Two questions about randomized in numerical linear algebra

1. ~~Multiway random walks for reducing variance~~ (arxiv.org/1608.04361)
2. An eigenvalue law for "triangle Laplacians"
3. An approximation bound for sampling zonotopes for clusters.

David F. Gleich

Purdue University

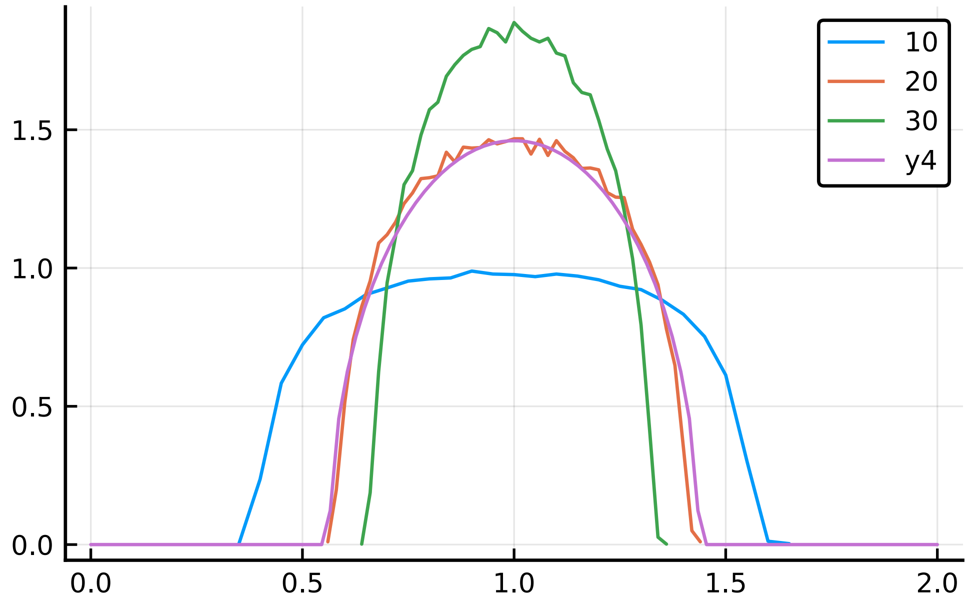
3 slides to pose a simple question.
I don’t have an answer ☺

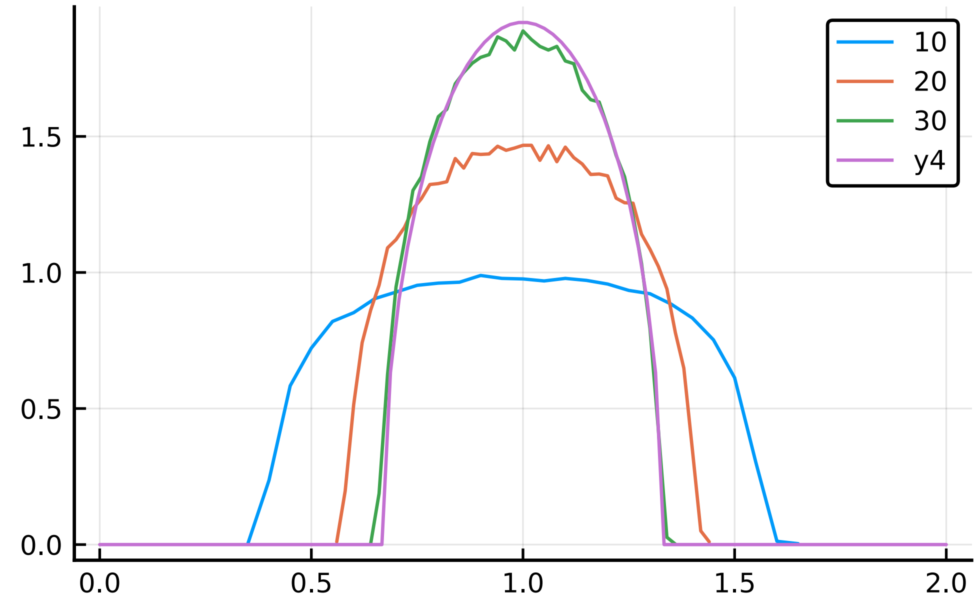
Eigenvalue laws of random ER graphs are well known

The adjacency matrix has a semi-circle law.

The normalized Laplacian matrix also has a semi-circle law with a specific gap

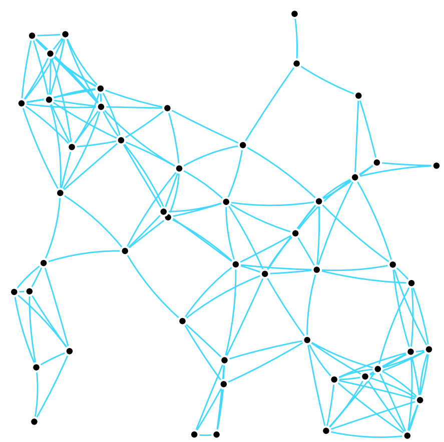
- C. Hoffman, M. Kahle, and E. Paquette. Spectral gaps of random graphs and applications to random topology. arXiv:1201.0425.

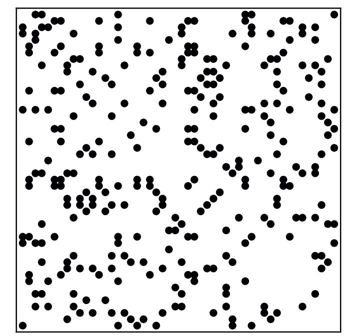




A idea based on our recent work

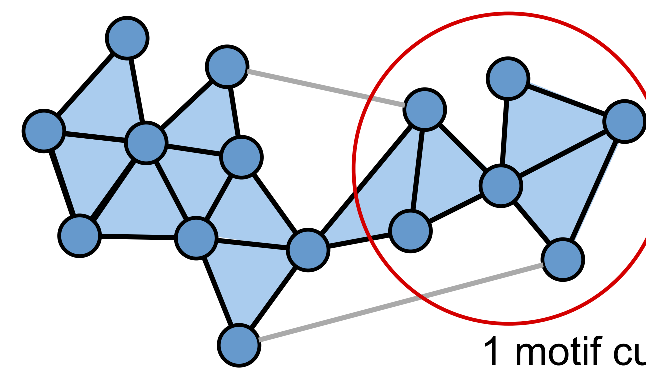
given a graph $G = (V, E)$ and its adjacency matrix $\boldsymbol{A}$





consider using the weighted matrix $\boldsymbol{W} = \boldsymbol{A}^2 \odot \boldsymbol{A}$

The matrix $\boldsymbol{W} = \boldsymbol{A}^2 \odot \boldsymbol{A}$ arises from our motif and higher-order clustering framework when using triangles as the motif.



1 motif cut
11 motif end points
motif conductance = 1 / 11

Benson, Gleich, Leskovec, Science 2016.

The eigenvalues of the normalized Laplacian of this weighted matrix do not follow a Marchenko–Pastur law

Method.

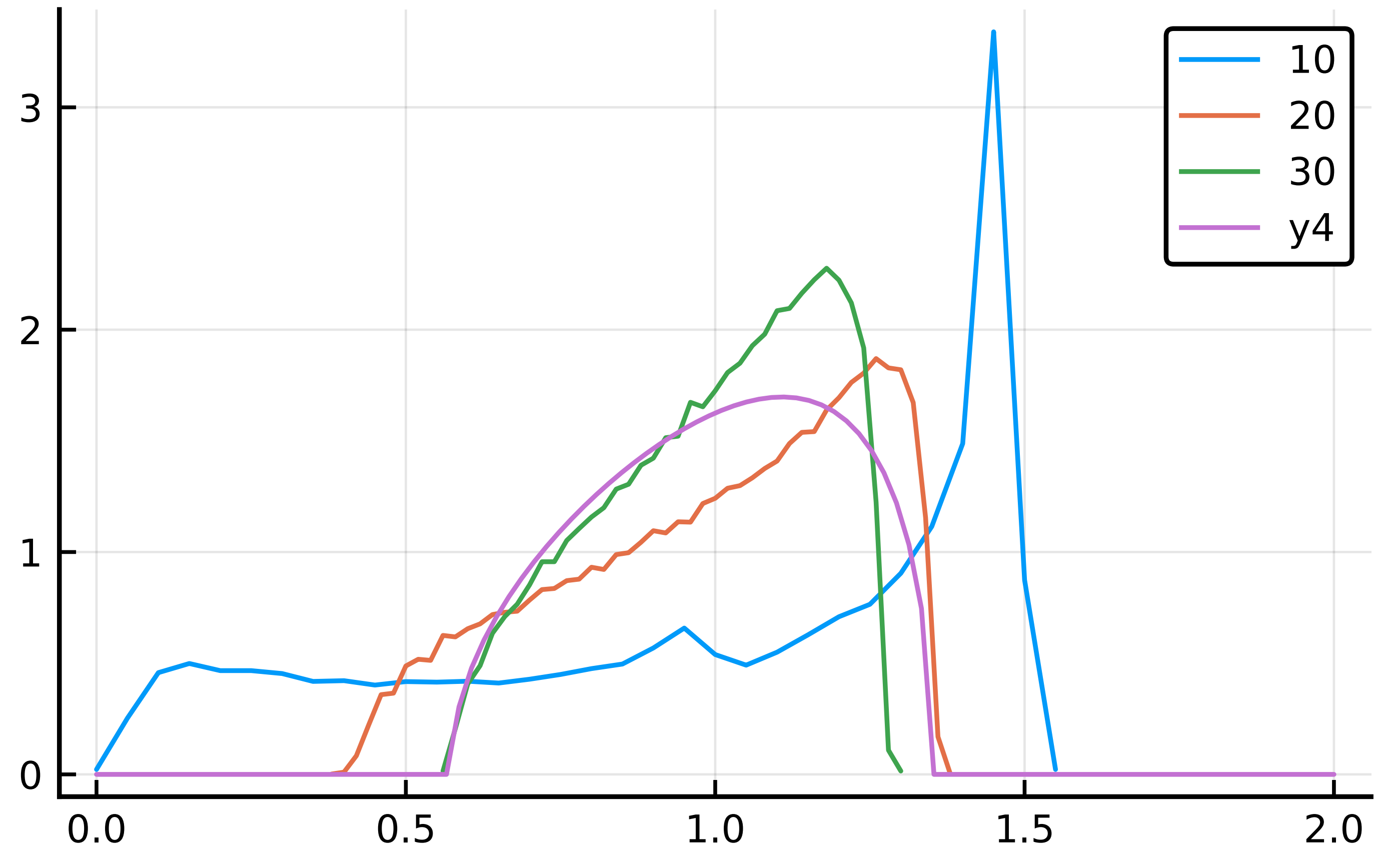
Generate a random ER graph, compute

$$\boldsymbol{W} = \boldsymbol{A}^2 \odot \boldsymbol{A}$$

Compute eigenvalues of the normalized Laplacian of the largest component of $\boldsymbol{W}$

Our question.

What is this distribution?



Code to reproduce https://gist.github.com/dgleich/4d4becc858e4a7d7952af6c66c99e7b9

Another problem where I don't know the solution.

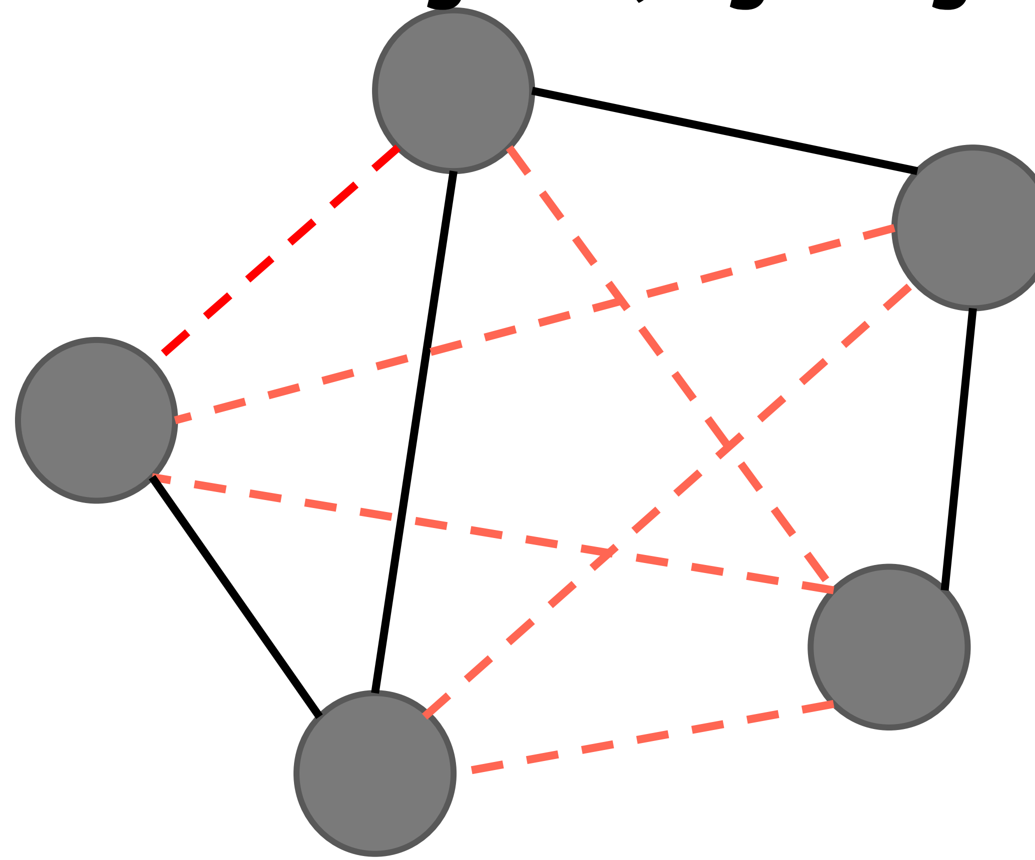
Motivation. Correlation Clustering of Low-Rank Marices.

Veldt, Wirth, Gleich, 2016. arXiv:1611.07305

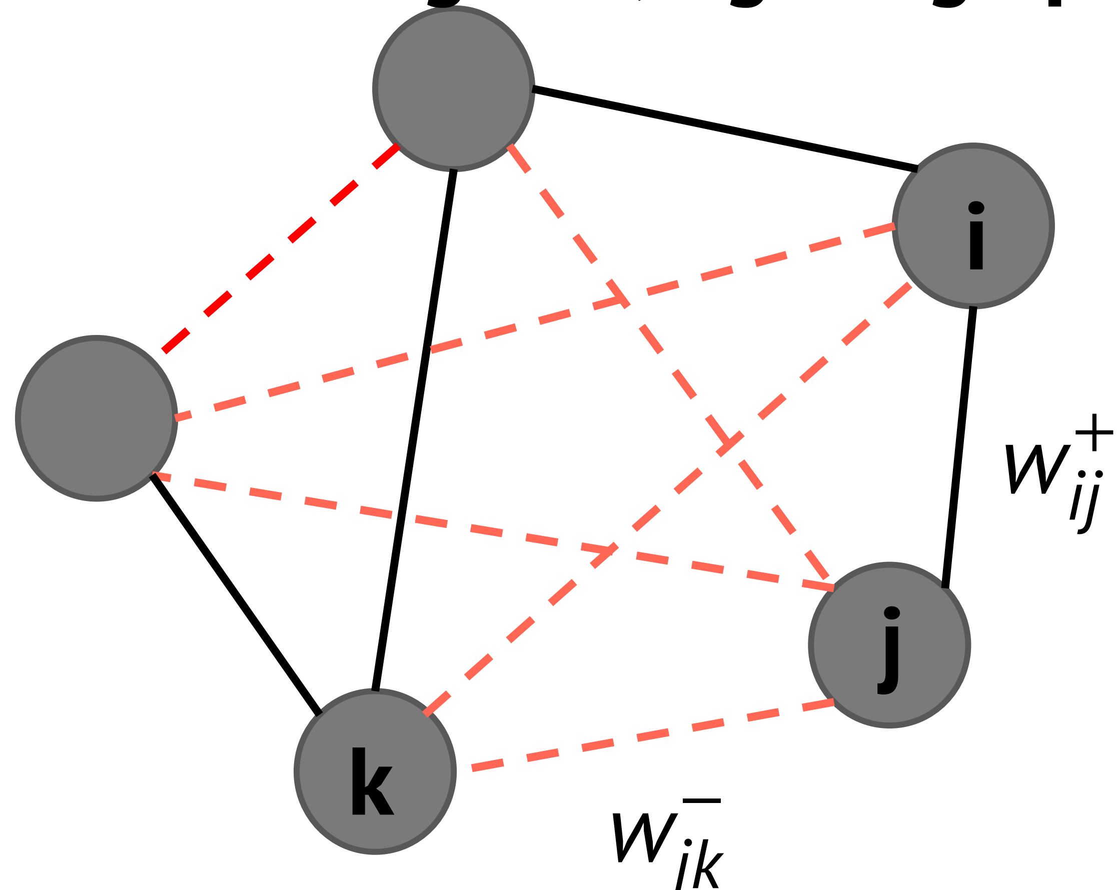
We have a heuristic randomized (and useful) algorithm that we don't have good analysis for ☺

Correlation clustering involves a weighted, signed graph

Edges in a signed graph indicate **similarity (+)** or **dissimilarity (-)**

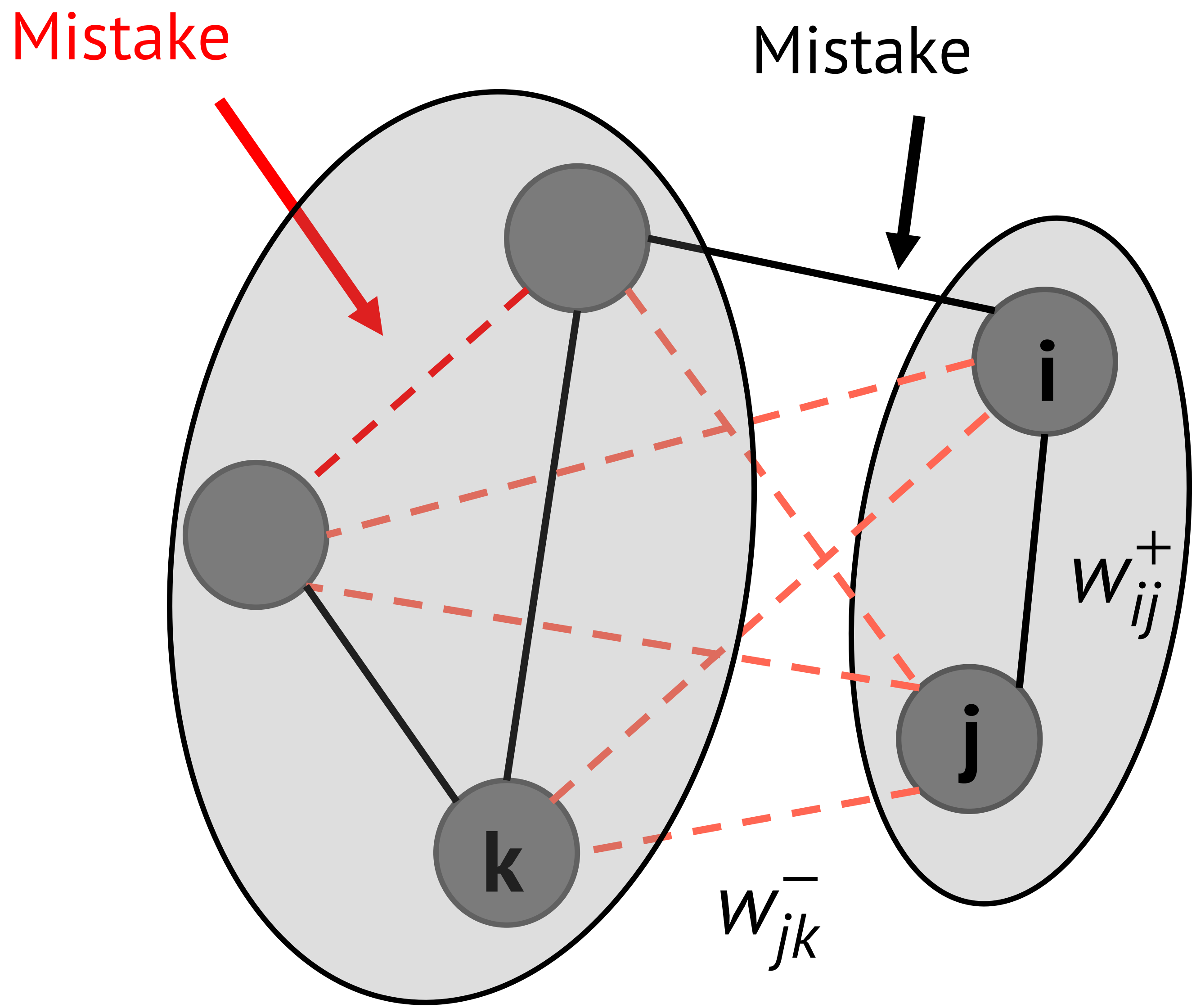


Correlation clustering involves a weighted, signed graph



Edges in a signed graph indicate **similarity (+)** or **dissimilarity (-)**

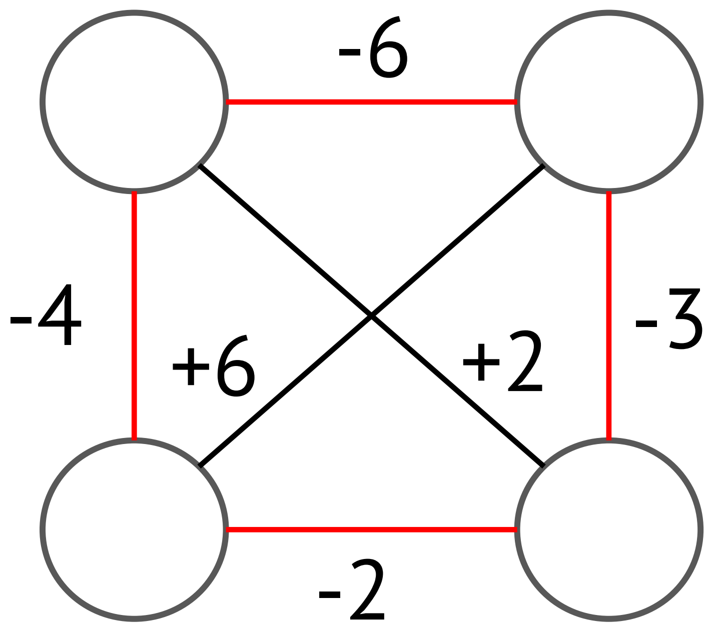
Edges can be weighted, but problems become harder.



Edges in a signed graph indicate **similarity (+)** or **dissimilarity (-)**

Objective: Minimize the weight of "mistakes"

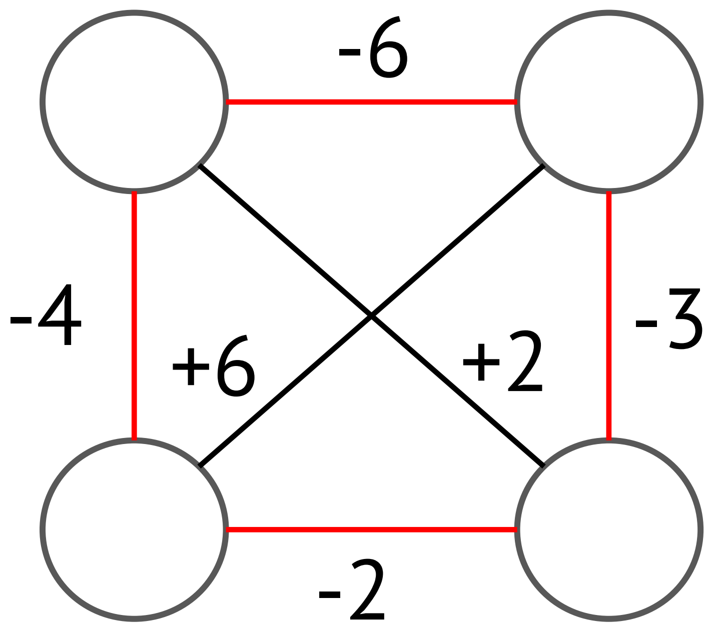
Edge weights can be stored in an adjacency matrix



$$\mathbf{A}_{ij} = w_{ij}^{+} - w_{ij}^{-}$$

$$\mathbf{A} = \begin{bmatrix} 0 & -6 & +2 & -4 \\ -6 & 0 & -3 & +6 \\ +2 & -3 & 0 & -2 \\ -4 & +6 & -2 & 0 \end{bmatrix}$$

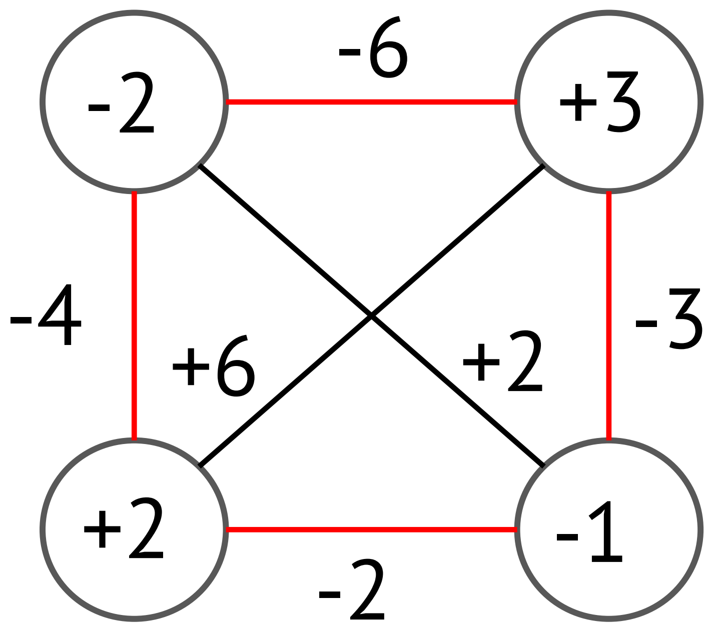
The rank-1 positive semidefinite case is very simple



$$\mathbf{A} = \mathbf{v}\mathbf{v}^T$$

$$\mathbf{A} = \begin{bmatrix} -2 \\ +3 \\ -1 \\ +2 \end{bmatrix} \begin{bmatrix} -2 & +3 & -1 & +2 \end{bmatrix}$$

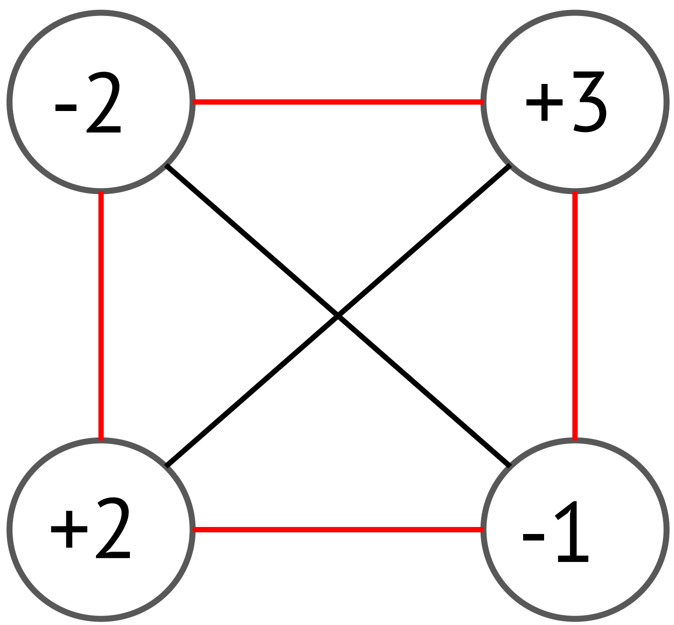
The rank-1 positive semidefinite case is very simple



$$\mathbf{A} = \mathbf{v}\mathbf{v}^T$$

$$\mathbf{A} = \begin{bmatrix} -2 \\ +3 \\ -1 \\ +2 \end{bmatrix} \begin{bmatrix} -2 & +3 & -1 & +2 \end{bmatrix}$$

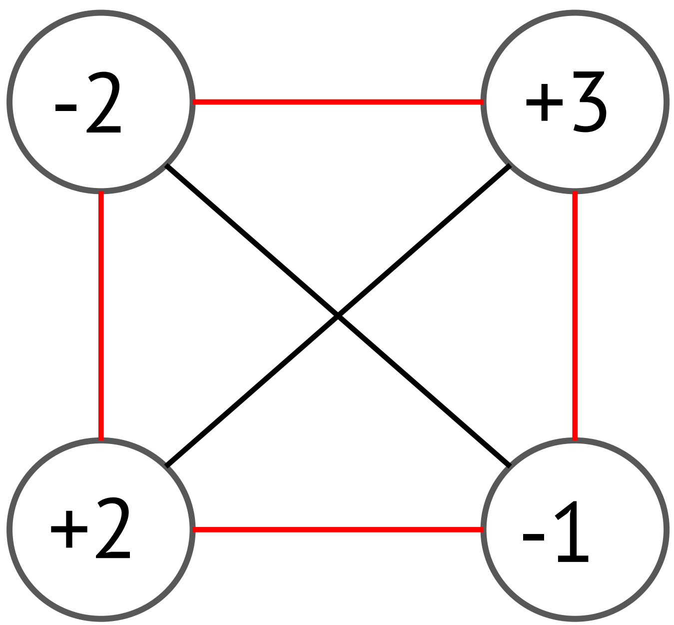
The rank-1 positive semidefinite case is very simple



$$\mathbf{A} = \mathbf{v}\mathbf{v}^T$$

$$\mathbf{A} = \begin{bmatrix} -2 \\ +3 \\ -1 \\ +2 \end{bmatrix} \begin{bmatrix} -2 & +3 & -1 & +2 \end{bmatrix}$$

The rank-1 positive semidefinite case is very simple

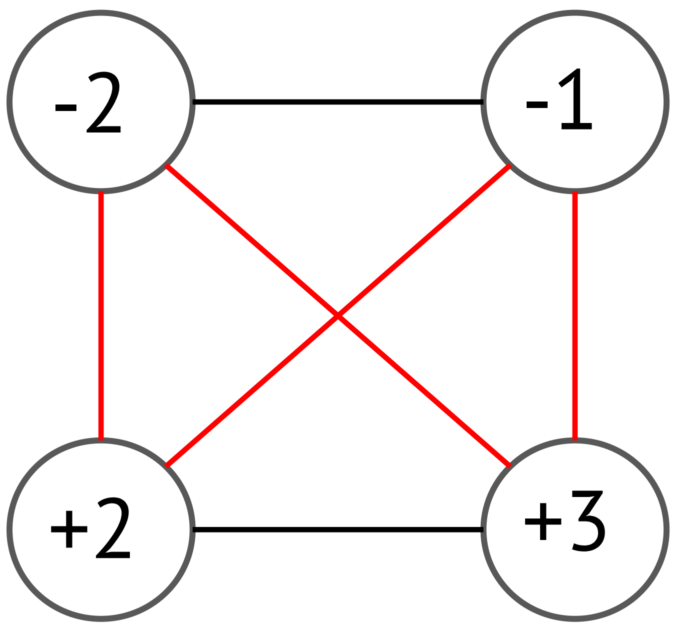


$$\mathbf{A} = \mathbf{v}\mathbf{v}^T$$

$$\begin{bmatrix} -2 \\ +3 \\ -1 \\ +2 \end{bmatrix} \longrightarrow \begin{bmatrix} -2 \\ -1 \\ +2 \\ +3 \end{bmatrix}$$

Ordering **v** gives a perfect clustering

The rank-1 positive semidefinite case is very simple

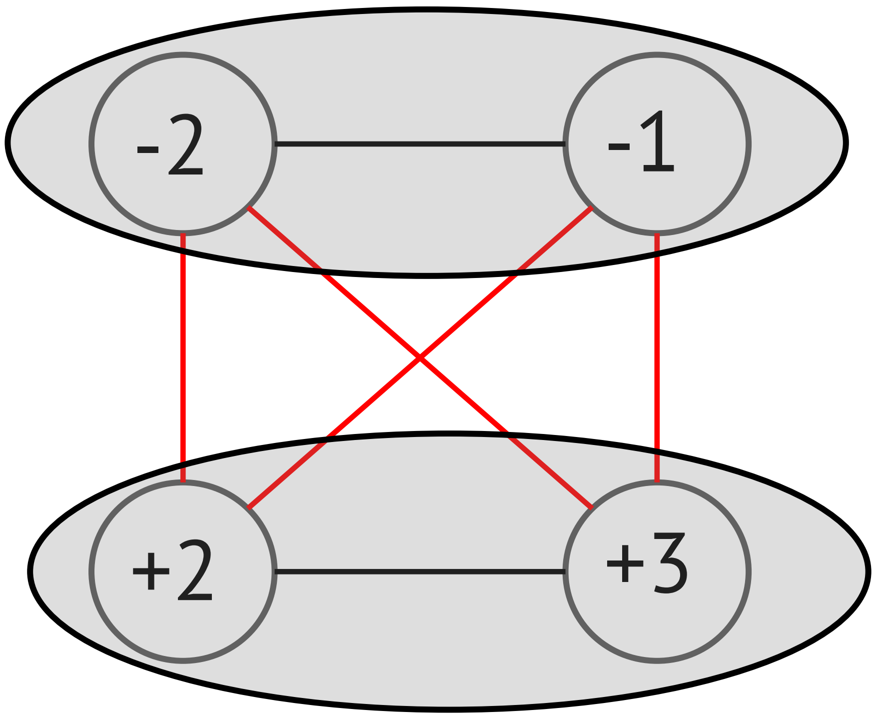


$$\mathbf{A} = \mathbf{v}\mathbf{v}^T$$

$$\begin{bmatrix} -2 \\ +3 \\ -1 \\ +2 \end{bmatrix} \longrightarrow \begin{bmatrix} -2 \\ -1 \\ +2 \\ +3 \end{bmatrix}$$

Ordering **v** gives a perfect clustering

The rank-1 positive semidefinite case is very simple



$$\mathbf{A} = \mathbf{v}\mathbf{v}^T$$

$$\begin{bmatrix} -2 \\ +3 \\ -1 \\ +2 \end{bmatrix} \longrightarrow \begin{bmatrix} -2 \\ -1 \\ +2 \\ +3 \end{bmatrix}$$

Ordering **v** gives a perfect clustering

What happens for other low-rank matrices?

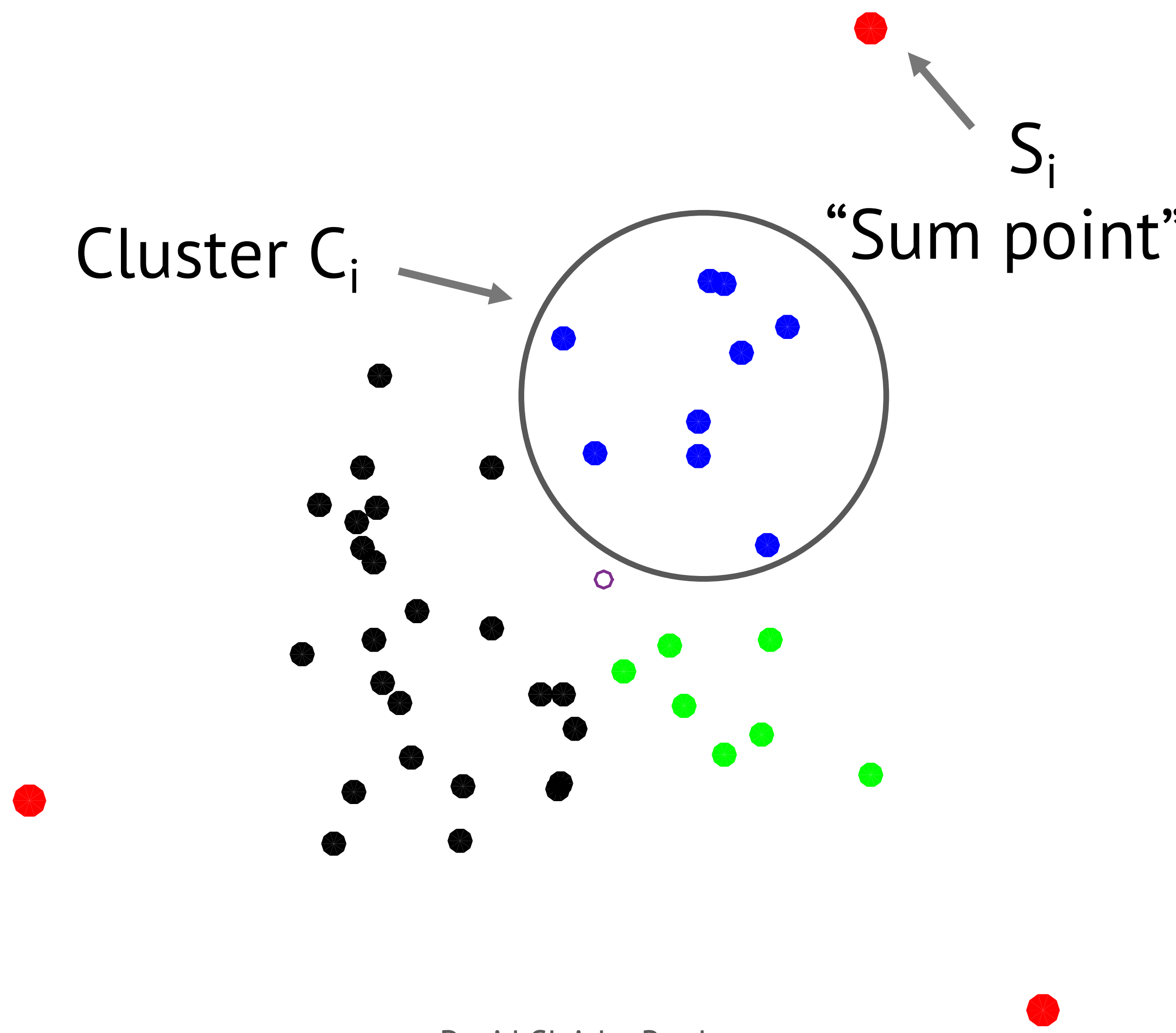
A simple solution for rank-1 positive semidefinite correlation clustering always exists.

Veldt, Gleich, Wirth arXiv:1611.07305

- A single negative eigenvalue makes the problem NP-hard ☹
- A rank-r pos def matrix can be solved in "polynomial" time $O(n^{r*r-1})$. Rank 3 = $O(n^8)$ time
- Algorithm is based on equiv between low-rank CC and vector partitioning *[Onn & Schulman 2001]* based on the vertices of a signing zonotope.
- In Stinson, Gleich, Constantine (arXiv:1602.06620) we proposed a randomized algorithm to find vertices of a zonotope.

Vector partitioning formulation

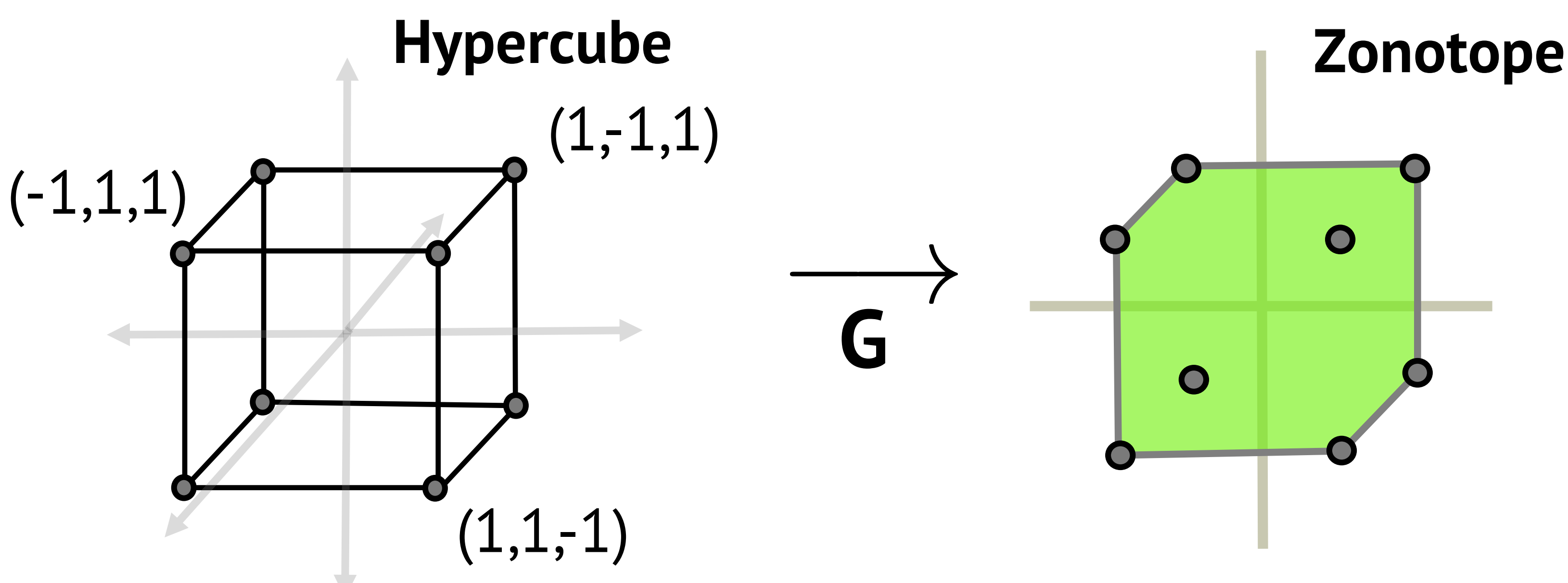
$$\max \sum_{k=1}^{\#\text{ clusters}} \|S_k\|^2$$



A zonotope is the linear projection of a hypercube into a lower dimension

$$\mathcal{Z}(\mathbf{G}) = \mathbf{conv}\{\mathbf{Gx} \mid \mathbf{x} \in \{\pm 1\}^n\}$$

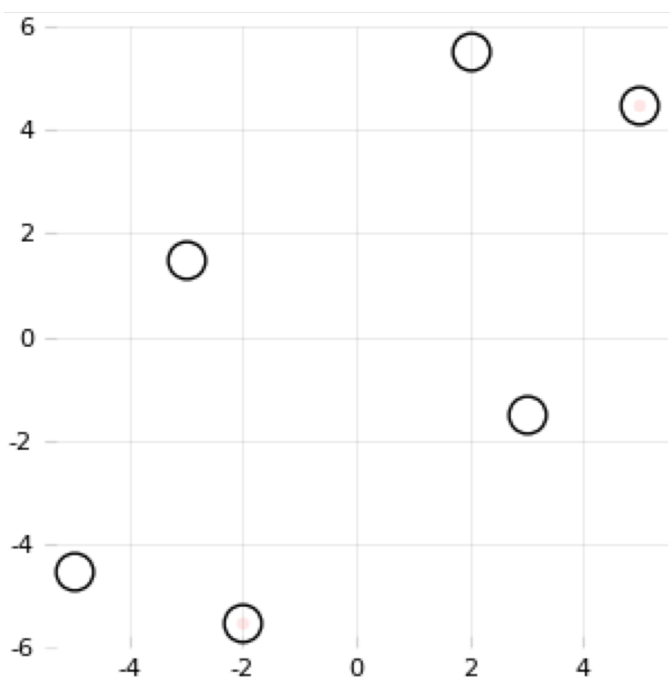
$\mathbf{G} \in \mathbb{R}^{D \times N}$ “generator” matrix

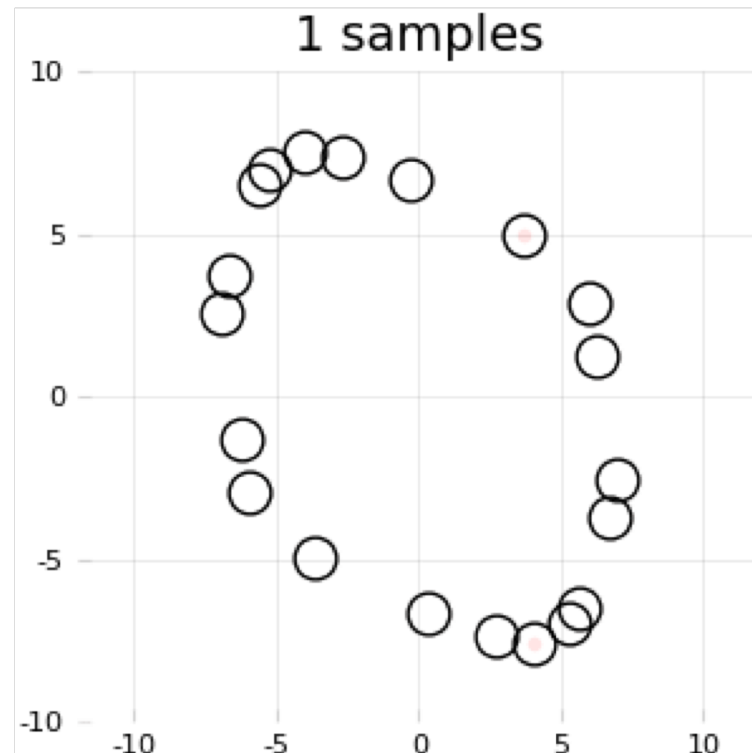


A really simple randomized algorithm to get a vertex of the zonotope

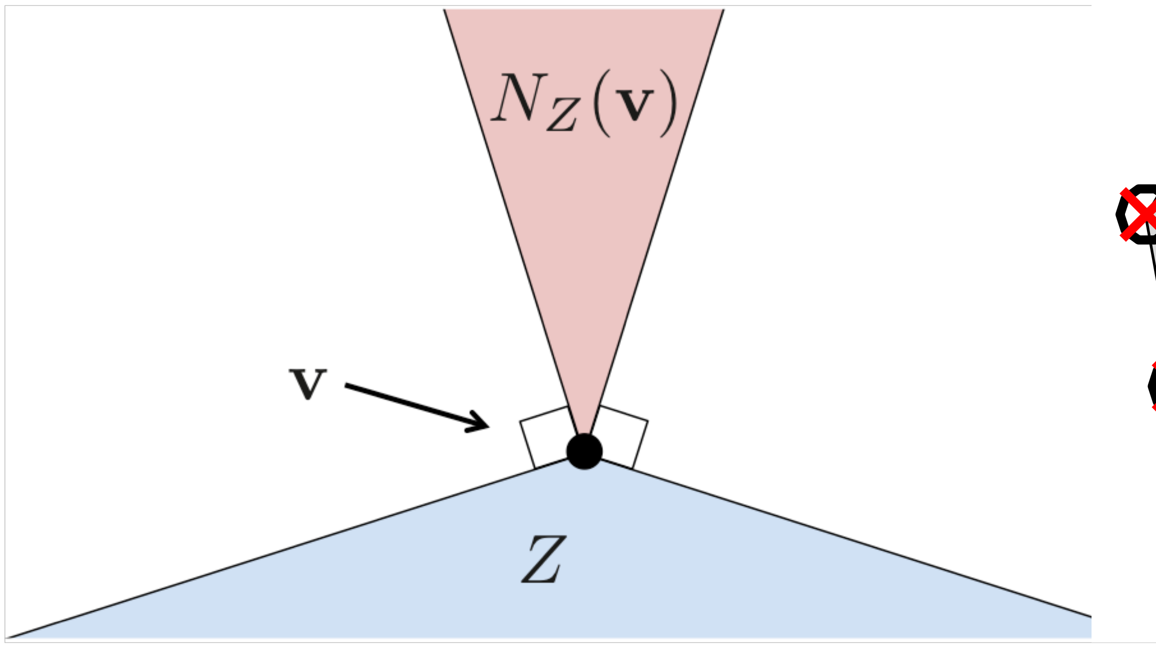
To generate a non-random vertex of a zonotope, compute

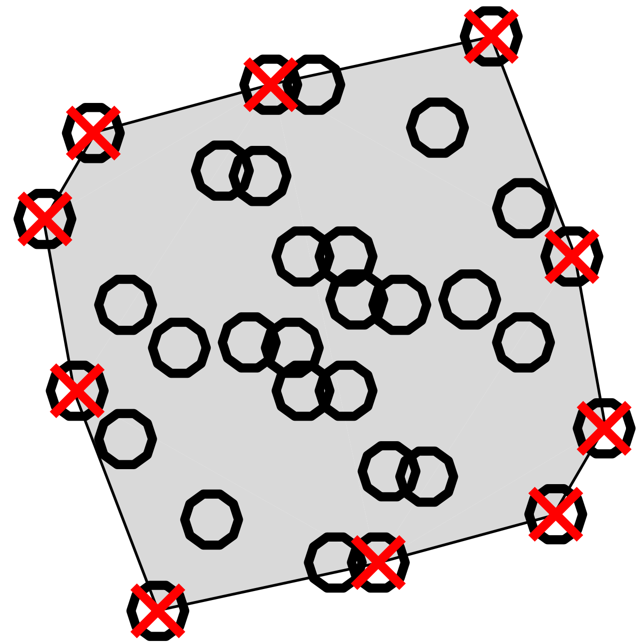
$$\boldsymbol{G}\text{sign}(\boldsymbol{G}^T\mathbf{x}) \text{ where } x_i \in N(0,1)$$

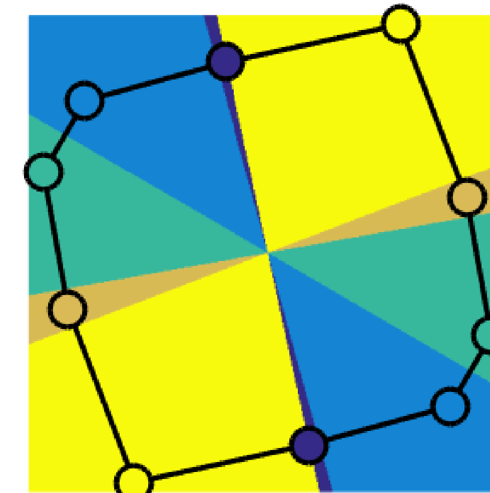




Prob. that a vertex is generated depends on normal cone

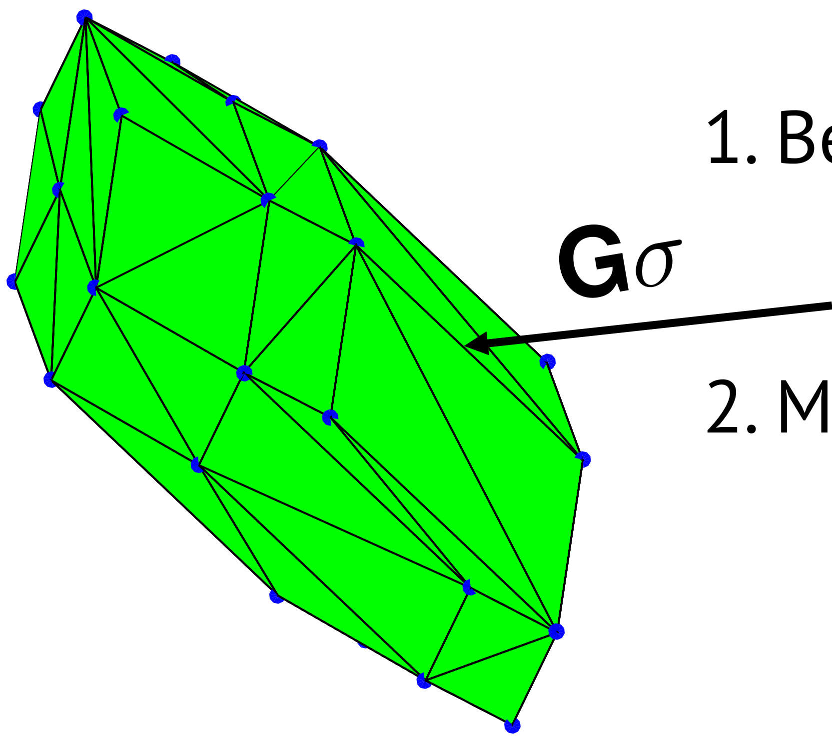






Applies theory of Damle & Sun arXiv:1405.4275

Vertices of the signing zonotope correspond to clusterings

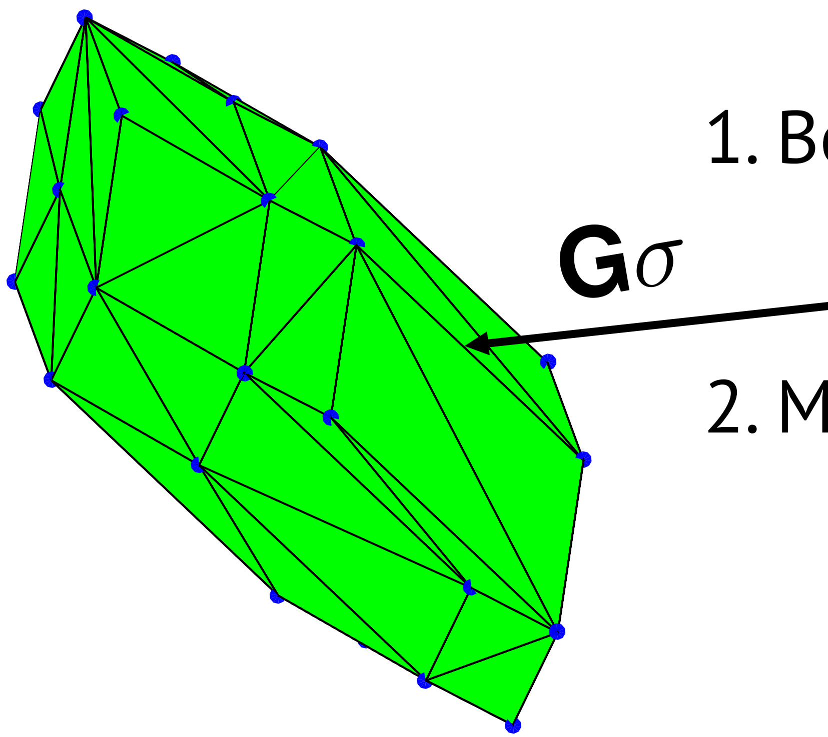


1. Begin with "signing" vector

$$\sigma = (\sigma^i_{r,s}) \in \{\pm 1\}^{n\binom{d+1}{2}}$$

2. Map signings into a zonotope

Vertices of the signing zonotope correspond to clusterings



1. Begin with "signing" vector

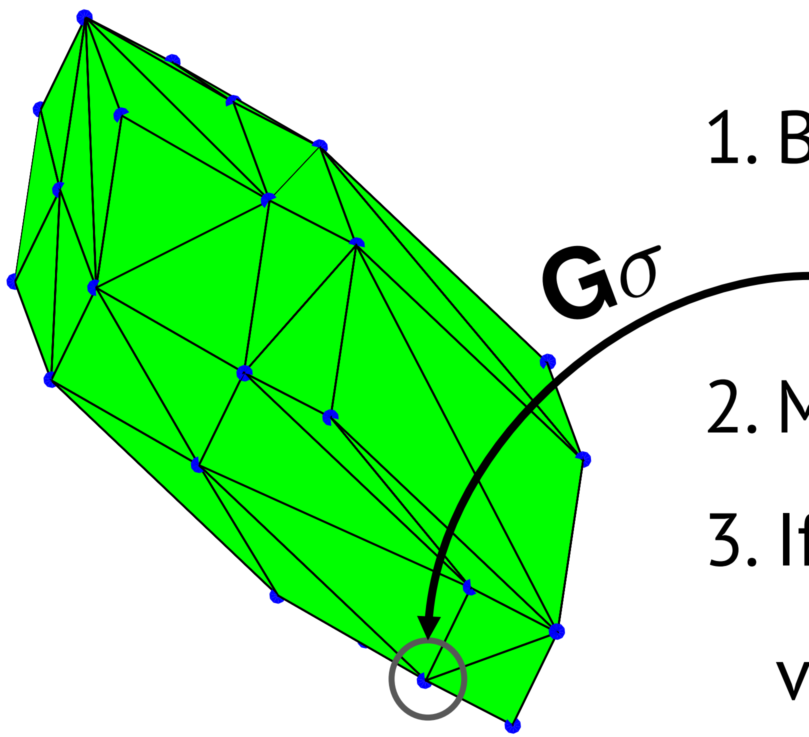
$$\sigma = (\sigma^i_{r,s}) \in \{\pm 1\}^{n\binom{d+1}{2}}$$

2. Map signings into a zonotope

Columns of **G** are vectorized outer products:

$$\mathbf{v}_\mathrm{i} \cdot (\mathbf{e}_\mathrm{r} - \mathbf{e}_\mathrm{s})^\mathrm{T}$$

Vertices of the signing zonotope correspond to clusterings



1. Begin with "signing" vector

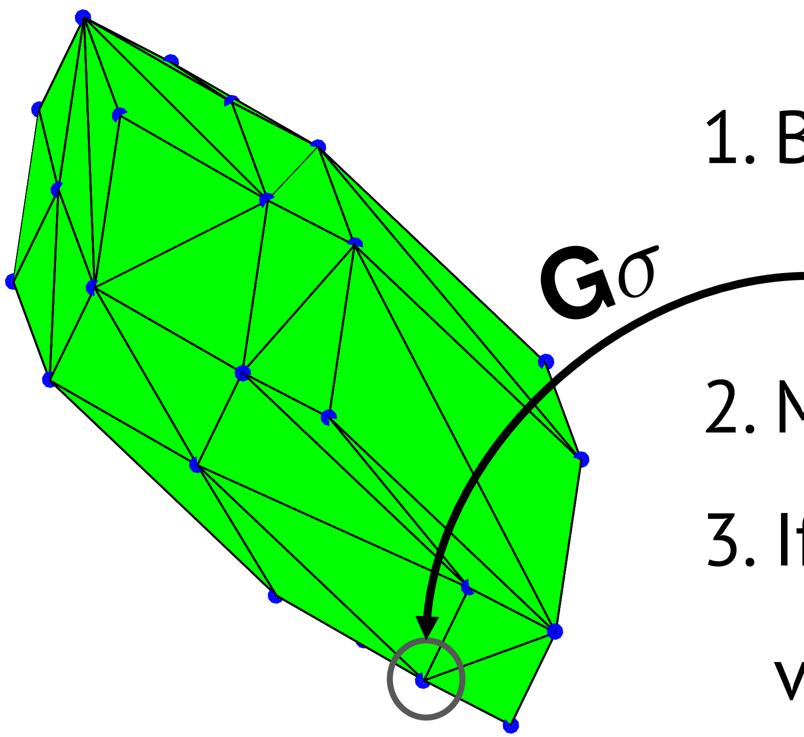
$$\sigma = (\sigma^{i}_{r,s}) \in \{\pm 1\}^{n\binom{d+1}{2}}$$

2. Map signings into a zonotope
3. If **σ** is mapped to a zonotope vertex, it defines a clustering

Columns of **G** are vectorized outer products:

$$\mathbf{v}_\mathrm{i} \cdot (\mathbf{e}_\mathrm{r} - \mathbf{e}_\mathrm{s})^\mathrm{T}$$

Vertices of the signing zonotope correspond to clusterings



$$\sigma^i_{r,s} = -1$$

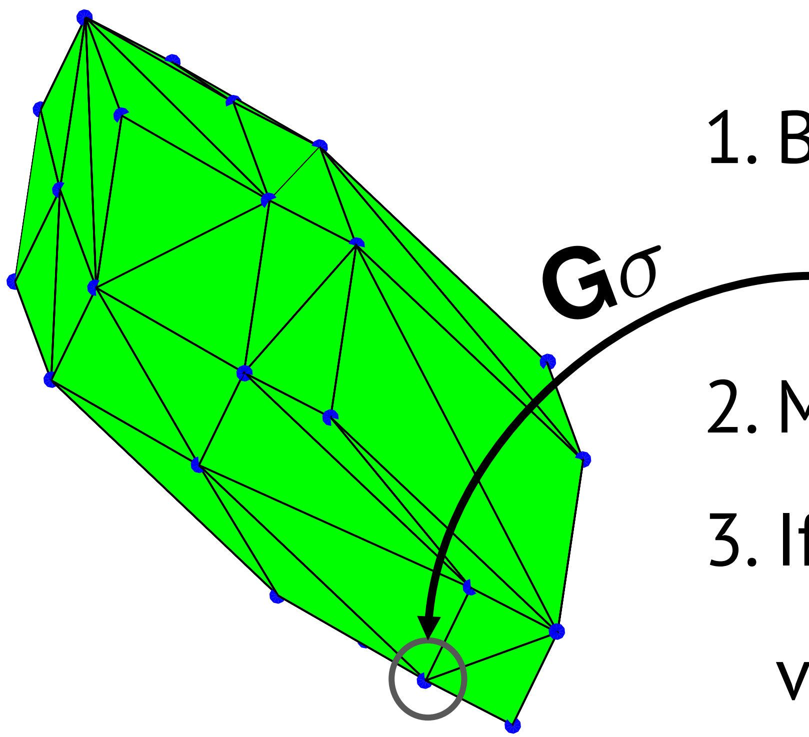
1. Begin with "signing" vector

$$\sigma = (\sigma^i_{r,s}) \in \{\pm 1\}^{n\binom{d+1}{2}}$$

2. Map signings into a zonotope
3. If **σ** is mapped to a zonotope vertex, it defines a clustering

"Node i is *not* in cluster r, but could be in cluster s"

Vertices of the signing zonotope correspond to clusterings



$$\sigma^{i}_{r,s} = -1$$

1. Begin with "signing" vector

$$\sigma = (\sigma^{i}_{r,s}) \in \{\pm 1\}^{n\binom{d+1}{2}}$$

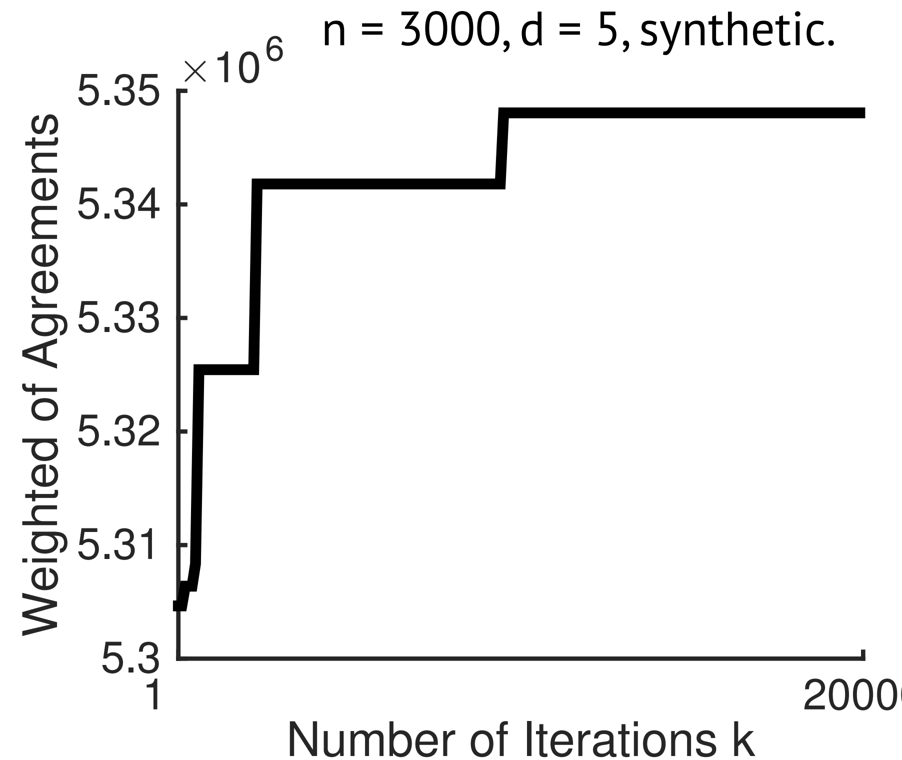
2. Map signings into a zonotope
3. If **σ** is mapped to a zonotope vertex, it defines a clustering
4. There are $O(n^{d^2-1})$ vertices

"Node i is *not* in cluster r, but could be in cluster s"

In practice we just sample vertices of the zonotope

ZonoCC Algorithm O(nk)

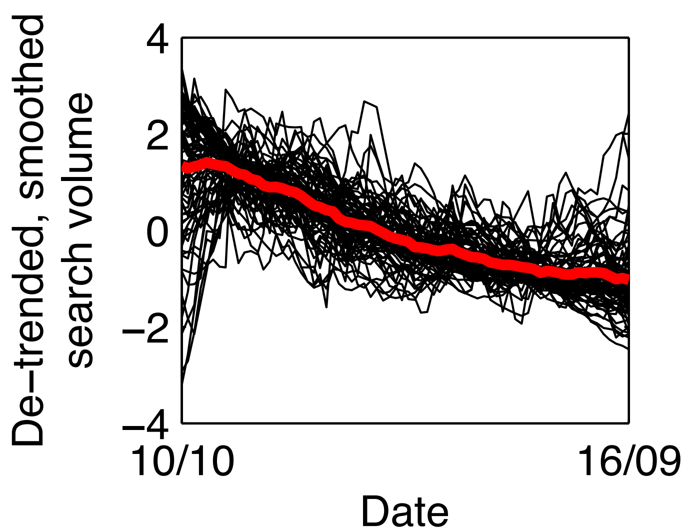
1. Generate k random extremal signings σ via random sampling
2. Find the clustering of each σ
3. Output clustering with highest objective score

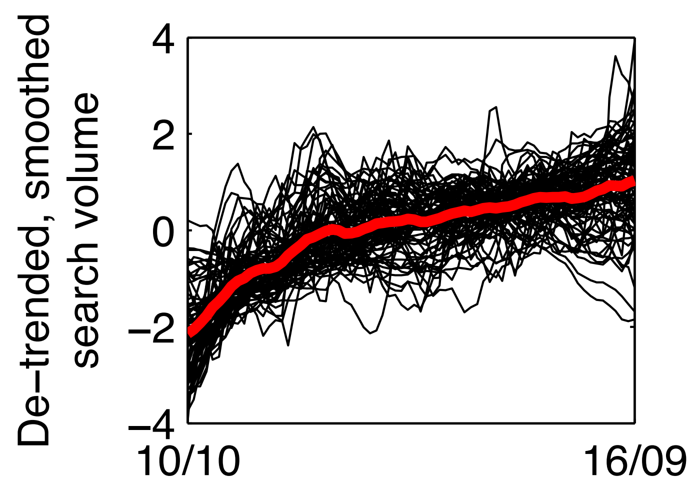


For a rank-3 problem on clustering conferences

Table 2: Objective scores and runtimes in seconds for correlation clustering algorithms on two real-world datasets. Due to the size of the stocks dataset, we can run only ZonoCC and Pivot on it.

Dataset		ZonoCC	Pivot	CGW	ILP
CS Conf.	Obj.	7540.0	7540.0	7540.0	7540.0
$n = 157$	Time	7	1	1380	52
Stocks	Obj.	5100.2	5099.5	—	—
$n = 497$	Time	40	20	—	—





Back to the question

Are there better ways of sampling these zonotopes?

$$\boldsymbol{G}\,\text{sign}(\boldsymbol{G}^T \mathbf{x}) \text{ where } x_i \in N(0,1)$$

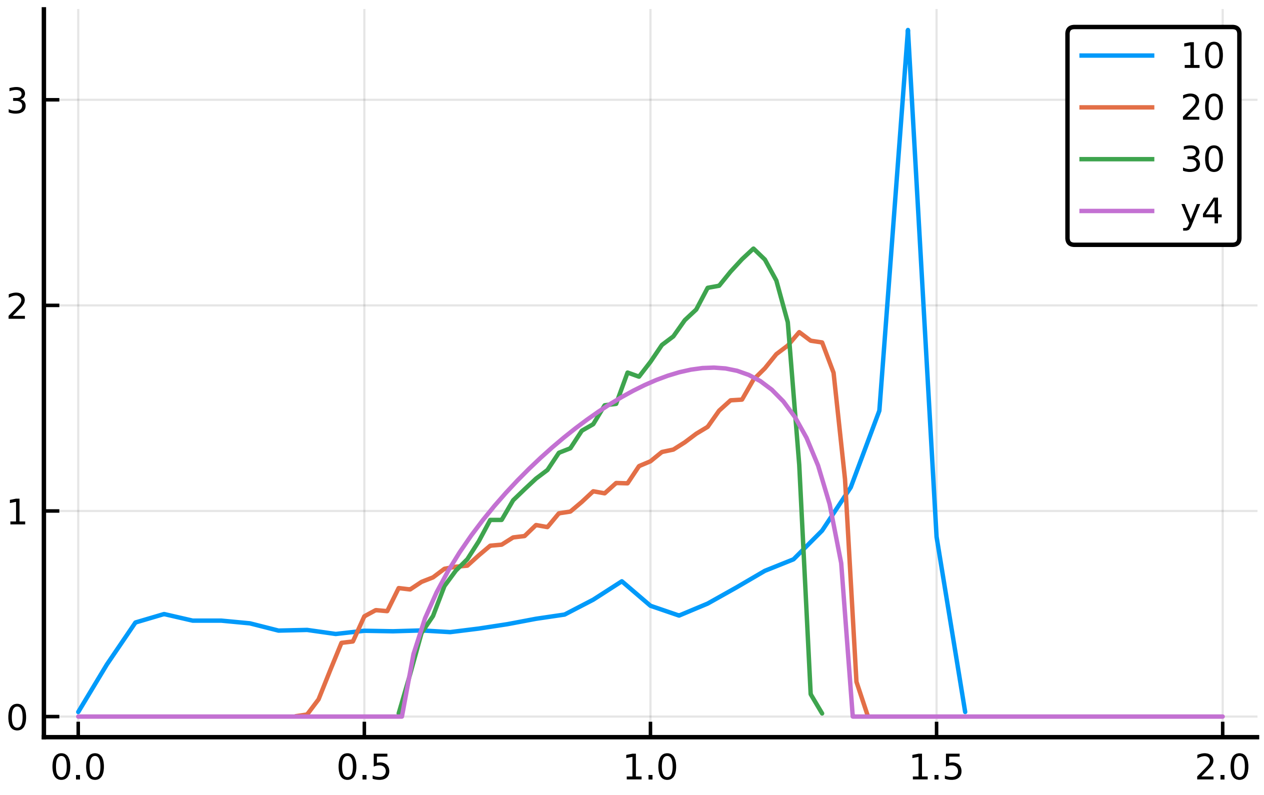
Is there anything like an approximation bound we can give for an algorithm based on sampling these in terms of the objective?

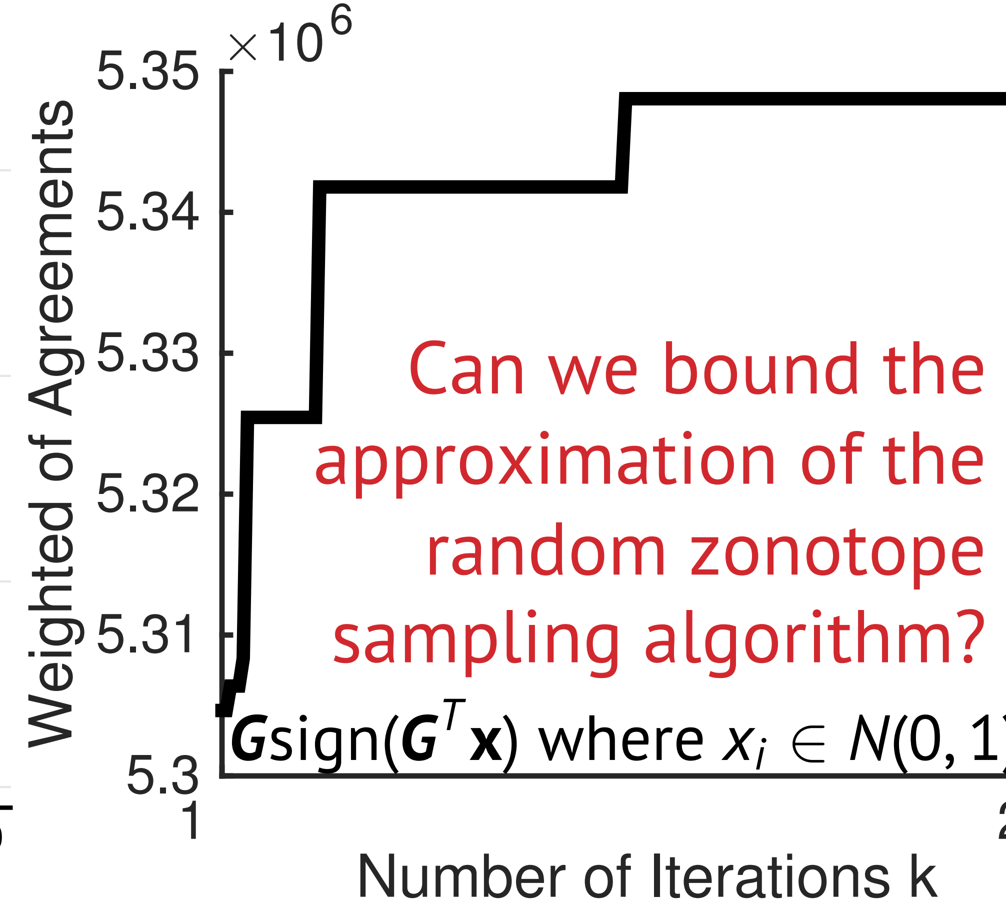
- [Ferrez, Fukuda, Libeling] reduce low-rank binary QP maximization to zonotope enumeration

$$\max \mathbf{x}^T \boldsymbol{A}\mathbf{x} \text{ such that } x_i \in \{0,1\}$$

Summary. Here are my two questions!

What is this distribution??





Can we bound the approximation of the random zonotope sampling algorithm?

$\boldsymbol{G}\text{sign}(\boldsymbol{G}^T\mathbf{x})$ where $x_i \in N(0,1)$

Code to reproduce https://gist.github.com/dgleich/4d4becc858e4a7d7952af6c66c99e7b9