On the Parameterized Intractability of Determinant Maximization

Naoto Ohsaka





Slides available https://todo314.github.io/ →

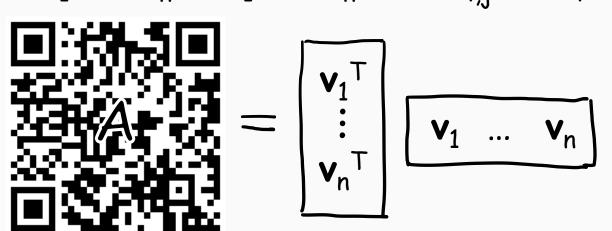
What is DETERMINANT MAXIMIZATION?

• Input: $n \times n$ positive semi-definite A in $\mathbb{Q}^{n \times n}$ & $k \in [n]$

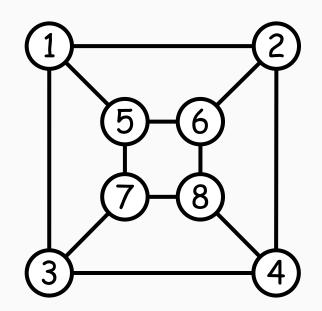
• Output: $S \in \binom{[n]}{k}$

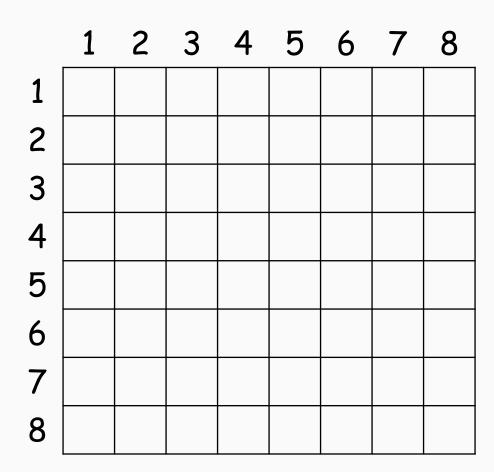
• Goal: maximize principal minor $det(A_s)$

A is typically given as Gram matrix for n vectors $\mathbf{v}_1, ..., \mathbf{v}_n$ in \mathbb{Q}^d $\mathbf{A} \stackrel{\text{def}}{=} [\mathbf{v}_1, ..., \mathbf{v}_n]^T [\mathbf{v}_1, ..., \mathbf{v}_n], \text{ or } \mathbf{A}_{i,j} \stackrel{\text{def}}{=} \langle \mathbf{v}_i, \mathbf{v}_j \rangle$

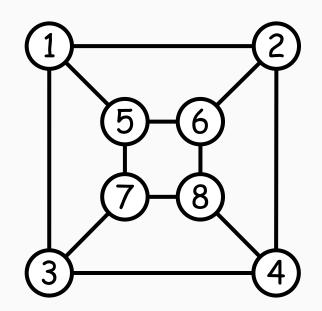


- $Q_3 = (V = [8], E)$: Hypercube graph
- $\mathbf{v}_i \in \{0,1\}^E$: $\mathbf{v}_i(e) \stackrel{\text{def}}{=} [[i \text{ is incident to } e]]$



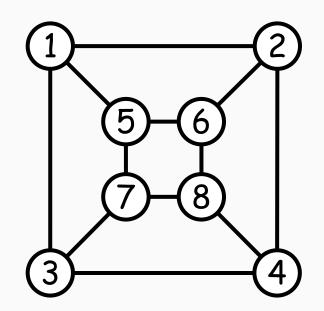


- $Q_3 = (V = [8], E)$: Hypercube graph
- $\mathbf{v}_i \in \{0,1\}^E$: $\mathbf{v}_i(e) \stackrel{\text{def}}{=} [[i] \text{ is incident to } e]]$



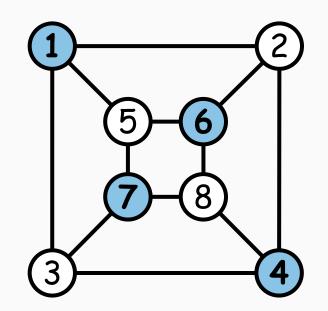
	1	2	3	4	5	6	7	8
1	3							
2		3						
3			თ					
4				3				
5					3			
6						თ		
7							3	
8								3

- $Q_3 = (V = [8], E)$: Hypercube graph
- $\mathbf{v}_i \in \{0,1\}^E$: $\mathbf{v}_i(e) \stackrel{\text{def}}{=} [[i] \text{ is incident to } e]]$



	1	2	3	4	5	6	7	8
1	3	1	1	0	1	0	0	0
2	1	3	0	1	0	1	0	0
3	1	0	3	1	0	0	1	0
4	0	1	1	3	0	0	0	1
5	1	0	0	0	3	1	1	0
6	0	1	0	0	1	3	0	1
7	0	0	1	0	1	0	3	1
8	0	0	0	1	0	1	1	3

- $Q_3 = (V = [8], E)$: Hypercube graph
- $\mathbf{v}_i \in \{0,1\}^E$: $\mathbf{v}_i(e) \stackrel{\text{def}}{=} [[i \text{ is incident to e }]]$



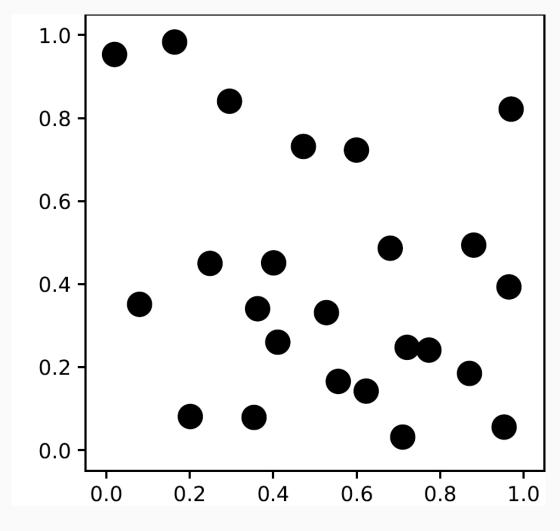
	1	2	3	4	5	6	7	8
1	3	1	1	0	1	0	0	0
2	1	3	0	1	0	1	0	0
3	1	0	3	1	0	0	1	0
4	0	1	1	3	0	0	0	1
5	1	0	0	0	က	1	1	0
6	0	1	0	0	1	3	0	1
7	0	0	1	0	1	0	3	1
8	0	0	0	1	0	1	1	3

 \bigcirc det(A_s) = 3^{|s|} → S is independent! e.g., S = {1,4,6,7}

Example 2: Selecting dispersed points

- $\mathbf{p}_1, ..., \mathbf{p}_n$: (random) points on \mathbb{R}^2
- Let $A_{i,j} \stackrel{\text{def}}{=} \exp(|\mathbf{p}_i \mathbf{p}_j|^2)$
 - Known as Gaussian/RBF kernel
 - A is positive semi-definite

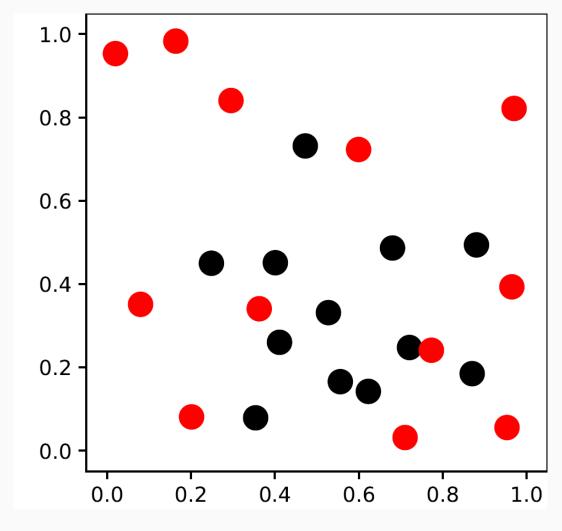
Q. What happens if $det(A_s)$ is max?



Example of n=24 & k=12

Example 2: Selecting dispersed points

- $\mathbf{p}_1, ..., \mathbf{p}_n$: (random) points on \mathbb{R}^2
- Let $A_{i,j} \stackrel{\text{def}}{=} \exp(|\mathbf{p}_i \mathbf{p}_j|^2)$
 - Known as Gaussian/RBF kernel
 - A is positive semi-definite
- Q. What happens if $det(A_s)$ is max?
- A. Select "dispersed" points



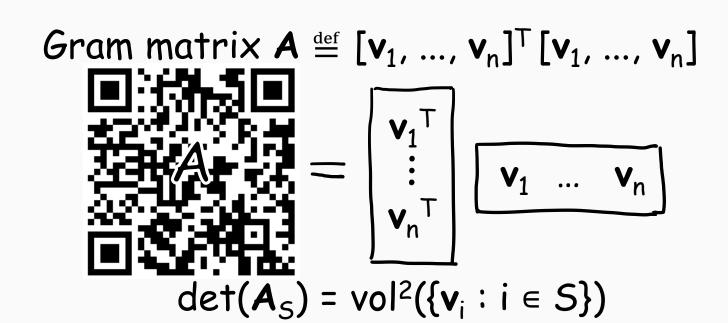
Example of n=24 & k=12

Why study DETERMINANT MAXIMIZATION?

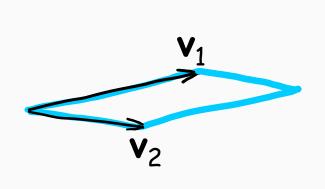
Various interpretations and applications

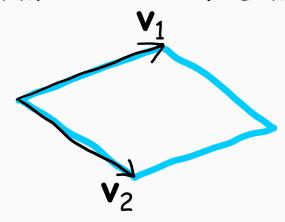
- Parallelepiped volume
- Diversity promotion in Machine Learning ... many applications! [Kulesza-Taskar. Found. Trends Mach. Learn. '12]
- Simplex volume [Nikolov. STOC'15]
- Maximum-entropy sampling [Ko-Lee-Queyranne. Oper. Res. '95]

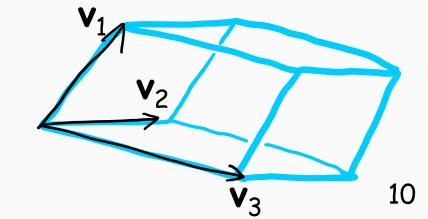
One interpretation: Parallelepiped volume



DETERMINANT MAXIMIZATION = VOLUME MAXIMIZATION







Known results in polynomial-time regime

- WP-hard [Ko-Lee-Queyranne. Oper. Res. '95] [Çivril & Magdon-Ismail. Theor. Comput. Sci. '09]
- DP-hard to 20(k)-approx. [Koutis. Inf. Process. Lett.'06]

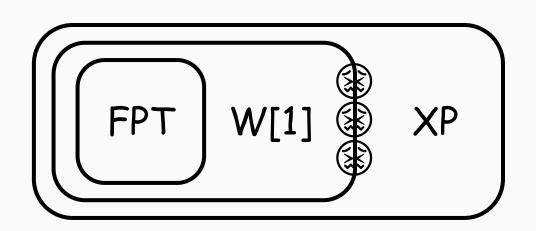
 [Çivril & Magdon-Ismail. Algorithmica'13]

 The nearly tight [Di Summa-Eisenbrand-Faenza-Moldenhauer. SODA'14]
- Can find e^k-approx. [Nikolov. STOC'15] k=|S| is the output size

Known results in parameterized regime

Measure complexity w.r.t. input size n & parameter k

- Fixed-parameter tractable (FPT): Solvable in $f(k)n^{O(1)}$ time
- $n^{O(k)}$ -time brute-force alg. \rightarrow XP w.r.t. k (very natural param.)
- But W[1]-hard w.r.t k [Ko-Lee-Queyranne. Oper. Res. '95] [Koutis. Inf. Process. Lett. '06]
 - → No FPT alg. unless Exponential Time Hypothesis is false (unlikely!)



Q. How can we make
DETERMINANT MAXIMIZATION tractable?

Three possible scenarios (we expect)

- 1. Structural restriction
- (Underlying graph of) A is very sparse
- e.g., PERMANENT is #P-hard in general, but FPT w.r.t. treewidth [Courcelle-Makowsky-Rotics. Discrete Appl. Math. '01] [Cifuentes-Parrilo. Linear Algebra Appl. '16]
- 2. Strong parameter
- rank(A) ≥ output size k (always!)
- Room for consideration of f(rank)n^{O(1)}-time FPT alg.
- 3. FPT approximation [Feldmann-Karthik-Lee-Manurangsi. Algorithms' 20]
- Some W[1]-hard problems is approximable in FPT time
- e.g., Partial Vertex Cover & Minimum K-Median [Har-Peled & Soham Mazumdar. STOC'04]

Three possible scenarios (we expect)

- 1. Structural restriction
- (Underlying graph of) A is very sparse
- e.g., PERMANENT is #P-hard in ger ral. but FDT It. treewidth

 [Courcelle-Makowsky-Ro] 1700 April 160

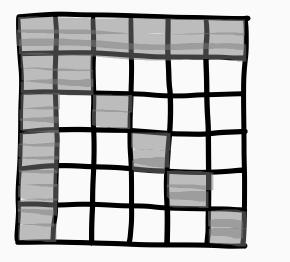
2. All hopes are dashed!

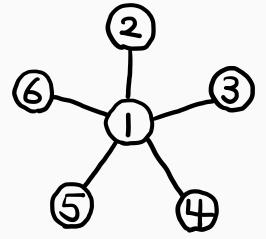
- rank(
- Room for consideration of f(rank)n^{O(1)}-time FPT alg.
- 3. FPT approximation [Feldmann-Karthik-Lee-Manurangsi. Algorithms' 20]
- Some W[1]-hard problems is approximable in FPT time
- e.g., Partial Vertex Cover & Minimum K-Median [Har-Peled & Soham Mazumdar. STOC'04]

Our first result:

Hardness on arrowhead matrices RR

Arrowhead = Star graph

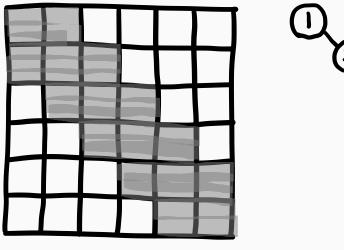


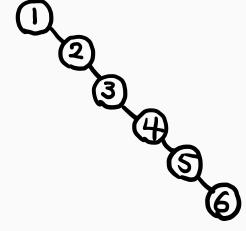




Treewidth & pathwidth = 1 vertex cover number = 1

Tridiagonal = Path graph





©Polytime solvable

[Al-Thani & Lee. LAGOS'21]

Structural sparsity is NOT very helpful

Our second & third results

- \rightarrow W[1]-hard w.r.t. output size k even if rank only depends on k
- \boxtimes W[1]-hard to $2^{O(\sqrt{k})}$ -approx. w.r.t. k under Parameterized Inapproximability Hypothesis

[Lokshtanov-Ramanujan-Saurab-Zehavi. SODA'20]

BINARY CONSTRAINT SATISFACTION PROBLEM
is W[1]-hard to approx.
w.r.t. # variables

Proof overview

(1) Proof overview on arrowhead matrices

(Thm) DETERMINANT MAXIMIZATION on arrowhead matrices is W[1]-hard

• k-Sum: Parameterized version of Subset Sum [Abboud-Lewi-Williams. ESA'14]



∆ Sophisticated construction of arrowhead matrix



• DETERMINANT MAXIMIZATION on arrowhead matrices

(1) Proof overview on W[1]-hardness on arrowhead matrices K-Sum [Abboud-Lewi-Williams. ESA'14] & reduction strategy

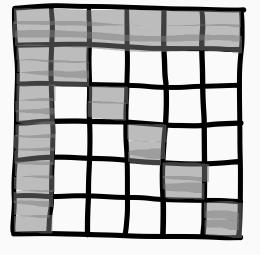
- Input: n integers $x_1, ..., x_n, t \in [0, n^{2k}], k \in [n]$
- Find: $S \in \binom{[n]}{k}$ s.t. $\sum_{i \in S} x_i = t$
- W[1]-complete w.r.t. k [Downey-Fellows. Theor. Comput. Sci. '95] [Abboud-Lewi-Williams. ESA'14]
- \bigcirc Construct n+1 vectors \mathbf{v}_0 , \mathbf{v}_1 , ..., \mathbf{v}_n s.t.
- Gram matrix in $\mathbb{R}^{[0..n]\times[0..n]}$ is arrowhead
- $det(A_S)$ s.t. $S \in \binom{[n]}{k+1}$ is maximum when $\sum_{i \in S \{0\}} x_i = t$ (if exists) i.e., \mathbf{v}_i corresponds to x_i

(1) Proof overview on W[1]-hardness on arrowhead matrices

Key finding on arrowhead matrices

• If **A** in $\mathbb{R}^{[0..n]\times[0..n]}$ is arrowhead and $0 \in S$:

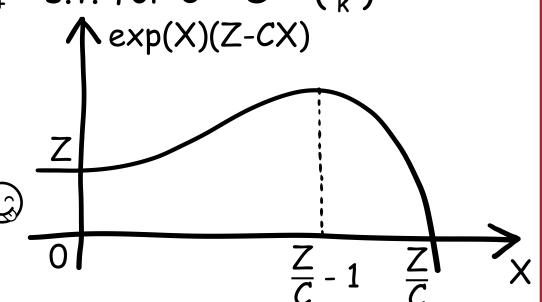
$$\det(\mathbf{A}_{S}) = \prod_{i \in S - \{0\}} A_{i,i} \cdot \left(A_{0,0} - \sum_{i \in S - \{0\}} \frac{A_{0,i} \cdot A_{0,i}}{A_{i,i}} \right)$$



(Lem) Carefully choose \mathbf{v}_0 , \mathbf{v}_1 , ..., $\mathbf{v}_n \in \mathbb{R}_+^{2n} s.t.$ for $0 \in S \in \binom{[n]}{k}$

$$det(A_S) = exp\left(\sum_{i \in S - \{0\}} x_i\right) \left(Z - C\sum_{i \in S - \{0\}} x_i\right)$$

Take max. at
$$\sum_{i \in S_{-0}} x_i = \frac{Z}{C} - 1$$
 set $t \bigcirc \underline{C}$



(1) Proof overview on W[1]-hardness on arrowhead matrices $Sketch\ of\ construction$

	1	i	n	n+1	n+i	n+n
v _O	$\gamma\sqrt{x_1}$	$\gamma \sqrt{x_i}$	$\gamma\sqrt{X_n}$			
\mathbf{v}_1	$\sqrt{\alpha e^{X_1}}$			$\sqrt{\beta} e^{X_1}$		
Vi		√a e ^X i			√B e ^{Xi}	
v _n			√a e ^X n			√B eXn

Parameterized by α , β , γ (to be determined appropriately)

Omitted details: We have to...

- efficiently approximate \mathbf{v}_0 , \mathbf{v}_1 , ..., \mathbf{v}_n using rationals
- ullet ensure that any optimal solution includes $oldsymbol{v}_0$

(2) Proof overview on W[1]-hardness by rank

(Thm) DETERMINANT MAXIMIZATION is W[1]-hard w.r.t. rank of A

• GRID TILING: W[1]-complete [Marx. FOCS'07]



 \triangle Can use only f(k)-dimensional vectors / f(k)-rank matrices e.g., vectors in \mathbb{Q}^n are not allowed



• DETERMINANT MAXIMIZATION parameterized by rank of A

(2) Proof overview on W[1]-hardness by rank GRID TILING [Marx. FOCS'07]

- Input: $S = (S_{i,j} \subseteq [n]^2 : i,j \in [k])$
- Find: Select (x,y) in $S_{i,j}$ for all (i,j) s.t.
 - Vertical neighbors agree in 1st coordinate
 - Horizontal neighbors agree in 2nd coordinate

- Equality constraints are SIMPLE ©
- Cells (i,j) are adjacent to FOUR cells ©

S _{1,1} (1,1) (3,1) (2,4)	S _{1,2} (5,1) (1,4) (5,3)	S _{1,3} (1,1) (2,4) (3,3)
S _{2,1} (2,2) (1,4)	S _{2,2} (3,1) (1,2)	S _{2,3} (2,2) (2,3)
S _{3,1} (1,3) (2,3) (3,3)	S _{3,2} (1,1) (1,3)	S _{3,3} (2,3) (5,3)

Example of k=3 & n=5
Taken from Fig. 14.2 of

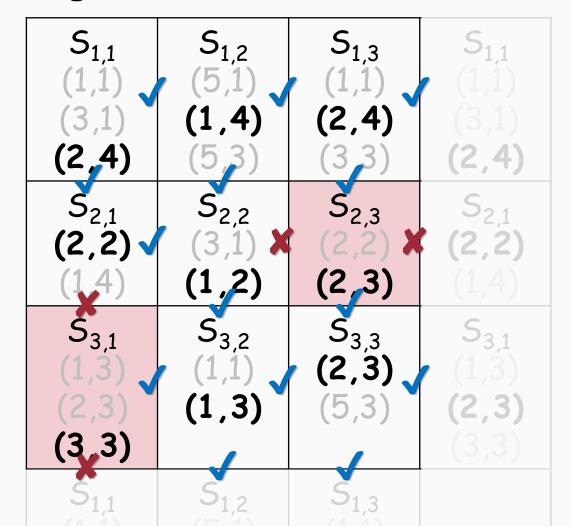
[Cygan-Fomin-Kowalik-Lokshtanov-Marx-Pilipczuk-Pilipczuk-Saurabh.]

(2) Proof overview on W[1]-hardness by rank GRID TILING [Marx. FOCS'07]

perfect consistency (2)

S _{1,1} (1,1)	S _{1,2} (5,1)	S _{1,3} (1,1)	S _{1,1} (1,1)
(3,1)	(1,4) (5,3)	(2,4) (3,3)	(3,1) (2,4)
(2,4) S _{2,1}	S _{2,2}	S _{2,3}	S _{2,1}
(2,2) (1,4)	(3,1) (1,2)	(2,2) (2,3)	(2,2) (1,4)
S _{3,1} (1,3)	S _{3,2} (1,1)	S _{3,3} (2,3)	S _{3,1} (1,3)
(2,3) (3,3)	(1,3)	(5,3)	(2,3) (3,3)
S _{1,1}	S _{1,2}	S _{1,3}	

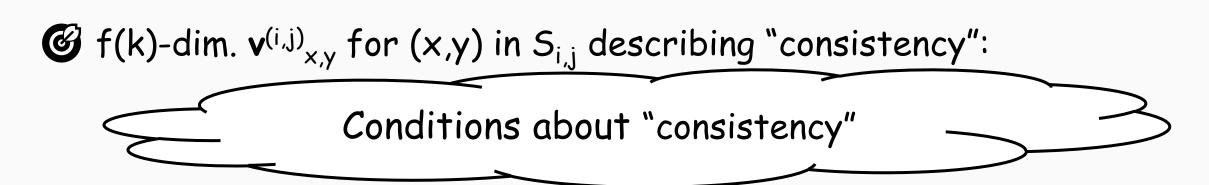
4 neighbors are inconsistent 🗟





(2) Proof overview on W[1]-hardness by rank Reduction from GRID TILING

- Input: $S = (S_{i,j} \subseteq [n]^2 : i,j \in [k])$
- Find: Select (x,y) in $S_{i,j}$ for all (i,j) s.t.
 - Vertical neighbors agree in 1st coordinate
 - Horizontal neighbors agree in 2nd coordinate



(2) Proof overview on W[1]-hardness by rank Reduction from GRID TILING

- Input: $S = (S_{i,j} \subseteq [n]^2 : i,j \in [k])$
- Find: Select (x,y) in $S_{i,j}$ for all (i,j) s.t.
 - Vertical neighbors agree in 1st coordinate
 - Horizontal neighbors agree in 2nd coordinate
- G f(k)-dim. $\mathbf{v}^{(i,j)}_{x,y}$ for (x,y) in $S_{i,j}$ describing "consistency":

- Same cell

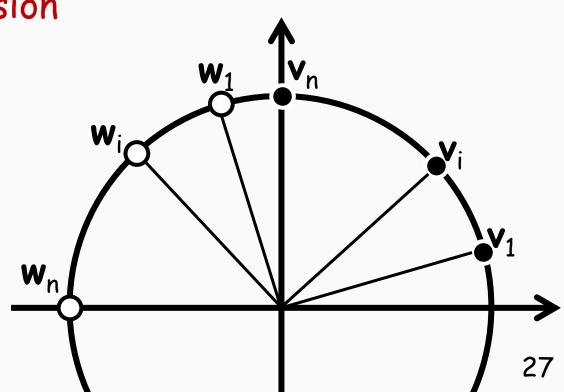
- Vertical nbr. $\langle \mathbf{v}^{(i,j)}_{x,y}, \mathbf{v}^{(i+1,j)}_{x',y'} \rangle = 0 \text{ iff } x=x'$ Horizontal nbr. $\langle \mathbf{v}^{(i,j)}_{x,y}, \mathbf{v}^{(i,j+1)}_{x',y'} \rangle = 0 \text{ iff } y=y'$ Focus in the next slide
 - $\langle \mathbf{v}^{(i,j)}_{\times \vee}, \mathbf{v}^{(i,j)}_{\times' \vee'} \rangle \neq 0$

- \bigcirc Gram matrix $A_{i,j,x,y,i',j',x',y'} \stackrel{\text{def}}{=} \langle \mathbf{v}^{(i,j)}_{x,y}, \mathbf{v}^{(i',j')}_{x',y'} \rangle$ satisfies...
- S is YES $\rightarrow \exists k^2 \times k^2$ diagonal submatrix ... select CORRECT $\mathbf{v}^{(i,j)}_{x,y}$ for each $(i,j) \in [k]^2$
- S is NO $\rightarrow \forall k^2 \times k^2$ submatrix is NOT diagonal

(2) Proof overview on W[1]-hardness by rank Represent "consistency" at lower dimensions?

- Want \mathbf{v}_1 , ..., \mathbf{v}_n , \mathbf{w}_1 , ..., \mathbf{w}_n in $\mathbb{Q}^{O(1)}$ s.t. $\langle \mathbf{v}_i, \mathbf{w}_i \rangle = 0$ iff i=j
 - How to construct?
- One-hot vectors require n-dimension [0,...,0,1,0,...,0]
- Use points on the unit circle:
- $\mathbf{v}_{i} \stackrel{\text{def}}{=} (\cos(\frac{\pi i}{2n}), \sin(\frac{\pi i}{2n}))$
- $\mathbf{w}_{j} \stackrel{\text{def}}{=} (\sin(\frac{\pi j}{2n}), -\cos(\frac{\pi j}{2n}))$

Use Pythagorean triples to get rational vectors



(3) Proof overview on inapproximability

(Thm) Under PIH, $\exists \delta$, Determinant Maximization is W[1]-hard w.r.t. output size k to approx. within $0.999^{\delta/k}$ -factor

Parameterized Inapproximability Hypothesis (PIH)

[Lokshtanov-Ramanujan-Saurab-Zehavi. SODA'20] I don't go into details in this talk



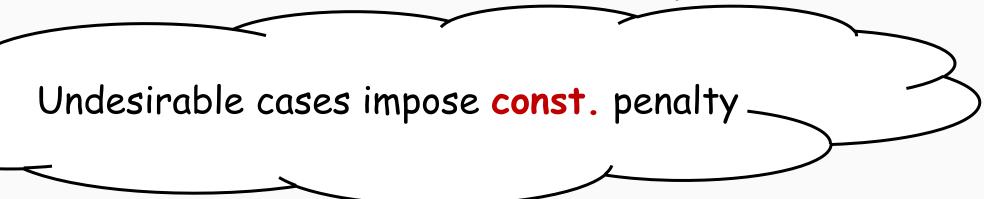
- Optimization version of GRID TILING: W[1]-hard to approx. w.r.t. k
 - ♣ Gap-preserving reduction (different from the last one)
- DETERMINANT MAXIMIZATION parameterized by k

(3) Proof overview on inapproximability
Optimization version of GRID TILING

- Input: $S \stackrel{\text{def}}{=} (S_{i,j} \subseteq [n]^2 : 1 \le i,j \le k)$ Output: Select (x,y) in $S_{i,j}$ for all (i,j)• Goal: maximize $(\# \text{ vertical nbr. agreeing in } 1^{st} \text{ coordinate})$ $+ (\# \text{ horizontal nbr. agreeing in } 2^{nd} \text{ coordinate})$ $\text{opt}(S) \stackrel{\text{def}}{=} \text{max. of } \mathbb{C}$
- (Lem) Under PIH, $\exists \delta$, it is W[1]-hard to distinguish between
- Completeness: opt(S) = $2k^2$...S is YES
- Soundness: opt(S) $\leq 2k^2 \delta k$... S is much worse than YES

(3) Proof overview on inapproximability Sketch of reduction from GRID TILING

Construct $\mathbf{v}^{(i,j)}_{x,y}$ in $\mathbb{Q}^{O(k^2n^2)}$ for each (x,y) of $S_{i,j}$ s.t. $|\mathbf{v}^{(i,j)}_{x,y}|^2 = 4$,



(3) Proof overview on inapproximability

Sketch of reduction from GRID TILING

Construct $\mathbf{v}^{(i,j)}_{x,y}$ in $\mathbb{Q}^{O(k^2n^2)}$ for each (x,y) of $S_{i,j}$ s.t. $|\mathbf{v}^{(i,j)}_{x,y}|^2 = 4$,
• Same cell $\langle \mathbf{v}^{(i,j)}_{x,y}, \mathbf{v}^{(i,j)}_{x',y'} \rangle \text{ is } \geq 2$ • Vertical nbr. $\langle \mathbf{v}^{(i,j)}_{x,y}, \mathbf{v}^{(i+1,j)}_{x',y'} \rangle \text{ is } \begin{cases} 0 & \text{if } x=x' \\ 1/2 & \text{otherwise} \end{cases}$ • Horizontal nbr. $\langle \mathbf{v}^{(i,j)}_{x,y}, \mathbf{v}^{(i,j+1)}_{x',y'} \rangle \text{ is } \begin{cases} 0 & \text{if } y=y' \\ 1/2 & \text{otherwise} \end{cases}$ • Otherwise

KEY: Gadget of [Çivril & Magdon-Ismail. Algorithmica'13]

(Lem) $det(A_S)$ exponentially decays in # duplicates & $2k^2$ -opt(S); so,

- Completeness: opt(S) = $2k^2$ \rightarrow max_{|S|=k×k} det(A_S) = $4^{k\times k}$
- Soundness: opt(S) $\leq 2k^2 \delta k$ \rightarrow max_{|S|=k×k} det(A_S) $\leq 4^{k\times k} \cdot 0.999^{\delta k}$

Some tractable cases (see the paper)

- 1. Polytime solvable on tridiagonal matrices [Al-Thani & Lee. LAGOS'21]
- Dynamic programming
- 2. Orthogonal vectors in \mathbb{Q}^d is FPT w.r.t. d for nonnegative vectors
- Reduce to SET PACKING

- 3. \(\epsilon\)-additive approximation (bounded entries) is FPT w.r.t. rank
- Use standard rounding technique

Conclusion and future work

• Study parameterized hardness of DETERMINANT MAXIMIZATION

1. Boundary between P vs. NP (or FPT vs. W[1])
 Tridiagonal & spider of bounded legs ...Polytime
 [Al-Thani & Lee. LAGOS'21]

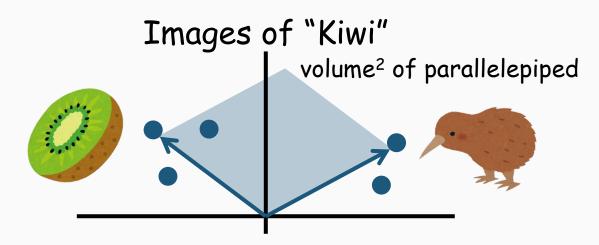
Tree of bounded degree ...?

Arrowhead ...NP-hard & W[1]-hard

- 2. Investigating stronger parameters
- 3. Strengthening inapprox. factor
 - W[1]-hardness of $2^{O(k)}$ -approx. ?



One application: Diversity promotion



Similar items have small determinant → Capture diversity

In Machine Learning, known as determinantal point processes

• Image search, video summarization, sensor placement, ... [Kulesza-Taskar. Found. Trends Mach. Learn. '12]

(3) Proof overview on inapproximability

Optimization version of GRID TILING

obj. val. = $18 = 2 \cdot 3^2$

S _{1,1} (1,1) (3,1) (2,4)	S _{1,2} (5,1) (1,4) (5,3)	S _{1,3} (1,1) (2,4) (3,3)	S _{1,1} (1,1) (3,1) (2,4)
S _{2,1} (2,2) (1,4)	S _{2,2} (3,1) (1,2)	S _{2,3} (2,2) (2,3)	S _{2,1} (2,2) (1,4)
S _{3,1} (1,3) (2,3) (3,3)	S _{3,2} (1,1) (1,3)	S _{3,3} (2,3) (5,3)	S _{3,1} (1,3) (2,3) (3,3)
S _{1,1}	S _{1,2}	5 _{1,3}	

obj. val. = 14

S _{1,1} (1,1) (3,1)	5 _{1,2} (5,1) (1,4)	S _{1,3} (1,1) (2,4)	S _{1,1} (1,1) (3,1)
(2,4)	(5,3)	(3,3)	(2,4)
S _{2,1} (2,2)	S _{2,2} (3,1) (1,2)	S _{2,3} (2,2) (2,3)	S _{2,1} (2,2) (1,4)
5 _{3,1} (1,3) (2,3) (3,3)	S _{3,2} (1,1) (1,3)	S _{3,3} (2,3) (5,3)	S _{3,1} (1,3) (2,3) (3,3)
51,1	S _{1,2}	S _{1,3}	