Computational lower bounds via almost orthonormal polynomials

Simone Maria Giancola

[1]Laboratoire de Mathématiques d'Orsay, Université Paris-Saclay

co-supervision with Alexandra Carpentier, Christophe Giraud, Nicolas Verzelen

Toy model: planted sub-matrix/sub-graph/random clique

For (λ, k) sample the $n \times n$ matrix:

$$X_{ij} = x_i x_j, \qquad x_i \overset{\text{i.i.d.}}{\sim} \sqrt{\lambda}\mathrm{Ber}\left(\frac{k}{n}\right)$$

and observe the $n \times n$ matrix:

$$Y_{ij} = \begin{cases} 1 & \text{with probability } \frac{1+X_{ij}}{2} \\ -1 & \text{with probability } \frac{1-X_{ij}}{2} \end{cases}.$$

Latent clique of (“strength”, “size”)$= (\lambda, k)$.

Toy model: planted sub-matrix/sub-graph/random clique

For (λ, k) sample the $n \times n$ matrix:

$$X_{ij} = x_i x_j, \qquad x_i \overset{\text{i.i.d.}}{\sim} \sqrt{\lambda}\mathrm{Ber}\left(\frac{k}{n}\right)$$

and observe the $n \times n$ matrix:

$$Y_{ij} = \begin{cases} 1 & \text{with probability } \frac{1+X_{ij}}{2} \\ -1 & \text{with probability } \frac{1-X_{ij}}{2} \end{cases}.$$

Latent clique of ("strength", "size")$= (\lambda, k)$.

Question: prove when cannot detect perturbations in poly-time

$$H_0 \ : \ Y \text{ structure is } (\lambda, k), \qquad H_1 : Y \text{ structure is } (\lambda + \eta, k).$$

Performance in hypothesis test

Statistical optimality

$$err_{IT}(\lambda, \eta, k) \coloneqq \inf_{t \text{ measurable test}} \mathbb{P}_{H_0}[t(Y) = H_1] + \mathbb{P}_{H_1}[t(Y) = H_0].$$

Performance in hypothesis test

Statistical optimality

$$err_{IT}(\lambda, \eta, k) := \inf_{t \text{ measurable test}} \mathbb{P}_{H_0}[t(Y) = H_1] + \mathbb{P}_{H_1}[t(Y) = H_0].$$

Main issue: could not be an algorithm, e.g. Likelihood/find large clique (NP-hard).

Performance in hypothesis test

Statistical optimality

$$err_{IT}(\lambda,\eta,k) := \inf_{t \text{ measurable test}} \mathbb{P}_{H_0}[t(Y)=H_1] + \mathbb{P}_{H_1}[t(Y)=H_0].$$

Main issue: could not be an algorithm, e.g. Likelihood/find large clique (NP-hard).

Poly-time optimality

$$err_{poly}(\lambda,\eta,k) := \inf_{t \text{ measurable poly-time test}} \mathbb{P}_{H_0}[t(Y)=H_1] + \mathbb{P}_{H_1}[t(Y)=H_0].$$

Statistical-to-computational gap [KWB19; BPW18]

Less studied **poly-time criterion is the important one in practice**.
If $1-\Omega(1) = err_{IT}(\lambda,\eta,k) < err_{poly}(\lambda,\eta,k) = 1-o(1)$ then η is hard.

Performance in hypothesis test

Poly-time optimality

$$err_{poly}(\lambda, \eta, k) := \inf_{t \text{ measurable poly-time test}} \mathbb{P}_{H_0}[t(Y) = H_1] + \mathbb{P}_{H_1}[t(Y) = H_0].$$

Performance in hypothesis test

Poly-time optimality

$$err_{poly}(\lambda, \eta, k) := \inf_{t \text{ measurable poly-time test}} \mathbb{P}_{H_0}[t(Y) = H_1] + \mathbb{P}_{H_1}[t(Y) = H_0].$$

Main issue: hard to capture algorithms, find surrogate condition.

Performance in hypothesis test

Poly-time optimality

$$err_{poly}(\lambda, \eta, k) := \inf_{t \text{ measurable poly-time test}} \mathbb{P}_{H_0}[t(Y) = H_1] + \mathbb{P}_{H_1}[t(Y) = H_0].$$

Main issue: hard to capture algorithms, find surrogate condition.

Computational bound, conjecture [Hop18]

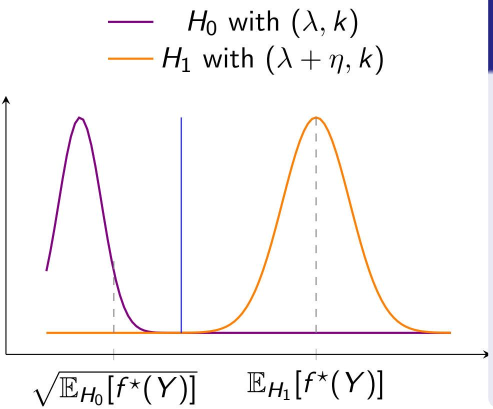
log n degree polynomials surrogate poly-time algorithms, and if:

$$Adv(\lambda, \eta, k) := \sup_{f : \deg(f) \lesssim \log n} \frac{\mathbb{E}_{H_1}[f(Y)]}{\sqrt{\mathbb{E}_{H_0}[f(Y)^2]}} = 1 + o(1),$$

and $err_{IT}(\lambda, \eta, k) \leq 1 - \Omega(1)$ then **statistical-to-computational gap**.

Intuition on advantage

When the advantage $Adv(\lambda, \eta, k) := \sup_{f:\deg(f) \lesssim \log n} \frac{\mathbb{E}_{H_1}[f(Y)]}{\sqrt{\mathbb{E}_{H_0}[f(Y)^2]}}$ is large:



If it is small for each polynomial of degree $\leq D$

no polynomial separates distributions! **Extrapolate hardness for polynomials, conjecture it for poly-time algorithms.**

Why? It works well in “nice” problems

- Originated from sum-of-squares [Bar+16; Hop18], then independent motivation [KWB19];

Why? It works well in “nice” problems

- Originated from sum-of-squares [Bar+16; Hop18], then independent motivation [KWB19];
- captures best known algorithms, but also polynomials and spectral methods [KWB19];

Why? It works well in "nice" problems

- Originated from sum-of-squares [Bar+16; Hop18], then independent motivation [KWB19];
- captures best known algorithms, but also polynomials and spectral methods [KWB19];
- applied to many problems for hypothesis testing, estimation, optimization [Wei25; EGV25a; EGV25b; EGV24];

Why? It works well in “nice” problems

- Originated from sum-of-squares [Bar+16; Hop18], then independent motivation [KWB19];
- captures best known algorithms, but also polynomials and spectral methods [KWB19];
- applied to many problems for hypothesis testing, estimation, optimization [Wei25; EGV25a; EGV25b; EGV24];
- linked to other techniques to claim/conjecture algorithmic hardness [Ban+22; BB20; MW22; Che+25; Wei25; GMZ22].

The orthonormal trick to bound *Adv*

Question: prove when cannot detect perturbations **in poly-time**,

$$H_0 \;:\; Y \text{ structure is } (\lambda, k), \qquad H_1 : Y \text{ structure is } (\lambda + \eta, k).$$

Basic idea: decomposition along basis of $Adv(\lambda, \eta, k)$.

The orthonormal trick to bound *Adv*

Question: prove when cannot detect perturbations **in poly-time**,

$$H_0 \ : \ Y \text{ structure is } (\lambda, k), \qquad H_1 : Y \text{ structure is } (\lambda + \eta, k).$$

Basic idea: decomposition along basis of $Adv(\lambda, \eta, k)$.

Imagine the basis is orthonormal in H_0 . . .

$$Adv(\lambda, \eta, k) = \sup_{f:\deg(f) \lesssim \log n} \frac{\mathbb{E}_{H_1}[f(Y)]}{\sqrt{\mathbb{E}_{H_0}[f(Y)^2]}}$$

Decompose in ψ_G orthonormal basis numerator and denominator.

$$f(Y) = \sum_{G \in \text{basis}} \alpha_G \psi_G, \qquad \alpha_G = \mathbb{E}_{H_0:(\lambda,k)}\left[f(Y)\psi_G\right].$$

The orthonormal trick to bound *Adv*

Question: prove when cannot detect perturbations **in poly-time**,

$$H_0 \;:\; Y \text{ structure is } (\lambda, k), \qquad H_1 : Y \text{ structure is } (\lambda + \eta, k).$$

Basic idea: decomposition along basis of $Adv(\lambda, \eta, k)$.

Imagine the basis is orthonormal in H_0 . . .

$$Adv(\lambda, \eta, k) = \sup_{\alpha} \frac{\mathbb{E}_{H_1}\left[\sum_{G \in basis} \alpha_G \psi_G\right]}{\sqrt{\mathbb{E}_{H_0}\left[\sum_{G, G' \in \text{basis}} \alpha_G \alpha_{G'} \psi_G \psi_{G'}\right]}}$$

$$= \sup_{\alpha} \frac{\mathbb{E}_{H_1}\left[\sum_{G \in \text{ basis}} \alpha_G \psi_G\right]}{\|\alpha\|_2} = LinAdv(\lambda, \eta, k),$$

Linear: **easy**!

Outside of pure noise, $(\lambda, k), \lambda, k \neq 0$

Question: prove when cannot detect perturbations **in poly-time**,

$$H_0 \ : \ Y \text{ structure is } (\lambda, k), \qquad H_1 : Y \text{ structure is } (\lambda + \eta, k).$$

at $(\lambda, k), \lambda, k \neq 0$ **not explicit ortho basis**!
In literature implicit recursive solutions [SW22; SW25].

Outside of pure noise, $(\lambda, k), \lambda, k \neq 0$

Question: prove when cannot detect perturbations **in poly-time**,

$$H_0 \ : \ Y \text{ structure is } (\lambda, k), \qquad H_1 : Y \text{ structure is } (\lambda + \eta, k).$$

at $(\lambda, k), \lambda, k \neq 0$ **not explicit ortho basis**!
In literature implicit recursive solutions [SW22; SW25].

Our solution: almost orthonormal basis

Find a collection of functions $(\widetilde{\psi}_G)_G$ forming a basis of H_0, H_1:

$$f(Y) = \sum_{G \in \text{ basis}} \alpha_G \widetilde{\psi}_G,$$

such that for some constants:

$$c \left\|\alpha\right\|_2^2 \leq \mathbb{E}_{H_0:(\lambda,k)} \left[\sum_{G, G' \in \text{basis}} \alpha_G \alpha_{G'} \widetilde{\psi}_G \widetilde{\psi}_{G'} \right] \leq C \left\|\alpha\right\|_2^2 .$$

Key proposition

Assume

The parameter (λ, k) to sample Y in H_0 is such that $\max\left\{\frac{k}{n}, \frac{\lambda k}{n}, \lambda\right\} \leq \mathrm{polylog}(n)$.

Key proposition

Assume

The parameter (λ, k) to sample Y in H_0 is such that $\max\left\{\frac{k}{n}, \frac{\lambda k}{n}, \lambda\right\} \leq \mathrm{polylog}(n)$.

Almost orthonormal basis exists, *Adv* simplifies

There is an almost orthonormal basis and for (λ, k):

$$Adv(\lambda, \eta, k) = \sup_{f:\deg(f)\lesssim \log n} \frac{\mathbb{E}_{H_1}[f(Y)]}{\sqrt{\mathbb{E}_{H_0}[f(Y)^2]}}$$

Decompose in $\widetilde{\psi}_G$ almost orthonormal basis numerator and denominator.

$$f(Y) = \sum_{G \in \text{basis}} \alpha_G \widetilde{\psi}_G.$$

Key proposition

Assume

The parameter (λ, k) to sample Y in H_0 is such that $\max\left\{\frac{k}{n}, \frac{\lambda k}{n}, \lambda\right\} \leq \mathrm{polylog}(n)$.

Almost orthonormal basis exists, *Adv* simplifies

There is an almost orthonormal basis and for (λ, k):

$$Adv(\lambda, \eta, k) = \sup_{\alpha} \frac{\mathbb{E}_{H_1}\left[\sum_{G \in basis} \alpha_G \widetilde{\psi}_G\right]}{\sqrt{\mathbb{E}_{H_0}\left[\sum_{G, G' \in \text{basis}} \alpha_G \alpha_{G'} \widetilde{\psi}_G \widetilde{\psi}_{G'}\right]}}$$

$$\leq \frac{1}{\sqrt{c}} \sup_{\alpha} \frac{\mathbb{E}_{H_1}\left[\sum_{G \in \text{ basis}} \alpha_G \widetilde{\psi}_G\right]}{\|\alpha\|_2} = \frac{1}{\sqrt{c}} LinAdv(\lambda, \eta, k),$$

Like orthonormal up to constants, explicit, linear: **easy**!

Main theorem

Assume

The parameter (λ, k) to sample Y in H_0 is such that
$\max\left\{\frac{k}{n}, \frac{\lambda k}{n}, \lambda\right\} \leq \mathrm{polylog}(n)$.
The perturbation η for H_1 is such that $\eta\frac{k^2}{n} \leq \mathrm{polylog}(n)$.

Main theorem

Assume

The parameter (λ, k) to sample Y in H_0 is such that $\max\left\{\frac{k}{n}, \frac{\lambda k}{n}, \lambda\right\} \leq \mathrm{polylog}(n)$.
The perturbation η for H_1 is such that $\eta\frac{k^2}{n} \leq \mathrm{polylog}(n)$.

Main Theorem

Under the conditions above for the planted sub-matrix model:

$$Adv(\lambda, \eta, k) \leq \frac{1}{\sqrt{c}} LinAdv(\lambda, \eta, k) = 1 + o(1)$$

so we conjecture $err_{poly}(\lambda, \eta, k)$ is large.

The region of the assumption and the perturbation is hard for poly-time algorithms.

Takeaways and extensions

Statistical-to-computational gap

The region of the assumptions is not hard for all functions so **there is a gap**. Conjecturally $1 - \Omega(1) = err_{IT}(\lambda, \eta, k) \ll err_{poly}(\lambda, \eta, k) = 1 - o(1)$.

Takeaways and extensions

Statistical-to-computational gap

The region of the assumptions is not hard for all functions so **there is a gap**. Conjecturally $1 - \Omega(1) = err_{IT}(\lambda, \eta, k) \ll err_{poly}(\lambda, \eta, k) = 1 - o(1)$.

Novelty

- proof technique via almost orthonormal basis;
- more explicit ;
- potentially sharper.

Takeaways and extensions

Statistical-to-computational gap

The region of the assumptions is not hard for all functions so **there is a gap**. Conjecturally $1-\Omega(1) = err_{IT}(\lambda, \eta, k) \ll err_{poly}(\lambda, \eta, k) = 1 - o(1)$.

Novelty

- proof technique via almost orthonormal basis;
- more explicit ;
- potentially sharper.

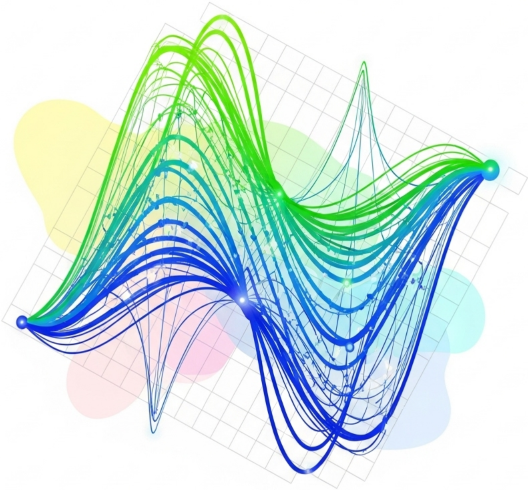
Extensions

Can let D vary, perturb k instead of λ, and nonasympotic result.
Many more models admit an almost orthonormal basis: stochastic block model, allow for fixed latent size, estimation instead of testing.

Concluding

Thank you!

Image credit: Gemini; prompt: *Generate an image with the following prompt "Computational lower bounds via almost orthonormal polynomials". Do not put text, and make the background white.*



References I

[Ban+22] Afonso S. Bandeira et al. "The Franz-Parisi Criterion and Computational Trade-offs in High Dimensional Statistics". In: *Advances in Neural Information Processing Systems*. Oct. 2022. (Visited on 02/05/2025).

[Bar+16] Boaz Barak et al. *A Nearly Tight Sum-of-Squares Lower Bound for the Planted Clique Problem*. Apr. 2016. DOI: 10.48550/arXiv.1604.03084. arXiv: 1604.03084 [cs]. (Visited on 07/19/2023).

[BB20] Matthew Brennan and Guy Bresler. "Reducibility and Statistical-Computational Gaps from Secret Leakage". In: *Proceedings of Thirty Third Conference on Learning Theory*. PMLR, 2020, pp. 648–847. (Visited on 02/05/2025).

References II

[BPW18] Afonso S. Bandeira, Amelia Perry, and Alexander S. Wein. *Notes on Computational-to-Statistical Gaps: Predictions Using Statistical Physics*. Apr. 2018. arXiv: 1803.11132 [cs, stat]. (Visited on 06/27/2023).

[Che+25] Siyu Chen et al. *An Optimized Franz-Parisi Criterion and Its Equivalence with SQ Lower Bounds*. June 2025. DOI: 10.48550/arXiv.2506.06259. arXiv: 2506.06259 [math]. (Visited on 06/12/2025).

[EGV24] Bertrand Even, Christophe Giraud, and Nicolas Verzelen. "Computation-Information Gap in High-Dimensional Clustering". In: *Proceedings of Thirty Seventh Conference on Learning Theory*. PMLR, June 2024, pp. 1646–1712. (Visited on 02/05/2025).

References III

[EGV25a] Bertrand Even, Christophe Giraud, and Nicolas Verzelen. *Computational Barriers for Permutation-Based Problems, and Cumulants of Weakly Dependent Random Variables*. July 2025. DOI: 10.48550/arXiv.2507.07946. arXiv: 2507.07946 [math]. (Visited on 07/14/2025).

[EGV25b] Bertrand Even, Christophe Giraud, and Nicolas Verzelen. *Computational Lower Bounds in Latent Models: Clustering, Sparse-Clustering, Biclustering*. June 2025. DOI: 10.48550/arXiv.2506.13647. arXiv: 2506.13647 [math]. (Visited on 07/14/2025).

References IV

[GMZ22] David Gamarnik, Cristopher Moore, and Lenka Zdeborová. "Disordered Systems Insights on Computational Hardness". In: *Journal of Statistical Mechanics: Theory and Experiment* 2022.11 (Nov. 2022), p. 114015. ISSN: 1742-5468. DOI: 10.1088/1742-5468/ac9cc8. arXiv: 2210.08312 [cond-mat, stat]. (Visited on 07/20/2023).

[Hop18] Samuel Brink Klevit Hopkins. "Statistical Inference and the Sum of Squares Method". PhD thesis. Cornell University Library, 2018. (Visited on 07/19/2023).

[KWB19] Dmitriy Kunisky, Alexander S. Wein, and Afonso S. Bandeira. *Notes on Computational Hardness of Hypothesis Testing: Predictions Using the Low-Degree Likelihood Ratio*. July 2019. arXiv: 1907.11636 [cs, math, stat]. (Visited on 06/27/2023).

References V

[MW22] Andrea Montanari and Alexander S. Wein. *Equivalence of Approximate Message Passing and Low-Degree Polynomials in Rank-One Matrix Estimation*. Dec. 2022. arXiv: 2212.06996 [math, stat]. (Visited on 07/03/2023).

[SW22] Tselil Schramm and Alexander S. Wein. "Computational Barriers to Estimation from Low-Degree Polynomials". In: *The Annals of Statistics* 50.3 (June 2022). ISSN: 0090-5364. DOI: 10.1214/22-AOS2179. arXiv: 2008.02269 [cs, math, stat]. (Visited on 07/03/2023).

[SW25] Youngtak Sohn and Alexander S. Wein. *Sharp Phase Transitions in Estimation with Low-Degree Polynomials*. Feb. 2025. DOI: 10.48550/arXiv.2502.14407. arXiv: 2502.14407 [math]. (Visited on 04/01/2025).

References VI

[Wei25] Alexander S. Wein. *Computational Complexity of Statistics: New Insights from Low-Degree Polynomials*. June 2025. DOI: 10.48550/arXiv.2506.10748. arXiv: 2506.10748 [math]. (Visited on 06/18/2025).

Proof sketch

1. Clarifications on low-degree (if needed);
2. what is the candidate basis;
3. invariance ideas;
4. proving almost orthonormality in practice.

Performance in hypothesis test

Idea: minimize type I and type II errors over a *class of functions*.

Statistical optimality

$$err_{IT}(\lambda, \eta, k) := \inf_{t \text{ measurable test}} \mathbb{P}_{H_0}[t(Y) = H_1] + \mathbb{P}_{H_1}[t(Y) = H_0].$$

Poly-time optimality

$$err_{poly}(\lambda, \eta, k) := \inf_{t \text{ measurable poly-time test}} \mathbb{P}_{H_0}[t(Y) = H_1] + \mathbb{P}_{H_1}[t(Y) = H_0].$$

Performance in hypothesis test

Idea: minimize type I and type II errors over a *class of functions*.

Statistical optimality

$$err_{IT}(\lambda, \eta, k) := \inf_{t \text{ measurable test}} \mathbb{P}_{H_0}[t(Y) = H_1] + \mathbb{P}_{H_1}[t(Y) = H_0].$$

Poly-time optimality

$$err_{poly}(\lambda, \eta, k) := \inf_{t \text{ measurable poly-time test}} \mathbb{P}_{H_0}[t(Y) = H_1] + \mathbb{P}_{H_1}[t(Y) = H_0].$$

Statistical-to-computational gap [KWB19; BPW18]

Less studied **poly-time criterion is the important one in practice**.
If $1 - \Omega(1) = err_{IT}(\lambda, \eta, k) < err_{poly}(\lambda, \eta, k) = 1 - o(1)$ then η is hard.

The low-degree method & conjecture

Main issue: no clear idea of how to tackle poly-time tests so:

Conjecture [Hop18]

Poly-time test functions are less powerful than test functions thresholding polynomials of degree $D \lesssim \log n$:

$$err_{IT}(\lambda, \eta, k) \leq err_{LD}(\lambda, \eta, k) \overset{\text{conjecture}}{\leq} err_{poly}(\lambda, \eta, k),$$

$$err_{LD}(\lambda, \eta, k) := \inf_{t:\text{thresh. poly deg.} \lesssim \log n} \mathbb{P}_{H_0}[t(Y) = H_1] + \mathbb{P}_{H_1}[t(Y) = H_0].$$

The low-degree method & conjecture

Main issue: no clear idea of how to tackle poly-time tests so:

Conjecture [Hop18]

Poly-time test functions are less powerful than test functions thresholding polynomials of degree $D \lesssim \log n$:

$$err_{IT}(\lambda, \eta, k) \leq err_{LD}(\lambda, \eta, k) \overset{\text{conjecture}}{\leq} err_{poly}(\lambda, \eta, k),$$

$$err_{LD}(\lambda, \eta, k) := \inf_{t:\text{thresh. poly deg.} \lesssim \log n} \mathbb{P}_{H_0}[t(Y) = H_1] + \mathbb{P}_{H_1}[t(Y) = H_0].$$

Statistical-to-computational gap

If $err_{IT}(\lambda, \eta, k)$ is small and $err_{LD}(\lambda, \eta, k)$ is **large then so is** $err_{poly}(\lambda, \eta, k)$ and we have a gap.

Last simplification: bounding the advantage

Main issue: Working on $err_{LD}(\lambda, \eta, k)$ is hard, find a surrogate condition.

Advantage bound

If:

$$Adv(\lambda, \eta, k) := \sup_{f:\deg(f) \lesssim \log n} \frac{\mathbb{E}_{H_1}[f(Y)]}{\sqrt{\mathbb{E}_{H_0}[f(Y)^2]}} = 1 + o(1),$$

then **expect** $err_{LD}(\lambda, \eta, k) \geq 1 - o(1)$.

Statistical-to-computational gap

Show $Adv(\lambda, \eta, k)$ is bounded, **extrapolate** $err_{LD}(\lambda, \eta, k)$ is large, **conjecture** $err_{poly}(\lambda, \eta, k)$ is large.

The orthonormal trick

When (λ, k) is such that Y is "pure noise" there is an **explicit orthonormal basis**

The orthonormal trick

When (λ, k) is such that Y is "pure noise" there is an **explicit orthonormal basis**
In planted sub-matrix: $(\lambda, k) = (0, 0)$ is such that $Y_{ij} \overset{i.i.d.}{\sim} \mathrm{Rad}\left(\frac{1}{2}\right)$.

$$x_i \overset{i.i.d.}{\sim} \sqrt{0}\mathrm{Ber}(0) \text{ for all } i \in [n] \implies Y_{ij} = \begin{cases} 1 & \text{with probability } \frac{1+0}{2} \\ -1 & \text{with probability } \frac{1-0}{2} \end{cases}.$$

The orthonormal trick

When (λ, k) is such that Y is "pure noise" there is an **explicit orthonormal basis**
In planted sub-matrix: $(\lambda, k) = (0, 0)$ is such that $Y_{ij} \overset{i.i.d.}{\sim} \mathrm{Rad}\left(\frac{1}{2}\right)$.

$$x_i \overset{i.i.d.}{\sim} \sqrt{0}\mathrm{Ber}(0) \text{ for all } i \in [n] \implies Y_{ij} = \begin{cases} 1 & \text{with probability } \frac{1+0}{2} \\ -1 & \text{with probability } \frac{1-0}{2} \end{cases}.$$

Orthonormal basis for pure noise

$$Y^G = \prod_{(i,j)\in E} Y_{ij},$$

for all $G = (V, E)$ labelled sub-graphs of the $n \times n$ complete graph.

Correlations in canonical basis

$$
\begin{aligned}
\mathbb{E}_{(0,0)}\left[Y^G Y^{G'}\right] &= \mathbb{E}_{(\lambda,k)=(0,0)}\left[\prod_{(i,j)\in E,E'} \underbrace{Y_{ij}^2}_{=1\text{ a.s.}} \prod_{(i,j)\in E \text{ only}} Y_{ij} \prod_{(i,j)\in E' \text{ only}} Y_{ij}\right] \\
&= \mathbb{E}_{(0,0)}\left[\mathbb{E}\left[\prod_{(i,j)\in E \text{ only}} Y_{ij} \prod_{(i,j)\in E' \text{ only}} Y_{ij} \mid (x_i x_j)_{i,j\in[n]}\right]\right] \\
&= \mathbb{E}_{(0,0)}\left[\prod_{(i,j)\in E \text{ only}} X_{ij} \prod_{(i,j)\in E' \text{ only}} X_{ij}\right] \\
&= \delta_{G=G'}.
\end{aligned}
$$

The orthonormal trick to bound *Adv*

If we have an orthonormal basis in H_0 can **decompose functions along such basis**:

$$f(Y) = \sum_{G \in \text{basis}} \alpha_G Y^G, \qquad \alpha_G = \mathbb{E}_{(\lambda,k)=(0,0)}\left[f(Y)Y^G\right].$$

The orthonormal trick to bound *Adv*

If we have an orthonormal basis in H_0 can **decompose functions along such basis**:

$$f(Y) = \sum_{G \in \text{basis}} \alpha_G Y^G, \qquad \alpha_G = \mathbb{E}_{(\lambda,k)=(0,0)} \left[f(Y) Y^G \right].$$

Rewrite the advantage

$$Adv(\lambda = 0, \eta, k = 0) = \sup_{f: deg(f) \lesssim \log n} \frac{\mathbb{E}_{H_1}[f(Y)]}{\sqrt{\mathbb{E}_{H_0}[f(Y)^2]}}$$

Decompose in Y^G orthonormal basis numerator and denominator.

The orthonormal trick to bound *Adv*

If we have an orthonormal basis can **decompose functions along such basis**:

$$f(Y) = \sum_{G \in \text{basis}} \alpha_G Y^G, \qquad \alpha_G = \mathbb{E}_{(\lambda,k)=(0,0)} \left[f(Y) Y^G \right].$$

Rewrite the advantage

$$Adv(\lambda = 0, \eta, k = 0) = \sup_{\alpha} \frac{\mathbb{E}_{H_1}\left[\sum_{G \in basis} \alpha_G Y^G\right]}{\sqrt{\mathbb{E}_{H_0}\left[\sum_{G,G' \in \text{basis}} \alpha_G \alpha'_G Y^G Y^{G'}\right]}}$$

$$= \sup_{\alpha} \frac{\mathbb{E}_{H_1}\left[\sum_{G \in \text{basis}} \alpha_G Y^G\right]}{\|\alpha\|_2} = LinAdv((\eta, k), (\lambda, k))$$

by orthonormality the denominator simplifies and the **advantage is a linear function**.

Outside of pure noise $(\lambda, k), \lambda, k \neq 0$

Question: prove when cannot detect perturbations in poly-time

$$H_0 \ : \ Y \text{ structure is } (\lambda, k), \qquad H_1 : Y \text{ structure is } (\lambda + \eta, k).$$

Conjecturally hard when decomposition along basis of $Adv(\lambda, \eta, k)$ but at $(\lambda, k), \lambda, k \neq 0$ **not explicit ortho basis**!

Outside of pure noise $(\lambda, k), \lambda, k \neq 0$

Question: prove when cannot detect perturbations **in poly-time**

$$H_0 \; : \; Y \text{ structure is } (\lambda, k), \qquad H_1 : Y \text{ structure is } (\lambda + \eta, k).$$

Conjecturally hard when decomposition along basis of $Adv(\lambda, \eta, k)$ but at $(\lambda, k), \lambda, k \neq 0$ **not explicit ortho basis**!

Problem

When $(\lambda, k), \lambda, k \neq 0$ the basis $\{Y^G\}_G$ is **not orthonormal**!

$$\mathbb{E}_{H_0:(\lambda,k)}\left[Y^G Y^{G'}\right] = \lambda^{\#\text{edges in symm. diff.}} \left(\frac{k}{n}\right)^{\#\text{vertices symm. diff}}.$$

No explicit formula. No linearization of advantage. In literature complicated recursive implicit solutions [SW22; SW25].

Outside of pure noise, $(\lambda, k), \lambda, k \neq 0$

Question: prove when cannot detect perturbations **in poly-time**,

$$H_0 \ : \ Y \text{ structure is } (\lambda, k), \qquad H_1 : Y \text{ structure is } (\lambda + \eta, k).$$

Basic idea: decomposition along basis of $Adv(\lambda, \eta, k)$ but at $(\lambda, k), \lambda, k \neq 0$ **not explicit ortho basis**!
In literature implicit recursive solutions [SW22; SW25].

Outside of pure noise, $(\lambda, k), \lambda, k \neq 0$

Question: prove when cannot detect perturbations **in poly-time**,

$$H_0 \ : \ Y \text{ structure is } (\lambda, k), \qquad H_1 : Y \text{ structure is } (\lambda + \eta, k).$$

Basic idea: decomposition along basis of $Adv(\lambda, \eta, k)$ but at $(\lambda, k), \lambda, k \neq 0$ **not explicit ortho basis**!
In literature implicit recursive solutions [SW22; SW25].

Our solution: almost orthonormal basis

Find a collection of functions $(\psi_G)_G$ forming a basis of H_0, H_1:

$$f(Y) = \sum_{G \in \text{ basis}} \alpha_G \psi_G,$$

such that for some constants:

$$c \|\alpha\|_2^2 \leq \mathbb{E}_{(\lambda,k)} \left[\sum_{G, G' \in \text{basis}} \alpha_G \alpha_{G'} \psi_G \psi_{G'} \right] \leq C \|\alpha\|_2^2 .$$

Adjusting the orthonormal basis

As we said the basis:

$$(Y^G)_G, \qquad Y^G = \prod_{(i,j)\in E} Y_{ij},$$

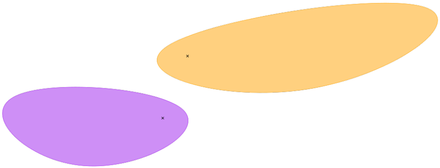
is **not orthonormal** when $(\lambda, k), \lambda, k \neq 0$, and has correlations dep. on symm. diff.:

$$\mathbb{E}_{H_0:(\lambda,k)}\left[Y^G Y^{G'}\right] = \lambda^{|E_{G\triangle G'}|}\left(\frac{k}{n}\right)^{V_{G\triangle G'}}.$$

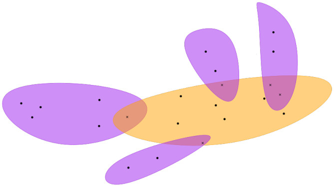
Rough intuition

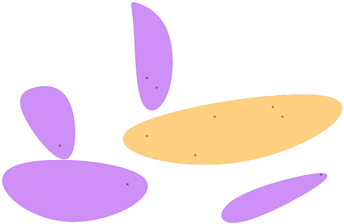
Adjust the Y^G basis to decrease correlations.
Use independence of random variables when G and G' are disconnected (latent Bernoullis $(x_i x_j)^G_{i,j\in[n]}, (x_i x_j)^{G'}_{i,j\in[n]}$ are indep.).

Some visuals with blobs of vertices



The two graphs correlate $\mathbb{E}_{H_0:(\lambda,k)}\left[Y^G Y^{G'}\right] \neq 0$.





The two graphs correlate $\mathbb{E}_{H_0:(\lambda,k)}\left[Y^G Y^{G'}\right] \neq 0$ in different ways.

Partial adjustment

Centered basis

The basis:

$$\widehat{Y}^G := Y^G - \mathbb{E}_{H_0}\left[Y^G\right],$$

correlates less than Y^G.

Indeed:

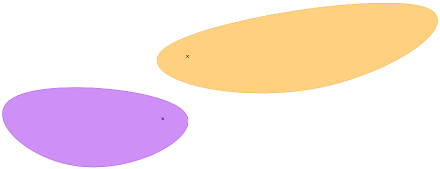
$$\mathbb{E}_{H_0:(\lambda,k)}\left[\widehat{Y}^G \widehat{Y}^{G'}\right] = \mathbb{E}_{H_0:(\lambda,k)}\left[Y^G Y^{G'}\right] - \mathbb{E}_{H_0:(\lambda,k)}\left[Y^G\right]\mathbb{E}_{H_0:(\lambda,k)}\left[Y^{G'}\right]$$

If G, G' are disconnected then $G \triangle G' = G \cup G'$ and the correlation is zero.

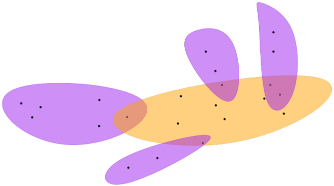
But this is not enough

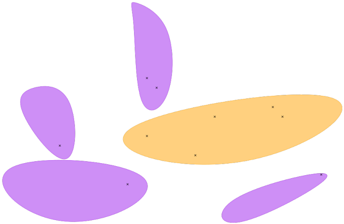
Can use more independence to zero out correlations.

Some visuals with blobs of vertices



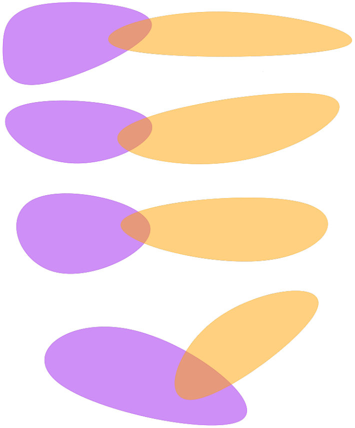
The two graphs **do not** correlate
$\mathbb{E}_{H_0:(\lambda,k)}\left[\widehat{Y}^G\widehat{Y}^{G'}\right] = 0.$



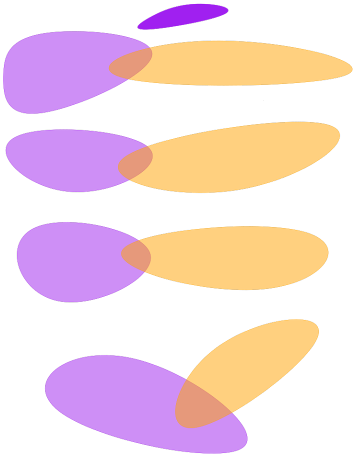


Above graphs correlate
$\mathbb{E}_{H_0:(\lambda,k)}\left[\widehat{Y}^G\widehat{Y}^{G'}\right] \neq 0$ below do not.

Some visuals with blobs of vertices



The two graphs correlate
$\mathbb{E}_{H_0:(\lambda,k)}\left[\widehat{Y}^{G}\widehat{Y}^{G'}\right] \neq 0.$



The two graphs correlate
$\mathbb{E}_{H_0:(\lambda,k)}\left[\widehat{Y}^{G}\widehat{Y}^{G'}\right] \neq 0$

Final fix

Basis proposal

The basis:

$$\overline{Y}^G := \prod_{\ell=1}^{m} Y^{G_\ell} - \mathbb{E}_{H_0:(\lambda,k)}\left[Y^{G_\ell}\right], \qquad G = (G_\ell)_{\ell=1}^m \text{ conn. comp.}$$

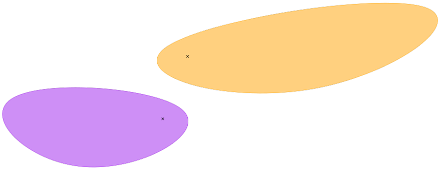
Correlates less than Y^G, $\widehat{Y}^G$.

Imagine G, G' have shared edges/vertices (so $\mathbb{E}_{H_0:(\lambda,k)}\left[\widehat{Y}^G \widehat{Y}^{G'}\right] \neq 0$), but one conn. comp. in G is isolated from all of G', then:

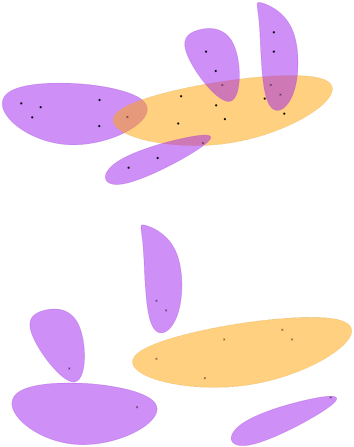
$$\mathbb{E}_{H_0:(\lambda,k)}\left[\overline{Y}^G \overline{Y}^{G'}\right] = \mathbb{E}_{H_0:(\lambda,k)}\left[\overline{Y}^{G\setminus G_{\ell^\star}} \overline{Y}^{G'}\right] \mathbb{E}_{H_0:(\lambda,k)}\left[\overline{Y}^{G_{\ell^\star}}\right] = 0,$$

since in the basis we center connected components.

Some visuals with blobs of vertices

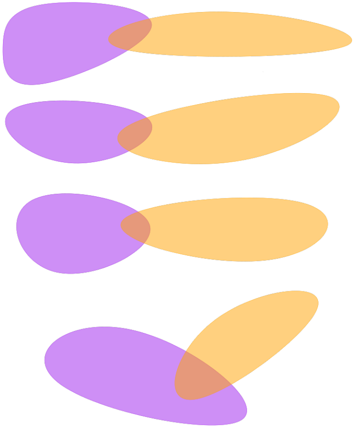


The two graphs **do not** correlate $\mathbb{E}_{H_0:(\lambda,k)}\left[\overline{Y}^{G}\overline{Y}^{G'}\right] = 0.$

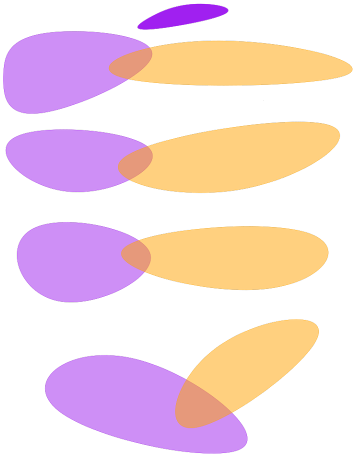


Above graphs correlate $\mathbb{E}_{H_0:(\lambda,k)}\left[\overline{Y}^{G}\overline{Y}^{G'}\right] \neq 0$ below do not.

Some visuals with blobs of vertices



The two graphs correlate
$\mathbb{E}_{H_0:(\lambda,k)}\left[\overline{Y}^{G}\overline{Y}^{G'}\right] \neq 0.$



The two graphs **do not** correlate
$\mathbb{E}_{H_0:(\lambda,k)}\left[\overline{Y}^{G}\overline{Y}^{G'}\right] = 0$

Making counting easier

The basis $(\overline{Y}^G)_G$ runs over **all labelled sub-graphs** of the complete n-graph that have less than $\lesssim \log n$ edges (degree constraint).

Counting matters

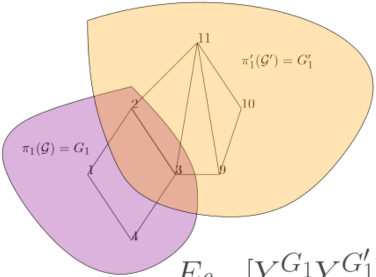
Enumerating such graphs is tedious, plus, the correlations depend on the symmetric difference:

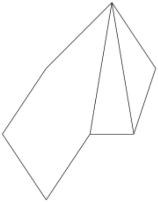
If two different labelled pairs (G_1, G_1'), (G_2, G_2') are such that $G_1 \simeq G_2$, $G_1' \simeq G_2'$ and they have the same symmetric difference then we count them twice.

Rough intuition

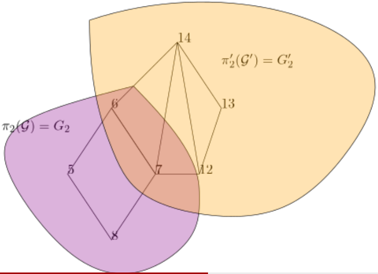
Group isomorphic graphs together and see how the equivalence classes generate different symmetric differences.

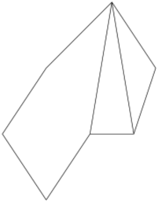
Visualization of symmetric difference





$$E_{\theta_{H_0}}[Y^{G_1}Y^{G'_1}] = E_{\theta_{H_0}}[Y^{G_2}Y^{G'_2}] = \lambda^9 \left(\frac{k}{n}\right)^7$$





Formalizing

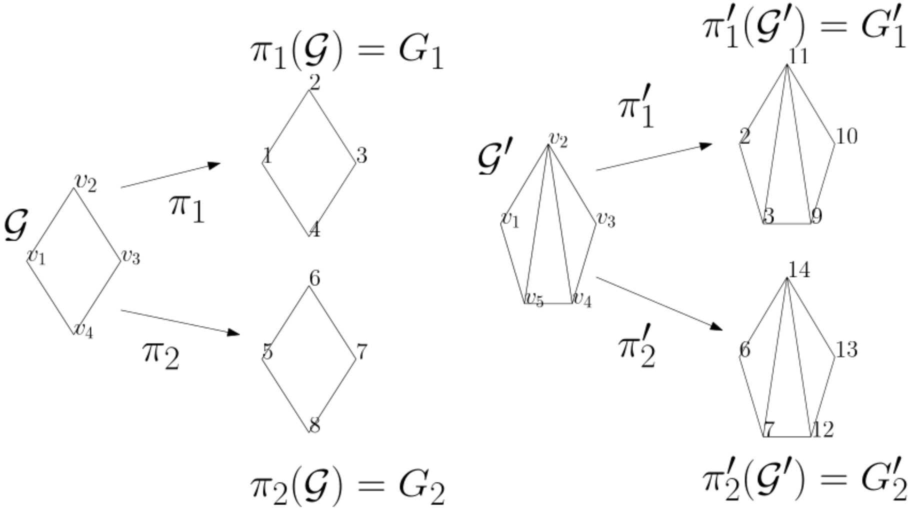
Collect G into labellings of graphs from an abstract space, i.e. $G = \pi(\mathcal{G})$ for some abstract $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, and $\pi : \mathcal{V} \mapsto [n]$ a **labelling**.
For two labellings $\pi(\mathcal{G}), \pi'(\mathcal{G})$ we have two graphs that come from the same "shape" in the abstract space.
Form the basis:

$$\overline{Y}^{\mathcal{G}} := \sum_{\pi \text{ labellings}} \overline{Y}^{\pi(\mathcal{G})}, \qquad \text{for all abstract graphs } \mathcal{G}.$$

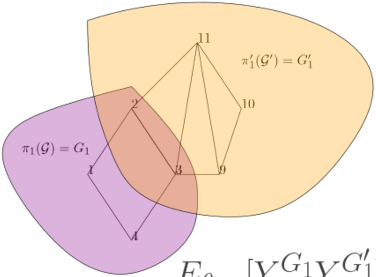
Double result

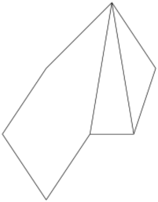
$\{\overline{Y}^{\mathcal{G}}\}_{\mathcal{G}}$ is a basis of **perm. invariant** polynomials, but the advantage $Adv(\lambda, \eta, k)$ is attained by a perm. invariant polynomial since H_0, H_1 are perm. invariant distributions.
No loss by working on this basis, plus simplified counting!

Visualization of labellings

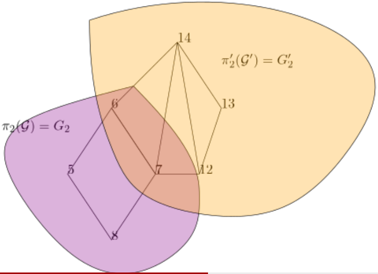


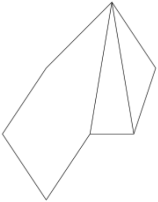
Visualization of symmetric difference





$$E_{\theta_{H_0}}[Y^{G_1}Y^{G'_1}] = E_{\theta_{H_0}}[Y^{G_2}Y^{G'_2}] = \lambda^9 \left(\frac{k}{n}\right)^7$$





Almost orthonormality exemplified

Recall that we want to write:

$$Adv(\lambda, \eta, k) = \sup_{f : deg(f) \lesssim \log n} \frac{\mathbb{E}_{H_1}[f(Y)]}{\sqrt{\mathbb{E}_{H_0}[f(Y)^2]}}$$

Use that Adv is attained by invariant polynomial.

Almost orthonormality exemplified

Recall that we want to write:

$$Adv(\lambda, \eta, k) = \sup_{\substack{f: deg(f) \lesssim \log n \\ f\ perm.invariant}} \frac{\mathbb{E}_{H_1}[f(Y)]}{\sqrt{\mathbb{E}_{H_0}[f(Y)^2]}}$$

Decompose along perm. invariant basis $\{\overline{Y}^{\mathcal{G}}\}$.

Almost orthonormality exemplified

Recall that we want to write:

$$Adv(\lambda, \eta, k) = \sup_{\alpha} \frac{\mathbb{E}_{H_1}\left[\sum_{\mathcal{G}} \alpha_G \overline{Y}^{\mathcal{G}}\right]}{\sqrt{\mathbb{E}_{H_0}\left[\sum_{\mathcal{G},\mathcal{G}'} \alpha_{\mathcal{G}} \alpha_{\mathcal{G}'} \overline{Y}^{\mathcal{G}} \overline{Y}^{\mathcal{G}'}\right]}}$$

Use almost orthonormality to linearize.

Almost orthonormality exemplified

Recall that we want to write:

$$
\begin{aligned}
Adv(\lambda,\eta,k) &= \sup_{\alpha} \frac{\mathbb{E}_{H_1}\left[\sum_{\mathcal{G}} \alpha_G \overline{Y}^{\mathcal{G}}\right]}{\sqrt{\mathbb{E}_{H_0}\left[\sum_{\mathcal{G},\mathcal{G}'} \alpha_{\mathcal{G}}\alpha_{\mathcal{G}'} \overline{Y}^{\mathcal{G}}\overline{Y}^{\mathcal{G}'}\right]}} \\
&\leq \frac{1}{\sqrt{c}} \sup_{\alpha} \frac{\mathbb{E}_{H_1}\left[\sum_{\mathcal{G}} \alpha_{\mathcal{G}} \overline{Y}^{\mathcal{G}}\right]}{\|\alpha\|} \\
&= \frac{1}{\sqrt{c}} LinAdv(\lambda,\eta,k).
\end{aligned}
$$

How to establish

$$
c\,\|\alpha\| \leq \mathbb{E}_{H_0:(\lambda,k)}\left[\sum_{\mathcal{G},\mathcal{G}'} \alpha_{\mathcal{G}}\alpha_{\mathcal{G}'} \overline{Y}^{\mathcal{G}}\overline{Y}^{\mathcal{G}'}\right] \leq C\,\|\alpha\|?
$$

Preliminary

The basis $\{\overline{Y}^{\mathcal{G}}\}$ is **not normalized** because we sum over many graphs, indeed:

$$\overline{Y}^{\mathcal{G}} = \underbrace{\sum_{\pi \text{ labellings}}}_{\textbf{exploding number of labellings}} \underbrace{\overline{Y}^{\pi(\mathcal{G})}}_{\text{norm order one}},$$

so we **need to normalize it**.

Technical

There exist a way to normalize the basis by rescaling $\overline{Y}^{\mathcal{G}}$ into $\widetilde{Y}^{\mathcal{G}} = \frac{\overline{Y}^{\mathcal{G}}}{\sqrt{\nu(\mathcal{G})}}$ such that:

$$\mathbb{E}_{H_0:(\lambda,k)}\left[(\widetilde{Y}^{\mathcal{G}})^2\right] = \mathbb{E}_{H_0:(\lambda,k)}\left[\frac{(\overline{Y}^{\mathcal{G}})^2}{\nu(\mathcal{G})}\right] \approx \text{constant order.}$$

Now all the variances are of the same size.

For the rescaled basis $\widetilde{Y}^{\mathcal{G}} = \frac{\overline{Y}^{\mathcal{G}}}{\sqrt{\nu(\mathcal{G})}}$ we then rewrite the denominator as a quadratic form:

$$\mathbb{E}_{H_0:(\lambda,k)}\left[\sum_{\mathcal{G},\mathcal{G}'} \alpha_{\mathcal{G}}\alpha_{\mathcal{G}'}\widetilde{Y}^{\mathcal{G}}\widetilde{Y}^{\mathcal{G}'}\right] = \alpha^{\top}\mathbb{E}_{H_0:(\lambda,k)}\left[\widetilde{Y}\widetilde{Y}^{\top}\right]\alpha,$$

where:

$$\mathbb{E}_{H_0:(\lambda,k)}\left[\widetilde{Y}\widetilde{Y}^{\top}\right] = \text{Gram matrix of correlations for } \{\widetilde{Y}^{\mathcal{G}}\}_{\mathcal{G}} \text{ basis.}$$

Aim

Show the eigenvalues of the Gram matrix are all constant.

Gershgorin criterion to the rescue

By Gershgorin circle theorem the eigenvalues of a Gram matrix are within the circles

$$\sup_i \left\{ G_{ii} \pm \sum_{j \neq i} |G_{ij}| \right\} = \sup_{\mathcal{G}} \left\{ \mathbb{E}_{H_0:(\lambda,k)} \left[\widetilde{Y}^{\mathcal{G}} \right] \pm \sum_{\mathcal{G}' \neq \mathcal{G}} \left| \mathbb{E}_{H_0:(\lambda,k)} \left[\widetilde{Y}^{\mathcal{G}} \widetilde{Y}^{\mathcal{G}'} \right] \right| \right\}.$$

So **show that it is a constant**.

Advantage of this view

Can go step-by-step, from correlations of labelled graphs $\pi(\mathcal{G})$, to correlations of abstract graphs $\sum_{\pi \in \text{ labellings}}$ and so on.

Steps for almost orthonormality via Gershgorin

1. the basis Y^G correlation is a symmetric difference, the candidate basis $\widetilde{Y}^{\pi(\mathcal{G})}$ correlations **approximate** symmetric differences;

Steps for almost orthonormality via Gershgorin

1. the basis Y^G correlation is a symmetric difference, the candidate basis $\widetilde{Y}^{\pi(\mathcal{G})}$ correlations **approximate** symmetric differences;
2. when summing over π labellings, two abstract graphs correlate as:

$$\mathbb{E}_{H_0:(\lambda,k)}\left[\widetilde{Y}^{\mathcal{G}}\widetilde{Y}^{\mathcal{G}'}\right] \lesssim (\log n)^{d(\mathcal{G},\mathcal{G}')},$$

for some proper distance d between graphs;

Steps for almost orthonormality via Gershgorin

1. the basis Y^G correlation is a symmetric difference, the candidate basis $\widetilde{Y}^{\pi(\mathcal{G})}$ correlations **approximate** symmetric differences;
2. when summing over π labellings, two abstract graphs correlate as:

$$\mathbb{E}_{H_0:(\lambda,k)}\left[\widetilde{Y}^{\mathcal{G}}\widetilde{Y}^{\mathcal{G}'}\right] \lesssim (\log n)^{d(\mathcal{G},\mathcal{G}')},$$

for some proper distance d between graphs;
3. summing over abstract graphs, the control by a distance is enough to show that the eigenvalues of the Gram matrix $\mathbb{E}_{H_0:(\lambda,k)}\left[\widetilde{Y}\widetilde{Y}^\top\right]$ are constant, and we have the almost orthonormality:

$$c\,\|\alpha\| \leq \mathbb{E}_{H_0:(\lambda,k)}\left[\sum_{\mathcal{G},\mathcal{G}'}\alpha_{\mathcal{G}}\alpha_{\mathcal{G}'}\widetilde{Y}^{\mathcal{G}}\widetilde{Y}^{\mathcal{G}'}\right] \leq C\,\|\alpha\|.$$

Bounding the advantage?

By almost orthonormality $\implies$ advantage takes linear form:

$$Adv(\lambda, \eta, k) \leq LinAdv(\lambda, \eta, k) = \frac{1}{\sqrt{c}} \sup_{\alpha} \frac{\mathbb{E}_{H_1}\left[\sum_{\mathcal{G}} \alpha_{\mathcal{G}} \widetilde{Y}^{\mathcal{G}}\right]}{\|\alpha\|},$$

which is easy to upper bound.