorems in Section 8. Section 9 returns to Conjecture 1.1, reformulates it using products of the polynomials J_G, and gives further special cases and computational evidence.

2 Background and notation

We work with finite simple graphs, giving each vertex weight one except when other weights are explicitly assigned. We write $G[S]$ for the induced subgraph on $S \subseteq V(G)$, $G - S$ for $G[V(G) \setminus S]$, and $\overline{G}$ for the complement. We denote disjoint union by $G \sqcup H$. In the join $G \odot H$, we also connect every vertex of one graph to every vertex of the other. We use stable set and independent set interchangeably for a vertex subset with no internal edges. The graph with no vertices has CSF and KSF both equal to one. The claw is $K_{1,3}$, and claw-free graphs exclude it as an induced subgraph. To say that the KSF determines a graph means that its isomorphism type is the only one realizing that KSF.

2.1 Colorings and symmetric functions

To assign positive integer weights to a graph, we specify (G, w) with $w : V(G) \to \mathbb{Z}_{>0}$. In a proper coloring $\kappa : V(G) \to \mathbb{Z}_{>0}$, the two endpoints of every edge have different images. The weighted CSF sums over these colorings:

$$X_{(G,w)}(\mathbf{x}) = \sum_{\kappa} \prod_{v \in V(G)} x_{\kappa(v)}^{w(v)} \qquad \text{[Sta95, CS20]}.$$

For a proper set coloring, the color assigned at each vertex is replaced by a finite, nonempty subset of $\mathbb{Z}_{>0}$. The sets assigned at the ends of an edge must be disjoint. Summing over proper set colorings gives the KSF:

$$\overline{X}_{(G,w)}(\mathbf{x}) = \sum_{\kappa} \prod_{v \in V(G)} \prod_{i \in \kappa(v)} x_i^{w(v)} \qquad \text{[CPS26]}. \tag{1}$$

Relabeling the colors leaves either sum unchanged, so both functions are symmetric. Let Λ be the symmetric-function ring over $\mathbb{Q}$, with degree-d part Λ_d. Although the CSF lies in Λ, the KSF

generally has no bound on its degree. We therefore use the completion

$$\overline{\Lambda} = \prod_{d \geq 0} \Lambda_d.$$

An element of this ring has one symmetric-function component in each degree. Positive vertex weights ensure that the component of any fixed degree is well defined, since only finitely many colors can occur in a term of that degree. The least possible degree is $\sum_v w(v)$, attained when every vertex receives just one color; this component is $X_{(G,w)}$. The bar in $\overline{X}_G$ denotes the Kromatic function, whereas the bar in $\overline{G}$ denotes graph complementation.

For a partition $\lambda = (\lambda_1 \geq \cdots \geq \lambda_\ell > 0)$, form m_λ by summing each monomial with exponent multiset λ once. Its augmented version is

$$\widetilde{m}_\lambda = m_\lambda \prod_{j \geq 1} r_j(\lambda)!,$$

where $r_j(\lambda)$ counts the parts of λ equal to j. If K_λ is the complete graph with vertex weights $\lambda_1, \ldots, \lambda_\ell$, then $X_{K_\lambda} = \widetilde{m}_\lambda$. The corresponding K-analogue is

$$\overline{\widetilde{m}}_\lambda = \overline{X}_{K_\lambda}.$$

Its lowest-degree part is $\widetilde{m}_\lambda$. Since the $\widetilde{m}_\lambda$ form a basis in each degree, we can express any element of $\overline{\Lambda}$ uniquely in terms of the $\overline{\widetilde{m}}_\lambda$ by working from lower degrees to higher ones. In particular, finite linear combinations of the $\overline{\widetilde{m}}_\lambda$ are unique. The empty partition gives the element 1.

In a *stable-set partition*, every block is stable. A *stable-set cover* instead allows overlap: it consists of distinct, nonempty stable sets and covers every vertex at least once. We denote the set of covers by $\mathrm{SSC}(G)$. For a cover C, sum the vertex weights within each member and arrange these sums in decreasing order to obtain $\lambda(C)$.

Proposition 2.1 (Stable-set expansions)**.** *For a weighted graph (G, w),*

$$X_{(G,w)} = \sum_{\pi} \widetilde{m}_{\lambda(\pi)}, \qquad \overline{X}_{(G,w)} = \sum_{C \in \mathrm{SSC}(G)} \overline{\widetilde{m}}_{\lambda(C)}, \tag{2}$$

where the first sum is over stable-set partitions [Sta95, CS20, CPS26].

Proof. For an ordinary coloring, the nonempty color classes give a stable-set partition. Assigning distinct colors to its blocks gives $\widetilde{m}_{\lambda(\pi)}$.

For a set coloring, associate each color with the vertices at which it occurs, and group together colors that give the same vertex set. The distinct vertex sets obtained in this way form a stable-set cover C. Each member of this cover receives a nonempty set of colors, and different members receive disjoint color sets. This is precisely a proper set coloring of the weighted complete graph whose vertex weights are the sums within the members of C. Its contribution is $\overline{\widetilde{m}}_{\lambda(C)}$. The grouping and its inverse are unique. □

2.2 Join multiplication

Following Tsujie [Tsu18], define a product on the augmented monomial basis $\widetilde{m}_\lambda$ by

$$\widetilde{m}_\lambda \odot \widetilde{m}_\mu = \widetilde{m}_{\lambda \sqcup \mu},$$

where $\lambda \sqcup \mu$ collects the parts of λ and μ into one partition. Distributivity extends the product to sums, degree by degree. Although we define it on $\widetilde{m}_\lambda$, Proposition 2.2 gives the corresponding rule for $\overline{\widetilde{m}}_\lambda$. Ordinary multiplication remains implicit. The symbol $\odot$ therefore denotes a product of symmetric functions or a join of graphs, according to its arguments.

Proposition 2.2. *For weighted graphs,*

$$\begin{aligned} X_{(G,w_G)\odot(H,w_H)} &= X_{(G,w_G)} \odot X_{(H,w_H)}, \\ \overline{X}_{(G,w_G)\odot(H,w_H)} &= \overline{X}_{(G,w_G)} \odot \overline{X}_{(H,w_H)}. \end{aligned} \tag{3}$$

Consequently $\overline{\widetilde{m}}_\lambda \odot \overline{\widetilde{m}}_\mu = \overline{\widetilde{m}}_{\lambda \sqcup \mu}$.

Proof. A stable set of a join lies entirely in one factor. The first equality follows from the stable-set partition expansion.

For the KSF identity, work with labeled vertices of G and let $\alpha \in \mathbb{Z}_{>0}^{V(G)}$ specify a positive integer at each vertex. Replace v by a clique of α_v vertices, all of weight $w_G(v)$, and join two replacement cliques when the original vertices were adjacent. Denote the resulting graph by $C_\alpha(G, w_G)$. Ordering the distinct colors at each original vertex gives

$$\overline{X}_{(G,w_G)} = \sum_{\alpha \in \mathbb{Z}_{>0}^{V(G)}} \frac{X_{C_\alpha(G,w_G)}}{\prod_v \alpha_v!}.$$

Indeed, a set coloring with α_v colors at v corresponds to exactly $\prod_v \alpha_v!$ ordinary colorings of the labeled replacement cliques. This identity is valid degree by degree.

The join of $C_\alpha(G, w_G)$ and $C_\beta(H, w_H)$ is the graph obtained from the join using the concatenated positive vector (α, β). Apply the first equality, multiply the two series, and group by these vectors. Their factorial denominators multiply, giving the second equality in (3). Taking $G = K_\lambda$ and $H = K_\mu$ proves the final assertion. □

For finite sums of the $\overline{\widetilde{m}}_\lambda$, the $\odot$ product can be identified with ordinary polynomial multiplication by setting

$$\overline{\widetilde{m}}_\lambda \longmapsto \prod_j y_{\lambda_j}.$$

Polynomial cancellation is therefore available for $\odot$. A finite graph has only finitely many stable-set covers, so its KSF is a finite sum of the $\overline{\widetilde{m}}_\lambda$ and belongs to this polynomial ring. This observation justifies the cancellations in the attachment formulas of Section 3.

2.3 Independence polynomials and induced copies

For an unweighted graph G, define

$$I_G(x) = \sum_{k \geq 0} i_k(G) x^k,$$

where $i_k(G)$ counts independent k-subsets. Thus $i_0(G) = 1$, $i_1(G) = |V(G)|$, and the degree is the independence number $\alpha(G)$. Let

$$\mathcal{I}(G) = \{\{I_{G[S]}(x) : S \subseteq V(G)\}\}.$$

The multiplicities count vertex subsets, so isomorphic induced subgraphs need not give distinct entries. The linear coefficient of an entry records the size of the subset that produced it.

Theorem 2.3 (Pierson [Pie25, Corollary 1.6])**.** *For finite unweighted graphs,* $\overline{X}_G = \overline{X}_H$ *if and only if* $\mathcal{I}(G) = \mathcal{I}(H)$.

Proof. We include a proof to keep the paper self-contained and to establish the change of variables used again in Section 9. We use stable-set covers. For an unweighted graph G, define

$$F_G(\mathbf{y}) = \sum_{C \in \mathrm{SSC}(G)} \prod_{S \in C} y_{|S|}.$$

Proposition 2.1 expresses the KSF using the same coefficients, and the $\overline{\widetilde{m}}_\lambda$ are linearly independent. Thus two graphs give the same cover polynomial exactly when their KSFs agree. The least degree also recovers the order n.

Apply inclusion–exclusion to the vertices left uncovered by an arbitrary collection of stable sets. A collection whose members all lie in $R \subseteq V(G)$ has generating polynomial $\prod_{k\geq 1}(1+y_k)^{i_k(G[R])}$, since each nonempty stable set is independently included or omitted. Hence

$$F_G(\mathbf{y}) = \sum_{R \subseteq V(G)} (-1)^{n-|R|} \prod_{k\geq 1} (1+y_k)^{i_k(G[R])}. \tag{4}$$

A collection covering a proper subset contributes zero in this sum, while a full cover contributes once.

Now substitute $y_1 = -z_1 - 1$ and $y_k = z_k - 1$ for $k \geq 2$. Since $i_1(G[R]) = |R|$, (4) becomes

$$(-1)^n \sum_{R \subseteq V(G)} \prod_{k\geq 1} z_k^{i_k(G[R])}.$$

Each monomial records exactly one independence polynomial, and its coefficient records the multiplicity. The substitution is invertible. Thus the KSF determines $\mathcal{I}(G)$ and conversely. The empty graph satisfies the same identity with $n = 0$. □

Corollary 2.4. *For any polynomial f, the KSF determines the number of subsets $S \subseteq V(G)$ for which $I_{G[S]} = f$. If H is the only graph up to isomorphism with independence polynomial I_H, then the KSF determines the number $n_H(G)$ of induced copies of H.*

Proof. The desired number is the multiplicity of f in $\mathcal{I}(G)$. In the second assertion, the subsets with polynomial I_H are exactly the induced copies of H. □

The small graphs in Figure 1 have induced-copy counts recoverable from the KSF [Pie24]. Some four-vertex graphs share an independence polynomial, so Corollary 2.4 alone does not recover every count. Theorem 6.2 separates the remaining cases using additional KSF data.

2.4 Notation

For reference, we collect the main notation below.

Symbol	Meaning
$G[S], G - S$	Induced subgraphs on S and on $V(G) \setminus S$.
$G \sqcup H, G \odot H$	Disjoint union and join.
$\widetilde{m}_\lambda, \overline{\widetilde{m}}_\lambda$	Augmented monomial function and its Kromatic analogue.
$\overline{\widetilde{m}}_{1^j}, \overline{\widetilde{m}}_{2^j}$	The partition index has j parts equal to one or two; these are not ordinary powers.
$I_G, \mathcal{I}(G)$	Independence polynomial and the multiset of all induced independence polynomials.
$P_G(u,v)$	Counts induced subgraphs by vertices and edges, Section 6.
$E_G(s,t)$	Counts edge subsets by incident vertices and edges, Section 6.
J_G, U_G	Polynomials carrying the full KSF information, Section 7.
$\mathrm{Newt}(A)$	Convex hull of the exponent vectors of A, Section 8.
$F_G, \widehat{F}_G$	Cover polynomials using one variable per size or per stable set, Section 9.

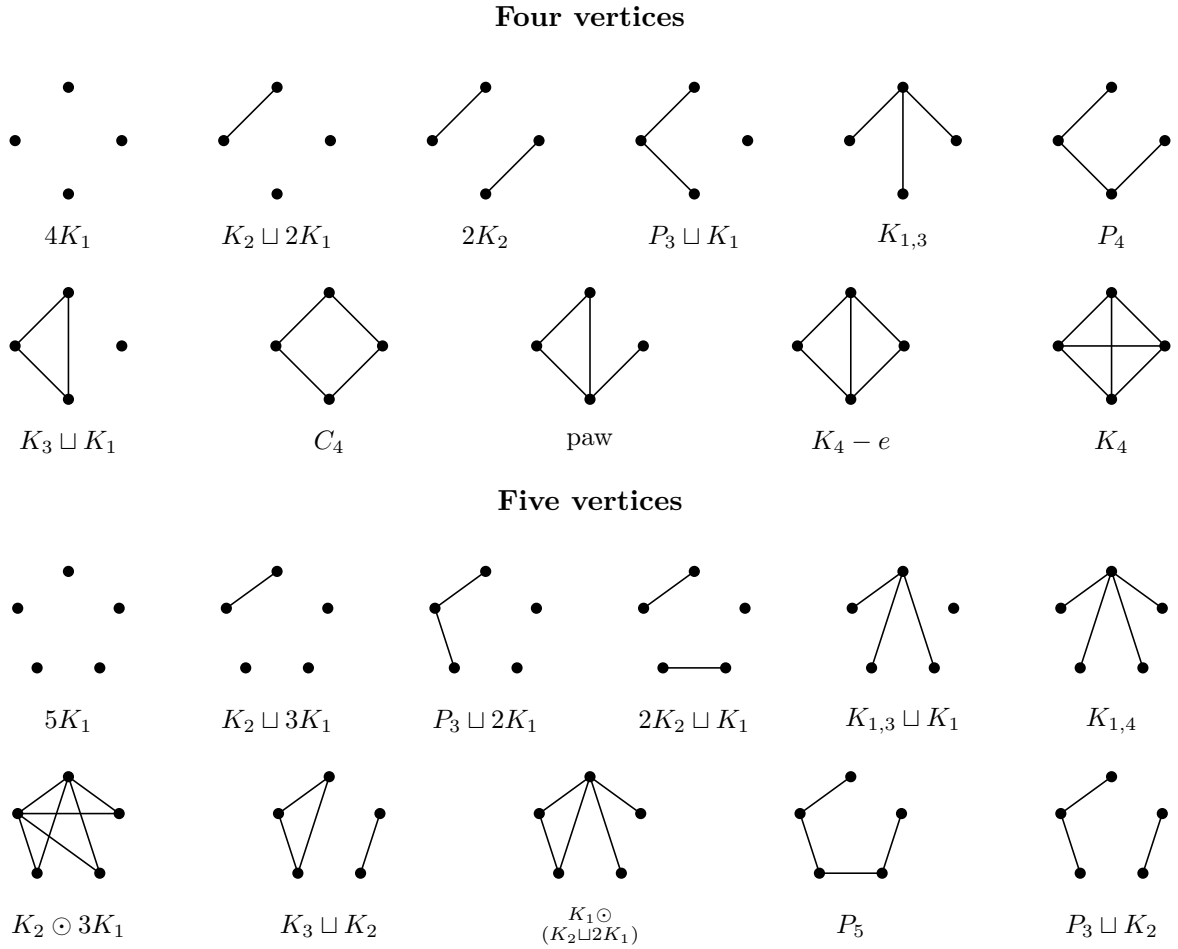


Figure 1: The eleven graphs on four vertices and the five-vertex examples of [Pie24, Theorems 3 and 4]. Their induced-copy counts are determined by the KSF. Theorem 6.2 recovers the counts for all eleven four-vertex graphs. The graph $K_4 - e$ is the diamond.

3 Graphs with equal Kromatic symmetric functions

3.1 The smallest examples

Figure 2 displays four pairs that disprove the reconstruction conjecture in [Pie24, Conjecture 2]. All eight graphs have order eight. The exhaustive comparison below shows that no smaller counterexample exists.

Proposition 3.1 (Computer check)**.** *The KSF distinguishes all graphs on at most seven vertices. Among the* 12 346 *isomorphism types on eight vertices, exactly four equivalence classes under equality of KSF contain more than one graph; each contains two graphs.*

Proof. For each graph in the catalogue, calculate the induced independence polynomials recursively:

$$I_{G[S]} = I_{G[S\setminus\{v\}]} + xI_{G[S\setminus N[v]]}, \qquad v \in S,$$

starting with $I_\varnothing = 1$. Record how many times each independence polynomial occurs, and compare the resulting tables of integer coefficients. Theorem 2.3 makes equality of these tables equivalent to equality of KSFs.

An exhaustive comparison of the unlabeled graphs through order seven gives no repeated table. At order eight there are 12 342 distinct tables, with precisely four double occurrences. Each repeated table is represented by the two nonisomorphic graphs in one row of Figure 2. The lists of graphs on seven and eight vertices are from McKay's graph data [McK26]. □

An induced P_4 occurs in every graph in Figure 2, placing all these examples outside the cograph family. The cograph computation reaches twelve vertices without a repeated KSF; see Proposition 9.10.

Proposition 3.2. *If* $\overline{X}_{G_1} = \overline{X}_{G_2}$ *and* $\overline{X}_{H_1} = \overline{X}_{H_2}$*, then*

$$\overline{X}_{G_1\sqcup H_1} = \overline{X}_{G_2\sqcup H_2}, \qquad \overline{X}_{G_1\odot H_1} = \overline{X}_{G_2\odot H_2}.$$

Proof. Proper set colorings on the two components can be chosen independently, so $\overline{X}_{G\sqcup H} = \overline{X}_G\overline{X}_H$. The join assertion follows from Proposition 2.2.

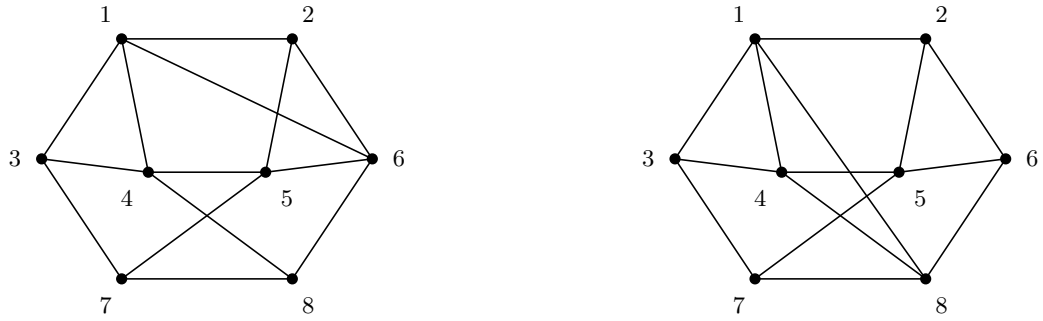
Alternatively, $I_{G\odot H} = I_G + I_H - 1$, since a nonempty independent set of a join lies in exactly one factor. Bijections between the equal induced-independence multisets in the two pairs combine to a bijection for their joins. Theorem 2.3 gives the same conclusion. □

If G_1 and G_2 are nonisomorphic, then adding the same graph H preserves this property. Indeed, an isomorphism $G_1\sqcup H \cong G_2\sqcup H$ would match the connected components. Canceling the components of H would give $G_1 \cong G_2$. Taking complements gives the same conclusion for joins. The displayed pairs therefore give infinite families of nonisomorphic graphs with equal KSF.

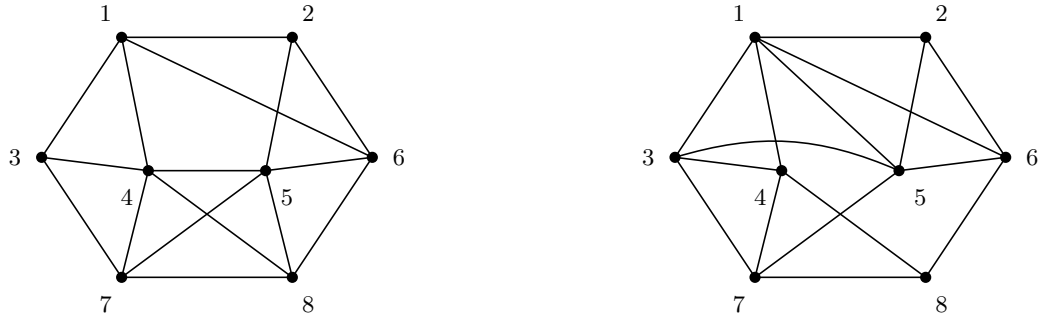
3.2 Attaching a graph outside a clique

Fix a k-vertex clique C of G. Starting from $G \sqcup H$, join each vertex of H to all vertices of G outside C, leaving C with no neighbors in H. Denote this graph by $G *_C H$. For $0 \le i \le k$, define

$$D_i(C,G) = \sum_{\substack{S\subseteq C\\ |S|=i}} \overline{X}_{G-S}. \tag{5}$$



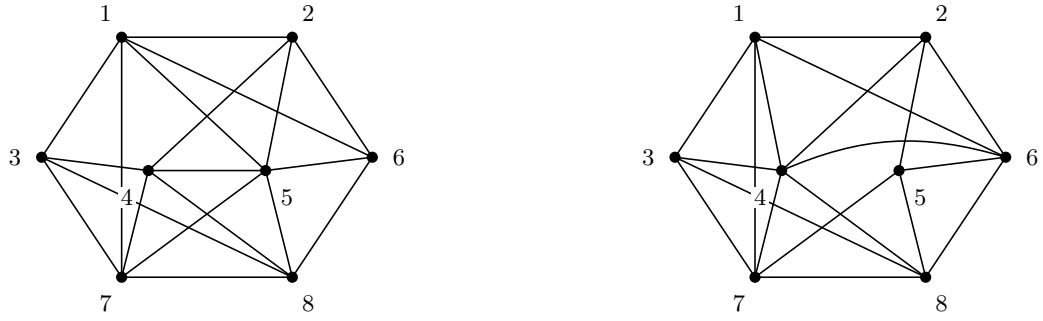
Pair 1



Pair 2



Pair 3



Pair 4

Figure 2: The four equal-KSF pairs of order eight. These are the smallest counterexamples to reconstruction. The vertex labels are used in the attachment examples of this section.

We will prove that these deletion sums and $\overline{X}_H$ determine $\overline{X}_{G*_C H}$.

Define $f(C, G)$ to be the sum of the $\overline{\widetilde{m}}$-terms of stable-set covers of G that do not use any singleton $\{v\}$ with $v \in C$. Every vertex of C is still covered, but only in sets of size at least two. Define

$$A_i(C, G) = \sum_{\substack{S \subseteq C \\ |S|=i}} f(C \setminus S, G - S).$$

We use the conventions

$$f(\varnothing, G) = \overline{X}_G, \qquad \overline{\widetilde{m}}_{1^0} = 1.$$

Here 1^j means a partition with j parts equal to one, so $\overline{\widetilde{m}}_{1^j} = \overline{X}_{K_j} = \overline{\widetilde{m}}_1^{\odot j}$, which is not an ordinary power.

Lemma 3.3. *With $A_i = A_i(C, G)$,*

$$\overline{X}_{G*_C H} = \sum_{i=0}^{k} A_i \odot \left(\sum_{j=i}^{k} \binom{k-i}{j-i} \overline{\widetilde{m}}_{1^j} \overline{X}_H \right). \tag{6}$$

The product inside parentheses is ordinary multiplication; the outer product is $\odot$.

Proof. Partition a stable-set cover into two collections. The first consists of sets meeting $V(G) \setminus C$. Such sets cannot meet H, and a set in this collection can contain at most one vertex of C. The second consists of the remaining sets, which lie in $C \sqcup H$; those not meeting H are singletons in C.

Let S' be the vertices of C not covered by the first collection, and let S be the vertices of C covered by the second. Covering all vertices requires $S' \subseteq S$. If $|S'| = i$, the first collection is precisely a cover counted by $f(C \setminus S', G - S')$. Conversely, every such cover consists of sets meeting $V(G) \setminus C$: a stable set entirely in the clique $C \setminus S'$ would be a forbidden singleton.

The second collection is a stable-set cover of $H \sqcup G[S]$. Since $G[S]$ is a clique of order $j = |S|$, its generating function is $\overline{\widetilde{m}}_{1^j} \overline{X}_H$ under ordinary multiplication. The two collections are disjoint, so their contributions combine by $\odot$. This construction is reversible because the two collections are distinguished by whether they meet $V(G) \setminus C$.

For a fixed S' of size i, there are $\binom{k-i}{j-i}$ supersets $S \subseteq C$ of size j. Summing over these choices gives (6). □

Lemma 3.4. *The deletion data $D_0(C, G), \ldots, D_k(C, G)$ determine all $A_i(C, G)$.*

Proof. Apply Lemma 3.3 to $G - T$, with exceptional clique $C \setminus T$ and $H = \varnothing$, and sum over $T \subseteq C$ of size i. If the further vertices missed by the first collection form $S \setminus T$, with $|S| = j$, then T can be chosen inside S in $\binom{j}{i}$ ways. The remaining optional singleton choices give the binomial expansion in $\odot$. Thus

$$D_i(C, G) = \sum_{j=i}^{k} \binom{j}{i} A_j(C, G) \odot \overline{\widetilde{m}}_{1^{j-i}} \odot (1 + \overline{\widetilde{m}}_1)^{\odot(k-j)}. \tag{7}$$

At $i = k$ this gives $A_k = D_k$. Now induct downward from $i = k - 1$ to $i = 0$. All summands with $j > i$ are already known, and the coefficient of A_i is the nonzero element $(1 + \overline{\widetilde{m}}_1)^{\odot(k-i)}$. As shown in Section 2, we can cancel a nonzero factor under $\odot$. Thus there is at most one possible value of A_i. The cover definition already gives a value satisfying the equation, so A_i is determined. Repeating this step recovers every A_i. □

Theorem 3.5. *Let $C_1 \subseteq V(G_1)$ and $C_2 \subseteq V(G_2)$ be cliques of the same order. If*

$$D_i(C_1, G_1) = D_i(C_2, G_2) \quad (0 \le i \le |C_1|), \qquad \overline{X}_{H_1} = \overline{X}_{H_2},$$

then $\overline{X}_{G_1 *_{C_1} H_1} = \overline{X}_{G_2 *_{C_2} H_2}$.

Proof. Lemma 3.4 gives matching A_i, and Lemma 3.3 then gives matching KSFs for the attached graphs. □

Corollary 3.6. *If $\overline{X}_{G_1} = \overline{X}_{G_2}$, $\overline{X}_{G_1 - v_1} = \overline{X}_{G_2 - v_2}$, and $\overline{X}_{H_1} = \overline{X}_{H_2}$, attaching H_i to every vertex of G_i except v_i gives equal KSFs.*

Proof. Take $C_i = \{v_i\}$ in Theorem 3.5. The two required deletion sums are exactly the stated KSFs. □

For a singleton clique the triangular relation has the particularly simple form

$$\overline{X}_G = f(\{v\}, G) \odot (1 + \widetilde{\overline{m}}_1) + \overline{X}_{G-v} \odot \widetilde{\overline{m}}_1.$$

It records the two possibilities: v is covered by another stable set, in which case its singleton is optional, or v is covered only by its singleton.

Example 3.7. Using the labels in Figure 2, the following choices satisfy the hypotheses. In the middle column a pair (a, b) denotes $v_1 = a, v_2 = b$; in the last column each entry gives the two exceptional cliques.

Pair	Singleton exceptions	Two-vertex clique exceptions
1	$(1,1),(3,2)$	None
2	$(5,1),(5,5),(6,4),(6,7)$	$(\{1,3\},\{4,8\}),(\{1,3\},\{7,8\})$
3	None	None
4	$(2,2),(4,1),(6,2),(8,1)$	$(\{4,5\},\{4,8\}),(\{5,8\},\{4,8\})$

The singleton conditions can be checked by isomorphism of the seven-vertex deletions. For a two-vertex clique $\{a, b\}$, the required data are $\overline{X}_G$, $\overline{X}_{G-a} + \overline{X}_{G-b}$, and $\overline{X}_{G-\{a,b\}}$. For each listed pair of cliques, these three expressions agree between the two graphs.

3.3 Attaching at one vertex

We next join every vertex of H to a single chosen vertex v of G. The stable sets containing v lead to another formula in terms of deletion sums. Suppose that

$$C = V(G) \setminus N_G[v]$$

is a clique of size k, and let $G \diamond_v H$ denote the resulting graph. The only stable sets containing v are $\{v\}$ and the pairs $\{v, c\}$ for $c \in C$; none can meet H. Define

$$E_j(v, G) = \sum_{\substack{S \subseteq C \\ |S| = j}} \overline{X}_{G - (\{v\} \cup S)}, \qquad 0 \le j \le k.$$

Theorem 3.8. *The KSF of $G \diamond_v H$ is*

$$\begin{aligned}&(\overline{X}_H E_0(v,G)) \odot [(1+\overline{\widetilde{m}}_1) \odot (1+\overline{\widetilde{m}}_2)^{\odot k} - 1] \\ &\quad + \sum_{j=1}^{k} (\overline{X}_H E_j(v,G)) \odot \overline{\widetilde{m}}_{2^j} \odot (1+\overline{\widetilde{m}}_1) \odot (1+\overline{\widetilde{m}}_2)^{\odot(k-j)}.\end{aligned} \tag{8}$$

Consequently, it suffices to match all E_j and $\overline{X}_H$ to preserve equality under this attachment.

Proof. From a stable-set cover remove all sets containing v. Let $S \subseteq C$ consist of vertices that are no longer covered. The remaining collection is a cover of $(G - (\{v\} \cup S)) \sqcup H$, and every missing vertex $c \in S$ must be covered by its pair $\{v, c\}$. These $j = |S|$ pairs are required and give $\overline{\widetilde{m}}_{2^j}$. Each of the other $k - j$ pairs is optional, as is the singleton $\{v\}$. This gives the summand for $j \geq 1$.

When S is empty, the remaining collection already covers every vertex except v. At least one set containing v must then be selected. The generating function for all optional choices is $(1+\overline{\widetilde{m}}_1) \odot (1+\overline{\widetilde{m}}_2)^{\odot k}$, and subtracting one removes the empty choice. This gives the first line.

Every cover determines S uniquely, and the sets removed are exactly those containing v. Thus the construction counts each cover once. □

Corollary 3.9. *For the second pair (G_1, G_2) of Figure 2, attaching equal-KSF graphs H_1, H_2 only to vertex 5 again gives equal KSFs.*

Proof. The nonneighbors of vertex 5, other than itself, are $C_1 = \{1, 3\}$ and $C_2 = \{4, 8\}$. Both are cliques. The deletions $G_1 - 5$ and $G_2 - 5$ are isomorphic, as are $G_1 - \{1, 3, 5\}$ and $G_2 - \{4, 5, 8\}$. Thus E_0 and E_2 agree.

Apply (8) with $H = \varnothing$. Since $\overline{X}_{G_1} = \overline{X}_{G_2}$, subtracting the two formulas and canceling the matching E_0, E_2 terms leaves

$$(E_1(5, G_1) - E_1(5, G_2)) \odot \overline{\widetilde{m}}_2 \odot (1+\overline{\widetilde{m}}_1) \odot (1+\overline{\widetilde{m}}_2) = 0.$$

The last three factors are nonzero in the $\odot$ polynomial ring. Therefore $E_1(5, G_1) = E_1(5, G_2)$ as well, and Theorem 3.8 applies to arbitrary H_1, H_2 with equal KSF. □

4 The Orellana–Scott construction

We now consider the construction of Orellana and Scott [OS14] of pairs with equal CSF. Their construction starts with four vertices inducing the path $u - z - w - v$. We will show that, under an additional condition, the two graphs have different numbers of induced claws and therefore different KSFs.

Theorem 4.1 (Orellana–Scott [OS14, Theorem 4.2])**.** *Suppose $uz, zw, wv \in E(G)$ and $uw, vz, uv \notin E(G)$. If $G - zw$ has an automorphism φ satisfying*

$$\varphi(\{u, w\}) = \{v, z\}, \qquad \varphi(\{v, z\}) = \{u, w\},$$

then $H = G + uw$ and $J = G + vz$ satisfy $X_H = X_J$.

Here $G + ab$ means that the nonedge ab is added. Figure 3 gives an example of the construction. For distinct vertices a, b, c, define



Figure 3: An Orellana–Scott pair with equal CSF. The added edges, uw in H and vz in J, are drawn in bold; all other edges have the standard line weight.

$$r_G(a; b, c) = |N_G(a) \setminus (N_G[b] \cup N_G[c])|.$$

A vertex counted here is a neighbor of a outside both closed neighborhoods of b and c. If a is the center of an induced path with endpoints b, c, adjoining such a vertex produces an induced claw. Let $c_{\text{claw}}(G) = n_{K_{1,3}}(G)$.

Theorem 4.2. *Under the hypotheses of Theorem 4.1,*

$$\begin{aligned} c_{\text{claw}}(H) - c_{\text{claw}}(J) = {} & r_G(w; u, v) + r_G(w; z, v) \\ & - r_G(z; u, v) - r_G(z; u, w). \end{aligned} \tag{9}$$

A nonzero value implies $\overline{X}_H \neq \overline{X}_J$.

Proof. Write $Q = \{u, v, w, z\}$. A four-set containing neither changed pair $\{u, w\}$ nor $\{v, z\}$ induces the same graph in H and J and contributes zero to the difference.

Now take a remaining four-set with exactly two vertices in Q. Those two vertices must form one of the pairs whose edge status changes. The automorphism φ interchanges these pairs. Neither the set nor its image contains both z and w. Consequently, neither induced subgraph uses the edge zw. The automorphism of $G - zw$ carries the added pair uw to vz, so it gives a bijection between these claws in H and J. Panels (a), (b), (e), and (f) of Figure 4 show the possible centers.

For a four-set meeting Q in three vertices, those three vertices must induce a path of length two or a stable triple if the full graph is a claw. In H, the triples of Q that can occur are $\{u, v, w\}$ and $\{z, v, w\}$, both with center w. Each of the other triples is either a triangle or an edge together with an isolated vertex. The possible fourth vertices are therefore counted by $r_G(w; u, v)$ and $r_G(w; z, v)$. In J, the contributing triples are $\{u, v, z\}$ and $\{u, w, z\}$, with center z, giving the two subtracted terms. In each case the unused vertex of Q is excluded by the closed-neighborhood conditions. The whole four-set Q is not a claw in either graph. These four counts are shown in panels (c), (d), (g), and (h). This proves (9).

Finally, the claw has independence polynomial $1 + 4x + 3x^2 + x^3$, and no other four-vertex graph has this polynomial. The table in Theorem 6.2 confirms this uniqueness, so Corollary 2.4 recovers the claw count from the KSF. □

Attaching rooted trees at u, z, v, w gives further applications of (9). The formula requires only the setwise symmetry in Theorem 4.1. For the next construction, we impose particular vertex swaps, allowing the trees on the two pairs to differ.

Corollary 4.3. *Let G_0 satisfy Theorem 4.1, with an automorphism of $G_0 - zw$ such that*

$$\varphi(u) = z, \qquad \varphi(z) = u, \qquad \varphi(v) = w, \qquad \varphi(w) = v.$$

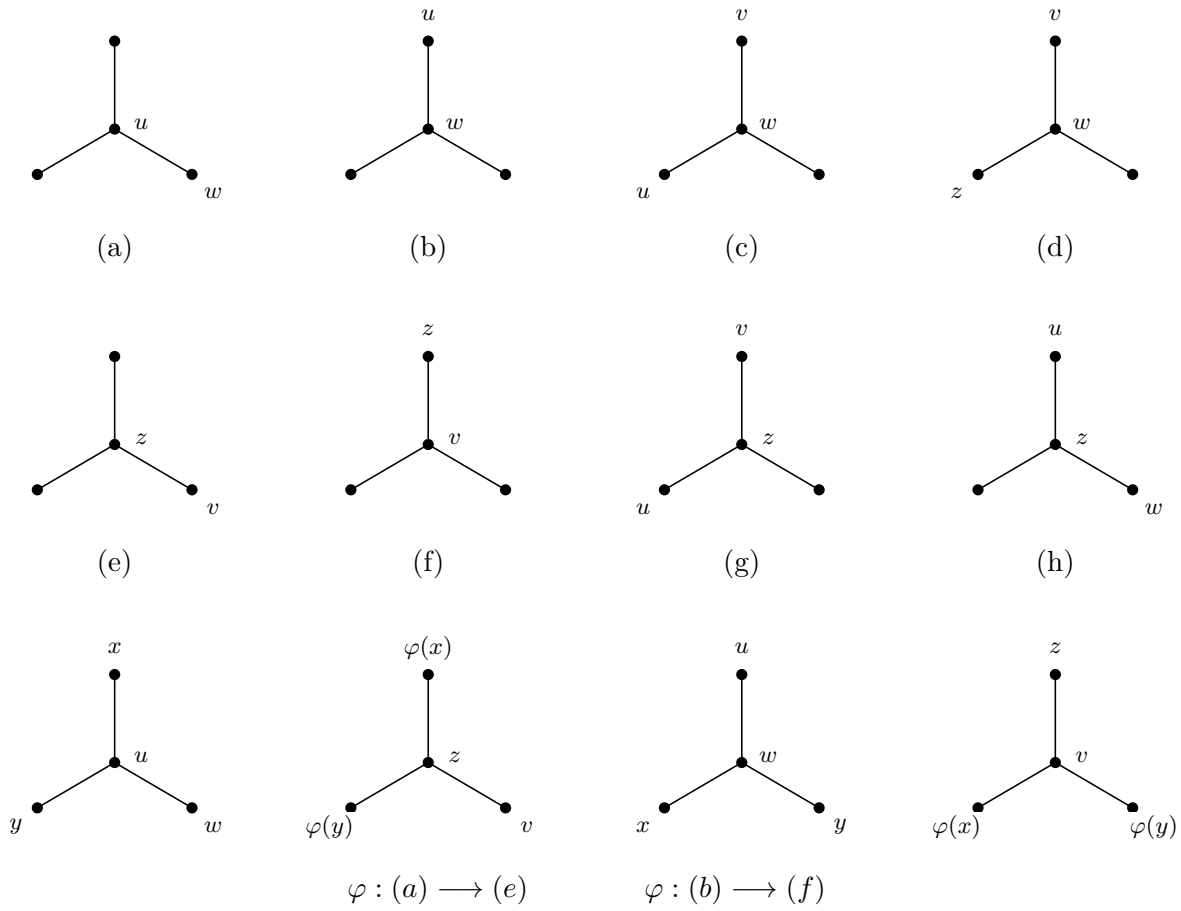


Figure 4: The possible claws in H, panels (a)–(d), and in J, panels (e)–(h), all in the same orientation. The lower diagrams show the two bijections for claws using exactly two vertices of $\{u, v, w, z\}$. The claws using three vertices give the four terms in (9).

Take four rooted trees, pairwise vertex-disjoint and disjoint from G_0, and identify their roots with u, z, v, w, respectively. Make no other identifications or attachments. Suppose the trees at u, z are isomorphic as rooted trees, and the trees at v, w are also isomorphic as rooted trees. Let G be the resulting graph, and let b_z, b_w be the numbers of new neighbors at z, w. Define

$$\delta_0 = r_{G_0}(w; u, v) + r_{G_0}(w; z, v) - r_{G_0}(z; u, v) - r_{G_0}(z; u, w).$$

Then $X_{G+uw} = X_{G+vz}$ and

$$c_{\text{claw}}(G + uw) - c_{\text{claw}}(G + vz) = \delta_0 + 2(b_w - b_z). \tag{10}$$

If the right-hand side is nonzero, the two KSFs differ. For the base path $G_0 = u - z - w - v$, this holds whenever $b_w \neq b_z$.

Proof. Choose rooted isomorphisms between the trees at u, z and between the trees at v, w. Together with φ on G_0, these isomorphisms and their inverses define an automorphism of $G - zw$. Thus Theorems 4.1 and 4.2 apply.

An added vertex contributes to $r_G(w; u, v)$ exactly when it is a new neighbor of w. It has no neighbors in G_0 other than w, so it is adjacent to neither endpoint. The same statement holds for the other term centered at w, and both terms centered at z gain b_z. The old vertices keep all their adjacencies. Each of the four counts is therefore its old value plus the corresponding b value, proving (10). On the base path all four old counts vanish, so $\delta_0 = 0$. □

For a fixed base graph, at most one integer value of $b_w - b_z$ makes (10) zero. Rooted stars realize every integer difference, so the construction always gives infinitely many pairs with equal CSF and different KSF.

5 The split-graph construction

Next we consider the construction of Aliste-Prieto, Crew, Spirkl, and Zamora [ACSZ21], which uses split graphs as in [LS19]. Given a simple graph L, form $\text{sp}(L)$ with vertex set $V(L) \sqcup E(L)$. Make $V(L)$ a clique, make $E(L)$ a stable set, and join each edge-vertex to its two endpoints. We call $V(L)$ the *core*. Figure 5 gives an example.

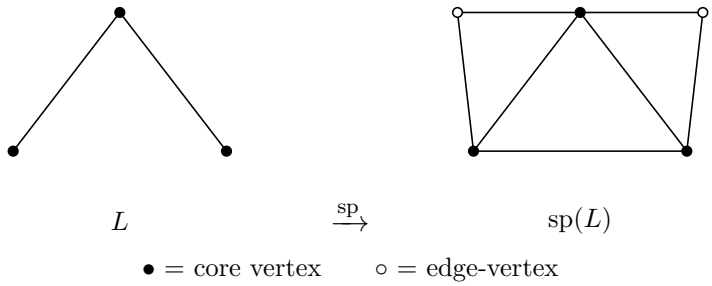


Figure 5: A graph L and its split graph $\text{sp}(L)$. The three core vertices form a clique; the two edge-vertices are each joined to their endpoints.

Theorem 5.1 (Aliste-Prieto–Crew–Spirkl–Zamora [ACSZ21, Lemma 8])**.** *Let ab and $a'b'$ be nonedges of G, each with distinct endpoints. Suppose an automorphism of G sends a to a', and an automorphism, possibly different, sends b to b'. Then*

$$X_{\mathrm{sp}(G+ab)} = X_{\mathrm{sp}(G+a'b')}. \tag{11}$$

The two pairs may share vertices. The two assumed automorphisms may be different, so there need not be a single automorphism carrying the nonedge ab to $a'b'$.

For a nonedge ab, define

$$W_G(a,b) = \sum_{w \in N_G(a) \cap N_G(b)} \big(d_G(w) + d_G(a) + d_G(b) - 4\big). \tag{12}$$

We use a seven-vertex polynomial whose two realizing graphs have placements that detect common neighbors of the added endpoints. Let $h(L)$ count the induced subsets of L with independence polynomial

$$f(x) = 1 + 7x + 15x^2 + 16x^3 + 9x^4 + 2x^5. \tag{13}$$

Theorem 5.2. *Under the hypotheses of Theorem 5.1, if $|V(G)| \geq 6$, then*

$$h(\mathrm{sp}(G+ab)) - h(\mathrm{sp}(G+a'b')) = W_G(a,b) - W_G(a',b'). \tag{14}$$

In particular, a nonzero right-hand side shows that the two split graphs have different KSFs.

We prove (14) in two steps: classify the graphs realizing f, then count their induced copies that use the new edge-vertex ab. The same count for $a'b'$ gives the required difference.

5.1 The two seven-vertex graphs

Let R be the tree with edges $15, 25, 35, 45, 56, 67$ (writing ij for the edge $\{i,j\}$). Let T have edges $25, 35, 45, 56, 57, 67$, with vertex 1 isolated. They are shown in Figure 6; the labels agree with the subsequent placement diagrams.

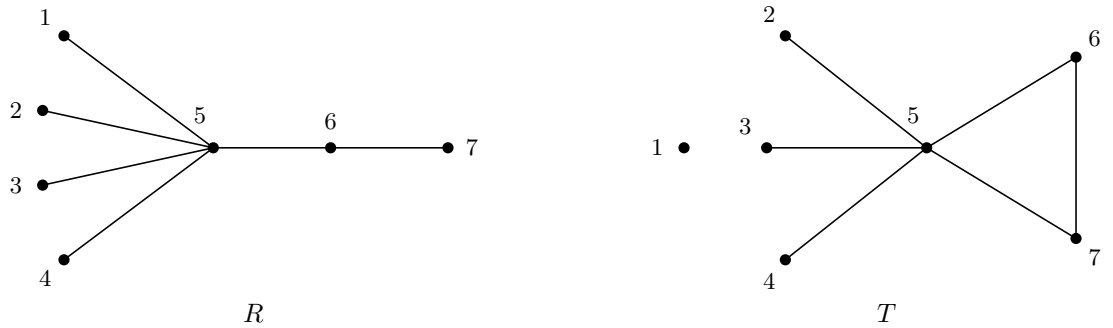


Figure 6: The two realizations of (13), denoted R and T. Their vertex labels will be used in the placement counts.

Lemma 5.3. *The graphs with independence polynomial (13) are exactly R and T, up to isomorphism.*

Proof. The linear coefficient and degree of (13) force order seven and independence number five. Fix a maximum independent set, leaving vertices p, q outside it. The pair $\{p,q\}$ meets every edge.

Among the five independent vertices, let z, r, s, t count those adjacent to neither cover vertex, only p, only q, or both. Interchanging p, q if necessary, assume $r \le s$. Let $\epsilon = \mathbf{1}_{pq \in E}$. The independence polynomial is

$$(1+x)^5 + x(1+x)^{z+r} + x(1+x)^{z+s} + (1-\epsilon)x^2(1+x)^z. \tag{15}$$

The four terms correspond to choosing neither cover vertex, only q, only p, or both. In particular, choosing p leaves the $z+s$ nonneighbors available. Let $\alpha_0 = z+r$ and $\beta_0 = z+s$, so $0 \le \alpha_0 \le \beta_0$ and $z+r+s+t = 5$.

If $\epsilon = 1$, absence of an x^6 term gives $\beta_0 \le 4$. The x^5 coefficient then forces $\beta_0 = 4$ and $\alpha_0 \le 3$. The x^4 coefficient is $5 + 4 + \binom{\alpha_0}{3} = 9$, so $\alpha_0 \le 2$. The x^3 coefficient is $10 + 6 + \binom{\alpha_0}{2} = 16$, so $\alpha_0 \le 1$. Finally, the x^2 coefficient gives $10 + 4 + \alpha_0 = 15$, hence $\alpha_0 = 1$. Since $0 \le z \le \alpha_0$, the possibilities are

$$(\epsilon, z, r, s, t) = (1, 0, 1, 4, 0), \qquad (1, 1, 0, 3, 1).$$

They represent R and T, respectively.

If $\epsilon = 0$, the x^2 coefficient gives $\alpha_0 + \beta_0 = 4$. The ordered possibilities are $(0,4), (1,3), (2,2)$. The coefficient of x^5 must satisfy

$$1 + \binom{\alpha_0}{4} + \binom{\beta_0}{4} + \binom{z}{3} = 2.$$

Since $z \le \alpha_0$, only $(\alpha_0, \beta_0) = (0,4)$ works; then $z = 0$. This gives $(\epsilon, z, r, s, t) = (0, 0, 0, 4, 1)$, another way to choose p, q in R. Substituting each of these three possibilities into (15) gives (13). The sizes of the four adjacency classes, together with the presence or absence of pq, specify the graph up to isomorphism. The classification is therefore complete. □

5.2 Copies of the tree R

Throughout the counting arguments below, G is the original graph used to build the split graph, not the split graph itself. Write $n = |V(G)|$, $m = e(G)$, $d_x = d_G(x)$, and $\epsilon_{pq} = \mathbf{1}_{pq \in E(G)}$. All degrees and neighborhoods refer to this G, before adding the chosen nonedge. The assumption $n \ge 6$ keeps the upper arguments $n-6$ in the counting formulas nonnegative. A binomial coefficient with nonnegative upper argument smaller than its lower argument is zero.

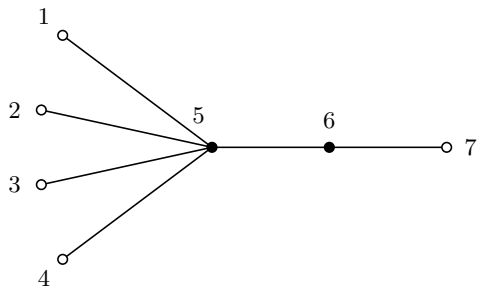
Copies in $\mathrm{sp}(G+ab)$ avoiding the new edge-vertex ab already occur in $\mathrm{sp}(G)$, and similarly for $a'b'$. It is enough to compare copies using the new edge-vertex.

Lemma 5.4. *The number of new induced copies of R in $\mathrm{sp}(G+ab)$ is*

$$\sum_{p \in \{a,b\}} \sum_{q \notin \{a,b\}} \left[\binom{d_p - \epsilon_{pq}}{3} (d_q - \epsilon_{pq}) + \binom{d_q - \epsilon_{pq}}{4} \right]. \tag{16}$$

This number is unchanged by replacing (a,b) with (a',b') under the automorphism assumptions of Theorem 5.1.

Proof. In a copy of R, vertex 5 has degree five and must belong to the core, since an edge-vertex has degree at most two. Vertex 6 is also in the core. Otherwise vertex 7 would be in the core and hence adjacent to vertex 5, contradicting the assumption that the copy is induced. Each remaining vertex misses at least one of these two core vertices and hence cannot belong to the core. This forces the placement shown in Figure 7.



R

$\bullet$ = core vertex $\quad$ $\circ$ = edge-vertex

Vertex	(a)	(b)	(c)	(d)
1	uv	$u'v'$	uv	$u'v'$
2	uw	$u'\varphi_u(w)$	uw	$u'\varphi_u(w)$
3	ux	$u'\varphi_u(x)$	ux	$u'\varphi_u(x)$
4	uy	$u'\varphi_u(y)$	uy	$u'\varphi_u(y)$
5	u	u'	u	u'
6	z	$\varphi_u(z)$	$\varphi_u^{-1}(v')$	v
7	zt	$\varphi_u(z)\varphi_u(t)$	$\varphi_u^{-1}(v')s$	vt

Figure 7: Copies of R in which the new edge-vertex is a leaf at vertex 5 (shown as vertex 1). Columns (a), (b) and (c), (d) are the two paired choices for the added edges. The labels u, v, u', v' mean a, b, a', b' in the text; other letters are local to these diagrams. Filled nodes are core vertices, and open nodes are edge-vertices.

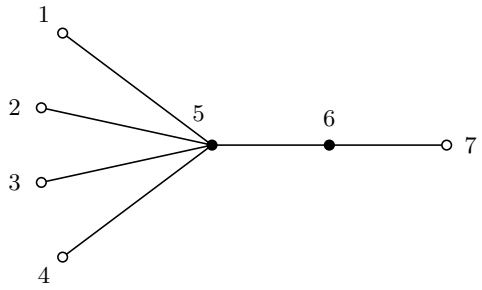
The new edge-vertex has two possible roles. If it is one of the four leaves at vertex 5, then that core vertex is an endpoint $p \in \{a, b\}$. The other core vertex is $q \notin \{a, b\}$. Choose three old edges incident to p but not q, and one old edge incident to q but not p. There are $\binom{d_p - \epsilon_{pq}}{3}(d_q - \epsilon_{pq})$ choices. The four leaves form an unordered set, so no extra ordering factor occurs.

If the new edge-vertex is vertex 7, its neighboring core vertex is an endpoint p, while vertex 5 is a vertex q different from both endpoints. Choose four old edges incident to q but not p, giving $\binom{d_q - \epsilon_{pq}}{4}$. These constructions give all copies uniquely and prove (16).

To show that the count is unchanged, extend the inner sum to all $q \neq p$. An automorphism sending p to another vertex preserves this extended sum, because it preserves adjacency and degrees. Its extra term is at the other endpoint $r \in \{a, b\} \setminus \{p\}$, for which $\epsilon_{pr} = 0$; that term is $\binom{d_p}{3}d_r + \binom{d_r}{4}$. It depends only on the two endpoint degrees. The assumed automorphisms preserve the two extended sums and the endpoint degrees. They therefore preserve the removed terms as well, which proves the claim. □

5.3 Copies of T with two core vertices

Vertex 5 of T has degree five, so it lies in the core. Since vertices 6 and 7 are adjacent, they cannot both be edge-vertices. Each of vertices 1–4 misses whichever of 6 and 7 lies in the core, forcing all four to be edge-vertices. This leaves two possibilities: the copy uses two core vertices or three. Figure 9 shows the first case.

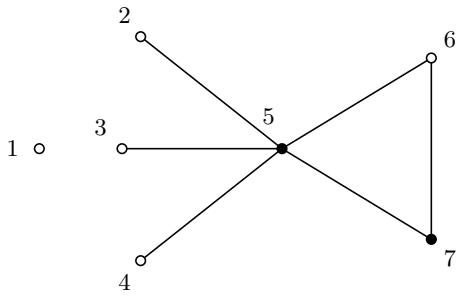


R

$\bullet$ = core vertex $\quad\circ$ = edge-vertex

Vertex	(a)	(b)	(c)	(d)
1	zw	$\varphi_u(z)\varphi_u(w)$	e_1	e'_1
2	zx	$\varphi_u(z)\varphi_u(x)$	e_2	e'_2
3	zy	$\varphi_u(z)\varphi_u(y)$	e_3	e'_3
4	zt	$\varphi_u(z)\varphi_u(t)$	e_4	e'_4
5	z	$\varphi_u(z)$	$\varphi_u^{-1}(v')$	v
6	u	u'	u	u'
7	uv	$u'v'$	uv	$u'v'$

Figure 8: Copies of R in which the new edge-vertex is vertex 7. Columns (a), (b) and (c), (d) are the paired choices. Together with Figure 7, these give all copies counted in (16). The endpoint notation and node symbols are the same as in Figure 7.



T

$\bullet$ = core vertex $\quad\circ$ = edge-vertex

Vertex	C_0	C_1	C_2
1	uv	e	e
2	wy	uv	ux
3	wz	uy	uy
4	wt	uz	uz
5	w	u	u
6	wx	ux	uv
7	x	x	v

Figure 9: A copy of T with two core vertices. The columns give the three placements: the new edge-vertex is isolated (C_0), a leaf at vertex 5 (C_1), or in the triangle (C_2). These cases give (17)–(19).

Lemma 5.5. *The number of new copies of T using two core vertices is the sum of*

$$C_0(a,b) = 4\sum_{w\neq a,b}\binom{d_w}{4} - \sum_{w\sim a}\binom{d_w-1}{3} - \sum_{w\sim b}\binom{d_w-1}{3}, \tag{17}$$

$$C_1(a,b) = \sum_{p\in\{a,b\}}\sum_{q\sim p}\binom{d_p-1}{2}(m-d_p-d_q+1), \tag{18}$$

$$C_2(a,b) = \left(\binom{d_a}{3}+\binom{d_b}{3}\right)(m-d_a-d_b). \tag{19}$$

Each expression is unchanged when we replace (a,b) with (a',b') under the assumptions of Theorem 5.1.

Proof. The new edge-vertex is either the isolated vertex, a leaf at vertex 5, or the edge-vertex in the triangle.

In the first case let w be vertex 5. Choose four edges incident to w, one to the other core vertex and three to the leaves. This initially gives $4\binom{d_w}{4}$ choices. We must have $w\neq a,b$, and the other core vertex must also avoid a,b. If $w\sim a$, there are $\binom{d_w-1}{3}$ choices whose other core vertex is a; subtract these, and similarly for b. The excluded cases are disjoint. This proves (17).

In the second case vertex 5 is an endpoint p of the new edge. Its other core vertex q must be a neighbor of p in G, because an old edge-vertex forms the third vertex of the triangle. Choose two additional leaf-edges at p other than pq. The isolated edge can be any old edge incident to neither p nor q, giving $m-d_p-d_q+1$ choices. The $+1$ restores the edge pq, subtracted twice. Summing over p,q proves (18).

In the third case the two core vertices are a,b. Choose the center of degree five, say a, then choose its three old leaf-edges. The isolated edge avoids both endpoints and has $m-d_a-d_b$ choices, since ab was a nonedge. Interchanging a,b gives (19).

An automorphism sending a to a' preserves d_a and every sum over its neighbors, and the same holds for b,b'. In (17), the sum over $w\neq a,b$ is the sum over all vertices with the two endpoint terms removed. Thus all terms in (17) and (18) are preserved. Formula (19) depends only on d_a,d_b,m, so it is preserved as well. □

5.4 Copies of T with three core vertices

With three core vertices, the entire triangle lies in the core. The new edge-vertex must occupy either the isolated position or a leaf position at vertex 5, as listed in Figure 10. Write

$$c(a,b) = |N_G(a)\cap N_G(b)|, \qquad t_\triangle(a) = n_{K_3}(G \text{ containing } a).$$

We abbreviate summation over common neighbors of the endpoints by $\sum_{w\sim a,b}$.

Lemma 5.6. *The number of new copies of T using three core vertices is $A(a,b)+B(a;b)+B(b;a)$, where*

$$\begin{aligned} A(a,b) = {} & \binom{n-6}{2}\sum_{w\neq a,b}\binom{d_w}{3} \\ & + (n-6)\sum_{w\sim a}\binom{d_w-1}{2} + (n-6)\sum_{w\sim b}\binom{d_w-1}{2} \\ & + \sum_{w\sim a,b}(d_w-2), \end{aligned} \tag{20}$$

Use the numbering of T in Figure 9, with core $\{5, 6, 7\}$.

Vertex	Isolated new edge-vertex	New leaf at vertex 5
1	uv	wx
2	wx	uv
3	wy	uy
4	wz	uz
5	w	u
6	s	s
7	t	t

Figure 10: The three-core-vertex placement of T. The common drawing and numbering are in Figure 9; here all three triangle vertices belong to the core. The two columns give the cases counted in (20) and (21). The labels u, v mean a, b in the formulas.

$$\begin{aligned} B(a;b) = &\binom{d_a}{2}(m-d_a)\binom{n-6}{2} \\ &+ (d_a-1)(n-6)\sum_{x\sim a}(d_x-1) + t_{\triangle}(a) \\ &+ \binom{d_a}{2}d_b(n-6) + (d_a-1)c(a,b). \end{aligned} \tag{21}$$

Proof. Suppose first that the isolated edge-vertex is ab. Choose the degree-five core vertex $w \neq a, b$ and its three leaf-edges wx, wy, wz. The remaining two core vertices must avoid all endpoints of the chosen edge-vertices and must differ from w. If a, b, w, x, y, z are distinct, there are $\binom{n-6}{2}$ choices. Summing this initial count gives the first term of (20).

Some of these six named vertices may coincide: one or both of a, b may occur among x, y, z. We call each such coincidence an identification. For $r \in \{0, 1, 2\}$ identifications, the correct number of choices is

$$\binom{n-6+r}{2} = \binom{n-6}{2} + r(n-6) + \binom{r}{2}. \tag{22}$$

Selecting a as a leaf neighbor requires $w \sim a$ and leaves $\binom{d_w-1}{2}$ choices for the other two leaf-edges; its linear correction is the second term. The same reasoning gives the third term for b. If both are selected, w is a common neighbor and the remaining leaf has $d_w - 2$ choices. The last term of (22) is then one, giving the final sum in (20).

Now let the new edge-vertex be a leaf and let its core endpoint be a. Choose two old leaf-edges ay, az, and choose an isolated old edge cd not incident to a. There are $\binom{d_a}{2}(m - d_a)$ choices before choosing the two other core vertices. The initially forbidden vertices are a, b, y, z, c, d, giving the first term of (21).

The only possible identifications are endpoints of cd with members of $\{b, y, z\}$. They are distinct within each group, and b cannot equal y or z because ab is a nonedge. Therefore there are at most two identifications and (22) applies again. Count each individual identification first:

- If an endpoint of cd equals a selected neighbor y of a, there are $d_y - 1$ choices for the isolated edge and $d_a - 1$ choices for the other leaf neighbor. Summing gives $(d_a - 1)\sum_{y\sim a}(d_y - 1)$ choices.
- If an endpoint equals b, there are $d_b\binom{d_a}{2}$ choices.

Multiplying these counts by $n-6$ gives the second and fourth terms of (21). A configuration with two identifications is counted twice here, once for each identification, as required by the linear term in (22).

When two identifications occur together, (22) requires one more choice beyond the linear correction. If the isolated edge joins the two chosen neighbors y, z, it completes a triangle through a; there are $t_{\triangle}(a)$ possibilities. If it joins b to one chosen neighbor y, then y is a common neighbor of a, b, and the other leaf neighbor has $d_a - 1$ choices. This gives $(d_a - 1)c(a, b)$. These are the only pairs of identifications, proving (21). The case where the core endpoint is b is the same, with the endpoints interchanged. □

Proof of Theorem 5.2. By Lemma 5.3, h is the sum of the induced-copy counts of R and T. Copies avoiding the new edge-vertex agree. Lemma 5.4 cancels the new copies of R, and Lemma 5.5 cancels the copies of T using two core vertices.

For copies using three core vertices, the assumed automorphisms preserve every term except the terms involving common neighbors. In particular, the global sum over $w \neq a, b$ differs from a graph-wide sum only by terms depending on d_a, d_b. The remaining contribution is

$$\sum_{w \sim a,b} (d_w - 2) + (d_a + d_b - 2)c(a, b) = W_G(a, b).$$

Subtracting the analogous expression for a', b' proves (14). The final assertion follows from Corollary 2.4. □

Corollary 5.7. *Assume the hypotheses of Theorem 5.1 and $n \geq 6$. Suppose a, b have a common neighbor while a', b' have none. If some common neighbor of a, b has degree at least three, or $d_a \geq 2$, or $d_b \geq 2$, then*

$$\overline{X}_{\mathrm{sp}(G+ab)} \neq \overline{X}_{\mathrm{sp}(G+a'b')}.$$

Proof. For each common neighbor w, we have $d_w \geq 2$ and $d_a, d_b \geq 1$, so every summand in $W_G(a, b)$ is nonnegative. The extra hypothesis makes at least one positive, whereas $W_G(a', b') = 0$. Apply Theorem 5.2. □

Common neighbors may occur for both endpoint pairs; Theorem 5.2 still separates the KSFs whenever their W_G values differ.

6 Information determined by the KSF

We now turn from comparing particular pairs to recovering graph data. Using the multiset in Theorem 2.3, we recover degrees and counts of small induced subgraphs, cliques, and matchings. These are the ingredients needed for cograph reconstruction in Section 8.

From this point onward all vertex weights are one. Write $n = |V(G)|$, $I_G(x) = \sum_{k \geq 0} i_k(G)x^k$, and

$$\mathcal{I}(G) = \{\{I_{G[S]}(x) : S \subseteq V(G)\}\}.$$

We can tell which entries of $\mathcal{I}(G)$ come from subsets of a given size because $[x]I_{G[S]} = |S|$. Recall also that

$$I_{G \sqcup H} = I_G I_H, \qquad I_{G \odot H} = I_G + I_H - 1. \tag{23}$$

The first identity follows by choosing independent sets separately in the two components. For the second, a nonempty independent set must stay on one side of the join; the subtraction avoids counting the empty set twice.

6.1 Vertex data

For $v \in V(G)$, let $N[v]$ consist of v and all its neighbors, and let $a_k(v)$ count independent k-sets containing v. A vertex is universal if it is adjacent to every other vertex.

Theorem 6.1. *The following data determine one another, and all are determined by $\overline{X}_G$:*

(i) I_G *together with* $\{\{I_{G-v} : v \in V(G)\}\}$*;*

(ii) $\{\{I_{G-N[v]} : v \in V(G)\}\}$*;*

(iii) $\{\{(a_1(v), a_2(v), \ldots) : v \in V(G)\}\}$*.*

In particular, the KSF determines the degree sequence and the number of universal vertices.

Proof. Separating independent sets according to whether they contain v gives

$$I_G - I_{G-v} = xI_{G-N[v]} = \sum_{k\geq 1} a_k(v)x^k. \tag{24}$$

This proves that (i) determines (ii) and (iii), and that (ii) and (iii) are equivalent. Conversely, counting an independent set once for each of its vertices gives

$$\sum_v I_{G-N[v]} = I'_G. \tag{25}$$

Thus (ii) gives the derivative of I_G. Integrating and using $I_G(0) = 1$ recovers I_G itself. Equation (24) then gives the multiset of vertex-deletion polynomials, proving (i).

To obtain (i) from the KSF, select the entries of $\mathcal{I}(G)$ whose coefficients of x are n and $n-1$. These entries correspond to the whole graph and to single-vertex deletions. The relation $a_2(v) = n - 1 - d_G(v)$ then recovers each degree, with multiplicity. Universal vertices are precisely those of degree $n-1$. □

6.2 Induced subgraphs on four vertices

Let $n_H(G)$ count induced copies of H, and let $\tau(G) = n_{K_3}(G)$. The paw is a triangle with one pendant vertex; the diamond is K_4 with one edge removed.

Theorem 6.2. *The KSF determines every induced-subgraph count on at most four vertices. Consequently, it determines whether a graph is a cograph, claw-free, threshold, or trivially perfect.*

Proof. Graphs of order at most three have distinct independence polynomials. For the eleven four-vertex graphs, direct counting gives the following complete list.

Graph	Independence polynomial
$4K_1$	$1 + 4x + 6x^2 + 4x^3 + x^4$
$K_2 \sqcup 2K_1$	$1 + 4x + 5x^2 + 2x^3$
$2K_2$	$1 + 4x + 4x^2$
$P_3 \sqcup K_1$	$1 + 4x + 4x^2 + x^3$
$K_{1,3}$	$1 + 4x + 3x^2 + x^3$
P_4, $K_3 \sqcup K_1$	$1 + 4x + 3x^2$
C_4, paw	$1 + 4x + 2x^2$
Diamond	$1 + 4x + x^2$
K_4	$1 + 4x$

Thus $\mathcal{I}(G)$ determines seven four-vertex counts individually and the sums for these two pairs. Let $s_3 = \sum_v \binom{d_G(v)}{3}$. Counting a center and three neighbors, and then a triangle and a fourth vertex, gives

$$s_3 = n_{K_{1,3}} + n_{\text{paw}} + 2n_{\text{diamond}} + 4n_{K_4},$$
$$(n-3)\tau = n_{K_3 \sqcup K_1} + n_{\text{paw}} + 2n_{\text{diamond}} + 4n_{K_4}.$$

In the first identity, a claw or paw has one possible center, a diamond has two, and K_4 has four. None of the remaining four-vertex types admits such a center. The second identity counts each selected four-set once for each triangle it contains. The first identity therefore recovers the paw and C_4 counts; the second recovers the remaining pair. For $n < 4$ all four-vertex counts vanish. Membership in the stated families can now be read from their forbidden induced subgraphs. Cographs have no induced P_4, and claw-free graphs have no induced $K_{1,3}$. Threshold graphs exclude $2K_2, C_4, P_4$, while trivially perfect graphs exclude C_4, P_4 [CLS81, Tsu18]. Each of these conditions can now be checked from the KSF. □

Corollary 6.3. *If G is a cograph and $\overline{X}_G = \overline{X}_H$, then H is a cograph. For $n \geq 4$, if m_4 is the multiplicity of $1 + 4x + 3x^2$ in $\mathcal{I}(G)$, then*

$$n_{P_4}(G) = m_4 - (n-3)\tau(G) + s_3(G) - n_{K_{1,3}}(G).$$

Proof. Subtract the two counting identities above and use $m_4 = n_{P_4} + n_{K_3 \sqcup K_1}$. A graph is a cograph exactly when this count is zero. □

6.3 The induced-edge polynomial

We can also count induced subgraphs by their numbers of vertices and edges. Define

$$P_G(u, v) = \sum_{S \subseteq V(G)} u^{|S|} v^{e(G[S])}. \tag{26}$$

This is a specialization of Kotek's three-variable Ising polynomial

$$Z_G(x, y, z) = \sum_{S \subseteq V(G)} x^{e(G[S])} y^{|S|} z^{e(G[V(G) \setminus S])},$$

namely $P_G(u, v) = Z_G(v, u, 1)$ [Kot12]. Kotek studies what graph information can be found from such polynomials, including a relation between matchings and a different bivariate specialization. Tittmann, Averbouch, and Makowsky [TAM11] also define an induced-subgraph polynomial, but theirs records vertices and connected components, whereas P_G counts vertices and edges. We give direct proofs of the properties of P_G that we will need.

Theorem 6.4. *The KSF determines P_G. The polynomial P_G determines the order, size, degree sequence, independence polynomial, and numbers of cliques and matchings of every size. For all graphs G, H,*

$$P_G = P_H \iff P_{\overline{G}} = P_{\overline{H}}.$$

If $\alpha(G), \alpha(H) \leq 2$, then we also have

$$\overline{X}_G = \overline{X}_H \iff P_G = P_H. \tag{27}$$

Proof. Each entry $I_{G[S]}$ gives $|S| = [x]I_{G[S]}$ and $e(G[S]) = \binom{|S|}{2} - [x^2]I_{G[S]}$, so $\mathcal{I}(G)$ gives P_G. The top u-term is $u^n v^m$, and the coefficient of $u^{n-1}v^e$ counts vertices of degree $m-e$. Also $I_G(u) = P_G(u,0)$, and the number of k-cliques is $[u^k v^{\binom{k}{2}}]P_G$.

To count matchings, we first count edge subsets by their number of edges and the number of vertices they use. Define

$$E_G(s,t) = \sum_{F \subseteq E(G)} s^{|V(F)|} t^{|F|},$$

where $V(F)$ is the set of endpoints of the edges in F. Write $v = 1 + (v-1)$ in each term of P_G. Expanding then amounts to choosing a subset F of the edges in $G[S]$. For a fixed F, the set S must contain $V(F)$ and may contain any of the other vertices. Its total contribution is

$$u^{|V(F)|}(1+u)^{n-|V(F)|}(v-1)^{|F|}.$$

Summing over F gives the first identity below; solving for E_G gives the second:

$$\begin{aligned} P_G(u,v) &= (1+u)^n E_G\left(\frac{u}{1+u}, v-1\right), \\ E_G(s,t) &= (1-s)^n P_G\left(\frac{s}{1-s}, 1+t\right). \end{aligned} \tag{28}$$

Thus P_G determines E_G. A set of q edges uses $2q$ vertices exactly when no two edges share an endpoint. Hence $[s^{2q}t^q]E_G$ counts the matchings of size q.

If $\alpha(G) \le 2$, every induced independence polynomial is $1 + |S|x + (\binom{|S|}{2} - e(G[S]))x^2$. Hence P_G recovers $\mathcal{I}(G)$ in this case. Finally, complementation replaces each monomial $u^k v^e$ by $u^k v^{\binom{k}{2}-e}$, an involution, proving the last equivalence. □

6.4 Edge-subset counts and edge covers

Extracting coefficients from the change of variables in Theorem 6.4 gives explicit conversion formulas between the two polynomials. Edge-cover counts will follow from the same formulas.

Corollary 6.5. *Write $p_{k,e} = [u^k v^e]P_G$ and $\eta_{q,r} = [s^r t^q]E_G$. With binomial coefficients outside their usual range taken to be zero,*

$$\eta_{q,r} = \sum_{k=0}^{r} (-1)^{r-k} \binom{n-k}{r-k} \sum_{e \ge q} \binom{e}{q} p_{k,e}, \tag{29}$$

$$p_{k,e} = \sum_{q \ge e} (-1)^{q-e} \binom{q}{e} \sum_{r} \binom{n-r}{k-r} \eta_{q,r}. \tag{30}$$

Both sums are finite.

Proof. In (28), a term $p_{k,e}u^k v^e$ becomes

$$p_{k,e} s^k (1-s)^{n-k} (1+t)^e.$$

Taking the coefficient of $s^r t^q$ gives (29). Similarly, in (28), a term of E_G becomes

$$\eta_{q,r} u^r (1+u)^{n-r} (v-1)^q.$$

Taking the coefficient of $u^k v^e$ gives (30). □

For example, the first few coefficients count familiar edge sets:

$$\eta_{0,0} = 1, \qquad \eta_{1,2} = m,$$
$$\eta_{2,3} = \sum_v \binom{d_G(v)}{2}, \qquad \eta_{2,4} = \binom{m}{2} - \sum_v \binom{d_G(v)}{2},$$
$$\eta_{3,3} = \tau(G).$$

The two-edge formulas separate pairs meeting at a vertex from pairs with disjoint endpoints. Three edges on three vertices form a triangle. These counts concern selected edges: other edges between their endpoints may be omitted.

The same polynomial also counts edge covers.

Corollary 6.6. *The KSF determines the number of edge covers of each size, and whether the graph has an isolated vertex.*

Proof. An edge cover is an edge set F with $V(F) = V(G)$. The coefficient of t^q in the polynomial below counts the edge covers with q edges:

$$[s^n]E_G(s,t) = \sum_q \eta_{q,n} t^q.$$

An isolated vertex cannot be covered by any edge. If no vertex is isolated, the full edge set is already a cover. For the empty graph, the empty edge set gives its sole cover. Theorem 6.4 supplies all the required coefficients. □

6.5 Reconstruction of basic graph families

Proposition 6.7. *Threshold graphs, disjoint unions of cliques, and complete multipartite graphs are determined by P_G, and hence by $\overline{X}_G$, against arbitrary graphs.*

Proof. For threshold graphs, the degree sequence suffices. Let a nonempty threshold graph G and an arbitrary graph H have matching degree sequences. A threshold graph has an isolated or universal vertex; the corresponding degree forces a vertex of the same kind in H. Remove one such vertex from each graph. If the deleted vertices are universal, decrease every remaining degree by one; if they are isolated, the remaining degrees do not change. The degree sequences still match after this deletion, and the remaining part of G is threshold. Induction identifies the smaller graphs. Restoring the deleted vertices then gives $G \cong H$.

A graph is a union of cliques exactly when it has no induced P_3. The number of induced copies of P_3 is $\sum_v \binom{d_G(v)}{2} - 3\tau(G)$, which is determined by P_G. Thus P_G tells us whether the graph is a union of cliques. Each K_a component contributes exactly a vertices of degree $a-1$, so these degree multiplicities recover all component sizes. Complementation gives the complete multipartite case. □

6.6 The limits of vertex-deletion data

Theorem 6.1 gives the independence polynomial of G and the multiset obtained by deleting one vertex at a time. These data do not always distinguish cographs. The following two eleven-vertex cographs have the same data, but different numbers of five-vertex cliques, so their KSFs are different.

Example 6.8. Let $L = 2K_1$ and define $G = G_1 \odot G_2$, $H = H_1 \odot H_2$, where

$$G_1 = K_1 \sqcup \big(K_1 \odot (K_1 \sqcup K_2)\big), \quad G_2 = K_1 \sqcup (K_1 \odot L \odot L),$$
$$H_1 = K_1 \sqcup (L \odot L), \qquad\qquad H_2 = K_1 \sqcup \big(K_2 \odot (K_1 \sqcup K_2)\big).$$

Both have independence polynomial $1 + 11x + 13x^2 + 4x^3$. In their common deletion multiset, the pairs (a, b) in $1 + 10x + ax^2 + bx^3$ are

$$(8, 2),\ (9, 2),\ (10, 2),\ (11, 3)^{\times 6},\ (12, 4)^{\times 2}.$$

These follow from (23), replacing the factor containing the deleted vertex. The clique polynomial of a graph has coefficient of t^k equal to its number of k-cliques, including the empty clique for $k = 0$. It multiplies under join, so the clique polynomials of G and H are, respectively,

$$(1 + 5t + 4t^2 + t^3)(1 + 6t + 8t^2 + 4t^3),$$
$$(1 + 5t + 4t^2)(1 + 6t + 8t^2 + 5t^3 + t^4).$$

The coefficients of t^5 are 24 and 25. Thus $G \ncong H$ and $P_G \neq P_H$, although all their local data agree.

7 Reconstructing connected components

The KSF also recovers the multiset of component KSFs. To prove this, we encode $\mathcal{I}(G)$ in two polynomial forms: one turns joins into products, and the other turns disjoint unions into products.

For the first polynomial, replace each independence polynomial by a monomial whose exponents are its coefficients. Define

$$J_G(\mathbf{z}) = \sum_{S \subseteq V(G)} \prod_{k \geq 1} z_k^{i_k(G[S])}. \tag{31}$$

For example, an induced subgraph with independence polynomial $1 + 4x + 3x^2$ contributes $z_1^4 z_2^3$ to J_G. If several subsets have this same polynomial, their contributions add.

For the second encoding, factor each independence polynomial over $\mathbb{Q}$ into irreducible polynomials, which admit no further factorization into nonconstant polynomials. Normalize each factor to have constant term one. With this normalization, only the order of the factors remains arbitrary.

Denote the set of normalized irreducible polynomials by $\mathcal{P}$. Write $I_H = \prod_{f \in \mathcal{P}} f^{\nu_f(H)}$, where $\nu_f(H)$ counts how many times f occurs, and define

$$U_G(\mathbf{Y}) = \sum_{S \subseteq V(G)} \prod_{f \in \mathcal{P}} Y_f^{\nu_f(G[S])}. \tag{32}$$

Thus the factorization $1 + 4x + 3x^2 = (1 + x)(1 + 3x)$ contributes $Y_{1+x} Y_{1+3x}$ to U_G. Only finitely many variables occur for a fixed graph.

Proposition 7.1. *Equality of any one of $\overline{X}_G$, $\mathcal{I}(G)$, J_G, or U_G is equivalent to equality of the others. Moreover,*

$$J_{G \odot H} = J_G J_H, \qquad U_{G \sqcup H} = U_G U_H. \tag{33}$$

In particular, if the same graph is joined to both sides of a KSF equality, it can be canceled. The same is true for a common disjoint-union factor.

Proof. The exponents of a monomial in J_G give the coefficients of an independence polynomial. For U_G, the variable labels and exponents give its irreducible factorization instead. In either encoding, the monomial coefficient is the multiplicity of that independence polynomial. Both therefore recover $\mathcal{I}(G)$, proving the first assertion by Theorem 2.3.

For joins, the coefficients of $x, x^2, \dots$ in the independence polynomials add. This follows from $I_{G\odot H} = I_G + I_H - 1$ and gives $J_{G\odot H} = J_G J_H$. Under disjoint union, multiplication of independence polynomials adds the multiplicities of their irreducible factors. The monomials in U_G consequently multiply as required. Finally, these are ordinary polynomial rings, where a common nonzero factor can be canceled. □

7.1 Private factors of independence polynomials

Factoring U_G will recover the component KSFs once we know that connected graphs give irreducible polynomials. The key is a factor of I_G that occurs in none of the independence polynomials of proper induced subgraphs. We locate it through the first positive root of $I_G(-t)$, using the root comparison below; see also Csikvári [Csi13]. Write

$$Q_G(t) = I_G(-t), \qquad \beta(G) = \min\{t > 0 : Q_G(t) = 0\},$$

and let $\beta(\varnothing) = +\infty$. The existence of $\beta(G)$ for every nonempty graph follows from the next lemma.

Lemma 7.2 (The first positive root)**.** *Every nonempty graph G has $0 < \beta(G) \le 1$. If G is connected, then*

(i) $\beta(G) < \beta(H)$ for every proper induced subgraph H of G;

(ii) $Q'_G(\beta(G)) < 0$, so $\beta(G)$ is a simple root.

For a disjoint union, $\beta(G \sqcup H) = \min\{\beta(G), \beta(H)\}$.

Proof. We use induction on the number of vertices. For K_1, we have $Q_{K_1}(t) = 1 - t$, so $\beta(K_1) = 1$. If the graph is disconnected, Q_G factors over its components. It vanishes for the first time when one of those component polynomials does, so the disconnected case follows by induction.

Let G be connected of order at least two. Choose v so that $L = G - v$ is connected, and let $M = G - N[v]$. We may choose a leaf of a spanning tree, since deleting it leaves a spanning tree on the remaining vertices. Since v has a neighbor, M is a proper induced subgraph of L. By induction, $Q_M(\beta(L)) > 0$. Therefore

$$Q_G(\beta(L)) = Q_L(\beta(L)) - \beta(L) Q_M(\beta(L)) < 0.$$

The value $Q_G(0) = 1$ has the opposite sign, so continuity forces a zero in $(0, \beta(L))$. Hence the first positive zero of Q_G precedes $\beta(L)$.

Now take any nonempty proper induced subgraph H and a component C attaining $\beta(C) = \beta(H)$. Grow C inside G, adding one adjacent vertex at a time, until it lies in a connected induced subgraph L of order $|V(G)| - 1$. The preceding sign argument applies to this L, and induction gives $\beta(L) \le \beta(C)$. Thus $\beta(G) < \beta(H)$.

Finally, (25) gives

$$Q'_G(\beta(G)) = -\sum_{v \in V(G)} Q_{G-N[v]}(\beta(G)) < 0,$$

since every summand is positive by the strict inequality just proved. The bound $\beta(G) \le 1$ follows from the order-one case and the same inequalities. □

Definition 7.3. An irreducible factor $f \in \mathcal{P}$ of I_G is *private* to G if it divides no independence polynomial $I_{G[S]}$ with $S \subsetneq V(G)$.

This condition concerns only proper induced subgraphs of G. The same factor may occur in the independence polynomials of other graphs.

Proposition 7.4. *Every nonempty connected graph has a private irreducible factor of multiplicity one in its independence polynomial.*

Proof. Let f be the normalized irreducible factor of I_G having $-\beta(G)$ as a root. Simplicity in Lemma 7.2 shows that f occurs once. If f divided $I_{G[S]}$ for a proper subset S, then $I_{G[S]}(-\beta(G)) = 0$, contradicting $\beta(G[S]) > \beta(G)$. □

7.2 Irreducibility and component reconstruction

Lemma 7.5. *Let K be a field, let M be a monomial in variables $\mathbf{Y}$, and let $B \in K[\mathbf{Y}]$ have constant term one. Then*

$$MZ + B$$

is irreducible in $K[\mathbf{Y}, Z]$.

Proof. Since B has constant term one, no variable dividing M can divide B. Thus M and B have no common nonconstant factor.

Suppose $MZ + B$ factors into two polynomials. Its degree in Z is one, so one factor must be independent of Z. Comparing coefficients in Z shows that this factor divides both M and B. Their coprimeness forces it to be constant, proving irreducibility. □

Theorem 7.6. *For every nonempty graph G,*

$$G \text{ is connected} \iff U_G \text{ is irreducible over } \mathbb{Q}.$$

For connected G, the same result holds over any field, provided that we keep the variables Y_f defined from the rational factors in (32).

Proof. Suppose G is connected, and choose the factor f from Proposition 7.4. The variable Y_f occurs only in the monomial contributed by $S = V(G)$, and its exponent there is one. Consequently

$$U_G = M_G Y_f + B_G, \tag{34}$$

where

$$M_G = \prod_{g \neq f} Y_g^{\nu_g(G)}, \qquad B_G = \sum_{S \subsetneq V(G)} \prod_{g} Y_g^{\nu_g(G[S])}.$$

The polynomial M_G is a monomial with coefficient one. The polynomial B_G has constant term one, coming from the empty subset. Lemma 7.5 therefore proves that U_G is irreducible. The same proof works over any field: the two coefficients equal to one remain nonzero, and the variables keep their original labels.

For disconnected G, split its components into two nonempty groups. The identity $U_{G_1 \sqcup G_2} = U_{G_1} U_{G_2}$ then gives a nontrivial factorization of U_G. □

Example 7.7. Let $G = K_1 \odot (K_2 \sqcup K_2)$. Its independence polynomial is reducible:

$$I_G(x) = 1 + 5x + 4x^2 = (1 + x)(1 + 4x).$$

This example shows why we do not need I_G itself to be irreducible. Write

$$a = Y_{1+x}, \quad b = Y_{1+2x}, \quad c = Y_{1+3x}, \quad d = Y_{1+4x},$$

and $q = Y_{1+3x+x^2}$, $r = Y_{1+4x+2x^2}$. Counting the induced subsets gives

$$U_G = 1 + 5a + 6b + 4a^2 + 4ab + b^2 + 2c + 4q + 4r + ad. \tag{35}$$

The factor $1 + 4x$ is private, so the variable d appears only in the term ad. The remaining terms have constant term one. The form required by Lemma 7.5 is therefore present, proving irreducibility of U_G despite the factorization of I_G.

By Theorem 7.6, factoring U_G gives one irreducible factor for each connected component. Since each of these factors determines the KSF of its component, we obtain the following result.

Theorem 7.8 (Component reconstruction)**.** *If $G = \bigsqcup_{i=1}^{r} C_i$ and $H = \bigsqcup_{j=1}^{s} D_j$ are connected-component decompositions, then $\overline{X}_G = \overline{X}_H$ if and only if $r = s$ and, after reordering, $\overline{X}_{C_i} = \overline{X}_{D_i}$ for every i.*

Proof. If $\overline{X}_G = \overline{X}_H$, then Proposition 7.1 gives

$$\prod_i U_{C_i} = \prod_j U_{D_j}.$$

Every displayed factor is irreducible by Theorem 7.6. Unique factorization therefore matches the factors on the two sides, apart from possible constant multiples. Comparing constant terms removes this ambiguity, since every factor has constant term one. It follows that $r = s$ and, after reordering, $U_{C_i} = U_{D_i}$. Proposition 7.1 then gives $\overline{X}_{C_i} = \overline{X}_{D_i}$.

Conversely, if the component KSFs match, multiply them using the disjoint-union identity to obtain $\overline{X}_G = \overline{X}_H$. □

Corollary 7.9. *The KSF determines the order, size, degree sequence, independence polynomial, and induced-edge polynomial of each connected component, grouped component by component. If every member of a class of connected graphs is determined by its KSF among all graphs, then the same is true of every finite disjoint union of members of that class. In particular, Conjecture 1.1 is equivalent to its restriction to connected cographs.*

Proof. Apply Theorems 6.1 and 6.4 to the recovered component KSFs. For reconstruction, match components by Theorem 7.8 and apply the assumed connected case. □

8 Reconstructing cographs

By Theorem 7.8, to distinguish cographs it is enough to distinguish the connected ones. Every connected cograph with at least two vertices is a join of smaller cographs. Since $J_{G \odot H} = J_G J_H$, we would like to determine these smaller graphs from a product of their polynomials.

The *canonical cotree* records how a cograph is built. A leaf represents a vertex; an internal node combines its children by union or join. Every internal node has at least two children, and successive internal nodes on a path use opposite operations. At a union node the children are connected components, while at a join node they correspond to components of the complement.

These decompositions make the cotree unique up to reordering children. In particular, the factors in the maximal join decomposition are single vertices or disconnected cographs.

We begin with join factors that are disjoint unions of cliques. Their Newton polytopes will show that nonisomorphic factors give relatively prime polynomials. This lets us recover the graph factors from the product, whether or not the individual polynomials are irreducible.

8.1 Newton polytopes of clique unions

To compare these polynomials, it helps to look at their exponents as points. For example, the monomial $z_1^3 z_2^2$ gives the point $(3, 2)$. For a nonzero polynomial A in r variables, the convex hull of all these points in $\mathbb{R}^r$ is called its *Newton polytope*, written $\mathrm{Newt}(A)$.

We will also use the sum of two polytopes, defined by adding every point of one to every point of the other:

$$P + Q = \{p + q : p \in P,\ q \in Q\}.$$

This operation is the Minkowski sum. Its role in polynomial factorization comes from the following classical identity; see Gao and Lauder [GL01]:

$$\mathrm{Newt}(AB) = \mathrm{Newt}(A) + \mathrm{Newt}(B). \tag{36}$$

To see why, choose any linear functional ℓ on the exponent vectors. In the product AB, the terms where ℓ is largest come from multiplying the terms where it is largest in A and in B. This product is nonzero over any field. Thus the largest value of ℓ is the same on both sides of (36). Since this holds for every ℓ, the two polytopes are equal.

Let $e_k(t_1, \ldots, t_r)$ denote the elementary symmetric polynomial of degree k, and define

$$\Phi_r(t_1, \ldots, t_r) = (e_1(\mathbf{t}), e_2(\mathbf{t}), \ldots, e_r(\mathbf{t})).$$

For a union of r cliques, the Newton polytope will be a simplex of dimension r. Such a polytope has $r + 1$ affinely independent vertices, giving a triangle in dimension two and a tetrahedron in dimension three.

We now describe the relevant exponent vectors. For positive real numbers $a_1 \le \cdots \le a_r$, let $w_0 = 0$. Let w_j be the list of coefficients of $x, x^2, \ldots, x^r$ in

$$\prod_{\ell=r-j+1}^{r} (1 + a_\ell x), \qquad 1 \le j \le r. \tag{37}$$

Lemma 8.1. *For $0 \le t_i \le a_i$, the vector $\Phi_r(\mathbf{t})$ lies in*

$$\Delta_{\mathbf{a}} = \mathrm{conv}\{w_0, w_1, \ldots, w_r\}.$$

The displayed vertices are affinely independent.

Proof. We use induction on r. For $r = 1$, the claim simply says that t_1 lies in the interval from 0 to a_1.

For larger r, first hold all but one coordinate fixed. Each elementary symmetric polynomial has degree at most one in the remaining coordinate, so its value between 0 and a_i is a weighted average of its values at the two endpoints. Repeating this for every coordinate writes $\Phi_r(\mathbf{t})$ as a convex combination of the values at the corners of the box. More explicitly, if $q_i = t_i/a_i$, then

$$\Phi_r(\mathbf{t}) = \sum_{S \subseteq [r]} \left(\prod_{i \in S} q_i \prod_{i \notin S} (1 - q_i) \right) \Phi_r(a_i \mathbf{1}_{i \in S} : i \in [r]).$$

The coefficients are nonnegative and sum to one.

We therefore only need to check the corners. The corner $(a_1, \ldots, a_r)$ gives w_r, and the all-zero corner gives w_0. Every other corner selects k of the sizes, where $0 < k < r$. Write the selected sizes as $b_1 \leq \cdots \leq b_k$. Since these are selected from the ordered list $a_1, \ldots, a_r$, we have $b_i \leq a_{r-k+i}$. Apply the induction hypothesis to the k-dimensional box with upper bounds $a_{r-k+1}, \ldots, a_r$. The induction hypothesis puts this corner in the convex hull of $w_0, \ldots, w_k$, with zeros in the coordinates after the kth. This is contained in $\Delta_{\mathbf{a}}$, as required.

Finally, the jth coordinate of w_j is the positive product of its j selected sizes, whereas all coordinates numbered greater than j vanish. The matrix with columns $w_1, \ldots, w_r$ is triangular with nonzero diagonal. These vectors are linearly independent. □

Theorem 8.2 (Clique-union simplex)**.** *If $D = K_{a_1} \sqcup \cdots \sqcup K_{a_r}$, where $1 \leq a_1 \leq \cdots \leq a_r$, then*

$$J_D(\mathbf{z}) = \sum_{0 \leq t_i \leq a_i} \left(\prod_{i=1}^{r} \binom{a_i}{t_i} \right) \prod_{k=1}^{r} z_k^{e_k(t_1, \ldots, t_r)}, \qquad \operatorname{Newt}(J_D) = \Delta_{\mathbf{a}}. \tag{38}$$

In particular, this Newton polytope is an r-dimensional simplex.

Proof. An induced subset meeting the ith clique in t_i vertices has independence polynomial $\prod_i (1 + t_i x)$, and there are $\prod_i \binom{a_i}{t_i}$ such subsets. This proves the formula. Every coefficient is positive, so every listed exponent vector occurs. Lemma 8.1 puts all these vectors in $\Delta_{\mathbf{a}}$. Conversely, each w_j occurs by choosing all vertices from the j largest cliques. Their convex hull is therefore exactly $\Delta_{\mathbf{a}}$. Affine independence in Lemma 8.1 establishes the dimension. □

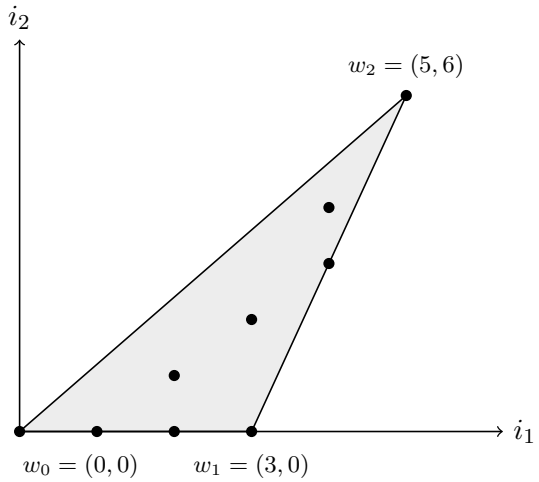


Figure 11: The Newton polytope of $J_{K_2 \sqcup K_3}$. The black points are the distinct vectors $(t_1 + t_2, t_1 t_2)$ for $0 \leq t_1 \leq 2$ and $0 \leq t_2 \leq 3$. Their convex hull is the triangle $\Delta_{(2,3)}$.

Figure 11 shows the Newton polytope for $K_2 \sqcup K_3$. The fact that these polytopes are simplices is useful because a simplex has very few possible Minkowski summands. As the next lemma shows, each summand is either a point or a scaled and translated copy of the simplex.

Lemma 8.3 (Summands of a simplex)**.** *Let Δ be an r-dimensional simplex with $r \geq 2$. If $P + Q = \Delta$, then there are a vector t and a number $0 \leq \lambda \leq 1$ such that*

$$P = t + \lambda\Delta, \qquad Q = -t + (1 - \lambda)\Delta.$$

In particular, each of P and Q is either a point or has the same dimension as Δ.

Proof. Write the vertices of Δ as $p_0, \dots, p_r$. Choose a linear functional whose maximum on Δ occurs only at p_i. Its maximum must also occur at a single point a_i of P and a single point b_i of Q; any additional maximizing point would produce another maximum in Δ. Hence $p_i = a_i + b_i$. This is the only way to write p_i as a sum of a point of P and a point of Q, since both points must maximize the chosen functional.

Next expose the edge $[p_i, p_j]$ of Δ with a linear functional. The maximizing parts of P and Q add to this edge. Each must therefore be a parallel segment or a point. Comparing their endpoints gives

$$a_j - a_i = \lambda_{ij}(p_j - p_i), \qquad 0 \le \lambda_{ij} \le 1.$$

On a triangular face with vertices p_i, p_j, p_k, add the three edge differences. The vectors $p_j - p_i$ and $p_k - p_i$ are independent, so the resulting equation forces $\lambda_{ij} = \lambda_{ik} = \lambda_{jk}$. Edges sharing a vertex lie in a common triangular face. If two edges have no common vertex, an edge joining one endpoint of each links them through two triangular faces. Hence there is a common value λ, and $a_i = t + \lambda p_i$ for every i.

To finish, we check that the list a_i includes every vertex a of P. Choose a linear functional whose maximum on P occurs only at a. There is a positive gap between its value at a and at every other vertex of P. This gap remains positive after a sufficiently small change in the functional. We can avoid the finitely many hyperplanes where two vertices of Δ have equal values, so its maximum on Δ is also unique, say at p_i. The definition of a_i then gives $a = a_i$. Thus $P = \operatorname{conv}\{a_i\} = t + \lambda\Delta$. The same argument gives the formula for Q. □

8.2 Recovering the join factors

Lemma 8.4. *For $r, s \ge 2$, two clique-union simplices $\Delta_{\mathbf{a}}$ and $\Delta_{\mathbf{b}}$ are positive scalar multiples of each other if and only if $r = s$ and their ordered clique-size lists agree.*

Proof. Scaling does not change dimension, so first $r = s$. Now set coordinates $3, \dots, r$ equal to zero and look at the resulting two-dimensional faces. Since all coordinates of the vertices are nonnegative, these faces are exactly the triangles with vertices w_0, w_1, w_2. Their vertices in the first two coordinates are

$$(0,0), \qquad (a_r, 0), \qquad (a_{r-1} + a_r, a_{r-1}a_r),$$

and similarly for $\mathbf{b}$. The edge between the two nonzero vertices has slope a_r. Scaling all coordinates by the same positive number preserves this slope, so $a_r = b_r$. Since the nonzero horizontal vertex is $(a_r, 0)$ in both triangles, the scale factor must be one.

For each j, intersecting with the coordinate subspace where coordinates above j vanish gives $\operatorname{conv}\{w_0, \dots, w_j\}$. Its only vertex outside the preceding subspace is w_j. Thus equality of the simplices identifies every w_j. The difference between the first coordinates of w_j and w_{j-1} is a_{r-j+1}, recovering all remaining sizes. □

Theorem 8.5 (Pairwise coprimeness)**.** *Let D and E be nonisomorphic graphs, each a disjoint union of at least two nonempty cliques. Then J_D and J_E are relatively prime over $\mathbb{C}$. Moreover, $1 + z_1$ is relatively prime to every such J_D.*

Proof. Use a polynomial ring containing all the variables that occur in J_D and J_E. Suppose $J_D = AB$. Setting $z_1 = 0$ gives $J_D(0, z_2, \dots) = 1$, because only the empty subset has no vertices. The product of the two substituted factors is therefore one, so each is a nonzero constant. Multiply

the factors by constants so that both substituted values are one. Each factor's Newton polytope then contains the origin, and it has no other point with first coordinate zero.

By (36) and Lemma 8.3, a nonconstant factor of J_D has Newton polytope $\lambda\Delta_{\mathbf{a}}$ for some $\lambda > 0$. There is no translation here: the origin is the only point with first coordinate zero in both the original simplex and the factor polytope. A point Newton polytope corresponds to a single monomial. The nonzero constant term then forces that monomial to have degree zero.

If a nonconstant polynomial divided both J_D and J_E, its Newton polytope would therefore be a positive scalar multiple of both simplices. Lemma 8.4 would give $D \cong E$, a contradiction. Finally, $\mathrm{Newt}(1 + z_1)$ is a segment, which cannot be a nonpoint summand of a simplex of dimension at least two. □

We next use Theorem 8.5 to distinguish a family of connected cographs. Recall that a graph is a disjoint union of cliques if and only if it has no induced P_3. Indeed, in a connected component that is not a clique, a shortest path between two nonadjacent vertices begins with an induced P_3. The next lemma characterizes when the join factors have this form.

Lemma 8.6. *A connected cograph is $(P_3 \sqcup K_1)$-free if and only if each nonsingleton factor in its maximal join decomposition is a disjoint union of cliques.*

Proof. Each nonsingleton join factor is disconnected. If one of its components is not a clique, an induced P_3 in that component and a vertex in another component give an induced $P_3 \sqcup K_1$.

Conversely, an induced $P_3 \sqcup K_1$ cannot meet two join factors: every induced subgraph meeting two factors is connected. A clique-union factor has no induced P_3, so it cannot contain the required copy. A singleton factor is also too small. □

Theorem 8.7. *If G is a connected $(P_3 \sqcup K_1)$-free cograph and H is any graph with $\overline{X}_G = \overline{X}_H$, then $G \cong H$.*

Proof. Suppose G belongs to the stated class and $\overline{X}_G = \overline{X}_H$. The four-vertex counts in Theorem 6.2 show that H is also a $(P_3 \sqcup K_1)$-free cograph. Theorem 7.8 shows that H is connected. By Lemma 8.6, the polynomials of their maximal join decompositions have the form

$$J_G = (1 + z_1)^s \prod_D J_D^{m_D}, \qquad J_H = (1 + z_1)^t \prod_D J_D^{n_D},$$

where we take one D from each isomorphism type of disjoint union of at least two cliques. The exponents m_D, n_D count how many times that graph occurs as a join factor. Only finitely many are nonzero.

Choose any D that occurs, and let q_D be an irreducible factor of J_D over $\mathbb{C}$. By Theorem 8.5, it is not a factor of any other J_E or of $1 + z_1$. Suppose it occurs $e_D > 0$ times in J_D. It then occurs $m_D e_D$ times on the left and $n_D e_D$ times on the right. Unique factorization gives $m_D e_D = n_D e_D$, so $m_D = n_D$. Counting the factors $1 + z_1$ also gives $s = t$. Thus the graph factors in the two joins agree, and $G \cong H$. □

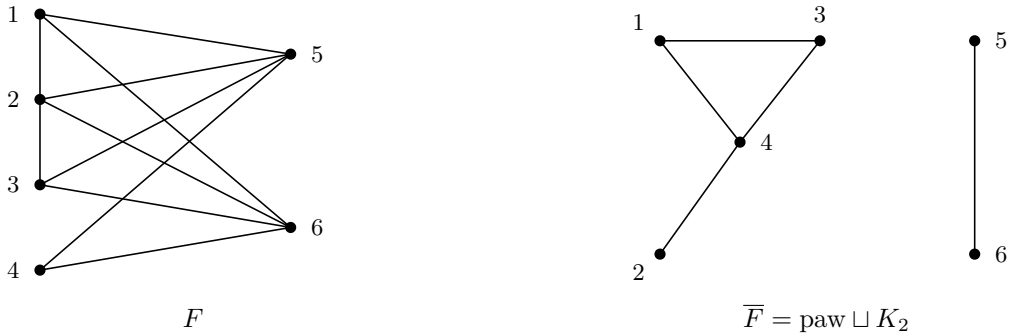
8.3 The forbidden-subgraph theorem

We can obtain a larger family from Theorem 8.7 by taking disjoint unions and adding universal vertices. To describe this family, let

$$F = (P_3 \sqcup K_1) \odot 2K_1. \tag{39}$$

This six-vertex cograph is connected and has no universal vertex. Figure 12 shows F alongside its complement, a paw disjoint from K_2. Excluding F will be enough for reconstruction.

Theorem 8.8. *If G is an F-free cograph and H is any graph with $\overline{X}_G = \overline{X}_H$, then $G \cong H$. Equivalently, the KSF determines every graph with no induced P_4 or F.*



F $\qquad$ $\overline{F} = \text{paw} \sqcup K_2$

Figure 12: The graph $F = (P_3 \sqcup K_1) \odot 2K_1$ and its complement. In F, vertices 1–3 form the path P_3, vertex 4 is isolated from these three, and vertices 5 and 6 are each joined to all four.

Proof. First consider a connected F-free cograph G without universal vertices. Its canonical join decomposition has at least two factors, each disconnected. Suppose one factor is not a disjoint union of cliques. It contains an induced $P_3 \sqcup K_1$, by the first part of the proof of Lemma 8.6. A different factor is disconnected, so we can choose two nonadjacent vertices in it. Together with the four already chosen, they induce F, a contradiction. Every join factor must therefore be a union of cliques, placing G within Theorem 8.7.

For a general F-free cograph G satisfying $\overline{X}_G = \overline{X}_H$, proceed by induction on the order. The empty graph and K_1 provide the base cases. If G is disconnected, Theorem 7.8 matches its component KSFs with those of H. Each component is a smaller F-free cograph, so induction reconstructs every component.

If G is connected and has $s > 0$ universal vertices, Theorem 6.1 gives the same number in H. Write $G = K_s \odot G^\circ$ and $H = K_s \odot H^\circ$. The product identity for J gives

$$(1 + z_1)^s J_{G^\circ} = (1 + z_1)^s J_{H^\circ}.$$

Cancel $(1 + z_1)^s$. Proposition 7.1 then gives $\overline{X}_{G^\circ} = \overline{X}_{H^\circ}$. The graph G° is a smaller F-free cograph, so induction applies. The connected case without universal vertices was established in the first paragraph. This covers every case. □

Proposition 8.9. *The class of F-free cographs is the smallest class containing the empty graph and all connected $(P_3 \sqcup K_1)$-free cographs, and closed under disjoint union and adjoining a universal vertex.*

Proof. The starting graphs contain no induced F, since F contains $P_3 \sqcup K_1$. Taking a disjoint union cannot create an induced F, because F is connected. Adding a universal vertex cannot create one either: any induced copy using the new vertex would have a universal vertex, but F has none. Both operations also keep the graph a cograph.

To show that these operations produce every F-free cograph, use induction on the number of vertices. A disconnected F-free cograph reduces to its smaller components, and one with a universal vertex reduces by deleting that vertex. The remaining connected graphs have no universal vertices; the first paragraph of the proof of Theorem 8.8 places them in the initial class. □

Corollary 8.10. *Every claw-free cograph is determined by its KSF. Every connected claw-free cograph is determined by P_G against arbitrary graphs.*

Proof. The structural observation below also appears in Tsujie [Tsu18, Lemma 4.1]. In a connected claw-free cograph, a stable triple inside one join factor together with a vertex of another factor would induce a claw. Each disconnected factor therefore has independence number at most two. There can be only two components, since choosing one vertex from each of three components would give a stable triple. Both components are cliques as well: a nonadjacent pair in one and a vertex in the other would again form a stable triple. The graph has a unique representation

$$G = K_s \odot \bigodot_{i=1}^{r} (K_{a_i} \sqcup K_{b_i}), \qquad a_i, b_i \geq 1, \tag{40}$$

and $\alpha(G) \leq 2$. Theorem 8.7 applies. If $P_G = P_H$, then $I_G = I_H$, so $\alpha(H) \leq 2$ as well. Theorem 6.4 converts this equality to equality of KSFs and gives $G \cong H$. The disconnected assertion follows from component reconstruction. □

Corollary 8.11. *Every cograph whose canonical cotree has height at most three is determined by its KSF, where height is the maximum number of edges in a root-to-leaf path.*

Proof. If the root is a join node, its nonsingleton children are union nodes whose connected children are cliques. Thus Theorem 8.7 applies. If the root is a union node, apply the same conclusion to its connected components and then use Theorem 7.8. □

Theorem 8.8 also applies to cographs with arbitrarily large cotree height. For example, begin with K_1 and repeatedly replace G by $K_1 \odot (G \sqcup G)$. Proposition 8.9 keeps every graph in the reconstructed class, while each iteration adds two levels to its canonical cotree.

Proposition 8.12. *For every graph, the KSF determines the KSF of the graph remaining after all universal vertices are removed.*

Proof. Read the number s from the degree sequence. Removing the factor $(1 + z_1)^s$ from J_G gives the polynomial of the remaining graph, whose KSF is recovered by Proposition 7.1. □

9 The cograph conjecture

The F-free case is now settled. To handle all cographs, we must recover the join factors from the product of their polynomials. We formulate this remaining question precisely, prove additional irreducibility results, and examine the effect of assigning one variable to all stable sets of a given size.

9.1 An equivalent factorization question

Theorem 9.1. *The following statements are equivalent.*

(i) The KSF distinguishes all cographs.

(ii) Whenever we have an equality

$$(1 + z_1)^s \prod_{i=1}^{r} J_{D_i} = (1 + z_1)^t \prod_{j=1}^{q} J_{E_j}, \tag{41}$$

where s, t are nonnegative integers and the graphs D_i, E_j are cographs with at least two connected components, we must have $s = t$, $r = q$, and, after reordering, $D_i \cong E_i$ for every i.

Proof. Assume (i). Take the join of $D_1, \ldots, D_r$ and s single vertices, and do the same with $E_1, \ldots, E_q$ and t single vertices. Equation (41) says that the resulting graphs have the same J polynomial, hence the same KSF. By (i), they are isomorphic.

None of the displayed graph factors can be split further as a join. For a disconnected graph, this follows because its complement is connected; it is also clear for a single vertex. Thus these are the maximal join decompositions of the two graphs. Their uniqueness matches the factors and proves (ii). The argument also covers a single disconnected factor or an empty join.

Conversely, suppose (ii), and let G, H be cographs with equal KSFs. Decompose each graph into its maximal join factors. Every factor is a singleton or a disconnected cograph, so equality $J_G = J_H$ is an instance of (41). Statement (ii) matches the graph factors, so $G \cong H$. By Theorem 6.2, any graph with the same KSF as a cograph is itself a cograph. Thus we also obtain the result when H was not assumed to be a cograph. □

The property in (ii) is called *multiplicative independence*. It would follow if the polynomials of different join factors had no common nonconstant factor, as we proved for disjoint unions of cliques in Theorem 8.5. The following stronger irreducibility conjecture would also suffice.

Conjecture 9.2 (Join irreducibility)**.** *If D is a cograph with at least two connected components, then J_D is irreducible over $\mathbb{Q}$.*

Theorem 9.3. *If Conjecture 9.2 holds, then $\overline{X}_G = \overline{X}_H$ implies $G \cong H$ whenever G is a cograph and H is any graph.*

Proof. Suppose G is a cograph and $\overline{X}_G = \overline{X}_H$. By Corollary 6.3, H is also a cograph. We use induction on their common number of vertices. The result is clear for zero or one vertex. If the graphs are disconnected, Theorem 7.8 matches their component KSFs. Induction identifies each matched pair of components, since their orders are smaller.

Otherwise write the canonical join decompositions $G = \bigodot_i A_i$ and $H = \bigodot_j B_j$. Every factor is a singleton or a disconnected cograph, so every J_{A_i} and J_{B_j} is irreducible by the conjecture and $J_{K_1} = 1 + z_1$. Their products agree by (33). Unique factorization matches the polynomial factors. There are no extra constant multiples to consider because all constant terms are one. Proposition 7.1 gives equal KSFs for the matching graph factors. These graphs have fewer vertices than G, so induction shows that they are isomorphic. Their joins are therefore isomorphic as well. □

9.2 Irreducibility for two cliques

We next prove Conjecture 9.2 for several disjoint unions of two cliques. Section 8 already shows that their polynomials can be distinguished in a product. Here we ask the stronger question of when the polynomials themselves are irreducible. For positive integers a, b, define

$$B_{a,b}(u, v) = \sum_{i=0}^{a} \sum_{j=0}^{b} \binom{a}{i} \binom{b}{j} u^{i+j} v^{ij} = P_{K_{a,b}}(u, v) = J_{K_a \sqcup K_b}(u, v). \tag{42}$$

The last equality holds because a subset meeting the cliques in i and j vertices has independence polynomial $1 + (i + j)x + ijx^2$.

Lemma 9.4. *For $a \le b$, let $d = \gcd(a, b)$. The Newton polygon of $B_{a,b}$ is*

$$T_{a,b} = \operatorname{conv}\{(0,0), (b,0), (a+b, ab)\}.$$

In any nontrivial complex factorization $B_{a,b} = AB$, normalize the factors to have constant term one. For some integer $1 \le k < d$,

$$\operatorname{Newt}(A) = \frac{k}{d} T_{a,b}, \qquad \operatorname{Newt}(B) = \frac{d-k}{d} T_{a,b}.$$

In particular, $\deg_u A(u, 0) = kb/d > 0$, and similarly for B.

Proof. A Newton polygon is the two-dimensional version of a Newton polytope. Here the exponent vectors are $(x, y) = (i + j, ij)$. They satisfy

$$y \ge 0, \qquad y \le \frac{ab}{a+b} x, \qquad y \ge b(x - b),$$

and the three vertices occur. The middle inequality follows from $1/i + 1/j \ge 1/a + 1/b$ when $ij > 0$; the last is $(b - i)(b - j) \ge 0$.

Newton polygons add when we multiply polynomials. Since $B_{a,b}(0, v) = 1$, the Newton polygons of both normalized factors contain the origin. Lemma 8.3 shows that they are scaled copies of $T_{a,b}$, with scale factors λ and $1 - \lambda$ for some $0 < \lambda < 1$.

Their vertices have integer coordinates. The horizontal edge has first-coordinate difference λb, and the edge parallel to $(1, b)$ has first-coordinate difference λa. Both are integers. Since $d = \gcd(a, b)$, this forces $\lambda = k/d$ for an integer $1 \le k < d$. Setting $v = 0$ keeps exactly the terms on the horizontal edge, so $\deg_u A(u, 0) = \lambda b = kb/d$. □

Theorem 9.5. *The polynomial $B_{a,b}$ is irreducible over $\mathbb{Q}$ if $a = b$, if $\gcd(a, b)$ is a power of two, or if $(1 + u)^a + (1 + u)^b - 1$ is irreducible over $\mathbb{Q}$. When $\gcd(a, b) = 1$, it is also irreducible over $\mathbb{C}$.*

Proof. If $\gcd(a, b) = 1$, Lemma 9.4 leaves no possible factorization, since there is no integer k with $1 \le k < 1$. Next suppose $(1 + u)^a + (1 + u)^b - 1$ is irreducible. Any two nonconstant factors of $B_{a,b}$ would remain nonconstant after setting $v = 0$. This would contradict the irreducibility of

$$B_{a,b}(u, 0) = (1 + u)^a + (1 + u)^b - 1.$$

When $a = b$, the polynomial on the right is $2(1 + u)^a - 1$. Let $x = 1 + u$ and reverse the coefficients. This gives $2 - x^a$, which is irreducible by Eisenstein's criterion at 2. Reversing these steps proves the equal-size case.

For the power-of-two case, let $g = \gcd(a + b, ab)$. The only exponent vectors on the upper edge are its endpoints, so the terms on this edge give

$$1 + u^{a+b} v^{ab} = 1 + t^g, \qquad t = u^{(a+b)/g} v^{ab/g}.$$

The two exponents in t have greatest common divisor one. Thus the integer points on the ray along this edge correspond to integer powers of t. By Lemma 9.4, a factorization of $B_{a,b}$ over $\mathbb{Q}$ would give a factorization of $1 + t^g$ into polynomials of positive degrees kg/d and $(d - k)g/d$. Write $a = dA$, $b = dB$ with $\gcd(A, B) = 1$. Since $\gcd(A + B, AB) = 1$, $g = d \gcd(A + B, d)$. Thus g is a power of two when d is. But $1 + t^g$ is irreducible over $\mathbb{Q}$: replacing t by $t + 1$ gives Eisenstein's criterion at 2. This excludes the factorization. □

9.3 Beyond disjoint unions of cliques

The first disconnected cograph beyond clique unions is $D_0 = P_3 \sqcup K_1$. Its Newton polytope is no longer a simplex. Nevertheless, the triangular faces impose enough restrictions to prove irreducibility.

Proposition 9.6. *Let $(u, v, w) = (z_1, z_2, z_3)$. Then*

$$\begin{aligned} J_{D_0}(u, v, w) = 1 + 4u + 2u^2 + 4u^2v + u^3v + 2u^3v^2 \\ + u^3v^3w + u^4v^4w. \end{aligned} \tag{43}$$

Its Newton polytope N has the six vertices

$$\begin{aligned} v_0 = (0, 0, 0), \quad v_1 = (2, 0, 0), \quad v_2 = (3, 1, 0), \\ v_3 = (3, 2, 0), \quad v_4 = (3, 3, 1), \quad v_5 = (4, 4, 1). \end{aligned}$$

If $N = P + Q$ for polytopes P, Q, then both summands are translated scalar multiples of N. If both summands have integer vertices, one of them is a point. Consequently, J_{D_0} is irreducible over $\mathbb{C}$.

Proof. There are four singletons, two adjacent pairs, and four nonadjacent pairs. Among the four triples, one induces P_3, two induce $K_2 \sqcup K_1$, and one is independent. The whole graph has independence polynomial $1 + 4x + 4x^2 + x^3$. These counts give (43).

The two exponent vectors not in the displayed vertex list are

$$(1, 0, 0) = \tfrac{1}{2}v_0 + \tfrac{1}{2}v_1, \qquad (2, 1, 0) = \tfrac{1}{4}v_0 + \tfrac{1}{4}v_1 + \tfrac{1}{2}v_3.$$

Thus $N = \operatorname{conv}\{v_0, \ldots, v_5\}$. The following four triangles are exposed faces of N. The table gives a supporting plane for each, with coordinates denoted by (x, y, z).

Face	Supporting plane
$\operatorname{conv}\{v_0, v_4, v_5\}$	$x - y = 0$
$\operatorname{conv}\{v_0, v_3, v_5\}$	$2x - 3y + 4z = 0$
$\operatorname{conv}\{v_2, v_3, v_5\}$	$x - z = 3$
$\operatorname{conv}\{v_0, v_1, v_4\}$	$y - 3z = 0$

For each row, exactly the three listed points lie on the plane, and the other three lie strictly on one side. The listed points are not collinear. Every v_i therefore is a vertex. Also v_0, v_1, v_2, v_4 are affinely independent, so N is three-dimensional. With six vertices, it is not a simplex.

Now suppose $N = P + Q$. The supporting-face argument in Lemma 8.3 assigns to each vertex a unique decomposition $v_i = a_i + b_i$, with $a_i \in P$ and $b_i \in Q$. Along any edge v_iv_j, we have

$$a_j - a_i = \lambda_{ij}(v_j - v_i), \qquad 0 \le \lambda_{ij} \le 1.$$

On a triangular face, the three edge scales agree. The first triangle in the table shares an edge with the second and fourth triangles, and the second shares an edge with the third. Thus all their edge scales have one common value λ. These triangles contain all six vertices and their edge graph is connected. It follows that

$$a_i = t + \lambda v_i, \qquad b_i = -t + (1 - \lambda)v_i \qquad (0 \le i \le 5).$$

Every vertex of a summand occurs among these a_i or b_i: choose a linear functional uniquely maximized at that vertex and perturb it so that it also has a unique maximum on N, as in Lemma 8.3. Hence $P = t + \lambda N$ and $Q = -t + (1 - \lambda)N$.

If P, Q have integer vertices, then $a_4 - a_0 = \lambda(3,3,1)$ has integer coordinates. Its third coordinate gives $\lambda \in \mathbb{Z}$, so $\lambda = 0$ or 1. One summand is therefore a point. Finally, a factorization of J_{D_0} would give such an integer decomposition by (36). A point polytope means that the corresponding factor is a monomial; the constant term one forces that factor to be constant. This proves irreducibility over $\mathbb{C}$. □

There is also a short algebraic check of the last conclusion. Write

$$J_{D_0} = u^3v^3(1+uv)w + B(u,v),$$
$$B = 1 + 4u + 2u^2 + 4u^2v + u^3v + 2u^3v^2.$$

The constant term of B excludes the factors u, v, and $B(-v^{-1}, v) = (1-v^{-1})^2 \neq 0$ excludes $1+uv$. Thus the two coefficients of this linear polynomial in w have no common nonconstant factor. Any factor independent of w would have to divide both, so it is constant.

Corollary 9.7. *The polynomial J_{D_0} is relatively prime over $\mathbb{C}$ to $1+z_1$ and to J_D for every disjoint union D of at least two nonempty cliques.*

Proof. The polynomial J_{D_0} is irreducible and depends on z_3, so it cannot divide $1+z_1$. If it divided J_D, its Newton polytope would be a positive scalar multiple of the clique-union simplex, by the proof of Theorem 8.5. This is impossible because N is not a simplex. □

Thus J_{D_0} is irreducible even though its Newton polytope is not a simplex. In this example, the triangular faces give enough restrictions on the Minkowski summands. We do not yet know how to extend this argument to arbitrary disconnected cographs.

9.4 Identifying stable-set variables

There is another way to formulate the irreducibility question using stable-set covers. In F_G, each stable set contributes a variable depending only on its size. We now give every nonempty stable set S a separate variable t_S and compare F_G with the resulting polynomial $\widehat{F}_G$:

$$F_G(\mathbf{y}) = \sum_{C \in \mathrm{SSC}(G)} \prod_{S \in C} y_{|S|}, \qquad \widehat{F}_G(\mathbf{t}) = \sum_{C \in \mathrm{SSC}(G)} \prod_{S \in C} t_S.$$

The stable-set-cover expansion makes F_G another encoding of the KSF. The inclusion–exclusion identity (4) gives

$$\begin{aligned} F_G &= \sum_{R \subseteq V(G)} (-1)^{n-|R|} \prod_{k \geq 1} (1+y_k)^{i_k(G[R])} \\ &= (-1)^n J_G(-(1+y_1), 1+y_2, \ldots). \end{aligned} \tag{44}$$

We can therefore get either F_G or J_G from the other by changing the variables as shown and multiplying by $(-1)^n$. In particular, one is irreducible exactly when the other is.

Theorem 9.8. *For a nonempty graph G, the irreducible factors of $\widehat{F}_G$ over any field are exactly the cover polynomials of the factors in its canonical join decomposition. In particular, $\widehat{F}_G$ is irreducible if and only if $\overline{G}$ is connected, or equivalently, G cannot be written as a join of two nonempty graphs.*

Proof. Every stable set of a join lies in one graph factor. Thus choosing a cover of the join amounts to choosing a cover of each factor, which proves the product identity.

We now prove irreducibility of $\widehat{F}_G$ under the assumption that $\overline{G}$ is connected. Suppose instead that $\widehat{F}_G = AB$, with both factors nonconstant. Each variable t_S has degree at most one in $\widehat{F}_G$, because a cover uses a given stable set at most once. Over a field, $\deg_{t_S}(AB) = \deg_{t_S} A + \deg_{t_S} B$. Therefore no variable can occur in both A and B. It follows that each product monomial comes from only one pair of monomials, one from each factor, so their coefficients cannot cancel.

Assign each vertex v to V_A or V_B according to which factor contains the singleton variable $t_{\{v\}}$. The cover consisting of all singleton sets shows that A contains the product of the singleton variables for V_A, and B contains the corresponding product for V_B.

Suppose t_S occurs in A. Complete S to a cover by using one singleton for each vertex outside S. Its monomial shows that B contains the product of the singleton variables for $V_B \setminus S$. Multiply this by the product for V_A already known to occur in A. The resulting monomial occurs in $\widehat{F}_G$, so it must still cover every vertex. This is only possible if $S \cap V_B = \varnothing$. Similarly, every set whose variable occurs in B lies entirely in V_B.

Every edge of $\overline{G}$ is a stable pair in G. The endpoints must consequently belong to one part together. Since $\overline{G}$ is connected, either V_A or V_B is empty. No variable can belong to the factor associated with an empty part. That factor is constant, contrary to our assumption. The same argument applies to every factor in the maximal join decomposition. □

For example, $\widehat{F}_{K_2} = t_{\{1\}}t_{\{2\}}$ is reducible, consistent with the fact that $\overline{K_2}$ is disconnected. For a graph D with at least two connected components, the complement is connected, so $\widehat{F}_D$ is irreducible. To obtain F_D, we replace every variable t_S by $y_{|S|}$. Thus Conjecture 9.2 asks whether this replacement keeps the polynomial irreducible for every disconnected cograph.

This can fail for general graphs. For example, $\widehat{F}_{C_5}$ is irreducible, but $J_{C_5} = P_{C_5}$ has the factor $1 + uv$. To check this, pair each subset S of the five-cycle with its complement. Their edge counts satisfy

$$e(C_5[S]) - e(C_5[V \setminus S]) = 2|S| - 5.$$

After setting $u = -v^{-1}$, the two corresponding terms have the same power of v and opposite signs, so they cancel. Hence $1 + uv$ divides P_{C_5}, first over $\mathbb{Q}(v)[u]$ and then over $\mathbb{Q}[u, v]$ by Gauss's lemma. More explicitly, when $\alpha(G) \leq 2$, setting $(z_1, z_2) = (u, v)$ gives $J_G = P_{\overline{G}}$: the coefficient $i_2(G[S])$ counts edges in $\overline{G}[S]$. Apply this to C_5 and use $C_5 \cong \overline{C_5}$ to obtain $J_{C_5} = P_{C_5}$. The five-cycle lies outside the class in Conjecture 9.2. It shows, however, that identifying variables can destroy irreducibility, so that step requires its own argument.

Proposition 9.9. *A pair of nonisomorphic cographs of smallest order with equal KSFs, if one exists, consists of connected graphs without universal vertices. Both contain an induced copy of F from (39). Their canonical join factors are disconnected cographs, no factor on one side has the same KSF as a factor on the other, and at least one of their join-factor polynomials J_D is reducible.*

Proof. Choose a pair G, H of minimum order among nonisomorphic cographs with equal KSFs. If they were disconnected, Theorem 7.8 would pair their components by KSF. Minimality of G, H would then identify each pair of components, contradicting nonisomorphism. Both graphs are therefore connected.

If they had universal vertices, Proposition 8.12 would remove them and give a smaller pair with equal KSFs. Again this contradicts the choice of the pair. Theorem 8.8 also shows that both graphs must contain an induced F.

Since there are no universal vertices, every factor in either maximal join decomposition is disconnected. If a factor on one side had the same KSF as a factor on the other, we could cancel their J polynomials. The smaller matching factors and the remaining joins would then be isomorphic

by the choice of G, H, another contradiction. Finally, if all the J factors were irreducible, unique factorization would force just such a match. At least one must therefore be reducible. □

9.5 Computational evidence and further questions

Proposition 9.10 (Computer check)**.** *The polynomial P_G distinguishes all cographs on at most twelve vertices.*

Proof. We check all cographs by building them in order of their number of vertices. Every disconnected cograph is a disjoint union of smaller connected cographs. Taking complements then gives all connected cographs with at least two vertices. Uniqueness of the component decomposition ensures that each isomorphism type occurs once.

At a disjoint-union step, multiply the component polynomials to obtain P_G. For complements, we replace each term $u^k v^e$ by $u^k v^{\binom{k}{2}-e}$. We compare the resulting integer coefficients exactly. At every order through twelve, all the polynomials are different. The number of graphs checked at each order is

Order	1	2	3	4	5	6
Count	1	2	4	10	24	66
Order	7	8	9	10	11	12
Count	180	522	1532	4624	14136	43930

□

We also computed J_D for every disconnected cograph D on at most twelve vertices. All 32 515 of these polynomials are irreducible over $\mathbb{Q}$, so Conjecture 9.2 holds for every graph checked.

Since P_G can be found from the KSF, Proposition 9.10 also verifies Conjecture 1.1 through twelve vertices. It raises a stronger question: does P_G distinguish all cographs? We do not know whether this is true. For connected claw-free cographs, it follows from Corollary 8.10.

For the full cograph conjecture, Theorem 9.1 gives the property of products of the J_D that is needed. Proving Conjecture 9.2 would be sufficient, but it may be possible to prove this property without proving irreducibility. For a disconnected cograph, Theorem 9.8 gives irreducibility before stable sets of the same size are assigned the same variable; the question is whether it is preserved after this substitution. Proposition 9.9 describes the conditions that a smallest counterexample to the cograph conjecture would have to satisfy.

Acknowledgments

I would like to thank Laura Pierson for her mentorship on this project. This paper would not have been possible without her.

References

[ACSZ21] José Aliste-Prieto, Logan Crew, Sophie Spirkl, and José Zamora. "A vertex-weighted Tutte symmetric function, and constructing graphs with equal chromatic symmetric function." *The Electronic Journal of Combinatorics* 28.2 (2021). `https://www.combinatorics.org/ojs/index.php/eljc/article/view/v28i2p1/pdf`.

[AZ14] José Aliste-Prieto and José Zamora. "Proper caterpillars are distinguished by their chromatic symmetric function." *Discrete Mathematics* 315 (2014), 158–164. https://doi.org/10.1016/j.disc.2013.10.016.

[CLS81] Derek G. Corneil, H. Lerchs, and L. Stewart Burlingham. "Complement reducible graphs." *Discrete Applied Mathematics* 3.3 (1981), 163–174. https://doi.org/10.1016/0166-218X(81)90013-5.

[CPS26] Logan Crew, Oliver Pechenik, and Sophie Spirkl. "The Kromatic symmetric function: A K-theoretic analog of X_G." *Canadian Mathematical Bulletin*, First View (2026), 1–23. https://doi.org/10.4153/S0008439526101830.

[CS20] Logan Crew and Sophie Spirkl. "A deletion–contraction relation for the chromatic symmetric function." *European Journal of Combinatorics* 89 (2020), 103143. https://doi.org/10.1016/j.ejc.2020.103143.

[Csi13] Péter Csikvári. "Note on the smallest root of the independence polynomial." *Combinatorics, Probability and Computing* 22.1 (2013), 1–8. https://doi.org/10.1017/S0963548312000302.

[GL01] Shuhong Gao and Alan G. B. Lauder. "Decomposition of polytopes and polynomials." *Discrete & Computational Geometry* 26 (2001), 89–104. https://doi.org/10.1007/s00454-001-0024-0.

[Kot12] Tomer Kotek. "Complexity of Ising polynomials." *Combinatorics, Probability and Computing* 21.5 (2012), 743–772. https://doi.org/10.1017/S0963548312000259.

[LS19] Martin Loebl and Jean-Sébastien Sereni. "Isomorphism of weighted trees and Stanley's isomorphism conjecture for caterpillars." *Annales de l'institut Henri Poincaré D* 6.3 (2019), 357–384. https://ems.press/journals/aihpd/articles/16211.

[MMW08] Jeremy L. Martin, Matthew Morin, and Jennifer D. Wagner. "On distinguishing trees by their chromatic symmetric functions." *Journal of Combinatorial Theory, Series A* 115.2 (2008), 237–253. https://www.sciencedirect.com/science/article/pii/S0097316507000775.

[McK26] Brendan D. McKay. *Combinatorial Data: Graphs*. Unlabeled simple-graph catalogues. Accessed September 11, 2026. https://users.cecs.anu.edu.au/~bdm/data/graphs.html.

[Mor05] Matthew Morin. "Caterpillars, ribbons, and the chromatic symmetric function." PhD thesis, University of British Columbia, 2005. https://doi.org/10.14288/1.0079438.

[OS14] Rosa Orellana and Geoffrey Scott. "Graphs with equal chromatic symmetric functions." *Discrete Mathematics* 320 (2014), 1–14. https://doi.org/10.1016/j.disc.2013.12.006.

[Pie24] Laura Pierson. "Counting induced subgraphs with the Kromatic symmetric function." 2024. https://arxiv.org/abs/2403.15929.

[Pie25] Laura Pierson. "Power sum expansions for Kromatic symmetric functions using Lyndon heaps." *Annals of Combinatorics* (2025). https://doi.org/10.1007/s00026-025-00785-8.

[PS26] Laura Pierson and Soham Samanta. "On graphs with equal and different Kromatic symmetric functions." *Discrete Applied Mathematics* 387 (2026), 137–159. https://doi.org/10.1016/j.dam.2026.02.039.

[Sta95] Richard P. Stanley. "A symmetric function generalization of the chromatic polynomial of a graph." *Advances in Mathematics* 111.1 (1995), 166–194. https://doi.org/10.1006/aima.1995.1020.

[TAM11] Peter Tittmann, Ilia Averbouch, and Johann A. Makowsky. "The enumeration of vertex induced subgraphs with respect to the number of components." *European Journal of Combinatorics* 32.7 (2011), 954–974. https://arxiv.org/abs/0812.4147.

[Tsu18] Shuhei Tsujie. "The chromatic symmetric functions of trivially perfect graphs and cographs." *Graphs and Combinatorics* 34 (2018), 1037–1048. https://doi.org/10.1007/s00373-018-1928-2.