Algebraic Approaches to Colored Gaussian Graphical Models

Aida Maraj
University of Michigan & IMSI–University of Chicago

Symmetry, Stability, and Computations

Storyline

1. Colored Gaussian Graphical (CGG) models

2. CGG models with toric vanishing ideals

based on the paper *Symmetrically colored Gaussian graphical models with toric vanishing ideals*, SIAM Journal on Applied Algebra and Geometry 7.1 (2023): 133-158. with

Jane Ivy Coons (University of Oxford)

Pratik Misra (TU Munich)

Miruna Stefana Sorea (SISSA)

3. CGG models toric after a linear change of variables

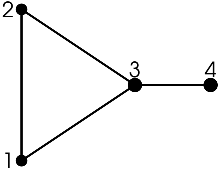
based on *Symmetry Lie algebras of varieties with applications to algebraic statistics*, arXiv:2309.10741, with

Arpan Pal (University of Idaho)

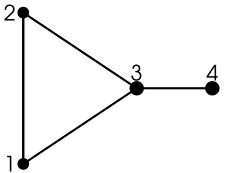
4. Ongoing and future work

An ideal $I \subseteq \mathbb{R}[x_1, \ldots, x_n]$ is **toric** if one of the equivalent properties holds:

I is a prime binomial ideal $\iff$ I is the kernel of a monomial map $\mathbb{R}[x_1, \ldots, x_n] \to \mathbb{R}[y_1^{\pm 1}, \ldots, y_m^{\pm 1}]$



$$\mathcal{L}_G = \left\{ K = \begin{bmatrix} k_{11} & k_{12} & k_{13} & 0 \\ k_{12} & k_{22} & k_{23} & 0 \\ k_{13} & k_{23} & k_{33} & k_{34} \\ 0 & 0 & k_{34} & k_{44} \end{bmatrix} \mid k_{11}, k_{12}, \ldots, k_{44} \in \mathbb{R} \right\}$$



$$\mathcal{L}_G = \left\{ K = \begin{bmatrix} k_{11} & k_{12} & k_{13} & 0 \\ k_{12} & k_{22} & k_{23} & 0 \\ k_{13} & k_{23} & k_{33} & k_{34} \\ 0 & 0 & k_{34} & k_{44} \end{bmatrix} \mid k_{11}, k_{12}, \ldots, k_{44} \in \mathbb{R} \right\}$$

A random vector X is distributed according to a **multivariate Gaussian distribution** $\mathcal{N}_n(\mathbf{0}, \Sigma)$ with mean $\mathbf{0}$ if it has probability density function

$$\varphi_\Sigma(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp(-\frac{1}{2}\mathbf{x}^T\Sigma^{-1}\mathbf{x}),$$

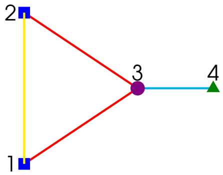
where Σ is a symmetric positive definite matrix.

- $\Sigma =$ **covariance matrix**
- $K := \Sigma^{-1} =$ **concentration matrix**

The **Gaussian graphical model** for G is the set of $\mathcal{N}_n(\mathbf{0}, \Sigma)$ with $\Sigma \in \mathcal{L}_G^{-1} \cap PD_4$

Here,

$$\mathcal{L}_G^{-1} = \overline{\{\Sigma \in \mathrm{Sym}_n(\mathbb{R}) \mid \Sigma^{-1} \in \mathcal{L}_G\}}.$$



$$\mathcal{L}_{\mathcal{G}} = \left\{ K = \begin{bmatrix} k_b & k_y & k_r & 0 \\ k_y & k_b & k_r & 0 \\ k_r & k_r & k_v & k_c \\ 0 & 0 & k_c & k_g \end{bmatrix} \mid k_b, k_y, \ldots, k_g \in \mathbb{R} \right\}$$

colors: b, v, g, y, r, c.

The **colored Gaussian graphical model** for $\mathcal{G}$ is the set of Gaussian distribution $\mathcal{N}_n(\mathbf{0}, \Sigma)$ such that $\Sigma \in \mathcal{L}_{\mathcal{G}}^{-1} \cap PD_4$.

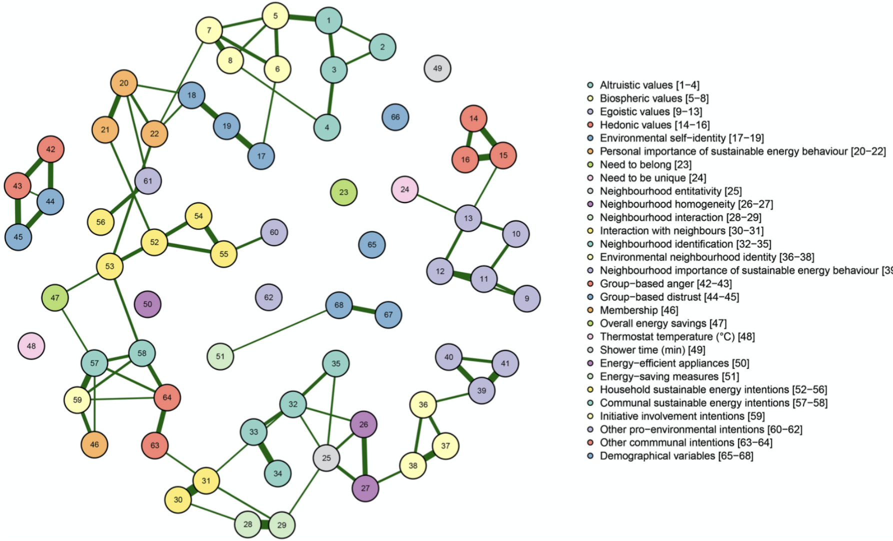
Høsgaard and Lauritzen. "Graphical Gaussian models with edge and vertex symmetries." *Journal of the Royal Statistical Society, Series B. Statistical Methodology* **70** (2008), no.5, 1005-1027.

Objective: $\mathcal{L}_{\mathcal{G}}^{-1}$ gives the set of covariance matrices. Understand its algebraic structure and contributions to applications.

$$\mathcal{L}_{\mathcal{G}}^{-1} = \{\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_{44} \end{bmatrix} \mid \sigma_{11} = \sigma_{22}, \sigma_{13} = \sigma_{23}, \sigma_{14} = \sigma_{24}, \sigma_{24}\sigma_{33} = \sigma_{23}\sigma_{34}\}$$

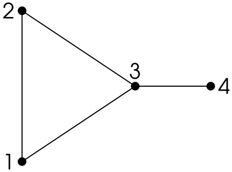
$$= V(\underbrace{\langle \sigma_{11} - \sigma_{22}, \sigma_{13} - \sigma_{23}, \sigma_{14} - \sigma_{24}, \sigma_{24}\sigma_{33} - \sigma_{23}\sigma_{34} \rangle}_{I_{\mathcal{G}}})$$

not linear, but toric

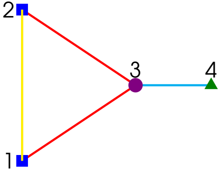


[1] Hiroyuki Toh and Katsuhisa Horimoto. "Inference of a genetic network by a combined approach of cluster analysis and graphical Gaussian modeling." Bioinformatics 18.2 (2002): 287-297.
[2] Ernst Wit and Antonino Abbruzzo. "Factorial graphical models for dynamic networks." Network Science 3.1 (2015): 37-57.
[3] Bhushan, Nitin, Florian Mohnert, Daniel Sloot, Lise Jans, Casper Albers, and Linda Steg. "Using a Gaussian graphical model to explore relationships between items and variables in environmental psychology research." Frontiers in psychology 10 (2019)

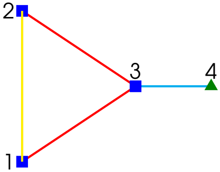
Objective: When is $\mathcal{L}^{-1}$ toric? Can we compute its defining binomials?



$$I_G = \langle \sigma_{24}\sigma_{33} - \sigma_{23}\sigma_{34}, \sigma_{14}\sigma_{33} - \sigma_{13}\sigma_{34} \rangle$$

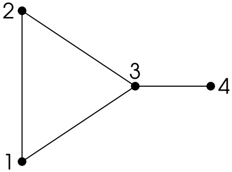


$$I_{\mathcal{G}} = \langle \sigma_{11} - \sigma_{22}, \sigma_{13} - \sigma_{23}, \sigma_{14} - \sigma_{24}, \sigma_{24}\sigma_{33} - \sigma_{23}\sigma_{34} \rangle$$

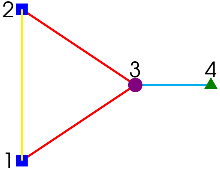


$$\begin{aligned} I_{\mathcal{G}'} = \langle &\sigma_{11} - \sigma_{22}, \sigma_{13} - \sigma_{23}, \sigma_{14} - \sigma_{24}, \sigma_{24}\sigma_{33} - \sigma_{23}\sigma_{34}, \\ &2\sigma_{12}\sigma_{24}^2 - 2\sigma_{22}\sigma_{24}^2 - \sigma_{23}\sigma_{24}\sigma_{34} + \sigma_{22}\sigma_{34}^2 + \\ &\sigma_{12}^2\sigma_{44} + \sigma_{22}^2\sigma_{44} - \sigma_{23}^2\sigma_{44} - \sigma_{22}\sigma_{33}\sigma_{44} \rangle \end{aligned}$$

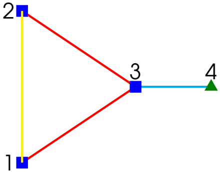
Objective: When is $\mathcal{L}^{-1}$ toric? Can we compute its defining binomials?



$$I_G = \langle \sigma_{24}\sigma_{33} - \sigma_{23}\sigma_{34}, \sigma_{14}\sigma_{33} - \sigma_{13}\sigma_{34} \rangle$$



$$I_{\mathcal{G}} = \langle \sigma_{11} - \sigma_{22}, \sigma_{13} - \sigma_{23}, \sigma_{14} - \sigma_{24}, \sigma_{24}\sigma_{33} - \sigma_{23}\sigma_{34} \rangle$$



$$\begin{aligned} I_{\mathcal{G}'} = \langle & \sigma_{11} - \sigma_{22}, \sigma_{13} - \sigma_{23}, \sigma_{14} - \sigma_{24}, \sigma_{24}\sigma_{33} - \sigma_{23}\sigma_{34}, \\ & 2\sigma_{12}\sigma_{24}^2 - 2\sigma_{22}\sigma_{24}^2 - \sigma_{23}\sigma_{24}\sigma_{34} + \sigma_{22}\sigma_{34}^2 + \\ & \sigma_{12}^2\sigma_{44} + \sigma_{22}^2\sigma_{44} - \sigma_{23}^2\sigma_{44} - \sigma_{22}\sigma_{33}\sigma_{44} \rangle \end{aligned}$$

The answer is in **symmetrically colored graphs**.

Equations vanishing in $\mathcal{L}_{\mathcal{G}}^{-1}$: a first recipe

The space $\mathcal{L}_{\mathcal{G}}^{-1}$ is the variety of the kernel $I_{\mathcal{G}}$ of the rational map

$$\rho_{\mathcal{G}} : \mathbb{R}[\Sigma] \to \mathbb{R}(K),\ \rho_{\mathcal{G}}(\sigma_{ij}) = \frac{(-1)^{i+j} K_{ij}}{\det(K)},$$ NOT a monomial map!!!

where K_{ij} is the ij-th minor of the symmetric matrix K.

$I_{\mathcal{G}}$ is the **vanishing ideal of the colored Gaussian graphical model for** $\mathcal{G}$.

Theorem [Misra & Sullivant, 2021][a]: Let G be a block graph. Then I_G is a toric ideal generated by quadratics.

[a] Misra and Sullivant. "Gaussian graphical models with toric vanishing ideals." *Annals of the Institute of Statistical Mathematics* **73** (2021), no. 4, 757-785.

Theorem Coons, M, Misra & Sorea, 2023: Let $\mathcal{G}$ be a block graph with RCOP coloring. Then $I_{\mathcal{G}}$ is a toric ideal. Moreover,

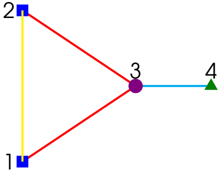
$$I_{\mathcal{G}} = I_G + I_{\overline{\mathcal{G}}},$$

where

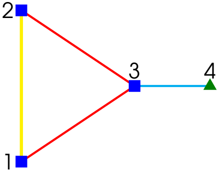
- I_G is the ideal of the uncolored Gaussian graphical model, and
- $I_{\overline{\mathcal{G}}} = \langle \sigma_{ij} - \sigma_{uv},\ \text{for Colors}(i \leftrightarrow j) = \text{Colors}(u \leftrightarrow v) \rangle$.

Let $\mathrm{Aut}(\mathcal{G})$ be the group of graph automorphisms of $\mathcal{G}$ that preserve its vertex and edge colors.

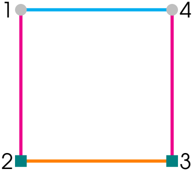
The colored graph $\mathcal{G}$ is an **RCOP graph*** if there exists a subgroup $\Gamma \subset \mathrm{Aut}(\mathcal{G})$ whose vertex and edge orbits are exactly the vertex and edge color classes of $\mathcal{G}$, respectively.



$\Gamma = \mathrm{Aut}(\mathcal{G}) = \{(), (12)\}$



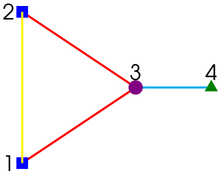
$\mathrm{Aut}(\mathcal{G}') = \{(), (12)\}$

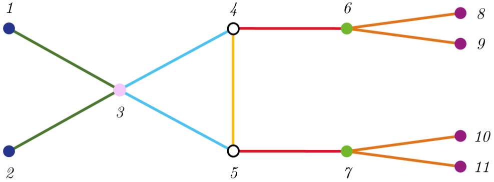


$\Gamma = \mathrm{Aut}(\mathcal{G}') = \{(), (14)(23)\}$

***** Høsgaard and Lauritzen. "Graphical Gaussian models with edge and vertex symmetries." *Journal of the Royal Statistical Society, Series B. Statistical Methodology* **70** (2008), no.5,

Intuitive idea: **Block graphs** are formed by inductively gluing complete graphs at a single vertex.



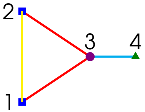


Key Property: Any two vertices in a block graph have a unique shortest path connecting them.

A key idea

The **shortest path map** $\varphi_{\mathcal{G}}$ for colored block graph $\mathcal{G}$ is

$$\begin{aligned}\varphi_{\mathcal{G}} : \mathbb{C}[\sigma_{ij} \mid 1 \le i \le j \le n] &\to \mathbb{C}[t_\lambda \mid \lambda \in \mathrm{Colors}(\mathcal{G})], \\ \sigma_{ij} &\mapsto t_{\lambda(i)} t_{\lambda(j)} \prod_{e \in E(i \leftrightarrow j)} t_{\lambda(e)}.\end{aligned}$$



$$\begin{aligned}\varphi_{\mathcal{G}}(\sigma_{11}) &= t_b{}^2 \\ \varphi_{\mathcal{G}}(\sigma_{12}) &= t_b{}^2 t_y \\ \varphi_{\mathcal{G}}(\sigma_{22}) &= t_b{}^2 \\ \varphi_{\mathcal{G}}(\sigma_{13}) &= t_b t_v t_r \\ \varphi_{\mathcal{G}}(\sigma_{23}) &= t_b t_v t_r\end{aligned}$$

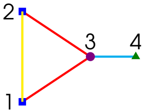
$$\begin{aligned}\varphi_{\mathcal{G}}(\sigma_{33}) &= t_v{}^2 \\ \varphi_{\mathcal{G}}(\sigma_{14}) &= t_b t_g t_r t_c \\ \varphi_{\mathcal{G}}(\sigma_{24}) &= t_b t_g t_r t_c \\ \varphi_{\mathcal{G}}(\sigma_{34}) &= t_v t_g t_c \\ \varphi_{\mathcal{G}}(\sigma_{44}) &= t_g{}^2\end{aligned}$$

$\ker \varphi_{\mathcal{G}} = \langle \sigma_{11} - \sigma_{22}, \sigma_{13} - \sigma_{23}, \sigma_{14} - \sigma_{24}, \sigma_{24}\sigma_{33} - \sigma_{23}\sigma_{34} \rangle$

A key idea

The **shortest path map** $\varphi_{\mathcal{G}}$ for colored block graph $\mathcal{G}$ is

$$\begin{aligned}\varphi_{\mathcal{G}} : \mathbb{C}[\sigma_{ij} \mid 1 \le i \le j \le n] &\to \mathbb{C}[t_\lambda \mid \lambda \in \mathrm{Colors}(\mathcal{G})], \\ \sigma_{ij} &\mapsto t_{\lambda(i)} t_{\lambda(j)} \prod_{e \in E(i \leftrightarrow j)} t_{\lambda(e)}.\end{aligned}$$



$$\varphi_{\mathcal{G}}(\sigma_{11}) = t_b{}^2$$
$$\varphi_{\mathcal{G}}(\sigma_{12}) = t_b{}^2 t_y$$
$$\varphi_{\mathcal{G}}(\sigma_{22}) = t_b{}^2$$
$$\varphi_{\mathcal{G}}(\sigma_{13}) = t_b t_v t_r$$
$$\varphi_{\mathcal{G}}(\sigma_{23}) = t_b t_v t_r$$

$$\varphi_{\mathcal{G}}(\sigma_{33}) = t_v{}^2$$
$$\varphi_{\mathcal{G}}(\sigma_{14}) = t_b t_g t_r t_c$$
$$\varphi_{\mathcal{G}}(\sigma_{24}) = t_b t_g t_r t_c$$
$$\varphi_{\mathcal{G}}(\sigma_{34}) = t_v t_g t_c$$
$$\varphi_{\mathcal{G}}(\sigma_{44}) = t_g{}^2$$

$$\ker \varphi_{\mathcal{G}} = \langle \sigma_{11} - \sigma_{22}, \sigma_{13} - \sigma_{23}, \sigma_{14} - \sigma_{24}, \sigma_{24}\sigma_{33} - \sigma_{23}\sigma_{34} \rangle = I_{\mathcal{G}}$$

Note: linear forms: $I_{\overline{\mathcal{G}}} = \langle \sigma_{ij} - \sigma_{k\ell}$, for $\mathrm{Colors}(i \leftrightarrow j) = \mathrm{Colors}(k \leftrightarrow \ell)\rangle$.
quadratic forms:
$I_G = \langle \sigma_{ij}\sigma_{k\ell} - \sigma_{i\ell}\sigma_{kj}$, for $\mathrm{Colors}(i \leftrightarrow j, k \leftrightarrow \ell) = \mathrm{Colors}(i \leftrightarrow \ell, k \leftrightarrow j)\rangle$.

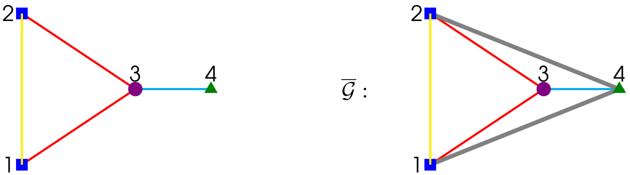
Theorem Coons, M, Misra & Sorea, 2023: Let $\mathcal{G}$ be a block graph with RCOP coloring. Then $I_{\mathcal{G}} = \ker \varphi_{\mathcal{G}} = I_G + I_{\overline{\mathcal{G}}}$.

Proof Idea:

1. $I_{\overline{\mathcal{G}}}$ is the vanishing of a complete colored graph $\overline{\mathcal{G}}$. It forms a Jordan algebra; that is, $\mathcal{L}_{\overline{\mathcal{G}}} = \mathcal{L}_{\overline{\mathcal{G}}}^{-1}$

$\mathcal{G}$:

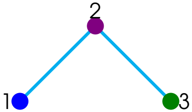
$\overline{\mathcal{G}}$:



$$\boldsymbol{I_{\overline{\mathcal{G}}}} = \langle \sigma_{11} - \sigma_{22}, \sigma_{13} - \sigma_{23}, \sigma_{14} - \sigma_{24} \rangle \subseteq \boldsymbol{I_{\mathcal{G}}}$$

2. $\ker \varphi_{\mathcal{G}} = I_G + I_{\overline{\mathcal{G}}}$ (*the hard work is here*)
3. $\ker \varphi_{\mathcal{G}} \subseteq I_{\mathcal{G}}$ and $\dim(\ker \varphi_{\mathcal{G}}) = \dim(I_{\mathcal{G}})$
4. Since $\ker \varphi_{\mathcal{G}}, I_{\mathcal{G}}$ are both prime ideals, (2)+(3) implies $\ker \varphi_{\mathcal{G}} = I_{\mathcal{G}}$

Some colored GGM are toric after a linear change of variables



$$I_{\mathcal{G}} = \langle \sigma_{23}\sigma_{11} - \sigma_{12}\sigma_{13},\ \sigma_{12}(\sigma_{23} + \sigma_{33}) - \sigma_{13}(\sigma_{22} + \sigma_{23}) \rangle.$$

The change of variables

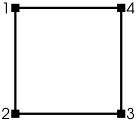
$$p_{22} = \sigma_{22} + \sigma_{23}$$
$$p_{33} = \sigma_{33} + \sigma_{23}$$
$$p_{ij} = \sigma_{ij} \text{ for all other } i, j$$

yields

$$I_{\mathcal{G}} = \langle p_{23}p_{11} - p_{12}p_{13},\ p_{22}p_{13} - p_{12}p_{23} \rangle.$$

Question: to what extend this happens?

... and some not under any linear change of variables



$$I_G = \langle$$

$$\sigma_{23}\sigma_{14}\sigma_{24} - \sigma_{13}\sigma_{24}^2 - \sigma_{22}\sigma_{14}\sigma_{34} + \sigma_{12}\sigma_{24}\sigma_{34} + \sigma_{22}\sigma_{13}\sigma_{44} - \sigma_{12}\sigma_{23}\sigma_{44},$$

$$\sigma_{13}\sigma_{23}\sigma_{14} - \sigma_{24}\sigma_{13}^2 - \sigma_{12}\sigma_{33}\sigma_{14} + \sigma_{11}\sigma_{33}\sigma_{24} + \sigma_{12}\sigma_{13}\sigma_{34} - \sigma_{11}\sigma_{23}\sigma_{34}\rangle$$

Sketch of the proof [M, Pal 2023]:

1. Determine a group action of $\mathrm{GL}_n(\mathbb{C})$ on $\mathbb{C}[x] = \mathbb{C}[x_1, \ldots, x_n]$ by

$$\text{for } g = \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix} \in \mathrm{GL}_n(\mathbb{C}), p \in \mathbb{C}[x], \quad g \cdot p(x_1, \ldots, x_n) = p(g_1 \cdot x, \ldots, g_n \cdot x).$$

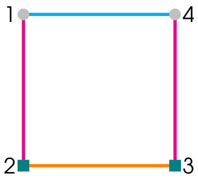
2. The stabilizer $G_I = \{g \in \mathrm{GL}_n(\mathbb{C}) \mid g \cdot p \in I, \forall p \in I\}$ of a prime homogeneous ideal in $I \subseteq R$ is a Lie group.
3. If I is toric, then the torus acting on $V(I)$ is embedded in G_I. So, $\dim(G_I) \geq \dim(I)$.

3′ If $\dim(G_I) < \dim(I)$, then I cannot be turned toric under any linear change of variables.

In our case $\dim(G_I) = 4 < 8 = \dim(I)$.

Project initiated in conversations with JM Landsberg at 2022 Texas Algebraic Geometry Symposium at Texas A&M

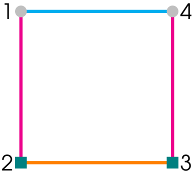
Good news



Its vanishing ideal is

$$
\begin{aligned}
I_{\mathcal{G}} = \langle & \sigma_{13} - \sigma_{24}, \sigma_{22} - \sigma_{33}, \sigma_{12} - \sigma_{34}, \sigma_{11} - \sigma_{44}, \\
& \sigma_{23}\sigma_{14}\sigma_{24} - \sigma_{24}^3 - \sigma_{33}\sigma_{14}\sigma_{34} + \sigma_{24}\sigma_{34}^2 + \sigma_{33}\sigma_{24}\sigma_{44} - \sigma_{23}\sigma_{34}\sigma_{44}\rangle \\
& = I_{\overline{\mathcal{G}}} + I_G
\end{aligned}
$$

Good news



Its vanishing ideal is

$$
\begin{aligned}
I_{\mathcal{G}} = \langle &\sigma_{13} - \sigma_{24}, \sigma_{22} - \sigma_{33}, \sigma_{12} - \sigma_{34}, \sigma_{11} - \sigma_{44}, \\
&\sigma_{23}\sigma_{14}\sigma_{24} - \sigma_{24}^3 - \sigma_{33}\sigma_{14}\sigma_{34} + \sigma_{24}\sigma_{34}^2 + \sigma_{33}\sigma_{24}\sigma_{44} - \sigma_{23}\sigma_{34}\sigma_{44}\rangle \\
&= I_{\overline{\mathcal{G}}} + I_G
\end{aligned}
$$

Theorem [Coons, M 2023+]: Let $\mathcal{G}$ be an RCOP graph. Then $I_{\mathcal{G}}$ is a minimal prime for $I_{\overline{\mathcal{G}}} + I_G$.

The algebra, especially its toric structure, of CGG models, contributes to:

the Maximum Likelihood (ML) estimates and ML degrees

ongoing work with Jane Ivy Coons

Høsgaard and Lauritzen. "Graphical Gaussian models with edge and vertex symmetries." *Journal of the Royal Statistical Society, Series B. Statistical Methodology* **70** (2008), no.5, 1005-1027.

the ML threeshold

leading a working group at IMSI with Roser Homs Poons

model selection and identifiability

unexplored

This work is partially supported by NSF DMS2306672 grant on Applications of Algebraic Geometry to Multivariate Gaussian Models

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

The algebra of CGG models, especially the toric structure, contributes to:

the Maximum Likelihood (ML) estimates and ML degrees

ongoing work with Jane Ivy Coons

Høsgaard and Lauritzen. "Graphical Gaussian models with edge and vertex symmetries." *Journal of the Royal Statistical Society, Series B. Statistical Methodology* **70** (2008), no.5, 1005-1027.

the ML threeshold

leading a reading group at IMSI with Roser Homs Poons

–long program "Algebraic Statistics and Our Changing World

Institute for Mathematical and Statistical Innovation

model selection and identifiability

unexplored

Thank you!

This work is partially supported by NSF DMS2306672 grant on Applications of Algebraic Geometry to Multivariate Gaussian Models

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.