Annals of the Institute of Statistical Mathematics (2021) 73:757–785
https://doi.org/10.1007/s10463-020-00765-0

Gaussian graphical models with toric vanishing ideals

Pratik Misra[1] · **Seth Sullivant**[1]

Received: 16 December 2019 / Revised: 2 September 2020 / Accepted: 4 September 2020 /
Published online: 2 November 2020

Abstract
Gaussian graphical models are semi-algebraic subsets of the cone of positive definite covariance matrices. They are widely used throughout natural sciences, computational biology and many other fields. Computing the vanishing ideal of the model gives us an implicit description of the model. In this paper, we resolve two conjectures given by Sturmfels and Uhler. In particular, we characterize those graphs for which the vanishing ideal of the Gaussian graphical model is generated in degree 1 and 2. These turn out to be the Gaussian graphical models whose ideals are toric ideals, and the resulting graphs are the 1-clique sums of complete graphs.

Keywords Clique sum · Toric ideals · SAGBI bases · Initial algebra

1 Introduction

Any positive definite $n \times n$ matrix Σ can be seen as the covariance matrix of a multivariate normal distribution in $\mathbb{R}^n$. The inverse matrix $K = \Sigma^{-1}$ is called the *concentration* matrix of the distribution, which is also positive definite. The statistical models where the concentration matrix K can be written as a linear combination of some fixed linearly independent symmetric matrices $K_1, K_2, \dots, K_d$ are called *linear concentration* models.

Let $\mathbb{S}^n$ denote the vector space of real symmetric matrices and let $\mathcal{L}$ be a linear subspace of $\mathbb{S}^n$ generated by $K_1, K_2, \dots, K_d$. The set $\mathcal{L}^{-1}$ is defined as

$$\mathcal{L}^{-1} = \{\Sigma \in \mathbb{S}^n \, : \, \Sigma^{-1} \in \mathcal{L}\}.$$

✉ Pratik Misra
pmisra@ncsu.edu

Seth Sullivant
smsulli2@ncsu.edu

[1] Department of Mathematics, North Carolina State University, 2108 SAS Hall Box 8205, Raleigh, NC 27695, USA

The homogeneous ideal of all the polynomials in $\mathbb{R}[\Sigma] = \mathbb{R}[\sigma_{11}, \sigma_{12}, \dots, \sigma_{nn}]$ that vanish on $\mathcal{L}^{-1}$ is denoted by $P_{\mathcal{L}}$. Note that $P_{\mathcal{L}}$ is prime because it is the vanishing ideal of $\mathcal{L}^{-1}$, which is the image of the irreducible variety $\mathcal{L}$ under the rational inversion map. In this paper, we study the problem of finding a generating set of $P_{\mathcal{L}}$ for the special case of Gaussian graphical models.

Gaussian graphical models are used throughout the natural sciences and especially in computational biology as seen in Koller and Friedman (2009) and Lauritzen (1996). These models explicitly capture the statistical relationships between the variables of interest in the form of a graph. The undirected Gaussian graphical model is obtained when the subspace $\mathcal{L}$ of $\mathbb{S}^n$ is defined by the vanishing of some off-diagonal entries of the concentration matrix K. We fix a graph $G = ([n], E)$ with vertex set $[n] = \{1, 2, \dots, n\}$ and edge set E, which is assumed to contain all self loops. The subspace $\mathcal{L}$ is generated by the set $\{K_{ij} | (i,j) \in E\}$ of matrices K_{ij} with 1 entry at the $(i,j)^{th}$ and $(j,i)^{th}$ position and 0 in all other positions. We denote the ideal $P_{\mathcal{L}}$ as P_G in this model.

One way to compute P_G is to eliminate the entries of an indeterminate symmetric $n \times n$ matrix K from the following system of equations:

$$\Sigma \cdot K = Id_n, \quad K \in \mathcal{L},$$

where Id_n is the $n \times n$ identity matrix. However, this elimination is computationally expensive, and we would like methods to identify generators of P_G directly in terms of the graph.

Various methods have been proposed for finding some generators in the ideal P_G and for trying to build P_G from smaller ideals associated to subgraphs. These approaches are based on separation criteria in the graph G.

Definition 1 Let $G = (V, E)$ be a graph.

- A set $C \subseteq V$ is called a *clique* of G if the subgraph induced by C is a complete graph.
- Let A, B, and C be disjoint subsets of the vertex set of G with $A \cup B \cup C = V$. Then, C *separates* A and B if for any $a \in A$ and $b \in B$, any path from a to b passes through a vertex in C.
- The graph G is said to be a *c-clique sum* of smaller graphs G_1 and G_2 if there exists a partition (A, B, C) of its vertex set such that

  (i) C is a clique with $|C| = c$,
  (ii) C separates A and B,
  (iii) G_1 and G_2 are the subgraphs induced by $A \cup C$ and $B \cup C$, respectively.

  In the case that G is a c-clique sum, we call the corresponding partition (A, B, C) a *c-clique partition* of G.

If G is a c-clique sum of G_1 and G_2, the ideal

$$P_{G_1} + P_{G_2} + \langle (c+1) \times (c+1)\text{-minors of } \Sigma_{A\cup C, B\cup C} \rangle \tag{1}$$

is contained in P_G. Here, $\Sigma_{A\cup C,B\cup C}$ denotes the submatrix of Σ obtained by taking all rows indexed by $A \cup C$ and columns indexed by $B \cup C$, and so

$$\langle (c+1) \times (c+1)\text{-minors of } \Sigma_{A\cup C, B\cup C} \rangle$$

is the conditional independence ideal associated to the conditional independence statement $A \perp\!\!\!\perp B | C$. Though the ideal (1) fails to equal P_G, (or even have the same radical as that of P_G) for $c \geq 2$, Sturmfels and Uhler (2010) conjectured it to be equal to P_G for $c = 1$.

Conjecture 1 (Sturmfels and Uhler 2010) Let G be a 1-clique sum of two smaller graphs G_1 and G_2. If (A, B, C) is the 1-clique partition of G where G_1 and G_2 are the subgraphs induced by $A \cup C$ and $B \cup C$, respectively, then

$$P_G = P_{G_1} + P_{G_2} + \langle 2 \times 2\text{-minors of } \Sigma_{A\cup C, B\cup C} \rangle.$$

In Sect. 2, we give counterexamples to this conjecture, and even a natural strengthening of it. However, the motivation for Conjecture 1 was to use it as a tool to prove a different conjecture characterizing the graphs for which the vanishing ideal P_G is generated in degree ≤ 2. To explain the details of this conjecture we need some further notions.

Let $X = (X_1, X_2, \ldots, X_n)$ be a Gaussian random vector. If $A, B, C \subseteq [n]$ are pairwise disjoint subsets, then from Proposition 4.1.9 of Sullivant (2018) we know that X_A is conditionally independent of X_B given X_C (i.e., $A \perp\!\!\!\perp B | C$) if and only if the submatrix $\Sigma_{A\cup C,B\cup C}$ of the covariance matrix Σ has rank $|C|$. The *Gaussian conditional independence ideal* for the conditional independence statement $A \perp\!\!\!\perp B | C$ is given by

$$J_{A\perp\!\!\!\perp B|C} = \langle (|C|+1) \times (|C|+1) \text{ minors of } \textstyle\sum_{A\cup C, B\cup C} \rangle.$$

If G is an undirected graph and (A, B, C) is a partition with C separating A from B, then the conditional independence statement $A \perp\!\!\!\perp B | C$ holds for all multivariate normal distributions where the covariance matrix Σ is obtained from G (by the global Markov property). The *conditional independence ideal* for the graph G is defined by

$$CI_G = \sum_{A\perp\!\!\!\perp B \,|\, C \text{ holds for } G} J_{A\perp\!\!\!\perp B|C}.$$

Proposition 1 *For any given graph G, $CI_G \subseteq P_G$.*

Proof As the rank of the submatrices $\Sigma_{A\cup C,B\cup C}$ of the covariance matrix Σ is $|C|$ for all partitions (A, B, C) of G, the generators of CI_G vanish on the matrices in $\mathcal{L}^{-1}$. □

Definition 2 A graph G is called a *1-clique sum of complete graphs* if there exists a partition (A, B, C) of its vertex set such that

(i) $|C| = 1$,
(ii) C separates A and B,
(iii) the subgraphs induced by $A \cup C$ and $B \cup C$ are either complete graphs or 1-clique sum of complete graphs.

The second conjecture in Sturmfels and Uhler (2010) which we prove in this paper is as follows:

Theorem 1 (Conjecture 4.4, Sturmfels and Uhler 2010) *The prime ideal P_G of an undirected Gaussian graphical model is generated in degree ≤ 2 if and only if each connected component of the graph G is a* 1*-clique sum of complete graphs*.

The "only if" part of the conjecture is proved in Sturmfels and Uhler (2010). That is, it is shown there that a graph that is not the 1-clique sum of complete graphs must have a generator of degree ≥ 3. Such a generator comes from a conditional independence statement with $\#C \geq 2$.

For 1-clique sum of complete graphs, the conditional independence ideal can be written as

$$CI_G = \langle \bigcup_{(A,B,C) \in \mathcal{C}_1(G)} 2 \times 2 \text{ minors of } \Sigma_{A\cup C, B\cup C} \rangle,$$

where $\mathcal{C}_1(G)$ denotes the set of all 1-clique partitions of G. In this paper, our main result will be a proof that $CI_G = P_G$ when G is a 1-clique sum of complete graphs.

The expression "1-clique sum of complete graphs" is somewhat cumbersome. We use the alternate expression *block graphs* for most of the paper, as that is a commonly used name in the literature. One important property of block graphs is that there is a unique locally shortest path between any pair of vertices in a connected component of a block graph.

Example 1 We illustrate the structure of Theorem 1 with an example. Let $G = ([6], E)$ be the block graph as shown in Fig. 1. This block graph G has four 1-clique partitions as follows:

Partition 1: $A = \{1, 2\}, B = \{4, 5, 6\}, C = \{3\}$, Partition 2: $A = \{1, 2, 3\}, B = \{5, 6\}, C = \{4\}$

Partition 3: $A = \{1, 2, 3, 5\}, B = \{6\}, C = \{4\}$, Partition 4: $A = \{1, 2, 3, 6\}, B = \{5\}, C = \{4\}$.

The associated matrices are as follows:

$$\text{For } 1 : \Sigma_{A\cup C, B\cup C} = \begin{bmatrix} \sigma_{13} & \sigma_{14} & \sigma_{15} & \sigma_{16} \\ \sigma_{23} & \sigma_{24} & \sigma_{25} & \sigma_{26} \\ \sigma_{33} & \sigma_{34} & \sigma_{35} & \sigma_{36} \end{bmatrix}, 2 : \Sigma_{A\cup C, B\cup C} = \begin{bmatrix} \sigma_{14} & \sigma_{15} & \sigma_{16} \\ \sigma_{24} & \sigma_{25} & \sigma_{26} \\ \sigma_{34} & \sigma_{35} & \sigma_{36} \\ \sigma_{44} & \sigma_{45} & \sigma_{46} \end{bmatrix}$$

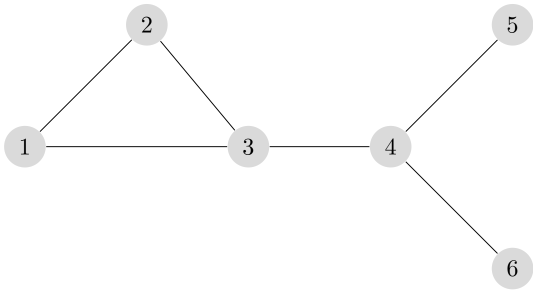


Fig. 1 A block graph with four 1-clique partitions

$$3: \Sigma_{A\cup C,B\cup C} = \begin{bmatrix} \sigma_{14} & \sigma_{16} \\ \sigma_{24} & \sigma_{26} \\ \sigma_{34} & \sigma_{36} \\ \sigma_{44} & \sigma_{46} \\ \sigma_{45} & \sigma_{56} \end{bmatrix}, \quad 4: \Sigma_{A\cup C,B\cup C} = \begin{bmatrix} \sigma_{14} & \sigma_{15} \\ \sigma_{24} & \sigma_{25} \\ \sigma_{34} & \sigma_{35} \\ \sigma_{44} & \sigma_{45} \\ \sigma_{46} & \sigma_{56} \end{bmatrix}.$$

The ideal $CI_G = P_G$ is the ideal generated by the 2×2 minors of all four matrices:

$$\begin{aligned} CI_G = \langle & \sigma_{13}\sigma_{24} - \sigma_{14}\sigma_{23}, \sigma_{13}\sigma_{25} - \sigma_{15}\sigma_{23}, \sigma_{13}\sigma_{26} \\ & - \sigma_{16}\sigma_{23}, \sigma_{14}\sigma_{25} - \sigma_{15}\sigma_{24}, \sigma_{23}\sigma_{34} - \sigma_{24}\sigma_{33}, \\ & \sigma_{23}\sigma_{35} - \sigma_{25}\sigma_{33}, \sigma_{23}\sigma_{36} - \sigma_{26}\sigma_{33}, \sigma_{24}\sigma_{35} \\ & - \sigma_{25}\sigma_{34}, \sigma_{24}\sigma_{36} - \sigma_{26}\sigma_{34}, \sigma_{25}\sigma_{36} - \sigma_{26}\sigma_{35}, \\ & \sigma_{13}\sigma_{34} - \sigma_{14}\sigma_{33}, \sigma_{13}\sigma_{35} - \sigma_{15}\sigma_{33}, \sigma_{13}\sigma_{36} \\ & - \sigma_{16}\sigma_{33}, \sigma_{14}\sigma_{35} - \sigma_{15}\sigma_{34}, \sigma_{14}\sigma_{36} - \sigma_{16}\sigma_{34}, \\ & \sigma_{15}\sigma_{36} - \sigma_{16}\sigma_{35}, \sigma_{14}\sigma_{45} - \sigma_{15}\sigma_{44}, \sigma_{14}\sigma_{46} \\ & - \sigma_{16}\sigma_{44}, \sigma_{15}\sigma_{46} - \sigma_{16}\sigma_{45}, \sigma_{24}\sigma_{45} - \sigma_{25}\sigma_{44}, \\ & \sigma_{24}\sigma_{46} - \sigma_{26}\sigma_{44}, \sigma_{25}\sigma_{46} - \sigma_{26}\sigma_{45}, \sigma_{34}\sigma_{45} \\ & - \sigma_{35}\sigma_{44}, \sigma_{34}\sigma_{46} - \sigma_{36}\sigma_{44}, \sigma_{35}\sigma_{46} - \sigma_{36}\sigma_{45}, \\ & \sigma_{14}\sigma_{56} - \sigma_{16}\sigma_{45}, \sigma_{24}\sigma_{56} - \sigma_{26}\sigma_{45}, \sigma_{34}\sigma_{56} \\ & - \sigma_{36}\sigma_{45}, \sigma_{44}\sigma_{56} - \sigma_{46}\sigma_{45}, \sigma_{14}\sigma_{56} - \sigma_{15}\sigma_{46}, \\ & \sigma_{24}\sigma_{56} - \sigma_{25}\sigma_{46}, \sigma_{34}\sigma_{56} - \sigma_{35}\sigma_{46}, \sigma_{44}\sigma_{56} \\ & - \sigma_{45}\sigma_{46}, \sigma_{14}\sigma_{26} - \sigma_{16}\sigma_{24}, \sigma_{15}\sigma_{26} - \sigma_{16}\sigma_{25}\rangle. \end{aligned}$$

The history of trying to characterize constraints on the covariance matrices in Gaussian graphical models goes back to Kelley (1935) and the discovery of the pentad constraints in the factor analysis model. Since then, the study of the constraints on Gaussian graphical models has seen many results including the deeper study of the factor analysis model in Drton et al. (2007), the study of directed graphical models and characterization of tree models in Sullivant (2008), and the complete characterization of the determinantal constraints that apply to Gaussian graphical models in Sullivant et al. (2010).

The study of the generators of the ideals P_G is an important problem for constraint-based inference for inferring the structure of the underlying graph from data. Elements of the vanishing ideal are tested to determine if the graph has certain underlying features, which are then used to reconstruct the entire graph. A prototypical example of this method is the TETRAD procedure in Spirtes et al. (2000) which specifically tests the degree 2 generators (tetrads) of the vanishing ideals of Gaussian graphical models for directed graphs. Our main result in this paper gives a characterization of which undirected graphs the tetrads are sufficient to characterize all distributions from the model, and is a key structural result for trying to use constraint based inference for undirected Gaussian graphical models. Developing characterizations of the vanishing ideals of Gaussian graphical models by higher order constraints (for example, determinantal constraints in Drton et al. (2008) and Sullivant et al. (2010)) has the potential to extend constraint-based inference beyond tetrad constraints.

This paper is organized as follows. We give two counterexamples to Conjecture 1 in Sect. 2. In Sect. 3 we define a rational map ρ and its pullback map ρ^*, whose kernel is the ideal P_G. We review properties of block graphs including the existence of a unique shortest path. Using this uniqueness property, we define the "shortest path map" ψ and the initial term map ϕ and show that the two maps have the same kernel. We prove that the kernel of ψ is equal to the ideal CI_G for block graphs with one central vertex in Sect. 4. This result is generalized for all block graphs in Sect. 5. Finally, in Sect. 6 we put all the pieces together to prove Theorem 1 using the results proved in the previous sections. We end the section by showing that the set F forms a SAGBI basis (Subalgebra Analog to Gröbner Basis for Ideals) using the initial term map.

2 Counterexamples to Conjecture 1

We first begin with some counterexamples to Conjecture 1. Initial counterexamples suggest a modification of Conjecture 1 might be true, but we show that that strengthened version is also false. This last counterexample suggests that it is unlike that there is a repair for the conjecture.

Example 2 Let $G = ([6], E)$ be the graph as shown in Fig. 2. Here, $A = \{1, 2\}, B = \{4, 5, 6\}$ and $C = \{3\}$. Computing the ideals P_G and $P_{G_1} + P_{G_2} + \langle 2 \times 2 \text{ minors of } \Sigma_{A\cup C, B\cup C}\rangle$, we get

$$\begin{aligned} P_G = \langle & \sigma_{14}\sigma_{25}\sigma_{46} - \sigma_{14}\sigma_{26}\sigma_{45} - \sigma_{15}\sigma_{24}\sigma_{46} + \sigma_{15}\sigma_{26}\sigma_{44} + \sigma_{16}\sigma_{24}\sigma_{45} - \sigma_{16}\sigma_{25}\sigma_{44}, \\ & \sigma_{24}\sigma_{45}\sigma_{56} - \sigma_{24}\sigma_{46}\sigma_{55} - \sigma_{25}\sigma_{44}\sigma_{56} + \sigma_{25}\sigma_{46}\sigma_{45} + \sigma_{26}\sigma_{44}\sigma_{55} - \sigma_{26}\sigma_{45}^2\rangle \\ & + P_{G_1} + P_{G_2} + \langle 2 \times 2 \text{ minors of } \Sigma_{A\cup C, B\cup C}\rangle. \end{aligned}$$

Note that even for some small block graphs Conjecture 1 is false.

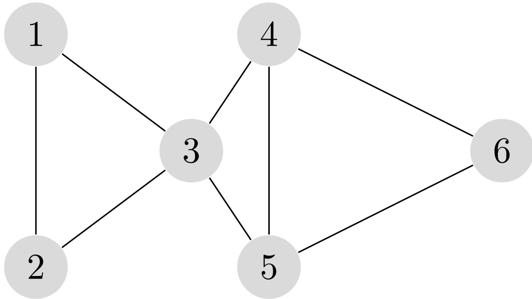


Fig. 2 A counterexample to Sturmfels-Uhler conjecture

Example 3 Consider the graph $G = ([4], E)$ which is a path of length 4. Taking $c = \{3\}$, we get a decomposition of G into G_1 and G_2 which are paths of length 3 and 2, respectively. A quick calculation in Macaulay2 [Grayson and Stillman (2017)] shows that $P_G = CI_G$ is generated by 5 quadratic binomials. However,

$$P_{G_1} + P_{G_2} + \langle 2 \times 2\text{-minors of } \Sigma_{\{1,2,3\},\{3,4\}} \rangle$$

has only 4 minimal generators.

Although P_G is not equal to $P_{G_1} + P_{G_2} + \langle 2 \times 2 \text{ minors of } \Sigma_{A \cup C, B \cup C} \rangle$ in these examples, we observe that the extra generators of P_G are also determinantal conditions arising from submatrices of Σ. Furthermore, they can be seen as being implied by the original rank conditions in P_{G_1} and P_{G_2} plus the rank conditions that are implied by $\langle 2 \times 2 \text{ minors of } \Sigma_{A \cup C, B \cup C} \rangle$.

For instance, in Example 3, the ideal $R_G = P_{G_1} + P_{G_2} + \langle 2 \times 2\text{-minors of } \Sigma_{\{1,2,3\},\{3,4\}} \rangle$ is generated by the 2×2 minors of the two matrices

$$\begin{pmatrix} \sigma_{12} & \sigma_{13} \\ \sigma_{22} & \sigma_{23} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \sigma_{13} & \sigma_{14} \\ \sigma_{23} & \sigma_{24} \\ \sigma_{33} & \sigma_{34} \end{pmatrix},$$

whereas the P_G is generated by the 2×2 minors of the two matrices.

$$\begin{pmatrix} \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{22} & \sigma_{23} & \sigma_{24} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \sigma_{13} & \sigma_{14} \\ \sigma_{23} & \sigma_{24} \\ \sigma_{33} & \sigma_{34} \end{pmatrix}.$$

However, we can take the generators R_G and, assuming that σ_{33} is not zero (which is valid since Σ is positive definite), we see that this implies that

$$\begin{pmatrix} \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{22} & \sigma_{23} & \sigma_{24} \end{pmatrix}$$

must be a rank 1 matrix.

Similarly, in Example 2, we know that $(\{3\},\{6\},\{4,5\})$ is a separating partition for the subgraph G_2. So, the ideal $J_{\{3\}\perp\!\!\!\perp\{6\}|\{4,5\}}$ is contained in P_{G_2}, which implies that rank of the submatrix $\Sigma_{\{3,4,5\},\{4,5,6\}}$ is 2. Similarly, $(\{1,2\},\{4,5,6\},\{3\})$ is a separating partition of G, which implies that rank of the submatrix $\Sigma_{\{1,2,3\},\{3,4,5,6\}}$ is 1. Now, as $\Sigma_{\{1,2,3\},\{4,5,6\}}$ is a submatrix of $\Sigma_{\{1,2,3\},\{3,4,5,6\}}$, we can say that $\Sigma_{\{1,2,3\},\{4,5,6\}}$ also has rank 1. Hence, from these two rank constraints and the added assumption that σ_{33} is not zero we can conclude that the submatrix $\Sigma_{\{1,2,4,5\},\{4,5,6\}}$ has rank 2.

The details of these examples suggest that a better version of the conjecture might be

$$P_G = \mathrm{Lift}(P_{G_1}) + \mathrm{Lift}(P_{G_2}) + \langle 2\times 2 \text{ minors of } \Sigma_{A\cup C,B\cup C}\rangle.$$

Here, $\mathrm{Lift}(P_{G_1})$ denotes some operation that takes the generators of P_{G_1} and extends them to the whole graph, analogous to how the toric fiber product in Sullivant (2007) lifts generators for reducible hierarchical models on discrete variables (Dobra and Sullivant 2004; Hoşten and Sullivant 2002). We do not make precise what this lifting operation could be, because if it preserves the degrees of generating sets the following example shows that no precise version of this notion could make this conjecture be true.

Example 4 Let $G = ([7], E)$ be the graph as shown in Fig. 3 and let (A, B, C) be the partition $(\{1,2,3\},\{5,6,7\},\{4\})$. Computing the vanishing ideal, we get $P_G = CI_G$, but that among the minimal generators of P_G is one degree 4 polynomial m where

$$\begin{aligned} m = {} & \sigma_{17}^2\sigma_{23}\sigma_{56} - \sigma_{13}\sigma_{17}\sigma_{27}\sigma_{56} - \sigma_{12}\sigma_{17}\sigma_{37}\sigma_{56} + \sigma_{11}\sigma_{27}\sigma_{37}\sigma_{56} - \sigma_{16}\sigma_{17}\sigma_{23}\sigma_{57} \\ & + \sigma_{13}\sigma_{16}\sigma_{27}\sigma_{57} + \sigma_{12}\sigma_{16}\sigma_{37}\sigma_{57} - \sigma_{11}\sigma_{26}\sigma_{37}\sigma_{57} - \sigma_{15}\sigma_{17}\sigma_{23}\sigma_{67} + \sigma_{13}\sigma_{15}\sigma_{27}\sigma_{67} \\ & + \sigma_{12}\sigma_{15}\sigma_{37}\sigma_{67} - \sigma_{11}\sigma_{25}\sigma_{37}\sigma_{67} - \sigma_{12}\sigma_{13}\sigma_{57}\sigma_{67} + \sigma_{11}\sigma_{23}\sigma_{57}\sigma_{67} + \sigma_{15}\sigma_{16}\sigma_{23}\sigma_{77} \\ & - \sigma_{13}\sigma_{15}\sigma_{26}\sigma_{77} - \sigma_{12}\sigma_{15}\sigma_{36}\sigma_{77} + \sigma_{11}\sigma_{25}\sigma_{36}\sigma_{77} + \sigma_{12}\sigma_{13}\sigma_{56}\sigma_{77} - \sigma_{11}\sigma_{23}\sigma_{56}\sigma_{77}. \end{aligned}$$

As both P_{G_1} and P_{G_2} are generated by polynomials of degree 3, this degree 4 polynomial could not be obtained from a degree preserving lifting operation.

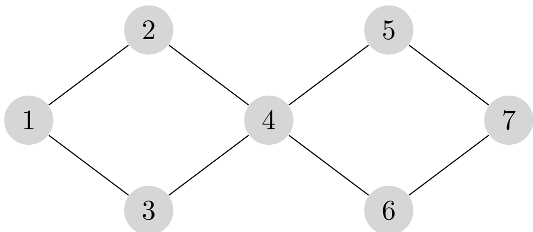


Fig. 3 A 1-clique sum of two graphs with a degree 4 generator

3 Shortest path in block graphs

Our goal for the rest of the paper is to prove Theorem 1. To do this, we need to phrase some parts in the language of commutative algebra. The vanishing ideal is the kernel of a certain ring homomorphism, or the presentation ideal of a certain $\mathbb{R}$-algebra. We will show that we can pass to a suitable initial algebra and analyze the combinatorics of the resulting toric ideal. This is proven in this section and those that follow.

We begin this section by giving an overview of toric ideals. We then define a rational map ρ such that the kernel of its pullback map gives us the ideal P_G. We also show the existence of a unique shortest path between any two vertices of a block graph. This property allows us to define the "shortest path map".

Let $\mathcal{A} = \{a_1, a_2, \dots, a_n\}$ be a fixed subset of $\mathbb{Z}^d$. We consider the homomorphism

$$\pi : \mathbb{N}^n \to \mathbb{Z}^d, \quad u = (u_1, \dots, u_n) \mapsto u_1 a_1 + \cdots + u_n a_n.$$

This map π lifts to a homomorphism of subgroup algebras:

$$\hat{\pi} : \mathbb{R}[x_1, \dots, x_n] \to \mathbb{R}[t_1, \dots, t_d, t_1^{-1}, \dots, t_d^{-1}], \quad x_i \mapsto t^{a_i}.$$

The kernel of $\hat{\pi}$ is called the *toric ideal* of $\mathcal{A}$. By Lemma 4.1 of Sturmfels (1996) we know that the toric ideal can be generated by the set of binomials of the form

$$\{x^u - x^v : u, v \in \mathbb{N}^n \text{ with } \pi(u) = \pi(v)\}.$$

From the construction above we observe that any monomial map can be written as $\hat{\pi}$ for some given set of vectors $\mathcal{A}$. This gives us that the kernel of every monomial map is a toric ideal.

Now, let $\mathbb{R}[K] = \mathbb{R}[k_{11}, k_{12}, \dots, k_{nn}]$ denote the polynomial ring in the entries of the concentration matrix K, and $\mathbb{R}(K)$ its fraction field.

We define the rational map $\rho : \mathcal{L} \dashrightarrow \mathcal{L}^{-1}$ as follows:

$$\rho(K) = \rho(k_{11}, k_{12}, \dots, k_{nn}) = (\rho_{11}(k_{11}, k_{12}, \dots, k_{nn}),$$
$$\rho_{12}(k_{11}, k_{12}, \dots, k_{nn}), \dots, \rho_{nn}(k_{11}, k_{12}, \dots, k_{nn})),$$

where $\rho_{ij} \in \mathbb{R}(K)$ is the (i, j) coordinate of K^{-1}. The rational map does not yield a well defined function from $\mathcal{L}$ to $\mathcal{L}^{-1}$ as every matrix in $\mathcal{L}$ is not invertible (chapter 3, Hassett 2007). Also note that the definition of ρ depends on the underlying graph G, since the zero pattern of K is determined by G.

The *pull-back* map of ρ is

$$\rho^* : \mathbb{R}[\Sigma] \to \mathbb{R}(K), \quad \sigma_{ij} \mapsto \rho_{ij}(K).$$

So, for each $p \in \mathbb{R}[\Sigma]$ and $K \in \mathcal{L}$,

$$\rho^*(p)(K) = p \circ \rho(K) = p(\rho_{11}(K), \rho_{12}(K), \dots, \rho_{nn}(K)).$$

Hence, we have

$$P_G = \mathcal{I}(\mathcal{L}^{-1}) = \ker(\rho^*).$$

For a given graph $G = ([n], E)$, let $f_{ij} \in \mathbb{R}[K]$ be the polynomial defined as $\det(K)$ times the (i, j) coordinate of the matrix K^{-1}. Let $F = \{f_{ij} : 1 \le i \le j \le n\}$. So, the map ρ^* can be written as

$$\rho^* : \mathbb{R}[\Sigma] \to \mathbb{R}(K) \quad \rho^*(\sigma_{ij}) = \frac{1}{\det(K)} \cdot f_{ij}.$$

As $1/\det(K)$ is a constant which is present in the image of every σ_{ij}, removing that factor from every image would not change the kernel of ρ^*. Hence, we change the map ρ^* as

$$\rho^* : \mathbb{R}[\Sigma] \to \mathbb{R}[F], \quad \rho^*(\sigma_{ij}) = f_{ij},$$

where $\mathbb{R}[F] = \mathbb{R}[f_{11}, f_{12}, \dots, f_{nn}] \subseteq \mathbb{R}[K]$.

Example 5 Let $G = ([4], E)$ be a graph with 4 vertices as shown in Fig 4. The matrices Σ and K for this graph are:

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_{44} \end{bmatrix}, \qquad K = \begin{bmatrix} k_{11} & k_{12} & k_{13} & 0 \\ k_{12} & k_{22} & k_{23} & 0 \\ k_{13} & k_{23} & k_{33} & k_{34} \\ 0 & 0 & k_{34} & k_{44} \end{bmatrix}.$$

The ideal P_G can be calculated by using the equation $\Sigma \cdot K = Id_4$ and eliminating the K variables.

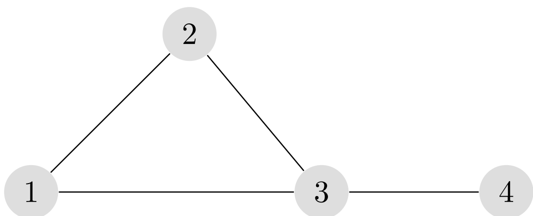
$$\begin{aligned} \langle \Sigma \cdot K - Id_4 \rangle = \langle & \sigma_{11}k_{11} + \sigma_{12}k_{12} + \sigma_{13}k_{13} - 1, \sigma_{11}k_{12} + \sigma_{12}k_{22} + \sigma_{13}k_{23}, \dots, \\ & \sigma_{14}k_{13} + \sigma_{24}k_{23} + \sigma_{34}k_{33} + \sigma_{44}k_{34}, \sigma_{34}k_{33} + \sigma_{44}k_{44} - 1 \rangle. \end{aligned}$$

Eliminating the K variables, we get

$$\begin{aligned} P_G = \langle \Sigma \cdot K - Id_4 \rangle \cap \mathbb{R}[\Sigma] = & \langle \sigma_{13}\sigma_{34} - \sigma_{14}\sigma_{33}, \sigma_{23}\sigma_{34} \\ & - \sigma_{24}\sigma_{33}, \sigma_{14}\sigma_{23} - \sigma_{13}\sigma_{24} \rangle. \end{aligned}$$

From the map ρ^*, we have

Fig. 4 A block graph with a single 1-clique sum decomposition



$$
\begin{aligned}
f_{11} &= \underline{k_{22}k_{33}k_{44}} - k_{22}k_{34}^2 - k_{23}^2k_{44} \\
f_{22} &= \underline{k_{11}k_{33}k_{44}} - k_{11}k_{34}^2 - k_{13}^2k_{44} \\
f_{33} &= \underline{k_{11}k_{22}k_{44}} - k_{44}k_{12}^2 \\
f_{44} &= \underline{k_{11}k_{22}k_{33}} - k_{11}k_{23}^2 - k_{12}^2k_{33} \\
&\quad + k_{12}k_{13}k_{23} + k_{13}k_{12}k_{23} - k_{13}^2k_{22}
\end{aligned}
\qquad
\begin{aligned}
f_{12} &= -\underline{k_{12}k_{33}k_{44}} - k_{12}k_{34}^2 - k_{23}k_{13}k_{44} \\
f_{13} &= -\underline{k_{13}k_{22}k_{44}} + k_{12}k_{23}k_{44} \\
f_{14} &= \underline{k_{13}k_{34}k_{22}} - k_{12}k_{23}k_{34} \\
f_{23} &= -\underline{k_{23}k_{11}k_{44}} + k_{12}k_{13}k_{44} \\
f_{24} &= \underline{k_{23}k_{34}k_{11}} - k_{34}k_{13}k_{12} \\
f_{34} &= -\underline{k_{34}k_{11}k_{22}} + k_{34}k_{12}^2
\end{aligned}
\tag{2}
$$

where f_{ij} is $\det(K)$ times the (i, j) coordinate of K^{-1}. Evaluating the kernel of ρ^*, we get

$$
\ker(\rho^*) \;=\; \langle \sigma_{13}\sigma_{34} - \sigma_{14}\sigma_{33}, \sigma_{23}\sigma_{34} - \sigma_{24}\sigma_{33}, \sigma_{14}\sigma_{23} - \sigma_{13}\sigma_{24} \rangle
$$

which is same as the ideal P_G. Note that G is a block graph with a single 1-clique sum decomposition. As the generators of P_G are the 2×2 minors of $\Sigma_{\{1,2,3\},\{3,4\}}$, the conjecture holds for this example.

Observe that in Example 5, each f_{ij} contains a monomial which corresponds to the shortest path from i to j in the graph G along with loops at the vertices not in the path. For example, f_{24} has the monomial $k_{23}k_{34}k_{11}$ where $k_{23}k_{34}$ corresponds to the shortest path from 2 to 4 and k_{11} corresponds to the loop at the vertex 1. In the (2), the underlined term is this special term.

This turns out to be important in our proofs, and we formalize this observation in Proposition 3. We now look at some properties of block graphs and 1-clique partitions in order to prove the existence of shortest paths.

Proposition 2 *If G is a block graph, then for any two vertices i and j there exists a unique shortest path in G connecting them. Further, if (A, B, C) is a 1-clique partition of G with $c \in C$ and if $i \in A$ and $j \in B$, then the unique shortest path from i to j can be decomposed into the unique shortest paths from i to c and c to j.*

Proof We prove this by applying induction on the number of vertices in G. If i and j are connected by a single edge, then that is the unique shortest path. If they are not connected by a single edge, then there exists a 1-clique partition (A, B, C) with $C = \{c\}$ which separates them. But as $A \cup C$ and $B \cup C$ are also block graphs and have fewer vertices than G, by induction there exist unique shortest paths from i to c and from c to j. But as any path from i to j must pass through c, the concatenation of the unique shortest paths from i to c and c to j would be the unique shortest path from i to j.

The second part follows from a property of unique shortest paths that if c is a point on the path, then the subpaths from i to c and c to j are the unique shortest paths from i to c and c to j, respectively. □

For the rest of the paper, we assume that G is a block graph and the shortest path from i to j in G is denoted by $i \leftrightarrow j$. We use $(i', j') \in i \leftrightarrow j$ to indicate that the edge (i', j') appears in the path $i \leftrightarrow j$. We let $\ell(i, j)$ denote the length of the shortest path from i to j. We now state a result from Jones and West (2005) which will be used to prove Proposition 3.

Theorem 2 (Theorem 1, Jones and West 2005) *Consider an n-dimensional multivariate normal distribution with a finite and non-singular covariance matrix Σ, with precision matrix $K = \Sigma^{-1}$. Let K determine the incidence matrix of a finite, undirected graph on vertices $\{1, \dots, n\}$, with nonzero elements in K corresponding to edges. The element of K corresponding to the covariance between variables x and y can be written as a sum of path weights over all paths in the graph between x and y:*

$$\sigma_{xy} = \sum_{P \in \mathscr{P}_{xy}} (-1)^{m+1} k_{p_1 p_2} k_{p_2 p_3} \dots k_{p_{m-1} p_m} \frac{\det(K_{\setminus P})}{\det(K)},$$

where $\mathscr{P}_{xy}$ represents the set of paths between x and y, so that $p_1 = x$ and $p_m = y$ for all $P \in \mathscr{P}_{xy}$ and $K_{\setminus P}$ is the matrix with rows and columns corresponding to variables in the path P omitted, with the determinant of a zero-dimensional matrix taken to be 1.

Proposition 3 *Let $G = ([n], E)$ be a block graph with the corresponding concentration matrix K. If f_{xy} denote* $\det(K)$ *times the (x, y) coordinate of K^{-1}, then f_{xy} has the monomial*

$$(-1)^{\ell(i,j)} \prod_{(x',y') \in x \leftrightarrow y} k_{x'y'} \prod_{t \notin x \leftrightarrow y} k_{tt}$$

as one of its terms. Furthermore, this term has the highest number of diagonal entries k_{tt} among all the monomials of f_{xy}.

Proof From Theorem 2, we have

$$f_{xy} = \det(K) \cdot \sigma_{xy} = \sum_{P \in \mathscr{P}_{xy}} (-1)^{m+1} k_{p_1 p_2} k_{p_2 p_3} \dots k_{p_{m-1} p_m} \det(K_{\setminus P}).$$

From Proposition 2 we know that if G is a block graph, then for any two vertices x and y, there exists a unique shortest path between x and y. If $z \in x \leftrightarrow y$ with $z \neq x, y$, then there exists a 1-clique partition (A, B, C) of G with $C = \{z\}$ and $x \in A, y \in B$. By the definition of 1-clique partition we know that any path from x to y must pass through z. As z is arbitrarily chosen, any path in G from x to y must pass through all the vertices in $x \leftrightarrow y$. This gives us that the unique shortest path has the least number of vertices among all the other paths from x to y. So, the matrix $K_{\setminus x \leftrightarrow y}$ has the highest dimension among all the other matrices $K_{\setminus P}, P \in \mathscr{P}_{xy}$.

Now, for any $P \in \mathscr{P}_{xy}$, $\det(K_{\setminus P})$ contains the monomial $\prod_{t \notin P} k_{tt}$ as G is assumed to have self loops. This monomial has the highest number of diagonals among all the monomials in $\det(K_{\setminus P})$ as the degree of $\det(K_{\setminus P})$ is same as the degree of $\prod_{t\notin P} k_{tt}$. So, the monomial

$$\prod_{(x',y')\in P} k_{x'y'} \prod_{t\notin P} k_{tt}$$

has the highest number of diagonal terms among all the monomials in $\prod_{(x',y')\in P} k_{x'y'} \det(K_{\setminus P})$. As $K_{\setminus x\leftrightarrow y}$ has the highest dimension, we can conclude that the monomial

$$\prod_{(x',y')\in x\leftrightarrow y} k_{x'y'} \prod_{t\notin x\leftrightarrow y} k_{tt}$$

has the maximum number of diagonal terms among all the monomials in f_{xy}. $\square$

We call the monomial defined above as the *shortest path monomial* of f_{ij}. As the shortest path monomial in each f_{ij} has the highest power of diagonals k_{tt} among all the other monomials in f_{ij}, we can define a weight order on $\mathbb{R}[K]$ where the weight of any monomial is the number of diagonal entries of the monomial. The initial term of f_{ij} in this order will be precisely the shortest path monomial.

Definition 3 Let G be a block graph. Define the $\mathbb{R}$-algebra homomorphism

$$\phi : \mathbb{R}[\Sigma] \to \mathbb{R}[K], \quad \sigma_{ij} \mapsto \prod_{(i',j')\in i\leftrightarrow j} k_{i'j'} \prod_{t\notin i\leftrightarrow j} k_{tt}.$$

This monomial homomorphism is called the *initial term map*.

The map ϕ is the initial term map of ρ^*, but with the sign $(-1)^{\ell(i,j)}$ omitted. We will use this to show that the set F forms a SAGBI basis of $\mathbb{R}[F]$ by using this term order, as part of our proof of Theorem 1. This appears in Sect. 6. To do this we must spend some time proving properties of ϕ and $\ker\phi$.

Note that the kernel of ϕ is the same with or without the signs $(-1)^{\ell(i,j)}$. This is because the monomials that appear are graded by the number of diagonal terms that appear, which is also counted by the $(-1)^{\ell(i,j)}$. Any binomial relation $\sigma^u - \sigma^v \in \ker\phi$ much also lead to the same power of negative one on both sides of the equation.

From the standpoint of proving results about this monomial map based on shortest paths in a block graph, it turns out to be easier to work with a related map that we call the shortest path map.

Definition 4 Let $G = ([n], E)$ be a block graph. The *shortest path map* ψ is defined as

$$\psi : \mathbb{R}[\Sigma] \to \mathbb{R}[a_1, \dots, a_n, k_{12}, \dots, k_{n-1,n}] = \mathbb{R}[A, K]$$
$$\psi(\sigma_{ij}) = \begin{cases} a_i a_j \prod_{(i',j')\in i\leftrightarrow j} k_{i'j'} & i \neq j \\ a_i^2 & i = j. \end{cases}$$

Example 6 Let G be the graph in Example 5. Let ψ be the shortest path map and ϕ the initial monomial map as given in Definitions 3 and 4. So for example,

$$\phi(\sigma_{11}) = k_{22}k_{33}k_{44}, \phi(\sigma_{12}) = k_{12}k_{33}k_{44}, \ldots$$

$$\psi(\sigma_{11}) = a_1^2, \psi(\sigma_{12}) = a_1a_2k_{12}, \ldots.$$

As is typical for monomial parametrizations, we can represent them by matrices whose columns are the exponent vectors of the monomials appearing in the parametrization. In this case, we get the following matrices corresponding to ϕ and ψ, respectively.

$$M_\phi = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

$$M_\psi = \begin{bmatrix} 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

The rows of M_ϕ are ordered as $\{k_{11}, k_{22}, k_{33}, k_{44}, k_{12}, k_{13}, k_{23}, k_{34}\}$ and the rows of M_ψ are ordered as $\{a_1, a_2, a_3, a_4, k_{12}, k_{13}, k_{23}, k_{34}\}$.

In fact, these two monomial maps have the same kernel for block graphs.

Proposition 4 *Let G be a block graph and let ϕ and ψ be the initial term map and the shortest path map, , respectively. Then* $\ker(\psi) = \ker(\phi)$.

Proof Both $\ker(\phi)$ and $\ker(\psi)$ are toric ideals. To show that they have the same kernel, it suffices to show that the associated matrices of exponent vectors have the same kernel, or equivalently, that they have the same row span. Let M_ϕ and M_ψ denote those matrices. As $\psi(\sigma_{ij}) = a_ia_j \prod_{(i',j')\in i\leftrightarrow j} k_{i'j'}$ and $\phi(\sigma_{ij}) = \prod_{(i',j')\in i\leftrightarrow j} k_{i'j'} \prod_{s\notin i\leftrightarrow j} k_{ss}$, the rows corresponding to k_{ij} with $i \neq j$ remain the same in both the matrices. So, we only need to write the k_{ii} rows of M_ϕ as a linear combination of the rows of M_ψ and vice versa.

The row vector corresponding to k_{ii} in M_ϕ is 1 at the σ_{pq} coordinates where $i \notin p \leftrightarrow q$ and is 0 elsewhere. Similarly, the row vector corresponding to a_i in M_ψ

is 2 at the σ_{ii} coordinate, 1 at the σ_{pq} coordinates where either of the end points is i (either $p = i$ or $q = i$) and 0 elsewhere.

We observe that the k_{ii} rows of M_ϕ can be written as a linear combination of the rows of M_ψ using the following relation:

$$2k_{ii} = \sum_{j \neq i} a_j - \sum_{s:i \leftrightarrow s \text{ is an edge}} k_{is}. \tag{3}$$

Here, we are using k_{ii} to denote the row vector of M_ϕ corresponding to the indeterminate k_{ii}, and similarly for a_j and k_{is}. We have

$$\begin{aligned}\sum_{j \neq i} a_j &= \text{ paths ending at } i + 2 \text{ (paths not ending at } i) - i \leftrightarrow i,\\ \sum_{s:i \leftrightarrow s \text{ is an edge}} k_{is} &= \text{ paths ending at } i + 2 \text{ (paths containing } i \text{ but not ending at } i) - i \leftrightarrow i.\end{aligned}$$

So,

$$\sum_{j \neq i} a_j - \sum_{s:i \leftrightarrow s \text{ is an edge}} k_{is} = 2 \text{ (paths not containing } i) = 2k_{ii}.$$

As this relation is true for any i, the row space of M_ϕ is contained in the row space of M_ψ. So, $\ker(\psi) \subseteq \ker(\phi)$.

To get the reverse containment, we need to write the a_i rows of M_ψ as a linear combination of the rows of M_ϕ. From (3), we get

$$\sum_{j \neq i} a_j = 2k_{ii} + \sum_{s:i \leftrightarrow s \text{ is an edge}} k_{is}.$$

Writing these n equations in the matrix form, we get an $n \times n$ matrix in the left-hand side which has 0 in its diagonal entries and 1 elsewhere. As this matrix is invertible for any $n > 1$, we can conclude that the row space of M_ψ is contained in the row space of A. Hence, $\ker(\psi) = \ker(\phi)$. □

Our goal in the next two sections will be to characterize the vanishing ideal of the shortest path map for block graphs.

Definition 5 Let G be a block graph. Let $SP_G = \ker(\psi) = \ker(\phi)$ be the kernel of the shortest path map. This ideal is called the *shortest path ideal*.

As the shortest path map is a monomial map, we know that the shortest path ideal is a toric ideal. We will eventually show that $SP_G = CI_G = P_G$, however we find it useful to have different notation for these ideals while we have not yet proven the equality.

4 Shortest path map for block graphs with 1 central vertex

In this section, we show that $SP_G = CI_G$ in the case that G is a block graph with only one central vertex. This will be an important special case and tool for proving that $SP_G = CI_G$ for all block graphs, which we do in Sect. 5. Our proof for graphs with only one central vertex depends on reducing the study of the ideal SP_G in this case to related notions of edge rings in DeLoera et al. (1995) and Herzog et al. (2018).

Definition 6 If G is a block graph, a vertex c in G is called a *central vertex* if there exists a 1-clique partition (A, B, C) of G such that $C = \{c\}$.

Example 7 Let G be the block graph with 5 vertices as in Fig. 5. There are three possible 1-clique partitions of G, $(\{1,2\},\{4,5\},\{3\}), (\{1,2,4\},\{5\},\{3\})$ and $(\{1,2,5\},\{4\},\{3\})$. We see that 3 is the only central vertex of G as $C = \{3\}$ for all the three partitions. Now computing SP_G for this graph, we get

$$\begin{aligned}\ker(\psi) = \langle & \sigma_{34}\sigma_{35} - \sigma_{33}\sigma_{45}, \sigma_{24}\sigma_{35} - \sigma_{23}\sigma_{45}, \sigma_{14}\sigma_{35} - \sigma_{13}\sigma_{45}, \sigma_{25}\sigma_{34} - \sigma_{23}\sigma_{45},\\ & \sigma_{15}\sigma_{34} - \sigma_{13}\sigma_{45}, \sigma_{25}\sigma_{33} - \sigma_{23}\sigma_{35}, \sigma_{24}\sigma_{33} - \sigma_{23}\sigma_{34}, \sigma_{15}\sigma_{33} - \sigma_{13}\sigma_{35},\\ & \sigma_{14}\sigma_{33} - \sigma_{13}\sigma_{34}, \sigma_{15}\sigma_{24} - \sigma_{14}\sigma_{25}, \sigma_{15}\sigma_{23} - \sigma_{13}\sigma_{25}, \sigma_{14}\sigma_{23} - \sigma_{13}\sigma_{24}\rangle.\end{aligned}$$

We observe that in Example 7, none of the generators of SP_G contain the terms $\sigma_{12}, \sigma_{11}, \sigma_{22}, \sigma_{44}$ and σ_{55}. These terms correspond to the edges in G which cannot be separated by any 1-clique partition of G. This property is true for all block graphs with one central vertex as we prove it in the next Lemma.

Lemma 1 *Let G be a block graph with one central vertex c and let D be the set of variables σ_{pq}, where the shortest path $p \leftrightarrow q$ does not intersect c. Then, none of the variables appearing in D appear in any of the minimal generators of the kernel of ψ.*

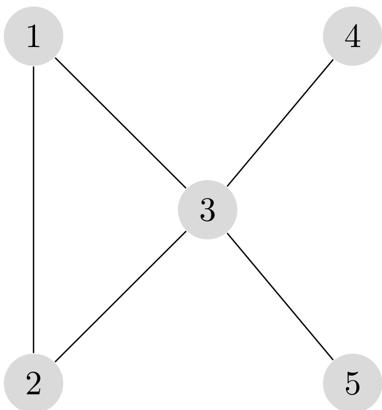


Fig. 5 A block graph with exactly one central vertex

Proof Since ψ is a monomial parametrization, the kernel of ψ is a homogeneous binomial ideal. Let

$$f = \sigma^u - \sigma^v$$

be an arbitrary binomial in any generating set for the kernel of SP_G. In particular, this implies that σ^u and σ^v have no common factors. Suppose by way of contradiction that σ_{pq} is some variable in D that divides one of the terms of f, say σ^u. Then, $\psi(\sigma^u)$ would have k_{pq} as a factor. But k_{pq} appears only in the image of σ_{pq} as no other shortest path between any two vertices in G contains the edge (p, q). This would imply that σ_{pq} is also a factor of σ^v contradicting the fact that σ^u and σ^v have no common factors.

Similarly, if σ_{pp} is a factor of σ^u where p is not the central vertex, then $\psi(\sigma^u)$ would have a_p^2 as a factor. In order to have a_p^2 as a factor of $\psi(\sigma^v)$, it would require two variables in σ^v to have p as one of their end points. As p is not a central vertex, we will have k_{cp}^2 as a factor of $\psi(\sigma^v)$. But then this means that there must be two variables in σ^u that touch vertex p. Which in turn forces another factor of a_p^2 to divide $\psi(\sigma^u)$. Which in turn forces another two variables in σ^v to touch vertex p, and so on. This process never terminates, showing that it is impossible that σ_{pp} is a factor of σ^u.

Hence, we can conclude that none of the variables in D appear in any of the generators of SP_G. □

Note that the proof of Lemma 1 also applies to any block graph with multiple central vertices. Hence, we can eliminate some of the variables in the computation of the shortest path ideal.

We let $\mathbb{R}[\Sigma \backslash D]$ denote the polynomial ring with the variables D eliminated. Here, we are always taking D to the be set of variables corresponding to paths that do not touch the central vertex x. Lemma 1 shows that it suffices to consider the problem of finding a generating set of SP_G inside of $\mathbb{R}[\Sigma \backslash D]$.

The next step in our analysis of SP_G for block graphs with one central vertex will be to relate this ideal to a simplified parametrization which we can then relate to edge ideals.

Let G be a block graph with one central vertex. Consider the map

$$\hat{\psi} : \mathbb{R}[\Sigma \backslash D] \to \mathbb{R}[a], \quad \sigma_{ij} \mapsto a_i a_j.$$

Proposition 5 *Let G be a block graph with one central vertex. Then* $\ker \hat{\psi} = \ker \psi$.

Proof Note that because we only consider $\sigma_{pq} \in \mathbb{R}[\Sigma \backslash D]$ then any time $\psi(\sigma_{pq})$ contains k_{pc} it will automatically contain a_p as well, and vice versa. Hence, the $a_p k_{pc}$ always occurs as a factor together in $\psi(\sigma_{pq})$. So we can eliminate the k_{pc} from the parametrization without affecting the kernel of the homomorphism. □

In order to analyze $SP_G = \ker \hat{\psi} = \ker \psi$, we find it useful to first extend the map to all of $\mathbb{R}[\Sigma]$, where the kernel is well understood. In particular, we associate an edge in the graph K_n° to each variable in $\mathbb{R}[\Sigma]$, where K_n° denotes the complete graph

K_n with a loop added to each vertex. We embed K_n° in the plane so that the vertices are arranged to lie on a circle. We consider the map

$$\hat{\psi} : \mathbb{R}[\Sigma] \to \mathbb{R}[a], \quad \sigma_{ij} = a_i a_j$$

and its kernel $SP_{K_n^\circ} = \ker \hat{\psi}$. We describe a Gröbner basis for this ideal, based on the combinatorics of the embedding of the graph K_n°. We consider a pair of edges (i, j), (k, l) to be *intersecting* if the two edges share a vertex or the edges intersect each other in the circular embedding of K_n°.

The *circular distance* between two vertices of K_n is defined as the length of the shorter path among the two paths present along the edges of the n-gon. We define the *weight* of the variable σ_{ij} as the number of edges of K_n° that do not intersect the edge (i, j). Let $\prec$ denote any term order that refines the partial order on monomials specified by these weights. Now, for any pair of non-intersecting edges (i, j), (k, l) of K_n°, one of the pairs (i, k), (j, l) or $(i, l)(j, k)$ is intersecting. If (i, k), (j, l) is the intersecting pair, we associate the binomial $\sigma_{ij}\sigma_{kl} - \sigma_{ik}\sigma_{jl}$ with the non-intersecting pair of edges (i, j), (k, l). We denote by S' the set of all binomials obtained in this way.

Lemma 2 *For any binomial $\sigma_{ij}\sigma_{kl} - \sigma_{ik}\sigma_{jl}$, where (i, j), (k, l) are non-intersecting edges and (i, k), (j, l) intersect, the initial term with respect to $\prec$ corresponds to the non-intersecting edges in K_n°.*

Proof We divide the set of vertices in K_n° into four different parts (excluding the vertices i, j, k and l). Let P_1 denote the set of vertices that are present in the path between i and j along the edges of the n-gon that do not contain k and l. Similarly, let P_2, P_3 and P_4 denote the set of vertices between j and k, k and l and l and i, respectively. Let the cardinality of each P_i be p_i for $i = 1, 2, 3, 4$. Then, the weight of the four variables are as follows:

$$w(\sigma_{ij}) = \sum_{i=1}^{4} \binom{p_i}{2} + p_2 p_3 + p_2 p_4 + p_3 p_4 + 2(p_2 + p_3 + p_4) + 1 + (n - 2)$$
$$w(\sigma_{kl}) = \sum_{i=1}^{4} \binom{p_i}{2} + p_1 p_2 + p_1 p_4 + p_2 p_4 + 2(p_1 + p_2 + p_4) + 1 + (n - 2)$$
$$w(\sigma_{ik}) = \sum_{i=1}^{4} \binom{p_i}{2} + p_1 p_2 + p_3 p_4 + p_1 + p_2 + p_3 + p_4 + (n - 2)$$
$$w(\sigma_{jl}) = \sum_{i=1}^{4} \binom{p_i}{2} + p_1 p_4 + p_2 p_3 + p_1 + p_2 + p_3 + p_4 + (n - 2).$$

This gives us

$$w(\sigma_{ij}) + w(\sigma_{kl}) - (w(\sigma_{ik}) + w(\sigma_{jl})) = 2p_2 p_4 + 2(p_2 + p_4) + 2 > 0.$$

Hence, the initial term of $\sigma_{ij}\sigma_{kl} - \sigma_{ik}\sigma_{jl}$ with respect to $\prec$ is $\sigma_{ij}\sigma_{kl}$. Further, if $k = l$ then we have the binomial $\sigma_{ij}\sigma_{kk} - \sigma_{ik}\sigma_{jk}$ where

$$w(\sigma_{kk}) = \binom{n-1}{2} + n - 1 \text{ and}$$
$$w(\sigma_{jk}) = \sum_{i=1}^{4} \binom{p_i}{2} + p_1p_4 + p_1p_3 + p_3p_4 + 2(p_1 + p_3 + p_4) + 1 + (n-2).$$

This gives us

$$w(\sigma_{ij}) + w(\sigma_{kk}) - (w(\sigma_{ik} + w(\sigma_{jk})) = \sum_{i=2}^{4} \frac{p_i}{2} + 2(p_2p_3 + p_2p_4) + \frac{3}{2}(p_2 + p_3 + p_4) + p_2 + 4 > 0.$$

So, the initial term of $\sigma_{ij}\sigma_{kk} - \sigma_{ik}\sigma_{jk}$ with respect to $\prec$ is $\sigma_{ij}\sigma_{kk}$. □

Lemma 3 *Let S' be the set of binomials obtained from all the pairs of non-intersecting edges of K_n°. Then S' is the reduced Gröbner basis of $SP_{K_n^\circ}$ with respect to $\prec$.*

Proof By Lemma 2 we know that for any binomial $\sigma_{ij}\sigma_{kl} - \sigma_{il}\sigma_{jk} \in S'$, where (i, j), (k, l) are non-intersecting edges and (i, l), (j, k) intersect, the initial term with respect to $\prec$ corresponds to the non-intersecting edges in K_n°. Clearly, $\sigma_{ij}\sigma_{kl} - \sigma_{il}\sigma_{jk} \in SP_{K_n^\circ}$.

The proof follows the basic outline as the proof of Theorem 9.1 in Sturmfels (1996). For any even closed walk $\Gamma = (i_1, i_2, \dots, i_{2k-1}, i_{2k}, i_1)$ in K_n° we associate the binomial

$$b_\Gamma := \prod_{l=1}^{k} \sigma_{i_{2l-1}, i_{2l}} - \prod_{l=1}^{k} \sigma_{i_{2l}, i_{2l+1}}$$

which belongs to $SP_{K_n^\circ}$. To prove that S' is a Gröbner basis, it is enough to prove that the initial monomial of any binomial b_Γ is divisible by some monomial $\sigma_{ij}\sigma_{kl}$ which is the initial term of some binomial in S', where (i, j) and (k, l) are a pair of non-intersecting edges. Let there exist a binomial $b_\Gamma = \sigma^u - \sigma^v \in SP_{K_n^\circ}$ with $in_\prec(b_\Gamma) = \sigma^u$ which contradicts the assertion. Then assuming that b_Γ has minimal weight, we can say that each pair of edges appearing in σ^v intersects.

The edges of the walk are labeled as even or odd, where even edges look like (i_{2r}, i_{2r+1}) and the odd edges are of the form (i_{2r-1}, i_{2r}). We pick an edge (s, t) of the walk Γ which has the least circular distance between s and t. The edge (s, t) separates the vertices of K_n° except s and t into two disjoint sets P and Q where $|P| \geq |Q|$. We start Γ at $(s, t) = (i_1, i_2)$. From our assertion on b_Γ we have that each pair of odd (resp. even) edges intersect. Also, it can be proved that if P contains an odd vertex i_{2r-1}, then it contains all the subsequent odd vertices $i_{2r+1}, i_{2r+3}, \dots, i_{2k-1}$. As the circular distance between s and t is the least, we need to have i_3 to be in P. So, all the odd vertices except i_1 lie in P and all the even vertices lie in $Q \cup \{i_1, i_2\}$. This gives us that the two even edges (i_2, i_3) and (i_{2k}, i_1) do not intersect, which is a contradiction. □

Our goal next is to use Lemma 3, to prove that $SP_G = CI_G$ for block graphs with one central vertex. Recall that the set D consisted of all pairs σ_{ij} such that in the graph G $i \leftrightarrow j$ does not touch the central vertex. As the σ_{ij} appearing in D do not appear in any generators of SP_G, let us construct an associated subgraph of K_n° without those edges. Specifically, let G° be the graph obtained by removing the edges (i, j) from K_n° such that $\sigma_{ij} \in D$. Note that we choose an embedding of G° so that each maximal clique minus c forms a contiguous block on the circle. The placement of c can be anywhere that is between the maximal blocks.

Figure 6 illustrates the construction of the graph G° in an example.

Example 8 Let G be a block graph with 5 vertices in Fig. 6. There are 3 possible 1-clique partitions of G, each of them having $C = \{3\}$. The edges in K_5° which cannot be separated by any 1-clique partition of G are $D = \{(1, 2), (1, 1), (2, 2), (4, 4), (5, 5)\}$. So we remove them from K_5° to get G°.

Lemma 4 *For any non-intersecting pair of edges (i, j), (k, l) in G°, there exists a 1-clique partition (A, B, C) of G such that $i, l \in A \cup C$ and $j, k \in B \cup C$.*

Proof We first prove this for the non-intersecting edges (i, j), (k, l) with $i, j, k, l \neq c$. Without loss of generality we can assume that $i < j < k < l$. We know that for each edge (i, j) in G° there exists a 1-clique partition (A, B, C) of G such that $i \in A \cup C$ and $j \in B \cup C$. This implies that i and j (similarly k and l) lie in different maximal cliques of G. As the vertices of G° are labeled counter-clockwise, there are only three ways how the vertices i, j, k, l can be placed:

$$\text{(i)} \quad i, l \in C_1, j, k \in C_2, \quad \text{(ii)} \; i, l \in C_1, j \in C_2, k \in C_3,$$
$$(iii) \; i \in C_1, j \in C_2, k \in C_3, l \in C_4,$$

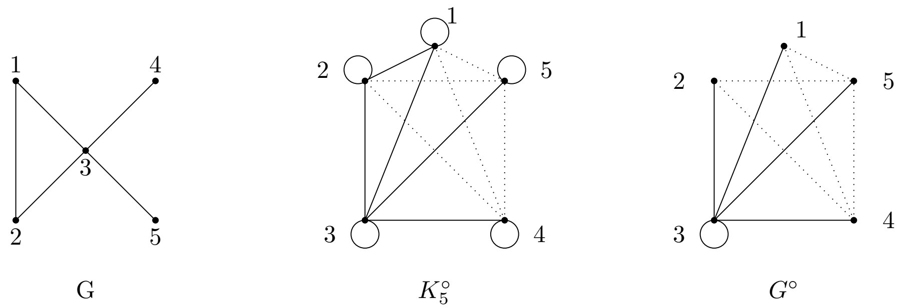


Fig. 6 Construction of the graph G°. The dark lines in K_5° correspond to the edges in G, whereas a dotted line between i and j tells us that there is no edge between i and j in G. The dotted line basically corresponds to the shortest path between the two vertices in G. Note that the addition of extra edges gives us K_5° and the deletion of some edges gives us G°

where C_i are the different maximal cliques of G. In all the three cases i and k (similarly j and l) are in different maximal cliques. Hence, there exists a 1-clique partition (A, B, C) such that $i, l \in A \cup C$ and $k, j \in B \cup C$.

A similar argument can be given for the non-intersecting edges (i, c), (k, l) and (c, c), (i, j). □

Lemma 5 *Let S' be the Gröbner basis for $SP_{K_n^\circ} \subseteq \mathbb{R}[\Sigma]$ as defined in Lemma 3. Then, the set $S' \cap \mathbb{R}[\Sigma \backslash D]$ forms a Gröbner basis for SP_G.*

Proof Let $g = \sigma^u - \sigma^v$ be an arbitrary binomial in $SP_G = \ker \hat{\psi}$. This implies that the initial term of g is contained in $\mathbb{R}[\Sigma \backslash D]$. Since S' is a Gröbner basis for $SP_{K_n^\circ}$ with respect to $\prec$, there must exist some $f \in S'$ such that $\mathrm{in}_\prec(f)$ divides $in_\prec(g)$. This gives us that the initial term of f is contained in $\mathbb{R}[\Sigma \backslash D]$.

So it is enough to show that for every $f \in S'$ whose leading term is in $\mathbb{R}[\Sigma \backslash D]$ is actually contained in $\mathbb{R}[\Sigma \backslash D]$. Let

$$f = \sigma_{ij}\sigma_{kl} - \sigma_{ik}\sigma_{jl}$$

be a binomial in S' whose leading term is contained in $\mathbb{R}[\Sigma \backslash D]$. Let $\sigma_{ij}\sigma_{kl}$ be the leading term. Then the edges (i, j), (k, l) are non-intersecting as the initial term of each binomial in S' corresponds to the non-intersecting edges. So by Lemma 4, there must exist a 1-clique partition (A, B, C) of G which separates the edges (i, j) and (k, l), that is, $i, l \in A \cup C$ and $j, k \in B \cup C$. This implies that (A, B, C) also separates the edges (i, k) and (j, l). Hence, we can say that $\sigma_{ik}, \sigma_{jl} \notin D$ and $\sigma_{ij}\sigma_{kl} - \sigma_{ik}\sigma_{jl} \in \mathbb{R}[\Sigma \backslash D]$. □

Now that we have all the required results, we prove the main result of this section.

Theorem 3 *Let G be a block graph with n vertices having only one central vertex. Then, the set of all 2×2 minors of $\Sigma_{A \cup C, B \cup C}$ for all possible 1-clique partitions (A, B, C) of G form a Gröbner basis for SP_G. In particular, $SP_G = CI_G$.*

Proof We rearrange the graph by placing the vertices in K_n° such that there is no intersection among the edges of G in $A \cup C$ and $B \cup C$ for any 1-clique partition (A, B, C) (with $C = \{c\}$). We complete the graph by drawing the remaining edges with dotted lines.

The complete graph K_n° gives us a partial term order on $\mathbb{R}[\Sigma]$ by defining the weight of the variable σ_{ij} as the number of edges of K_n° which do not intersect the edge (i, j). Let $\prec$ denote the term order that refines the partial order on monomials specified by the weights. Let S be the set of all 2×2 minors of $\Sigma_{A \cup C, B \cup C}$ for all possible 1-clique partitions of G. Any binomial in S has one of the three forms:

(i) $\sigma_{ij}\sigma_{kl} - \sigma_{ik}\sigma_{jl}$ with $i, l \in A \cup C$ and $j, k \in B \cup C$
(ii) $\sigma_{ij}\sigma_{kl} - \sigma_{il}\sigma_{jk}$ with $i, k \in A \cup C$ and $j, l \in B \cup C$
(iii) $\sigma_{il}\sigma_{jk} - \sigma_{ik}\sigma_{jl}$ with $i, j \in A \cup C$ and $k, l \in B \cup C$.

Here, (i, j), (k, l) and (i, l), (j, k) are the non-intersecting pairs of edges and $(i, k)(j, l)$ is the intersecting pair in G°. So any binomial in S of the form (i) or (iii) is contained in S'. If the binomial $\sigma_{ij}\sigma_{kl} - \sigma_{il}\sigma_{jk}$ (of form (ii)) is in S, then by Lemma 4 we know that the binomials $\sigma_{ij}\sigma_{kl} - \sigma_{ik}\sigma_{jl}$ and $\sigma_{il}\sigma_{jk} - \sigma_{ik}\sigma_{jl}$ are also in S. As

$$\sigma_{ij}\sigma_{kl} - \sigma_{il}\sigma_{jk} = \sigma_{ij}\sigma_{kl} - \sigma_{ik}\sigma_{jl} - (\sigma_{il}\sigma_{jk} - \sigma_{ik}\sigma_{jl}),$$

we can conclude that S and $S \cap S'$ generate the same ideal. Furthermore, the set $S \cap S'$ has the same initial terms as $S' \cap \mathbb{R}[\Sigma \backslash D]$ so this guarantees that S is a Gröbner basis for SP_G as well. □

5 The shortest path ideal for an arbitrary block graph

To generalize the statement in Theorem 3 for any arbitrary block graph, we further exploit the toric structure of the ideal SP_G. As SP_G is the kernel of a monomial map, it is a toric ideal, a prime ideal generated by binomials. Finding a generating set of SP_G is equivalent to finding a set of binomials that make some associated graphs connected. We use this perspective to prove that $SP_G = CI_G$.

From the shortest path map ψ, we can obtain the matrix M_ψ as shown in Example 6. So $SP_G = \ker(\psi)$ is the toric ideal of the matrix M_ψ as

$$\psi(\sigma^u) = t^{M_\psi u},$$

where $\sigma = (\sigma_{11}, \sigma_{12}, \dots, \sigma_{nn})$ and $t = (a_1, a_2, \dots, a_n, k_{12}, \dots, k_{n-1n})$.

Let $G = ([n], E)$ be a block graph. For any vector $b \in \mathbb{N}^{(n+|E|)}$, the *fiber* of M_ψ over b is defined as

$$M_\psi^{-1}(b) = \{u \in \mathbb{N}^{(n^2+n)/2} \, : \, M_\psi u = b\}.$$

As the columns of M_ψ are nonzero and nonnegative, $M_\psi^{-1}(b)$ is always finite for any $b \in \mathbb{N}^{(n+|E|)}$. Let $\mathcal{F}$ be any finite subset of $\ker_{\mathbb{Z}}(M_\psi)$. The *fiber graph* $M_\psi^{-1}(b)_{\mathcal{F}}$ is defined as follows:

(i) The nodes of this graph are the elements of $M_\psi^{-1}(b)$.
(ii) Two nodes u and u' are connected by an edge if $u - u' \in \mathcal{F}$ or $u' - u \in \mathcal{F}$.

The fundamental theorem of Markov bases connects the generating sets of toric ideals to connectivity properties of the fiber graphs. We state this explicitly in the case of the fiber graphs for the shortest path maps.

Theorem 4 (Thm 5.3, Sturmfels 1996) *Let $\mathcal{F} \subset \ker_{\mathbb{Z}}(M_\psi)$. The graphs $M_\psi^{-1}(b)_{\mathcal{F}}$ are connected for all $b \in \mathbb{N}M_\psi = \{\lambda_1 M_{\psi 1} + \cdots + \lambda_{n+|E|} M_{\psi n+|E|} \, : \, \lambda_i \in \mathbb{N}, M_{\psi i}$ are columns of $M_\psi\}$ if and only if the set $\{\sigma^{v^+} - \sigma^{v^-} \, : \, v \in \mathcal{F}\}$ generates the toric ideal SP_G.*

As we proved in Theorem 3 that the set of all 2×2 minors of $\Sigma_{A \cup C, B \cup C}$ for all possible 1-clique partitions of G form a Gröbner basis for $\ker(\psi)$ for all block graphs with one central vertex, by using Theorem 4 we can say that the graph $M_{\psi}^{-1}(b)_{\mathcal{F}}$ is connected for all $b \in \mathbb{N}M_{\psi}$. Here, $\mathcal{F}$ is the set of all 2×2 minors of $\Sigma_{A \cup C, B \cup C}$ in the vector form, for all possible 1-clique partitions of G.

So, to generalize the result in Theorem 3 for all block graphs, we need to show that $M_{\psi}^{-1}(b)_{\mathcal{F}}$ is connected for any $b \in \mathbb{N}M_{\psi}$. For a fixed b, let $u, v \in M_{\psi}^{-1}(b)_{\mathcal{F}}$. This implies that both $M_{\psi}u$ and $M_{\psi}v$ are equal to b, which gives us $\psi(\sigma^u - \sigma^v) = 0$. Therefore, it is enough to show that for any $f = \sigma^u - \sigma^v \in SP_G$, σ^u and σ^v are connected by the moves in $\mathcal{F}$.

Let G be a block graph with n vertices. Let $u \in \mathbb{N}^{(n^2+n)/2}$ which is a node in the graph of $M_{\psi}^{-1}(b)_{\mathcal{F}}$. We represent this u, or equivalently σ^u, as a graph in the following way: For each factor σ_{ij} of σ^u we draw the shortest path $i \leftrightarrow j$ along G with end points at i and j. For each σ_{ii} we draw a loop at the vertex i. Let $\deg_i(\sigma^u)$ denote the *degree* of a vertex i in σ^u which is defined to be the number of end points of paths in σ^u. We count the loops corresponding to σ_{ii} as having two endpoints at i.

If $f = \sigma^u - \sigma^v$ is a homogeneous binomial in SP_G, then $\psi(\sigma^u) = \psi(\sigma^v)$ if and only if the following conditions are satisfied:

(i) The graphs of σ^u and σ^v both have the same number of paths (as f is homogeneous),
(ii) The graphs of σ^u and σ^v have the same number of edges between any two adjacent vertices i and j (as the exponent of k_{ij} in $\psi(\sigma^u)$ gives the number of edges between i and j in the graph of σ^u),
(iii) The degree of any vertex in both the graphs is the same (as the exponent of a_i in $\psi(\sigma^u)$ gives us the degree of the vertex i in the graph of σ^u).

Next we show how to use the results from Sect. 4 to make moves that bring σ^u and σ^v closer together. This approach works by localizing the computations at each central vertex in the graph.

Let c be a central vertex in G. We define a map ρ_c between the set of vertices as follows:

$$\rho_c(i) = \begin{cases} c & i = c \\ i & i \text{ is adjacent to } c \\ i' & i' \text{ is adjacent to } c \text{ and lies in } i \leftrightarrow c. \end{cases}$$

Let G_c be the graph obtained by applying ρ_c to the vertices of G. Note that G can have multiple vertices mapped to a single vertex in G_c. The map ρ_c can also be seen as a map between $\mathbb{R}[\Sigma]$ to itself by the rule $\rho_c(\sigma_{ij}) = \sigma_{\rho_c(i)\rho_c(j)}$.

For a vector $u \in \mathbb{N}^{n(n+1)/2}$ and c a central vertex let u_c be the vector that extracts all the coordinates that correspond to shortest paths that touch c. That is,

$$u_c(ij) = \begin{cases} u(ij) & c \in i \leftrightarrow j \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 6 *Suppose that $\sigma^u - \sigma^v \in SP_G$ and let c be a central vertex of G. Then $\psi_{G_c}(\rho_c(\sigma^{u_c})) - \psi_{G_c}(\rho_c(\sigma^{v_c})) = 0$.*

Note that we use the notation ψ_{G_c} to denote that we use the ψ map associated to the graph G_c. However, the map ψ associated to G can be used since that will give the same result.

Proof We have

$$\rho_c(\sigma_{ij}) = \begin{cases} \sigma_{ij} & i,j \text{ are adjacent to } c \\ \sigma_{ic} & i \text{ is adjacent to } c, j = c \\ \sigma_{cj} & j \text{ is adjacent to } c, i = c \\ \sigma_{i'c} & i' \text{ is adjacent to } c \text{ and } i' \in i \leftrightarrow c, j = c \\ \sigma_{cj'} & j' \text{ is adjacent to } c \text{ and } j' \in j \leftrightarrow c, i = c \\ \sigma_{ij'} & i,j' \text{ are adjacent to } c \text{ and } j' \in c \leftrightarrow j \\ \sigma_{i'j} & i',j \text{ are adjacent to } c \text{ and } i' \in i \leftrightarrow c \\ \sigma_{i'j'} & i',j' \text{ are adjacent to } c \text{ and } i' \in i \leftrightarrow c, j' \in j \leftrightarrow c \\ \sigma_{i'i'} & i' \text{ is adjacent to } c \text{ and } i' \in i \leftrightarrow c \text{ and } j \leftrightarrow c. \end{cases}$$

We know that σ^u and σ^v have the same number of paths. Also, the degree of each vertex and the number of edges between any two adjacent vertices is the same. So, it is enough to show that $\rho_c(\sigma^{u_c})$ and $\rho_c(\sigma^{v_c})$ have the same number of paths and the degree of each vertex, number of edges between any two adjacent vertices is also the same.

$$\begin{aligned} \text{Number of paths in } \sigma^{u_c} = & \text{ number of paths in } \sigma^u \text{ ending at } c \\ + & \text{ number of paths containing } c \text{ but not ending at } c \\ = & \text{ degree of } a_c \text{ in } \psi(\sigma^u) + 1/2(\text{ number of variables of} \\ & \text{the form } k_{ic} \text{ in } \psi(\sigma^u) - \text{ degree of } a_c \text{ in } \psi(\sigma^u)) \\ = & \text{ number of paths in } \sigma^{v_c} \end{aligned}$$

The number of paths in σ^{u_c} and $\rho_c(\sigma^{u_c})$ are the same as ρ_c maps monomials of degree 1 to monomials of degree 1.

For any vertex s which adjacent to c, the degree of s in $\rho_c(\sigma^{u_c})$ is

$$\begin{aligned} \deg_s(\rho_c(\sigma^{u_c})) = & \text{ number of edges } s \leftrightarrow c \text{ in } \sigma^u \\ = & \text{ number of edges } s \leftrightarrow c \text{ in } \sigma^v \\ = & \, deg_s(\rho_c(\sigma^{v_c})). \end{aligned}$$

Now, for any two vertices i' and j' adjacent to c, the number of edges $i' \leftrightarrow j'$ in $\rho_c(\sigma^{u_c})$ is 0 as every path in $\rho_c(\sigma^{u_c})$ contains c. The number of edges $i' \leftrightarrow c$ in $\rho_c(\sigma^{u_c})$ is equal to the number of edges $i' \leftrightarrow c$ in σ^u, which is equal to the number of edges $i' \leftrightarrow c$ in σ^v.

Hence, we can conclude that $\psi_{G_c}(\rho_c(\sigma^{u_c})) - \psi_{G_c}(\rho_c(\sigma^{v_c})) = 0$. □

By Theorem 4 we know that we can reach from $\rho_c(\sigma^{u_c})$ to $\rho_c(\sigma^{v_c})$ by making a finite set of moves from the set of 2×2 minors of $\Sigma_{A\cup C, B\cup C}$, for all possible 1-clique partitions of G_c. But from the map ρ_c we have that for each move $\sigma_{i'j'}\sigma_{k'l'} - \sigma_{i'l'}\sigma_{k'j'}$ in G_c there exists a corresponding move $\sigma_{ij}\sigma_{kl} - \sigma_{il}\sigma_{kj}$ in G, where $i' \leftrightarrow j' \subseteq i \leftrightarrow j$ and $k' \leftrightarrow l' \subseteq k \leftrightarrow l$. In fact, there are many such corresponding moves corresponding to all the ways to pull back ρ_c.

Definition 7 Let G be a block graph and let c be a central vertex. We call two monomials σ^u and σ^v in the same fiber to be *similar at a vertex c* if the subgraph over c and its adjacent vertices is the same for both the monomials.

For a given block graph G and a central vertex c, let S_c denote the set of all 2×2 minors of all matrices $\Sigma_{A\cup C, B\cup C}$ where (A, B, C) is a separation condition that is valid for G with $C = \{c\}$.

Proposition 7 *If a sequence of moves in G_c take $\rho_c(\sigma^{u_c})$ to $\rho_c(\sigma^{v_c})$, then there exist a corresponding sequence of moves in S_c which takes σ^u to a monomial which is similar to σ^v at c.*

Proof We know that $\rho_c(\sigma^{u_c})$ and σ^u are similar at c by construction. So, it is enough to show that if m is a move in G_c and m' is the corresponding move in G, then m applied to $\rho_c(\sigma^{u_c})$ and m' applied to σ^u are similar at c. Let $m = \sigma_{i'j'}\sigma_{k'l'} - \sigma_{i'l'}\sigma_{k'j'}$ be a move in G_c acting on the paths $\sigma_{i'j'}, \sigma_{k'l'}$ in $\rho_c(\sigma^{u_c})$. Let $m' = \sigma_{ij}\sigma_{kl} - \sigma_{il}\sigma_{kj}$ be its corresponding move in S_c acting on the paths σ_{ij}, σ_{kl} in σ^u. As $i' \leftrightarrow j' \subseteq i \leftrightarrow j$, $k' \leftrightarrow l' \subseteq k \leftrightarrow l$ and $c \in i' \leftrightarrow j'$ and $k' \leftrightarrow l'$, m and m' make the same changes at c in both the graphs. So, we can conclude that m applied to $\rho_c(\sigma^{u_c})$ and m' applied to σ^u are similar at c. □

Once we have the set of moves which takes σ^u to a monomial which is similar to σ^v at c, we can apply the same procedure at the other central vertices as well. To show that this ends up producing two monomials that are similar at every central vertex it is necessary to check that the moves obtained for a different central vertex c' do not affect the structure previously obtained at c.

Proposition 8 *Let $m = \sigma_{ij}\sigma_{kl} - \sigma_{il}\sigma_{kj}$ be a move obtained from a partition with $C = \{c\}$. Let V be the set of vertices in G. Then σ^u and m applied to σ^u are similar at $V \backslash c$.*

Proof If s is any vertex which is not in $i \leftrightarrow j$ or $k \leftrightarrow l$, then σ^u and m applied to σ^u remain similar at s as the move does not make any change at s. If $s \neq c$ is a vertex in $i \leftrightarrow j$, we then consider 2 cases:

Case 1: $s \in i \leftrightarrow j$ and $s \notin k \leftrightarrow l$

Let $s \in i \leftrightarrow c$. As m converts $i \leftrightarrow c \leftrightarrow j$ to $i \leftrightarrow c \leftrightarrow l$, $i \leftrightarrow c$ is contained in $i \leftrightarrow l$. This implies that s and all the vertices in $i \leftrightarrow j$ adjacent to s are also present in $i \leftrightarrow l$. A similar argument applies for $s \in c \leftrightarrow j$.

Case 2: $s \in i \leftrightarrow j$ and $s \in k \leftrightarrow l$

Let $s \in i \leftrightarrow c$ and $s \in k \leftrightarrow c$. As m converts $i \leftrightarrow c \leftrightarrow j$ to $i \leftrightarrow c \leftrightarrow l$ and $k \leftrightarrow c \leftrightarrow l$ to $k \leftrightarrow c \leftrightarrow j$, $i \leftrightarrow c$ is contained in $i \leftrightarrow l$ and $k \leftrightarrow c$ is contained in $k \leftrightarrow j$. So s and all the vertices in $i \leftrightarrow j$ ($k \leftrightarrow l$) adjacent to s are present in $i \leftrightarrow l$ ($k \leftrightarrow j$). A similar argument applies for $s \in c \leftrightarrow j, c \leftrightarrow l$.

In both the cases, m preserves the structure of σ^u around the vertex s. Hence, σ^u and m applied to σ^u are similar at all the vertices in $V \backslash c$. □

Note an important key feature that follows from the proof of Proposition 8: If m can be obtained from two partitions (A_1, B_1, C_1) and (A_2, B_2, C_2) with different central vertices, then σ^u and m applied to σ^u are similar at the central vertices as well.

We now give a proof for the generalized version of Theorem 3.

Theorem 5 *Let G be a block graph. Then, the shortest path ideal SP_G is generated by the set of all 2×2 minors of $\Sigma_{A\cup C,B\cup C}$, for all possible 1-clique partitions of G, i.e., $SP_G = CI_G$.*

Proof Suppose that $c_1, \dots, c_k$ are the central vertices of G. Let $S_1, \dots S_k$ be the corresponding quadratic moves associated to each central vertex. Let $f = \sigma^u - \sigma^v \in SP_G$. By applying Propositions 7 and 8 together with Theorem 3, we can assume that σ^u and σ^v are similar at every vertex after applying moves from $S_1, \dots, S_k$.

We can assume that σ^u and σ^v have no variables in common, otherwise we could delete this variable from both monomials and do an induction on dimension. So consider an arbitrary path $i \leftrightarrow j$ in σ^u which is not present in σ^v. We select the path in σ^v which has the highest number of common edges with $i \leftrightarrow j$. Let that path be $i' \leftrightarrow j'$ and let $s \leftrightarrow t$ be the common path in both the paths. Let s_1 and t_1 be the vertices adjacent to s and t, respectively, in $i \leftrightarrow j$. Similarly, let s' and t' be the vertices adjacent to s and t, respectively, in $i' \leftrightarrow j'$. Let p be the vertex in $s \leftrightarrow t$ adjacent to t (see Fig. 7 for an illustration of the idea).

If we apply the map ρ_t on both the monomials, we get that there exists a path $p \leftrightarrow t_1$ in $\rho_t(\sigma^u)$ which is not in $\rho_t(\sigma^v)$. But as σ^u and σ^v are similar at t, there must exist a path $x \leftrightarrow y$ in σ^v containing $p \leftrightarrow t_1$. So, the move $m = \sigma_{i'j'}\sigma_{xy} - \sigma_{i'y}\sigma_{xj'}$ is a valid move as none of the vertices in $i' \leftrightarrow p$ can be adjacent to any vertex in $t_1 \leftrightarrow y$ (as it would form a closed circuit implying that $i' \leftrightarrow t$ is not the shortest path). Similarly, none of the vertices in $x \leftrightarrow p$ can be adjacent to any vertex in $t' \leftrightarrow j'$. Further,

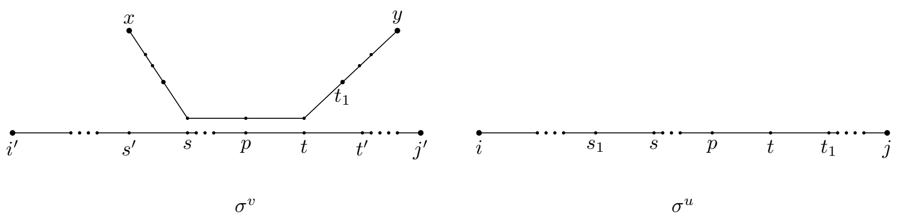


Fig. 7 Graphical illustration for the proof of Theorem 5

this move can be obtained from two different partitions with central vertices p and t, respectively. So, by Proposition 8 and the comment after its proof, we know that the move $\sigma_{i'j'}\sigma_{xy} - \sigma_{i'y}\sigma_{xj'}$ preserves the similarity of all the vertices.

Applying m on σ^v increases the length of the common path between $i \leftrightarrow j$ and $i' \leftrightarrow j'$ by at least 1, while keeping the monomials σ^u and m applied to σ^v similar at all the vertices. Repeating this process again, we can continue to shorten the length of the disagreement until the resulting monomials have a common monomial, in which case induction implies that we can use moves to connect these smaller degree monomials.

This implies that the set of binomials $S_1 \cup \cdots \cup S_k$ generates SP_G and hence $CI_G = SP_G$. □

6 Initial term map and SAGBI bases

In this section, we put all our previous results on shortest path maps together to prove Theorem 1. We also show that the set of polynomials $\{f_{ij} : 1 \leq i \leq j \leq n\}$ obtained from the inverse of K are a SAGBI basis for the $\mathbb{R}$-algebra they generate in the case of block graphs.

Proof of Theorem 1 We have already seen that $SP_G = CI_G \subseteq P_G$. We just need to show that $SP_G = P_G$ to complete the proof. Note that both SP_G and P_G are prime ideals so it suffices to show that they have the same dimension.

In both SP_G and P_G an upper bound on the dimension is equal to the number of vertices plus the number of edges in the graph. This follows because that is the number of free parameters in both parametrizations. In the case of P_G this upper bound is tight, because the map that sends $\Sigma \mapsto \Sigma^{-1}$ is the inverse map that recovers the entries of K. Since $SP_G \subseteq P_G$ we have the $\dim SP_G \geq \dim P_G$. Hence, they must have the same dimension. □

Finally, we can show the SAGBI basis property for the polynomials $\{f_{ij} : 1 \leq i \leq j \leq n\}$. Recall the definition of a SAGBI basis (which stands for Subalgebra Analogue of Gröbner Basis for Ideals). See Chapter 11 of Sturmfels (1996) for more details.

Definition 8 Let R be a finitely generated subalgebra of the polynomial ring $\mathbb{R}[K]$. Let $\prec$ be a term order on $\mathbb{R}[K]$. The *initial algebra* $\mathrm{in}_\prec(R)$ is defined as the $\mathbb{R}$-vector space spanned by $\{\mathrm{in}_\prec(f) : f \in R\}$. A finite set of polynomials $F \subseteq R$ is called a SAGBI basis for R if

(i) $R = \mathbb{R}[F]$, and
(ii) $\mathrm{in}_\prec(R) = \mathbb{R}[\{\mathrm{in}_\prec(f) : f \in F\}]$.

Let G be a block graph and let $F = \{f_{ij} : 1 \leq i \leq j \leq n\}$ be the polynomials appearing as the numerators in K^{-1}. To prove this, we will use some key result on SAGBI bases. Note that if $\prec$ is a term order on $\mathbb{R}[K]$ induced by a weight vector ω, then this induces a partial term order on $\mathbb{R}[\Sigma]$ by declaring that the weight of the variable σ_{ij} is the weight of the largest monomial appearing in f_{ij}. Denote by ω^* this induced weight order on $\mathbb{R}[\Sigma]$.

Both the algebras $\mathbb{R}[F]$ and $\mathbb{R}[\{\mathrm{in}_\prec(f) : f \in F\}]$ have presentation ideals in $\mathbb{R}[\Sigma]$. In the first case, this presentation ideal is exactly P_G, the vanishing ideal of the Gaussian graphical model. That is, $\mathbb{R}[F] = \mathbb{R}[\Sigma]/P_G$. In the second case, this presentation is exactly SP_G, the shortest path ideal, since that is the ideal of relations among the shortest path monomials. That is, $\mathbb{R}[\{\mathrm{in}_\prec(f) : f \in F\}] = \mathbb{R}[\Sigma]/SP_G$.

A fundamental theorem on SAGBI bases applied in the specific case of these ideals says the following.

Theorem 6 (Thm 11.4, Sturmfels 1996) *The set* $F \subseteq \mathbb{R}[K]$ *is a SAGBI basis if and only if* $\mathrm{in}_{\omega^*}(P_G) = SP_G$.

Corollary 1 *Let G be a block graph. Then, the set* $F \subseteq \mathbb{R}[K]$ *is a SAGBI basis of* $\mathbb{R}[F]$.

Proof We have already shown that $SP_G = P_G$. By construction, every one of the binomials in SP_G is homogeneous with respect to the weighting ω^*. Indeed, this weighting is exactly the weighting that counts the multiplicity of each edge of σ^u and the $\deg_i(\sigma^u)$ as used in Sect. 5. But then $\mathrm{in}_{\omega^*}(P_G) = \mathrm{in}_{\omega^*}(SP_G) = SP_G$ as desired. By Theorem 6, this shows that F is a SAGBI basis. □

Acknowledgements Pratik Misra and Seth Sullivant were partially supportedby the US National Science Foundation (DMS 1615660).

References

DeLoera, J. A., Sturmfels, B., Thomas, R. R. (1995). Gröbner bases and triangulations of the second hypersimplex. *Combinatorica*, *15*(3), 409–424.

Dobra, A., Sullivant, S. (2004). A divide-and-conquer algorithm for generating Markov bases of multiway tables. *Computational Statistics*, *19*(3), 347–366.

Drton, M., Massam, H., Olkin, I. (2008). Moments of minors of Wishart matrices. *The Annals of Statistics*, *36*(5), 2261–2283.

Drton, M., Sturmfels, B., Sullivant, S. (2007). Algebraic factor analysis: Tetrads, pentads and beyond. *Probability Theory and Related Fields*, *138*(3–4), 463–493.

Grayson, D. R., Stillman, M. E. (2017). Macaulay2, a software system for research in algebraic geometry. https://faculty.math.illinois.edu/Macaulay2/. Retrieved Sept 1, 2017

Hassett, B. (2007). *Introduction to algebraic geometry*. Cambridge: Cambridge University Press.

Herzog, J., Hibi, T., Ohsugi, H. (2018). *Binomial ideals. Graduate texts in mathematics*, Vol. 29. Cham: Springer.

Hoşten, S., Sullivant, S. (2002). Gröbner bases and polyhedral geometry of reducible and cyclic models. *Journal of Combinatorial Theory Series A*, *100*(2), 277–301.

Jones, B., West, M. (2005). Covariance decomposition in undirected Gaussian graphical models. *Biometrika*, *92*(4), 779–786.
Kelley, T. (1935). *Essential traits of mental life. Harvard studies in education*, Vol. 26. Cambridge, MA: Harvard University Press.
Koller, D., Friedman, N. (2009). *Probabilistic graphical models. Adaptive computation and machine learning*. Cambridge, MA: MIT Press.
Lauritzen, S. L. (1996). *Graphical models. Oxford statistical science series*, Vol. 17. New York: Oxford University Press.
Spirtes, P., Glymour, C., Scheines, R. (2000). *Causation, prediction, and search. Adaptive computation and machine learning*2nd ed. Cambridge, MA: MIT Press.
Sturmfels, B. (1996). *Gröbner bases and convex polytopes. University Lecture Series*, Vol. 8. Providence, RI: American Mathematical Society.
Sturmfels, B., Uhler, C. (2010). Multivariate Gaussian, semidefinite matrix completion, and convex algebraic geometry. *Annals of the Institute of Statistical Mathematics*, *62*(4), 603–638.
Sullivant, S. (2007). Toric fiber products. *Journal of Algebra*, *316*(2), 560–577.
Sullivant, S. (2008). Algebraic geometry of Gaussian Bayesian networks. *Advances in Applied Mathematics*, *40*(4), 482–513.
Sullivant, S. (2018). *Algebraic statistics. Graduate Studies in Mathematics*, Vol. 194. Providence, RI: American Mathematical Society.
Sullivant, S., Talaska, K., Draisma, J. (2010). Trek separation for Gaussian graphical models. *The Annals of Statistics*, *38*(3), 1665–1685.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.