On the Parameterized Intractability of Determinant Maximization

Naoto Ohsaka

CyberAgent AI Lab

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 $v_1, \ldots, v_n$ in $\mathbb{Q}^d$

$$A \overset{\text{def}}{=} [v_1, \ldots, v_n]^T [v_1, \ldots, v_n], \text{ or } A_{i,j} \overset{\text{def}}{=} \langle v_i, v_j \rangle$$
Example 1: Independent set

- \( Q_3 = (V = [8], E) \): Hypercube graph
- \( v_i \in \{0,1\}^E: v_i(e) \overset{\text{def}}{=} [i \text{ is incident to } e] \)
Example 1: Independent set

- $Q_3 = (V = [8], E)$: Hypercube graph
- $v_i \in \{0, 1\}^E$: $v_i(e) \overset{\text{def}}{=} \left[ i \text{ is incident to } e \right]$
Example 1: Independent set

- $Q_3 = (V = [8], E)$: Hypercube graph
- $v_i \in \{0,1\}^E$: $v_i(e) \overset{\text{def}}{=} \left[ i \text{ is incident to } e \right]$
Example 1: Independent set

- $Q_3 = (V = [8], E)$: Hypercube graph
- $v_i \in \{0,1\}^E$: $v_i(e) \overset{\text{def}}{=} \left[ i \text{ is incident to } e \right]$

$\det(A_S) = 3^{|S|} \Rightarrow S \text{ is independent!}$

e.g., $S = \{1,4,6,7\}$
Example 2: Selecting dispersed points

- \( p_1, \ldots, p_n \): (random) points on \( \mathbb{R}^2 \)
- Let \( A_{i,j} \overset{\text{def}}{=} \exp(|p_i - p_j|^2) \)
  - 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

- $p_1, \ldots, p_n$: (random) points on $\mathbb{R}^2$
- Let $A_{i,j} \overset{\text{def}}{=} \exp(|p_i - p_j|^2)$
  - Known as Gaussian/RBF kernel
  - $A$ is positive semi-definite

Q. What happens if $\det(A_S)$ is max?
A. Select “dispersed” points
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

Gram matrix $A \overset{\text{def}}{=} [v_1, ..., v_n]^T [v_1, ..., v_n]$

$$\det(A_S) = \text{vol}^2\{v_i : i \in S\}$$

**Determinant Maximization = Volume Maximization**
Known results in polynomial-time regime

- NP-hard [Ko-Lee-Queyranne. Oper. Res.'95]
- Greedy is $k!$-approx. [Çivril & Magdon-Ismail. Theor. Comput. Sci.'09]

- NP-hard to $2^{O(k)}$-approx.  
  \[ \uparrow \downarrow \text{nearly tight} \]  
  [Çivril & Magdon-Ismail. Algorithmica'13]  
  [Di Summa-Eisenbrand-Faenza-Moldenhauer. SODA'14]

- Can find $e^k$-approx. [Nikolov. STOC'15]
  \[ k=|S| \text{ is the output size} \]
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. $\Rightarrow$ said to be **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]  
  $\Rightarrow$ 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., $\text{PERMANENT}$ is $\#P$-hard in general, but $\text{FPT}$ w.r.t. treewidth
     
     Courcelle-Makowsky-Rotics. *Discrete Appl. Math. '01* [Cifuentes-Parrilo. *Linear Algebra Appl. '16*]

2. **Strong parameter**
   - $\text{rank}(A) \geq$ output size $k$ (always!)
   - Room for consideration of $f(\text{rank})n^{O(1)}$-time $\text{FPT}$ alg.

3. **FPT approximation** [Feldmann-Karthik-Lee-Manurangsi. *Algorithms'20*]
   - Some $W[1]$-hard problems are approximable in $\text{FPT}$ time
   - e.g., $\text{PARTIAL VERTEX COVER} \& \text{MINIMUM k-MEDIAN}$
     
     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 $\mathbb{P}$-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($A$) ≥ $\Omega$(output size $k$) (always!)
     - Room for consideration of $f(\text{rank})n^{\omega(1)}$-time FPT alg.
     - 😭 *All hopes are dashed!* 😭

3. **FPT approximation** [Feldmann-Karthik-Lee-Manurangsi. Algorithms'20]
   - Some W[1]-hard problems are 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**

Arrowhead = Star graph

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

- **W[1]-hard** when parameterized by rank of $A$
- **W[1]-hard** w.r.t. output size $k$ even if rank only depends on $k$

- **W[1]-hard** to $2^{O(\sqrt{k})}$-approx. w.r.t. $k$ under Parameterized Inapproximability Hypothesis

[{
Lokshtanov-Ramanujan-Saurab-Zehavi. SODA'20][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
Proof overview on $W[1]$-hardness on arrowhead matrices $k$-$\text{SUM}$ [Abboud-Lewi-Williams. ESA'14] & reduction strategy

- **Input:** $n$ integers $x_1, \ldots, x_n$, $t \in [0, n^{2k}]$, $k \in [n]$  
- **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]

- Construct $n+1$ vectors $v_0, v_1, \ldots, v_n$ 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., $v_i$ 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 \( 0 \in S \):

\[
\det(A_S) = \prod_{i \in S \setminus \{0\}} A_{i,i} \cdot \left( A_{0,0} - \sum_{i \in S \setminus \{0\}} \frac{A_{0,i} \cdot A_{0,i}}{A_{i,i}} \right)
\]

(Lem) Carefully choose \( v_0, v_1, ..., v_n \in \mathbb{R}^{+2n} \) s.t. for \( 0 \in S \in \binom{[n]}{k} \)

\[
\det(A_S) \propto \exp\left( \sum_{i \in S \setminus \{0\}} x_i \right) \cdot \left( Z - \sum_{i \in S \setminus \{0\}} x_i \right)
\]

Maximized at \( \sum_{i \in S \setminus \{0\}} x_i = Z - 1 \) - set \( t \)!
(1) Proof overview on $W[1]$-hardness on arrowhead matrices

**Sketch of construction**

|   | 1     | ... | $i$   | ... | $n$ | $n+1$ | ... | $n+i$ | ... | $n+n$ |
|---|-------|-----|-------|-----|-----|-------|-----|-------|-----|-------|
| $v_0$ | $\gamma \sqrt{x_1}$ | $\gamma \sqrt{x_i}$ | $\gamma \sqrt{x_n}$ |   |     |       |     |       |     |       |
| $v_1$ | $\sqrt{a \ e^{x_1}}$ |   |       |     |       | $\sqrt{\beta \ e^{x_1}}$ |     |       |     |       |
| $v_i$ | $\sqrt{a \ e^{x_i}}$ |   | $\sqrt{\beta \ e^{x_i}}$ |     |       |       |     |       |     |       |
| $v_n$ | $\sqrt{a \ e^{x_n}}$ |   |       |     |       | $\sqrt{\beta \ e^{x_n}}$ |     |       |     |       |

Parameterized by $\alpha$, $\beta$, $\gamma$ (to be determined appropriately)

**Omitted details:** We have to...

- efficiently approximate $v_0$, $v_1$, $\ldots$, $v_n$ using rationals
- ensure that any optimal solution includes $v_0$
(2) Proof overview on $W[1]$-hardness by rank

**(Thm)** $\text{Determinant Maximization}$ is $W[1]$-hard w.r.t. rank of $A$

- **Grid Tiling**: $W[1]$-complete [Marx. *FOCS'07*]

⚠️ Can use only $f(k)$-dimensional vectors / $f(k)$-rank matrices

e.g., vectors in $\mathbb{Q}^n$ are not allowed

- $\text{Determinant Maximization}$ 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 1\(^{st}\) coordinate
  - Horizontal neighbors agree in 2\(^{nd}\) coordinate

- Equality constraints are **SIMPLE 😊**
- Cells \((i,j)\) are adjacent to **FOUR** cells 😊

| \( S_{1,1} \) | \( S_{1,2} \) | \( S_{1,3} \) |
|--------------|--------------|--------------|
| (1,1)        | (5,1)        | (1,1)        |
| (3,1)        | (1,4)        | (2,4)        |
| (2,4)        | (5,3)        | (3,3)        |

| \( S_{2,1} \) | \( S_{2,2} \) | \( S_{2,3} \) |
|--------------|--------------|--------------|
| (2,2)        | (3,1)        | (2,2)        |
| (1,4)        | (1,2)        | (2,2)        |

| \( S_{3,1} \) | \( S_{3,2} \) | \( S_{3,3} \) |
|--------------|--------------|--------------|
| (1,3)        | (1,1)        | (2,3)        |
| (2,3)        | (1,3)        | (3,3)        |
| (3,3)        | (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*]

|   | $S_{1,1}$ | $S_{1,2}$ | $S_{1,3}$ | $S_{1,1}$ |
|---|-----------|-----------|-----------|-----------|
| 1 | (1,1)     | (5,1)     | (1,1)     | (1,1)     |
| 2 | (3,1)     | (1,4)     | (2,4)     | (3,1)     |
| 3 | (2,4)     | (5,3)     | (3,3)     | (2,4)     |

|   | $S_{2,1}$ | $S_{2,2}$ | $S_{2,3}$ | $S_{2,1}$ |
|---|-----------|-----------|-----------|-----------|
| 1 | (2,2)     | (3,1)     | (2,2)     | (2,2)     |
| 2 | (1,4)     | (1,2)     | (2,3)     | (1,4)     |
| 3 | (2,3)     | (2,3)     | (2,3)     | (2,3)     |

|   | $S_{3,1}$ | $S_{3,2}$ | $S_{3,3}$ | $S_{3,1}$ |
|---|-----------|-----------|-----------|-----------|
| 1 | (1,3)     | (1,1)     | (1,3)     | (1,3)     |
| 2 | (2,3)     | (2,3)     | (2,3)     | (2,3)     |
| 3 | (3,3)     | (5,3)     | (3,3)     | (3,3)     |

Perfect consistency 😊

4 neighbors are inconsistent 😖

|   | $S_{1,1}$ | $S_{1,2}$ | $S_{1,3}$ | $S_{1,1}$ |
|---|-----------|-----------|-----------|-----------|
| 1 | (1,1)     | (5,1)     | (1,1)     | (1,1)     |
| 2 | (3,1)     | (1,4)     | (2,4)     | (3,1)     |
| 3 | (2,4)     | (5,3)     | (3,3)     | (2,4)     |

|   | $S_{2,1}$ | $S_{2,2}$ | $S_{2,3}$ | $S_{2,1}$ |
|---|-----------|-----------|-----------|-----------|
| 1 | (2,2)     | (3,1)     | (2,2)     | (2,2)     |
| 2 | (1,4)     | (1,2)     | (2,3)     | (1,4)     |
| 3 | (2,3)     | (2,3)     | (2,3)     | (2,3)     |

|   | $S_{3,1}$ | $S_{3,2}$ | $S_{3,2}$ | $S_{3,1}$ |
|---|-----------|-----------|-----------|-----------|
| 1 | (1,3)     | (1,1)     | (1,1)     | (1,3)     |
| 2 | (2,3)     | (2,3)     | (2,3)     | (2,3)     |
| 3 | (3,3)     | (5,3)     | (3,3)     | (3,3)     |

|   | $S_{1,1}$ | $S_{1,2}$ | $S_{1,3}$ | $S_{1,1}$ |
|---|-----------|-----------|-----------|-----------|
| 1 | (1,1)     | (5,1)     | (1,1)     | (1,1)     |
| 2 | (3,1)     | (1,4)     | (2,4)     | (3,1)     |
| 3 | (2,4)     | (5,3)     | (3,3)     | (2,4)     |

|   | $S_{2,1}$ | $S_{2,2}$ | $S_{2,3}$ | $S_{2,1}$ |
|---|-----------|-----------|-----------|-----------|
| 1 | (2,2)     | (3,1)     | (2,2)     | (2,2)     |
| 2 | (1,4)     | (1,2)     | (2,3)     | (1,4)     |
| 3 | (2,3)     | (2,3)     | (2,3)     | (2,3)     |

|   | $S_{3,1}$ | $S_{3,2}$ | $S_{3,3}$ | $S_{3,1}$ |
|---|-----------|-----------|-----------|-----------|
| 1 | (1,3)     | (1,1)     | (1,3)     | (1,3)     |
| 2 | (2,3)     | (2,3)     | (2,3)     | (2,3)     |
| 3 | (3,3)     | (5,3)     | (3,3)     | (3,3)     |
(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

$\mathcal{G}$: $f(k)$-dim. $v^{(i,j)}_{x,y}$ for $(x,y)$ in $S_{i,j}$ describing "consistency":

**Conditions about "consistency"**
(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

$\mathsf{f(k)}$-dim. $v_{x,y}^{(i,j)}$ for $(x,y)$ in $S_{i,j}$ describing “consistency”:
- Vertical nbr. $\langle v_{x,y}^{(i,j)}, v_{x',y'}^{(i+1,j)} \rangle = 0$ iff $x = x'$
- Horizontal nbr. $\langle v_{x,y}^{(i,j)}, v_{x',y'}^{(i,j+1)} \rangle = 0$ iff $y = y'$
- Same cell $\langle v_{x,y}^{(i,j)}, v_{x',y'}^{(i,j)} \rangle \neq 0$

Gram matrix $A_{x,y,i,j,i',j',x',y'} \stackrel{\text{def}}{=} \langle v_{x,y}^{(i,j)}, v_{x,y'}^{(i',j')} \rangle$ satisfies...

- $S$ is YES $\Rightarrow \exists k^2 \times k^2$ diagonal submatrix... select CORRECT $v_{x,y}^{(i,j)}$ for each $(i,j) \in [k]^2$
- $S$ is NO $\Rightarrow \forall k^2 \times k^2$ submatrix is NOT diagonal
Proof overview on W[1]-hardness by rank

Represent "consistency" at lower dimensions?

- Want \( v_1, \ldots, v_n, w_1, \ldots, w_n \) in \( \mathbb{Q}^{O(1)} \) s.t. \( \langle v_i, w_j \rangle = 0 \) iff \( i=j \)
  😊 How to construct?

🚫 One-hot vectors require \( n \)-dimension \([0,\ldots,0,1,0,\ldots,0]\)

😊 Use points on the unit circle:
  - \( v_i \overset{\text{def}}{=} (\cos\left(\frac{\pi i}{2n}\right), \sin\left(\frac{\pi i}{2n}\right)) \)
  - \( w_j \overset{\text{def}}{=} (\sin\left(\frac{\pi j}{2n}\right), -\cos\left(\frac{\pi j}{2n}\right)) \)

Use Pythagorean triples to get rational vectors
(3) Proof overview on inapproximability

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

- Parameterized Inapproximability Hypothesis (PIH)  
  \textit{[Lokshtanov-Ramanujan-Saurab-Zehavi. SODA'20]} I don't go into details in this talk

- Optimization version of \textsc{Grid Tiling}: $W[1]$-hard to approx. w.r.t. $k$
  \textit{⚠️ Gap-preserving reduction (different from the last one)}

- \textsc{Determinant Maximization} parameterized by $k$
(3) Proof overview on inapproximability

**Optimization version of GRID TILING**

- **Input:** \( S \overset{\text{def}}{=} (S_{i,j} \subseteq [n]^2 : 1 \leq i,j \leq k) \)
- **Output:** Select \((x,y)\) in \(S_{i,j}\) for all \((i,j)\)
- **Goal:** maximize (\# vertical nbr. agreeing in 1\(^{\text{st}}\) coordinate) 
  + (\# horizontal nbr. agreeing in 2\(^{\text{nd}}\) coordinate)

\[ \text{opt}(S) \overset{\text{def}}{=} \max \text{ of } \]

(Lem) Under PIH, \( \exists \delta \), it is \( W[1]\)-hard to distinguish between

- **Completeness:** \( \text{opt}(S) = 2k^2 \) \( \ldots \) \( S \) is YES
- **Soundness:** \( \text{opt}(S) \leq 2k^2 - \delta k \) \( \ldots \) \( S \) is much worse than YES
(3) Proof overview on inapproximability

Sketch of reduction from GRID TILING

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

Undesirable cases impose **const.** penalty
(3) Proof overview on inapproximability

Sketch of reduction from GRID TILING

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

- Same cell
  \[ \langle v^{(i,j)}_{x,y}, v^{(i,j)}_{x',y'} \rangle \] is \( \geq 2 \)

- Vertical nbr.
  \[ \langle v^{(i,j)}_{x,y}, v^{(i+1,j)}_{x',y'} \rangle \] is \( \begin{cases} 0 & \text{if } x=x' \\ 1/2 & \text{otherwise} \end{cases} \)

- Horizontal nbr.
  \[ \langle v^{(i,j)}_{x,y}, v^{(i,j+1)}_{x',y'} \rangle \] is \( \begin{cases} 0 & \text{if } y=y' \\ 1/2 & \text{otherwise} \end{cases} \)

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

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

- Completeness: \( \text{opt}(S) = 2k^2 \) \( \Rightarrow \) \( \max_{|S|=k \times k} \det(A_S) = 4^{k \times k} \)

- Soundness: \( \text{opt}(S) \leq 2k^2-\delta k \) \( \Rightarrow \) \( \max_{|S|=k \times k} \det(A_S) \leq 4^{k \times k \cdot 0.999^{\delta k}} \)
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. $\varepsilon$-additive approximation (bounded entries) is FPT w.r.t. **rank**
   • Use standard rounding technique
Conclusion and future work

- Study parameterized hardness of \textsc{Determinant Maximization}

1. Boundary between P vs. NP (or FPT vs. W[1])
   - Tridiagonal & spider of bounded legs ... Polytime
   - \cite{Al-Thani & Lee. LAGOS'21}
   - Tree of bounded degree ...
   - Arrowhead ... NP-hard & W[1]-hard

2. Further strong parameters?

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