On the Parameterized Intractability of Determinant Maximization

Naoto Ohsaka
CyberAgent, Inc., Tokyo, Japan

Abstract

In the Determinant Maximization problem, given an $n \times n$ positive semi-definite matrix $\mathbf{A}$ in $\mathbb{Q}^{n\times n}$ and an integer k, we are required to find a $k \times k$ principal submatrix of $\mathbf{A}$ having the maximum determinant. This problem is known to be NP-hard and further proven to be W[1]-hard with respect to k by Koutis [26]; i.e., a $f(k)n^{\mathcal{O}(1)}$-time algorithm is unlikely to exist for any computable function f. However, there is still room to explore its parameterized complexity in the *restricted case*, in the hope of overcoming the general-case parameterized intractability. In this study, we rule out the fixed-parameter tractability of Determinant Maximization even if an input matrix is extremely sparse or low rank, or an approximate solution is acceptable. We first prove that Determinant Maximization is NP-hard and W[1]-hard even if an input matrix is an *arrowhead matrix*; i.e., the underlying graph formed by nonzero entries is a star, implying that the structural sparsity is not helpful. By contrast, we show that Determinant Maximization is solvable in polynomial time on *tridiagonal matrices*. Thereafter, we demonstrate the W[1]-hardness with respect to the *rank* r of an input matrix. Our result is stronger than Koutis' result in the sense that any $k \times k$ principal submatrix is singular whenever $k > r$. We finally give evidence that it is W[1]-hard to approximate Determinant Maximization parameterized by k within a factor of $2^{-c\sqrt{k}}$ for some universal constant $c > 0$. Our hardness result is conditional on the *Parameterized Inapproximability Hypothesis* posed by Lokshtanov, Ramanujan, Saurab, and Zehavi [30], which asserts that a gap version of Binary Constraint Satisfaction Problem is W[1]-hard. To complement this result, we develop an ε-additive approximation algorithm that runs in $\varepsilon^{-r^2} \cdot r^{\mathcal{O}(r^3)} \cdot n^{\mathcal{O}(1)}$ time for the rank r of an input matrix, provided that the diagonal entries are bounded.

2012 ACM Subject Classification Theory of computation → Parameterized complexity and exact algorithms

Keywords and phrases Determinant maximization, Parameterized complexity, Approximability

Digital Object Identifier 10.4230/LIPIcs.ISAAC.2022.46

Related Version *Full Version*: `https://arxiv.org/abs/2209.12519` [38]

1 Introduction

Background. We study the following Determinant Maximization problem: Given an $n \times n$ positive semi-definite matrix $\mathbf{A}$ in $\mathbb{Q}^{n\times n}$ and an integer k in $[n]$ denoting the solution size, find a $k \times k$ principal submatrix of $\mathbf{A}$ having the maximum determinant; namely, maximize $\det(\mathbf{A}_S)$ subject to $S \in \binom{[n]}{k}$. One motivating example for this problem is a *subset selection task*. Suppose we are given n items (e.g., images or products) associated with feature vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and required to select a "diverse" set of k items among them. We can measure the diversity of a set S of k items using the principal minor $\det(\mathbf{A}_S)$ of the Gram matrix $\mathbf{A}$ defined by feature vectors such that $A_{i,j} \triangleq \langle \mathbf{v}_i, \mathbf{v}_j \rangle$ for all $i, j \in [n]$, resulting in Determinant Maximization. This formulation is justified by the fact that $\det(\mathbf{A}_S)$ is equal to the squared volume of the parallelepiped spanned by $\{\mathbf{v}_i : i \in S\}$; that is, a pair of vectors at a large angle is regarded as more diverse. In artificial intelligence and machine learning communities, Determinant Maximization is also known as MAP inference on a *determinantal point process* [6, 31], and has found many applications over the past decade, including tweet timeline generation [43], object detection [29], change-point

33rd International Symposium on Algorithms and Computation (ISAAC 2022).
Editors: Sang Won Bae and Heejin Park; Article No. 46; pp. 46:1–46:16
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

detection [44], document summarization [8, 27], YouTube video recommendation [42], and active learning [5]. See the survey of Kulesza and Taskar [28] for further details. Though Determinant Maximization is known to be NP-hard to solve exactly [25], we can achieve an e^{-k}-factor approximation in polynomial time [36], which is nearly optimal because a 2^{-ck}-factor approximation for some constant $c > 0$ is impossible unless $\mathsf{P} = \mathsf{NP}$ [11, 14, 26].

Having known a nearly tight hardness-of-approximation result in the polynomial-time regime, we resort to *parameterized algorithms* [13, 16, 19]. We say that a problem is *fixed-parameter tractable* (FPT) with respect to a parameter $k \in \mathbb{N}$ if it can be solved in $f(k)|\mathcal{I}|^{\mathcal{O}(1)}$ time for some computable function f and instance size $|\mathcal{I}|$. One very natural parameter is the *solution size* k, which is expected to be small in practice. By enumerating all $k \times k$ principal submatrices, we can solve Determinant Maximization in $n^{k+\mathcal{O}(1)}$ time; i.e., it belongs to the class XP. Because $\mathsf{FPT} \subsetneq \mathsf{XP}$ [16], it is even more desirable if an FPT algorithm exists. Unfortunately, Koutis [26] has already proven that Determinant Maximization is W[1]-hard with respect to k. Therefore, under the widely-believed assumption that $\mathsf{FPT} \neq \mathsf{W[1]}$, an FPT algorithm for Determinant Maximization does not exist.

However, there is still room to explore the parameterized complexity of Determinant Maximization in the *restricted case*, in the hope of circumventing the general-case parameterized intractability. Here, we describe three possible scenarios. One can first assume an input matrix $\mathbf{A}$ to be *sparse*. Of particular interest is the structural sparsity of the *symmetrized graph* of $\mathbf{A}$ [9, 12] defined as the underlying graph formed by nonzero entries of $\mathbf{A}$, encouraged by numerous FPT algorithms for NP-hard graph-theoretic problems parameterized by the treewidth [13, 20]. For example, in change-point detection applications, Zhang and Ou [44] observed a small-bandwidth matrix and developed an efficient heuristic for Determinant Maximization. In addition, one may adopt a strong parameter. The *rank* of an input matrix $\mathbf{A}$ is such a natural candidate. We often assume that $\mathbf{A}$ is low-rank in applications; for instance, the feature vectors $\mathbf{v}_i$ are inherently low-dimensional [7] or the largest possible subset is significantly smaller than the ground set size n. Since any $k \times k$ principal submatrix of $\mathbf{A}$ is singular whenever $k > \mathrm{rank}(\mathbf{A})$, we can ensure that $k \leqslant \mathrm{rank}(\mathbf{A})$; namely, parameterization by $\mathrm{rank}(\mathbf{A})$ is considered stronger than that by k. Intriguingly, the partition function of product determinantal point processes is FPT with respect to rank while #P-hard in general [40]. The last possibility to be considered is *FPT-approximability.* Albeit W[1]-hardness of Determinant Maximization with parameter k, it could be possible to obtain an approximate solution in FPT time. It has been demonstrated that several W[1]-hard problems can be approximated in FPT time, such as Partial Vertex Cover and Minimum k-Median [22] (refer to the survey of Marx [34] and Feldmann, Karthik, Lee, and Manurangsi [17]). One may thus envision the existence of a $1/\rho(k)$-factor FPT-approximation algorithm for Determinant Maximization for a *small* function ρ. Alas, we refute the above possibilities under a plausible assumption in parameterized complexity.

Our Results. We improve the W[1]-hardness of Determinant Maximization due to Koutis [26] by showing that it is still W[1]-hard even if an input matrix is extremely sparse or low rank, or an approximate solution is acceptable, along with some tractable cases.

We first prove that Determinant Maximization is NP-hard and W[1]-hard with respect to k even if the input matrix $\mathbf{A}$ is an *arrowhead matrix* (Theorem 3.1). An arrowhead matrix is a square matrix that can include nonzero entries only in the first row, the first column, or the diagonal; i.e., its symmetrized graph is a star. Our hardness result implies that the "structural sparsity" of input matrices is not helpful; in particular, it follows from Theorem 3.1 that this problem is NP-hard even if the treewidth, pathwidth, and vertex cover number of the symmetrized graph are all 1. The proof is based on a parameterized reduction from k-Sum, which is a parameterized version of Subset Sum known to be W[1]-complete [1, 15], and involves a structural feature of the determinant of arrowhead matrices. On the other hand, we show that Determinant Maximization is solvable in polynomial time on *tridiagonal* matrices (Observation 3.9), whose symmetrized graph is a path graph.

Thereafter, we demonstrate that Determinant Maximization is W[1]-hard when parameterized by the *rank* of an input matrix (Corollary 4.3). In fact, we obtain the stronger result that it is W[1]-hard to determine whether an input set of n d-dimensional vectors includes k pairwise orthogonal vectors when parameterized by d (Theorem 4.2). Unlike the proof of Theorem 3.1, we are allowed to construct only a $f(k)$-dimensional vector in a parameterized reduction. Note that a straightforward parameterized reduction from a canonical W[1]-complete k-Clique problem fails (see Remark 4.4). Therefore, we reduce from a different W[1]-complete problem called Grid Tiling due to Marx [33, 35]. In Grid Tiling, we are given k^2 nonempty sets of integer pairs arranged in a $k \times k$ grid, and the task is to select k^2 integer pairs such that the vertical and horizontal neighbors agree respectively in the first and second coordinates (see Problem 4.5 for the precise definition). Grid Tiling is favorable for our purpose because the constraint consists of simple equalities, and each cell is adjacent to (at most) four cells. To express the consistency between adjacent cells using only a $f(k)$-dimensional vector, we exploit Pythagorean triples. It is essential in Theorem 4.2 that the input vectors can include *both* positive and negative entries in a sense that we can find k d-dimensional nonnegative vectors that are pairwise orthogonal in FPT time with respect to d (Observation 4.7).

Our final contribution is to give evidence that it is W[1]-hard to determine whether the optimal value of Determinant Maximization is equal to 1 or at most $2^{-c\sqrt{k}}$ for some universal constant $c > 0$; namely, Determinant Maximization is FPT-inapproximable within a factor of $2^{-c\sqrt{k}}$ (Theorem 5.1). Our result is conditional on the *Parameterized Inapproximability Hypothesis* (PIH), which is a conjecture posed by Lokshtanov, Ramanujan, Saurab, and Zehavi [30] asserting that a gap version of Binary Constraint Satisfaction Problem is W[1]-hard when parameterized by the number of variables. PIH can be thought of as a parameterized analogue of the PCP theorem [2,3]; e.g., Lokshtanov et al. [30] show that assuming PIH and FPT $\neq$ W[1], Directed Odd Cycle Transversal does not admit a $(1-\varepsilon)$-factor FPT-approximation algorithm for some $\varepsilon > 0$. The proof of Theorem 5.1 involves FPT-inapproximability of Grid Tiling under PIH, which is reminiscent of Marx's work [33] and might be of some independent interest. Because we cannot achieve an exponential gap by simply reusing the parameterized reduction from Grid Tiling of the second hardness result (as inferred from Observation 5.11 below), we apply a gadget invented by Çivril and Magdon-Ismail [11] to construct an $\mathcal{O}(k^2n^2)$-dimensional vector for each integer pair of a Grid Tiling instance. We further show that the same kind of hardness result does *not* hold when parameterized by the rank r of an input matrix. Specifically, we develop an ε-additive approximation algorithm that runs in $\varepsilon^{-r^2} \cdot r^{\mathcal{O}(r^3)} \cdot n^{\mathcal{O}(1)}$ time for any $\varepsilon > 0$, provided that the diagonal entries are bounded (Observation 5.11).

Owing to space limitations, proofs marked with $*$ are omitted and can be found in a full version of this paper [38].

More Related Work. Determinant Maximization is not only applied in artificial intelligence and machine learning but also in computational geometry [21] and discrepancy theory; refer to Nikolov [36] and references therein. On the negative side, Ko, Lee, and

Queyranne [25] prove that Determinant Maximization is NP-hard, and Koutis [26] proves that it is further W[1]-hard. NP-hardness of approximating Determinant Maximization has been investigated in [11, 14, 26, 39]. On the algorithmic side, a greedy algorithm achieves an approximation factor of $1/k!$ [10]. Subsequently, Nikolov [36] gives an e^{-k}-factor approximation algorithm; partition constraints [37] and matroid constraints [32] are also studied. Several #P-hard computation problems over matrices including permanents [9, 12], hyperdeterminants [9], and partition functions of product determinantal point processes [40] are efficiently computable if the treewidth of the symmetrized graph or the matrix rank is bounded.

2 Preliminaries

Notations and Definitions. For two integers $m, n \in \mathbb{N}$ with $m \leqslant n$, let $[n] \triangleq \{1, 2, \dots, n\}$ and $[m \mathinner{..} n] \triangleq \{m, m+1, \dots, n-1, n\}$. For a finite set S and an integer k, we write $\binom{S}{k}$ for the family of all size-k subsets of S. For a statement P, $[\![P]\!]$ is 1 if P is true, and 0 otherwise. The base of logarithms is 2. The Euclidean norm is denoted $\|\cdot\|$; i.e., $\|\mathbf{v}\| \triangleq \sqrt{\sum_{i \in [d]} (v(i))^2}$ for a vector $\mathbf{v} \in \mathbb{R}^d$. We use $\langle \cdot, \cdot \rangle$ for the standard inner product; i.e., $\langle \mathbf{v}, \mathbf{w} \rangle \triangleq \sum_{i \in [d]} v(i) \cdot w(i)$ for two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^d$. For an $n \times n$ matrix $\mathbf{A}$ and an index set $S \subseteq [n]$, we use $\mathbf{A}_S$ to denote the principal submatrix of $\mathbf{A}$ whose rows and columns are indexed by S. For an $m \times n$ matrix $\mathbf{A}$, the *spectral norm* $\|\mathbf{A}\|_2$ is defined as the square root of the maximum eigenvalue of $\mathbf{A}^\top \mathbf{A}$ and the *max norm* is defined as $\|\mathbf{A}\|_{\max} \triangleq \max_{i,j} |A_{i,j}|$. It is well-known that $\|\mathbf{A}\|_{\max} \leqslant \|\mathbf{A}\|_2 \leqslant \sqrt{mn} \cdot \|\mathbf{A}\|_{\max}$. The *symmetrized graph* [9, 12] of an $n \times n$ matrix $\mathbf{A}$ is defined as an undirected graph G that has each integer of $[n]$ as a vertex and an edge $(i, j) \in \binom{[n]}{2}$ if $A_{i,j} \neq 0$ or $A_{j,i} \neq 0$; i.e., $G = ([n], \{(i, j) \colon A_{i,j} \neq 0\})$. For a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, its *determinant* is defined as follows:

$$\det(\mathbf{A}) \triangleq \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) \prod_{i \in [n]} A_{i,\sigma(i)},$$

where $\mathfrak{S}_n$ denotes the symmetric group on $[n]$, and $\operatorname{sgn}(\sigma)$ denotes the sign of a permutation σ. We define $\det(\mathbf{A}_\emptyset) \triangleq 1$. For a collection $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of n vectors in $\mathbb{R}^d$, the *volume* of the parallelepiped spanned by $\mathbf{V}$ is defined as follows:

$$\operatorname{vol}(\mathbf{V}) \triangleq \|\mathbf{v}_1\| \cdot \prod_{2 \leqslant i \leqslant n} \mathrm{d}(\mathbf{v}_i, \{\mathbf{v}_1, \dots, \mathbf{v}_{i-1}\}). \tag{1}$$

Here, $\mathrm{d}(\mathbf{v}, \mathbf{P})$ denotes the distance of $\mathbf{v}$ to the subspace spanned by $\mathbf{P}$; i.e., $\mathrm{d}(\mathbf{v}, \mathbf{P}) \triangleq \|\mathbf{v} - \operatorname{proj}_{\mathbf{P}}(\mathbf{v})\|$, where $\operatorname{proj}_{\mathbf{P}}(\cdot)$ is an operator of orthogonal projection onto the subspace spanned by $\mathbf{P}$. We define $\operatorname{vol}(\emptyset) \triangleq 1$ for the sake of consistency to the determinant of an empty matrix (i.e., $\det([\,]) = 1 = \operatorname{vol}^2(\emptyset)$). If $\mathbf{A}$ is the Gram matrix defined as $A_{i,j} \triangleq \langle \mathbf{v}_i, \mathbf{v}_j \rangle$ for all $i, j \in [n]$, we have a simple relation between the principal minor and the volume of the parallelepiped that

$$\det(\mathbf{A}_S) = \operatorname{vol}^2(\{\mathbf{v}_i \colon i \in S\}) \tag{2}$$

for every $S \subseteq [n]$. We formally define the Determinant Maximization problem as follows.[1]

[1] Note that if we consider the decision version of Determinant Maximization, we are additionally given a target number τ and are required to decide if $\operatorname{maxdet}(\mathbf{A}, k) \geqslant \tau$.

▶ **Problem 2.1.** *Given a positive semi-definite matrix* $\mathbf{A}$ *in* $\mathbb{Q}^{n\times n}$ *and a positive integer* $k \in [n]$, Determinant Maximization *asks to find a set* $S \in \binom{[n]}{k}$ *such that the determinant* $\det(\mathbf{A}_S)$ *of a* $k \times k$ *principal submatrix is maximized. The optimal value is denoted* $\mathrm{maxdet}(\mathbf{A}, k) \triangleq \max_{S\in\binom{[n]}{k}} \det(\mathbf{A}_S)$.

Due to the equivalence between squared volume and determinant in Eq. (2), Determinant Maximization is equivalent to the following problem of volume maximization: Given a collection of n vectors in $\mathbb{Q}^d$ and a positive integer $k \in [n]$, we are required to find k vectors such that the volume of the parallelepiped spanned by them is maximized. We shall use the problem definition based on the determinant and the volume interchangeably.

Parameterized Complexity. Given a *parameterized problem* Π consisting of a pair $\langle \mathcal{I}, k\rangle$ of instance $\mathcal{I}$ and parameter $k \in \mathbb{N}$, we say that Π is *fixed-parameter tractable* (FPT) with respect to k if it is solvable in $f(k)|\mathcal{I}|^{\mathcal{O}(1)}$ time for some computable function f, and *slice-wise polynomial* (XP) if it is solvable in $|\mathcal{I}|^{f(k)}$ time; it holds that $\mathsf{FPT} \subsetneq \mathsf{XP}$ [16]. The value of parameter k may be independent of the instance size $|\mathcal{I}|$ and may be given by some computable function $k = k(\mathcal{I})$ on instance $\mathcal{I}$ (e.g., the rank of an input matrix). Our objective is to prove that a problem (i.e., Determinant Maximization) is unlikely to admit an FPT algorithm under plausible assumptions in parameterized complexity. The central notion for this purpose is a parameterized reduction, which is used to demonstrate that a problem of interest is hard for a particular class of parameterized problems that is believed to be a superclass of FPT. We say that a parameterized problem Π_1 is *parameterized reducible* to another parameterized problem Π_2 if (i) an instance $\mathcal{I}_1$ with parameter k_1 for Π_1 can be transformed into an instance $\mathcal{I}_2$ with parameter k_2 for Π_2 in FPT time and (ii) the value of k_2 only depends on the value of k_1. Note that a parameterized reduction may not be a polynomial-time reduction and vice versa. $\mathsf{W}[1]$ is a class of parameterized problems that are parameterized reducible to k-Clique, and it is known that $\mathsf{FPT} \subseteq \mathsf{W}[1] \subseteq \mathsf{XP}$. This class is often regarded as a parameterized counterpart to NP of classical complexity; in particular, the conjecture $\mathsf{FPT} \neq \mathsf{W}[1]$ is a widely-believed assumption in parameterized complexity [16, 19]. Thus, the existence of a parameterized reduction from a $\mathsf{W}[1]$-complete problem to a problem Π is a strong evidence that Π is not in FPT. In Determinant Maximization, a simple brute-force search algorithm that examines all $\binom{[n]}{k}$ subsets of size k runs in $n^{k+\mathcal{O}(1)}$ time; hence, this problem belongs to XP. On the other hand, it is proven to be $\mathsf{W}[1]$-hard [26].

3 W[1]-hardness and NP-hardness on Arrowhead Matrices

We first prove the $\mathsf{W}[1]$-hardness with respect to k and NP-hardness on arrowhead matrices. A square matrix $\mathbf{A}$ in $\mathbb{R}^{[0..n]\times[0..n]}$ is an *arrowhead matrix* if $A_{i,j} = 0$ for all $i, j \in [n]$ with $i \neq j$. In the language of graph theory, $\mathbf{A}$ is arrowhead if its symmetrized graph is a star $K_{1,n}$.

▶ **Theorem 3.1.** Determinant Maximization *on arrowhead matrices is* NP*-hard and* $\mathsf{W}[1]$*-hard when parameterized by* k.

The proof of Theorem 3.1 requires a reduction from k-Sum, a natural parameterized version of the NP-complete Subset Sum problem, whose membership of $\mathsf{W}[1]$ and $\mathsf{W}[1]$-hardness was proven by Abboud, Lewi, and Williams [1] and Downey and Fellows [15], respectively.

▶ **Problem 3.2** (k-Sum due to Abboud, Lewi, and Williams [1]). *Given* n *integers* $x_1, \ldots, x_n \in [0 .. n^{2k}]$, *a target integer* $t \in [0 .. n^{2k}]$, *and a positive integer* $k \in [n]$, *we are required to decide if there exists a size-*k *set* $S \in \binom{[n]}{k}$ *such that* $\sum_{i\in S} x_i = t$.

Here, we introduce a slightly-modified version of k-Sum such that the input numbers are *rational* and their sum is *normalized* to 1, without affecting its computational complexity.

▶ **Problem 3.3** (k-Sum modified from [1]). *Given n rational numbers $x_1, \ldots, x_n$ in $(0,1) \cap \mathbb{Q}_+$, a target rational number t in $(0,1) \cap \mathbb{Q}_+$, and a positive integer $k \in [n]$ such that x_i's are integer multiples of some rational number at least $\frac{1}{n^{2k+1}}$ and $\sum_{i \in [n]} x_i = 1$, k-Sum asks to decide if there exists a set $S \in \binom{[n]}{k}$ such that $\sum_{i \in S} x_i = t$.*

Hereafter, for any set $S \subseteq [0 .. n]$ including 0, we denote $S_{-0} \triangleq S \setminus \{0\}$.

3.1 Reduction from k-Sum and Proof of Theorem 3.1

In this subsection, we give a parameterized, polynomial-time reduction from k-Sum. We first use an explicit formula of the determinant of arrowhead matrices.

▶ **Lemma 3.4** (∗). *Let $\mathbf{A}$ be an arrowhead matrix in $\mathbb{R}^{[0..n] \times [0..n]}$ such that $A_{i,i} \neq 0$ for all $i \in [n]$. Then, for any set $S \subseteq [0 .. n]$, it holds that*

$$\det(\mathbf{A}_S) = \begin{cases} \prod_{i \in S_{-0}} A_{i,i} \cdot \left(A_{0,0} - \sum_{i \in S_{-0}} \frac{A_{0,i} \cdot A_{i,0}}{A_{i,i}} \right) & \text{if } 0 \in S, \\ \prod_{i \in S} A_{i,i} & \text{if } 0 \notin S. \end{cases}$$

Lemma 3.4 shows us a way to express the *product* of $\prod_{i \in S_{-0}} x_i$ and $1 - C \cdot \sum_{i \in S_{-0}} x_i$ for some constant C, which is a key step in proving Theorem 3.1. Specifically, given n rational numbers $x_1, \ldots, x_n$ and a target rational number t as a k-Sum instance, we construct $n+1$ $2n$-dimensional vectors $\mathbf{v}_0, \ldots, \mathbf{v}_n$ in $\mathbb{R}_+^{2n}$, each entry of which is defined as follows:

$$v_0(j) = \begin{cases} \gamma \cdot \sqrt{x_j} & \text{if } j \leqslant n, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and } v_i(j) = \begin{cases} \sqrt{\alpha \cdot \mathrm{e}^{x_i}} & \text{if } j = i, \\ \sqrt{\beta \cdot \mathrm{e}^{x_i}} & \text{if } j = i + n, \\ 0 & \text{otherwise,} \end{cases} \text{ for all } i \in [n], \tag{3}$$

where α, β, and γ are parameters, whose values are positive and will be determined later. We calculate the principal minor of the Gram matrix defined by $\mathbf{v}_0, \ldots, \mathbf{v}_n$ as follows.

▶ **Lemma 3.5** (∗). *Let $\mathbf{A}$ be the Gram matrix defined by $n+1$ vectors $\mathbf{v}_0, \ldots, \mathbf{v}_n$ that are constructed from an instance of k-Sum by Eq. (3). Then, $\mathbf{A}$ is an arrowhead matrix, and for any set $S \subseteq [0 .. n]$, it holds that*

$$\det(\mathbf{A}_S) = \begin{cases} (\alpha + \beta)^{|S|-1} \cdot \gamma^2 \cdot \exp\left(\sum_{i \in S_{-0}} x_i \right) \cdot \left(1 - \frac{\alpha}{\alpha + \beta} \sum_{i \in S_{-0}} x_i \right) & \text{if } 0 \in S, \\ (\alpha + \beta)^{|S|} \cdot \exp\left(\sum_{i \in S} x_i \right) & \text{if } 0 \notin S. \end{cases}$$

Moreover, if we regard the principal minor $\det(\mathbf{A}_S)$ in the case of $0 \in S$ as a function in $X \triangleq \sum_{i \in S_{-0}} x_i$, it is maximized when $X = \frac{\beta}{\alpha}$.

We now determine the values of α, β, and γ. Since Lemma 3.5 demonstrates that the principal minor for S including 0 is maximized when $\sum_{i \in S_{-0}} x_i = \frac{\beta}{\alpha}$, we fix $\alpha \triangleq 1$ and $\beta \triangleq t$. We define $\delta \triangleq \frac{1}{n^{2k+1}}$, denoting a lower bound on the *minimum possible absolute* *difference* between any sum of x_i's; i.e., $|\sum_{i \in S} x_i - \sum_{i \in T} x_i| \geqslant \delta$ for any $S, T \subseteq [n]$ whenever $\sum_{i \in S} x_i \neq \sum_{i \in T} x_i$. For the correctness of the value of δ, refer to the definition of Problem 3.3. We finally fix the value of γ as $\gamma \triangleq 5$, so that

$$(1+t)^2 \cdot \mathrm{e}^{1-t} \cdot \frac{1}{\mathrm{e}^{-\delta} \cdot (1+\delta)} \leqslant 2^2 \cdot \mathrm{e} \cdot \frac{1}{\mathrm{e}^{-\frac{1}{2}} \cdot 1} < 25 = \gamma^2. \tag{4}$$

The above inequality ensures that $\det(\mathbf{A}_S)$ is "sufficiently" small whenever $0 \notin S$, as validated in the following lemma.

▶ **Lemma 3.6** (∗). *Let $\mathbf{A}$ be the Gram matrix defined by $n+1$ vectors constructed according to Eq. (3), where $\alpha = 1$, $\beta = t$, and $\gamma = 5$. Define $\mathsf{OPT} \triangleq (1+t)^{k-1} \cdot \gamma^2 \cdot \mathrm{e}^t$. Then, for any set $S \in \binom{[0..n]}{k+1}$,*

$$\det(\mathbf{A}_S) \text{ is } \begin{cases} \text{equal to } \mathsf{OPT} & \text{if } 0 \in S \text{ and } \sum_{i \in S_{-0}} x_i = t, \\ \text{at most } \mathrm{e}^{-\delta}(1+\delta) \cdot \mathsf{OPT} & \text{otherwise.} \end{cases}$$

In particular, $\operatorname{maxdet}(\mathbf{A}, k+1)$ *is* OPT *if k-*Sum *has a solution, and is at most* $\mathrm{e}^{-\delta}(1+\delta) \cdot \mathsf{OPT} < \mathsf{OPT}$ *otherwise.*

We complete our reduction by approximating the Gram matrix $\mathbf{A}$ of $n+1$ vectors defined in Eq. (3) by a *rational* matrix $\mathbf{B}$ whose maximum determinant maintains sufficient information to solve k-Sum.

▶ **Lemma 3.7** (∗). *Let $\mathbf{B}$ be the Gram matrix in $\mathbb{Q}^{(n+1)\times(n+1)}$ defined by $n+1$ vectors $\mathbf{w}_0, \ldots, \mathbf{w}_n$ in $\mathbb{Q}^{2n}$, each entry of which is a $(1 \pm \varepsilon)$-factor approximation to the corresponding entry of $n+1$ vectors $\mathbf{v}_0, \ldots, \mathbf{v}_n$ defined by Eq. (3), where $\varepsilon = 2^{-\mathcal{O}(k \log(nk))}$. Then,*

$$\operatorname{maxdet}(\mathbf{B}, k+1) \text{ is } \begin{cases} \text{at least } \left(\frac{2}{3} + \frac{1}{3}\mathrm{e}^{-\delta}(1+\delta)\right) \cdot \mathsf{OPT} & \text{if } k\text{-Sum has a solution,} \\ \text{at most } \left(\frac{1}{3} + \frac{2}{3}\mathrm{e}^{-\delta}(1+\delta)\right) \cdot \mathsf{OPT} & \text{otherwise.} \end{cases}$$

Moreover, we can calculate $\mathbf{B}$ in polynomial time.

The crux of its proof is to approximate $\mathbf{A}$ within a factor of $\varepsilon = 2^{-\mathcal{O}(k \log(nk))}$. To this end, we use the following lemma.

▶ **Lemma 3.8** (cf. [4, page 107]). *For two complex-valued $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$, the absolute difference in the determinant of $\mathbf{A}$ and $\mathbf{B}$ is bounded from above by*

$$|\det(\mathbf{A}) - \det(\mathbf{B})| \leqslant n \cdot \max\{\|\mathbf{A}\|_2, \|\mathbf{B}\|_2\}^{n-1} \cdot \|\mathbf{A} - \mathbf{B}\|_2.$$

What remains to be done is to prove Theorem 3.1 using Lemma 3.7.

Proof of Theorem 3.1. Our parameterized reduction is as follows. Given n rational numbers $x_1, \ldots, x_n \in (0,1) \cap \mathbb{Q}$, a target rational number $t \in (0,1) \cap \mathbb{Q}$, and a positive integer $k \in [n]$ as an instance of k-Sum, we construct $n+1$ rational vectors $\mathbf{w}_0, \ldots, \mathbf{w}_n$ in $\mathbb{Q}_+^{2n}$, each of which is an entry-wise $(1 \pm \varepsilon)$-factor approximation to $\mathbf{v}_0, \ldots, \mathbf{v}_n$ defined by Eq. (3), where $\varepsilon = 2^{-\mathcal{O}(k \log(nk))}$. This construction requires polynomial time owing to Lemma 3.7. Thereafter, we compute the Gram matrix $\mathbf{B}$ in $\mathbb{Q}^{(n+1)\times(n+1)}$ defined by $\mathbf{w}_0, \ldots, \mathbf{w}_n$. Consider Determinant Maximization defined by $(\mathbf{B}, k+1)$ with parameter $k+1$. According to Lemma 3.7, the maximum principal minor $\operatorname{maxdet}(\mathbf{B}, k+1)$ is at least $(\frac{2}{3} + \frac{1}{3}\mathrm{e}^{-\delta}(1+\delta)) \cdot \mathsf{OPT}$ if and only if k-Sum has a solution. Moreover, if this is the case, the optimal solution S^* for Determinant Maximization satisfies that $\sum_{i \in S^*_{-0}} x_i = t$. The above discussion ensures the correctness of the parameterized reduction from k-Sum to Determinant Maximization, finishing the proof. ◀

3.2 Polynomial-time Algorithm for Tridiagonal Matrices

Here, we demonstrate that Determinant Maximization is polynomial-time solvable on tridiagonal matrices. Recall that a *tridiagonal matrix* is a square matrix $\mathbf{A}$ such that $A_{i,j} = 0$ whenever $|i - j| \geqslant 2$; i.e., its symmetrized graph is a path graph (and thus a linear forest). Our polynomial-time algorithm is based on dynamic programming and uses the observation that the removal of any pair of row and column from a tridiagonal matrix renders it block diagonal.

▶ **Observation 3.9** (∗). Determinant Maximization *on tridiagonal matrices can be solved in polynomial time.*

4 W[1]-hardness With Respect to Rank

We then prove the W[1]-hardness of Determinant Maximization when parameterized by the *rank* of an input matrix. In fact, we obtain the stronger hardness result on the problem of finding a set of pairwise orthogonal rational vectors, which is formally stated below.

▶ **Problem 4.1.** *Given n d-dimensional vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in $\mathbb{Q}^d$ and a positive integer $k \in [n]$, we are required to decide if there exists a set of k vectors that is pairwise orthogonal, i.e., a set $S \in \binom{[n]}{k}$ such that $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for all $i \neq j \in S$.*

▶ **Theorem 4.2.** *Problem 4.1 is* W[1]*-hard when parameterized by the dimension d of the input vectors. Moreover, the same hardness result holds even if every vector has the same Euclidean norm.*

The following is immediate from Theorem 4.2.

▶ **Corollary 4.3** (∗). Determinant Maximization *is* W[1]*-hard when parameterized by the rank of an input matrix.*

▶ Remark 4.4. We briefly describe a failed attempt to proving Theorem 4.2 using a straightforward reduction from k-Clique. Let $G = (V, E)$ be a graph on n vertices and k a parameter denoting the solution size. Consider constructing a $f(k)$-dimensional vector $\mathbf{v}_i$ for each vertex $i \in V$ such that $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ if and only if $(i, j) \in E$, which ensures that k-Clique has a solution if and only if Problem 4.1 has a solution. When G consists of a clique of size $n - 1$ and an isolated vertex o, this seems impossible: On one hand, every vector should be a nonzero vector because $\mathbf{v}_o$ and $\mathbf{v}_i$ for $i \neq o$ are nonorthogonal; on the other hand, for $n - 1$ vectors in $\{\mathbf{v}_i : i \neq o\}$ to be pairwise orthogonal, they must be (at least) $(n - 1)$-dimensional.

The key tool to bypass this difficulty is Grid Tiling introduced in the next subsection.

4.1 Grid Tiling and Pythagorean Triples

We first define Grid Tiling due to Marx [33].

▶ **Problem 4.5** (Grid Tiling due to Marx [33]). *For two integers n and k, given a collection $\mathcal{S}$ of k^2 nonempty sets $S_{i,j} \subseteq [n]^2$ called* cells *for each $i, j \in [k]$,* Grid Tiling *asks to find an assignment $\sigma : [k]^2 \to [n]^2$ with $\sigma(i, j) \in S_{i,j}$ such that*

1. *vertical neighbors agree in the first coordinate; i.e., if $\sigma(i, j) = (x, y)$ and $\sigma((i + 1) \bmod k, j) = (x', y')$, then $x = x'$, and*
2. *horizontal neighbors agree in the second coordinate; i.e., if $\sigma(i, j) = (x, y)$ and $\sigma(j, (i + 1) \bmod k) = (x', y')$, then $y = y'$,*

where we define $(k + 1) \bmod k \triangleq 1$, and hereafter omit the symbol mod *for modulo operator.*

Grid Tiling parameterized by k is proven to be W[1]-hard by Marx [33, 35]. We say that two cells (i_1, j_1) and (i_2, j_2) are *adjacent* if the Manhattan distance between them is 1. Let $\mathcal{I}$ be the set of all pairs of two adjacent cells; i.e.,

$$\mathcal{I} \triangleq \Big\{(i_1, j_1, i_2, j_2) \in [n]^4 : |i_1 - i_2| + |j_1 - j_2| = 1\Big\}. \tag{5}$$

Note that $|\mathcal{I}| = 2k^2$. Grid Tiling has the two useful properties that (i) the constraint to be satisfied is the equality on the first and second coordinates, which is pretty simple, and (ii) there are only k^2 cells and each cell is adjacent to (at most) *four* cells. By contrast, as there are $\mathcal{O}(n^2)$ candidates for the assignment of integer pairs, we need to represent the *consistency* between adjacent cells using only $f(k)$-dimensional vectors. For this purpose, we exploit rational points on the unit circle; i.e., Pythagorean triples. A *Pythagorean triple* is a triple of three positive integers (a, b, c) such that $a^2 + b^2 = c^2$; e.g., $(a, b, c) = (3, 4, 5)$. It is further said to be *primitive* if (a, b, c) are coprime; i.e., $\gcd(a, b) = \gcd(b, c) = \gcd(c, a) = 1$. We assume for a while that we have n primitive Pythagorean triples, denoted $(a_1, b_1, c_1), \ldots, (a_n, b_n, c_n)$.

4.2 Reduction from Grid Tiling and Proof of Theorem 4.2

We are now ready to describe a parameterized reduction from Grid Tiling to Problem 4.1. Given an instance $\mathcal{S} = (S_{i,j})_{i,j \in [k]}$ of Grid Tiling, we define a rational vector for each $(x, y) \in S_{i,j}$, whose dimension is bounded by some function in k. Each vector consists of $|\mathcal{I}| = 2k^2$ blocks (indexed by an element of $\mathcal{I}$), each of which is two dimensional and is either a rational point on the unit circle or the origin O. Hence, each vector is of dimension $2|\mathcal{I}| = 4k^2$. Let $\mathbf{v}^{(i,j)}_{x,y}$ denote the vector for an element $(x, y) \in S_{i,j}$ of cell $(i, j) \in [k]^2$, let $\mathbf{v}^{(i,j)}_{x,y}(i_1, j_1, i_2, j_2)$ denote the block of $\mathbf{v}^{(i,j)}_{x,y}$ corresponding to each pair of adjacent cells $(i_1, j_1, i_2, j_2) \in \mathcal{I}$. Each bock is defined as follows:

$$\mathbf{v}^{(i,j)}_{x,y}(e) \triangleq \begin{cases} \left[-{}^{b_x}\!/_{c_x}, {}^{a_x}\!/_{c_x}\right] & \text{if } e = (i-1, j, i, j), \\ \left[{}^{a_x}\!/_{c_x}, {}^{b_x}\!/_{c_x}\right] & \text{if } e = (i, j, i+1, j), \\ \left[-{}^{b_y}\!/_{c_y}, {}^{a_y}\!/_{c_y}\right] & \text{if } e = (i, j-1, i, j), \\ \left[{}^{a_y}\!/_{c_y}, {}^{b_y}\!/_{c_y}\right] & \text{if } e = (i, j, i, j+1), \\ [0, 0] & \text{otherwise.} \end{cases} \tag{6}$$

Because each vector contains exactly four points on the unit circle, its squared norm is equal to 4. We denote by $\mathbf{V}^{(i,j)}$ the set of vectors corresponding to the elements of $S_{i,j}$; i.e., $\mathbf{V}^{(i,j)} \triangleq \{\mathbf{v}^{(i,j)}_{x,y} : (x, y) \in S_{i,j}\}$. We now define an instance $(\mathbf{V}, K)$ of Problem 4.1 as $\mathbf{V} \triangleq \bigcup_{i,j \in [k]} \mathbf{V}^{(i,j)}$ and $K \triangleq k^2$. Note that $\mathbf{V}$ consists of $N \triangleq \sum_{i,j \in [k]} |S_{i,j}|$ vectors. We prove that the existence of a set of pairwise orthogonal k^2 vectors yields the answer of Grid Tiling. The key property of the above construction is that $[-{}^{b_x}\!/_{c_x}, {}^{a_x}\!/_{c_x}]$ and $[{}^{a_{x'}}\!/_{c_{x'}}, {}^{b_{x'}}\!/_{c_{x'}}]$ are orthogonal if and only if $x = x'$.

▶ **Lemma 4.6** (∗)**.** *Let $\mathbf{V}$ be the set of vectors constructed from an instance $\mathcal{S} = (S_{i,j})_{i,j \in [k]}$ of* Grid Tiling *according to Eq.* (6)*. Then,* Grid Tiling *has a solution if and only if Problem 4.1 has a solution.*

Proof of Theorem 4.2. Our parameterized reduction is as follows. Given an instance $\mathcal{S} = (S_{i,j})_{i,j \in [k]}$ of Grid Tiling, we first generate n primitive Pythagorean triples (a_1, b_1, c_1), $\ldots, (a_n, b_n, c_n)$. This can be done in $\mathcal{O}(n^2)$ time, e.g., using Fibonacci's method [18]. We then construct a set $\mathbf{V}$ of N $4k^2$-dimensional rational vectors from $\mathcal{S}$ according to Eq. (6)

in polynomial time, where $N \triangleq \sum_{i,j\in[k]} |S_{i,j}|$. According to Lemma 4.6, $\mathcal{S}$ has a solution of Grid Tiling if and only if there exists a set of k^2 pairwise orthogonal vectors in $\mathbf{V}$. Since Grid Tiling is W[1]-hard with respect to k, Problem 4.1 is also W[1]-hard when parameterized by dimension $d(= 4k^2)$. Note that every vector is of squared norm 4, completing the proof. ◀

4.3 Problem 4.1 on Nonnegative Vectors is FPT

We note that Problem 4.1 is FPT with respect to the dimension if the input vectors are *nonnegative*. Briefly speaking, Problem 4.1 on nonnegative vectors is equivalent to Set Packing parameterized by the size of the universe, which is easily shown to be FPT.

▶ **Observation 4.7** (∗). *Problem 4.1 is FPT with respect to the dimension if every input vector is entry-wise nonnegative.*

5 W[1]-hardness of Approximation

Our final result is FPT-inapproximability of Determinant Maximization as stated below.

▶ **Theorem 5.1.** *Under the Parameterized Inapproximability Hypothesis, it is* W[1]*-hard to approximate* Determinant Maximization *within a factor of* $2^{-c\sqrt{k}}$ *for some universal constant* $c > 0$ *when parameterized by the number* k *of vectors to be selected. Moreover, the same hardness result holds even if the diagonal entries of an input matrix are restricted to* 1.

Since the above result relies on the Parameterized Inapproximability Hypothesis, Section 5.1 begins with its formal definition.

5.1 Inapproximability of Grid Tiling under Parameterized Inapproximability Hypothesis

We first introduce Binary Constraint Satisfaction Problem, for which the Parameterized Inapproximability Hypothesis asserts FPT-inapproximability. For two integers n and k, we are given a set $V \triangleq [k]$ of k variables, an alphabet $\Sigma \triangleq [n]$ of size n, and a set of constraints $\mathcal{C} = (C_{i,j})_{i,j\in V}$ such that $C_{i,j} \subseteq \Sigma^2$.[2] Each variable $i \in V$ may take a value from Σ. Each constraint $C_{i,j}$ specifies the pairs of values that variables i and j can take simultaneously, and it is said to be *satisfied* by an assignment $\psi : V \to \Sigma$ of values to the variables if $(\psi(i), \psi(j)) \in C_{i,j}$.

▶ **Problem 5.2.** *Given a set* V *of* k *variables, an alphabet set* Σ *of size* n*, and a set of constraints* $\mathcal{C} = (C_{i,j})_{i,j\in V}$*,* Binary Constraint Satisfaction Problem *(*BCSP*) asks to find an assignment* $\psi : V \to \Sigma$ *that satisfies the maximum fraction of constraints.*

It is well known that BCSP parameterized by the number k of variables is W[1]-complete from a standard parameterized reduction from k-Clique. Lokshtanov et al. [30] posed a conjecture asserting that a constant-factor gap version of BCSP is also W[1]-hard.

▶ **Hypothesis 5.3** (Parameterized Inapproximability Hypothesis (PIH) [30]). *There exists some universal constant* $\varepsilon \in (0, 1)$ *such that it is* W[1]*-hard to distinguish between* BCSP *instances that are promised to either be satisfiable, or have a property that every assignment violates at least* ε*-fraction of the constraints.*

---

[2] Though each constraint is actually indexed by an unordered pair of variables $\{i, j\}$, we use the present notation $C_{i,j}$ for sake of clarity and assume that $C_{i,j} = C_{j,i}$ without loss of generality.

Here, we prove that an *optimization version* of Grid Tiling is FPT-inapproximable assuming PIH. Given an instance $\mathcal{S} = (S_{i,j})_{i,j\in[k]}$ of Grid Tiling and an assignment $\sigma : [k]^2 \to [n]^2$, $\sigma(i,j)$ and $\sigma(i',j')$ for a pair of adjacent cells $(i,j,i',j') \in \mathcal{I}$ are said to be *consistent* if they agree on the first coordinate when $j = j'$ or on the second coordinate when $i = i'$, and *inconsistent* otherwise. The *consistency* of σ, denoted $\text{cons}(\sigma)$, is defined as the number of pairs of adjacent cells that are consistent; namely,

$$\text{cons}(\sigma) \triangleq \sum_{(i_1,j_1,i_2,j_2)\in\mathcal{I}} \left[\!\!\left[\sigma(i_1,j_1) \text{ and } \sigma(i_2,j_2) \text{ are consistent}\right]\!\!\right].$$

The *inconsistency* of σ is defined as the number of inconsistent pairs of adjacent cells. The optimization version of Grid Tiling asks to find an assignment σ such that $\text{cons}(\sigma)$ is maximized.[3] Note that the maximum possible consistency is $|\mathcal{I}| = 2k^2$. We will use $\text{opt}(\mathcal{S})$ to denote the optimal consistency among all possible assignments. We now demonstrate that Grid Tiling is FPT-inapproximable in an *additive sense* under PIH, whose proof is reminiscent of [33].

▶ **Lemma 5.4** (∗). *Under PIH, there exists some universal constant $\delta \in (0,1)$ such that it is* W[1]*-hard to distinguish* Grid Tiling *instances between the following cases:*

- *Completeness: the optimal consistency is $2k^2$.*
- *Soundness: the optimal consistency is at most $2k^2 - \delta k$.*

It should be noted that we may not be able to significantly improve the additive term $\mathcal{O}(k)$ owing to a polynomial-time εk^2-additive approximation algorithm for any constant $\varepsilon > 0$:

▶ **Observation 5.5** (∗). *Given an instance of* Grid Tiling *and an error tolerance parameter $\varepsilon > 0$, we can find an assignment whose consistency is at least $\text{opt}(\mathcal{S}) - \varepsilon k^2$ in $\varepsilon^2 k^2 n^{\mathcal{O}(1/\varepsilon^2)}$ time.*

Our technical result is a gap-preserving parameterized reduction from Grid Tiling to Determinant Maximization, whose proof is presented in the subsequent subsection.

▶ **Lemma 5.6.** *There is a polynomial-time, parameterized reduction from an instance $\mathcal{S} = (S_{i,j})_{i,j\in[k]}$ of* Grid Tiling *to an instance $(\mathbf{A}, k^2)$ of* Determinant Maximization *such that all diagonal entries of $\mathbf{A}$ are 1 and the following conditions are satisfied:*

- *Completeness: If $\text{opt}(\mathcal{S}) = 2k^2$, then $\text{maxdet}(\mathbf{A}, k^2) = 1$.*
- *Soundness: If $\text{opt}(\mathcal{S}) \leqslant 2k^2 - \delta k$ for some $\delta > 0$, then $\text{maxdet}(\mathbf{A}, k^2) \leqslant 0.999^{\delta k}$.*

Using Lemma 5.6, we can prove Theorem 5.1.

Proof of Theorem 5.1. Our gap-preserving parameterized reduction is as follows. Given an instance $\mathcal{S} = (S_{i,j})_{i,j\in[k]}$ of Grid Tiling, we construct an instance $(\mathbf{A} \in \mathbb{Q}^{N\times N}, K \triangleq k^2)$ of Determinant Maximization in polynomial time according to Lemma 5.6, where $N \triangleq \sum_{i,j\in[k]} |S_{i,j}|$. The diagonal entries of $\mathbf{A}$ are 1 by definition. Since K is a function only in k, this is a parameterized reduction. According to Lemmas 5.4 and 5.6, it is W[1]-hard to determine whether $\text{maxdet}(\mathbf{V}, K) = 1$ or $\text{maxdet}(\mathbf{V}, K) \leqslant 0.999^{\delta k}$ under PIH, where $\delta \in (0,1)$ is a constant appearing Lemma 5.4. In particular, Determinant Maximization is W[1]-hard to approximate within a factor better than $0.999^{\delta k} = 2^{-c\sqrt{K}}$ when parameterized by K, where $c \in (0,1)$ is some universal constant. This completes the proof. ◀

[3] Our definition is different from Marx [33] in that the latter seeks a partial assignment such that the number of defined cells is maximized while the former requires maximizing the number of consistent adjacent pairs.

5.2 Gap-preserving Reduction from Grid Tiling and Proof of Lemma 5.6

To prove Lemma 5.6, we describe a gap-preserving parameterized reduction from Grid Tiling to Determinant Maximization. Before going into its details, we introduce a convenient gadget due to Çivril and Magdon-Ismail [11].

▶ **Lemma 5.7** (Çivril and Magdon-Ismail [11, Lemma 13]). *For any positive even integer ℓ, we can construct a set of 2^ℓ rational vectors $\mathbf{B}^{(\ell)} = \{\mathbf{b}_1, \dots, \mathbf{b}_{2^\ell}\}$ of dimension $2^{\ell+1}$ in $\mathcal{O}(4^\ell)$ time such that the following conditions are satisfied:*

- *Each entry of vectors is either 0 or $2^{-\frac{\ell}{2}}$; $\|\mathbf{b}_i\| = 1$ for all $i \in [2^\ell]$.*
- *$\langle \mathbf{b}_i, \mathbf{b}_j \rangle = \frac{1}{2}$ for all $i, j \in [2^\ell]$ with $i \neq j$.*
- *$\langle \mathbf{b}_i, \overline{\mathbf{b}_j} \rangle = \frac{1}{2}$ for all $i, j \in [2^\ell]$ with $i \neq j$, where $\overline{\mathbf{b}_j} \triangleq 2^{-\frac{\ell}{2}} \cdot \mathbf{1} - \mathbf{b}_j$.*

By definition of $\mathbf{B}^{(\ell)}$, we further have the following:

$$\langle \mathbf{b}_i, \overline{\mathbf{b}_i} \rangle = 2^{-\frac{\ell}{2}} \langle \mathbf{1}, \mathbf{b}_i \rangle - \langle \mathbf{b}_i, \mathbf{b}_i \rangle = 0,$$
$$\langle \overline{\mathbf{b}_i}, \overline{\mathbf{b}_j} \rangle = \langle 2^{-\frac{\ell}{2}} \mathbf{1} - \mathbf{b}_i, 2^{-\frac{\ell}{2}} \mathbf{1} - \mathbf{b}_j \rangle = 2^{-\ell} \langle \mathbf{1}, \mathbf{1} \rangle - 2^{-\frac{\ell}{2}} \langle \mathbf{1}, \mathbf{b}_i + \mathbf{b}_j \rangle + \langle \mathbf{b}_i, \mathbf{b}_j \rangle = \langle \mathbf{b}_i, \mathbf{b}_j \rangle.$$

Our reduction strategy is very similar to that of Theorem 4.2. Given an instance $\mathcal{S} = (S_{i,j})_{i,j \in [k]}$ of Grid Tiling, we construct a rational vector $\mathbf{v}_{x,y}^{(i,j)}$ for each element $(x, y) \in S_{i,j}$ of cell $(i, j) \in [k]^2$. Each vector consists of $|\mathcal{I}| = 2k^2$ blocks indexed by $\mathcal{I}$, each of which is either a vector in the set $\mathbf{B}^{(2\lceil \log n \rceil)}$ or the zero vector $\mathbf{0}$. Hence, the dimension of the vectors is $2k^2 \cdot 2^{2\lceil \log n \rceil + 1} = \mathcal{O}(k^2 n^2)$. Let $\mathbf{v}_{x,y}^{(i,j)}(i_1, j_1, i_2, j_2)$ denote the block of $\mathbf{v}_{x,y}^{(i,j)}$ corresponding to a pair of adjacent cells $(i_1, j_1, i_2, j_2) \in \mathcal{I}$. Each block is subsequently defined as follows:

$$\mathbf{v}_{x,y}^{(i,j)}(e) \triangleq \begin{cases} \overline{\mathbf{b}_x} & \text{if } e = (i-1, j, i, j), \\ \mathbf{b}_x & \text{if } e = (i, j, i+1, j), \\ \overline{\mathbf{b}_y} & \text{if } e = (i, j-1, i, j), \\ \mathbf{b}_y & \text{if } e = (i, j, i, j+1), \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

Hereafter, two vectors $\mathbf{v}_{x,y}^{(i,j)}$ and $\mathbf{v}_{x',y'}^{(i',j')}$ are said to be *adjacent* if (i, j) and (i', j') are adjacent, and two adjacent vectors are said to be *consistent* if (x, y) and (x', y') are consistent (i.e., $x = x'$ whenever $j = j'$ and $y = y'$ whenever $i = i'$) and *inconsistent* otherwise. Since each vector contains exactly four vectors chosen from $\mathbf{B}^{(2\lceil \log n \rceil)}$, its squared norm is equal to 4. In addition, $\mathbf{v}_{x,y}^{(i,j)}$ and $\mathbf{v}_{x',y'}^{(i',j')}$ are orthogonal whenever (i, j) and (i', j') are not identical or adjacent. Observe further that if two cells are adjacent, the inner product of two vectors in $\mathbf{V}$ is calculated as follows:

$$\left\langle \mathbf{v}_{x,y}^{(i,j)}, \mathbf{v}_{x',y'}^{(i+1,j)} \right\rangle = \langle \mathbf{b}_x, \overline{\mathbf{b}_{x'}} \rangle = \begin{cases} 0 & \text{if they are consistent } x = x', \\ \frac{1}{2} & \text{otherwise (i.e., } x \neq x'), \end{cases} \tag{7}$$

$$\left\langle \mathbf{v}_{x,y}^{(i,j)}, \mathbf{v}_{x',y'}^{(i,j+1)} \right\rangle = \langle \mathbf{b}_y, \overline{\mathbf{b}_{y'}} \rangle = \begin{cases} 0 & \text{if they are consistent } y = y' \\ \frac{1}{2} & \text{otherwise (i.e., } y \neq y'). \end{cases} \tag{8}$$

On the other hand, the inner product of two vectors in the same cell is as follows:

$$\left\langle \mathbf{v}_{x,y}^{(i,j)}, \mathbf{v}_{x',y'}^{(i,j)} \right\rangle = 2 \cdot \langle \mathbf{b}_x, \mathbf{b}_{x'} \rangle + 2 \cdot \langle \mathbf{b}_y, \mathbf{b}_{y'} \rangle = \begin{cases} 4 & \text{if } x = x' \text{ and } y = y', \\ 3 & \text{if } x = x' \text{ xor } y = y', \\ 2 & \text{if } x \neq x' \text{ and } y \neq y'. \end{cases} \tag{9}$$

We denote by $\mathbf{V}^{(i,j)}$ the set of vectors corresponding to the elements of $S_{i,j}$; i.e., $\mathbf{V}^{(i,j)} \triangleq \{\mathbf{v}_{x,y}^{(i,j)} : (x,y) \in S_{i,j}\}$ for each $i,j \in [k]$. We now define an instance $(\mathbf{V}, K)$ of Determinant Maximization as $\mathbf{V} \triangleq \bigcup_{i,j\in[k]} \mathbf{V}^{(i,j)}$ and $K \triangleq k^2$. Note that $\mathbf{V}$ contains $N \triangleq \sum_{i,j\in[k]} |S_{i,j}|$ vectors.

We now proceed to the proof of (the soundness argument of) Lemma 5.6. Let $\mathbf{S}$ be a set of k^2 vectors from $\mathbf{V}$. Define $\mathbf{S}^{(i,j)} \triangleq \mathbf{V}^{(i,j)} \cap \mathbf{S} = \{\mathbf{v}_{x,y}^{(i,j)} \in \mathbf{S} : (x,y) \in S_{i,j}\}$ for each $i,j \in [k]^2$. Denote by $\mathrm{cov}(\mathbf{S})$ the number of cells $(i,j) \in [k]^2$ such that $\mathbf{S}$ includes $\mathbf{v}_{x,y}^{(i,j)}$ for some (x,y); i.e., $\mathrm{cov}(\mathbf{S}) \triangleq \{(i,j) \in [k]^2 : \mathbf{S}^{(i,j)} \neq \emptyset\}$, and we also define $\mathrm{dup}(\mathbf{S}) \triangleq \{(i,j) \in [k]^2 : \mathbf{S}^{(i,j)} = \emptyset\}$. It follows from the definition that $\mathrm{cov}(\mathbf{S}) + \mathrm{dup}(\mathbf{S}) = k^2$ and $\mathrm{dup}(\mathbf{S})$ counts the total number of "duplicate" vectors in the same cell. We first present an upper bound on the volume of $\mathbf{S}$ in terms of $\mathrm{dup}(\mathbf{S})$, implying that we cannot select many duplicate vectors from the same cell.

▶ **Lemma 5.8** ($*$)**.** *If* $\mathrm{dup}(\mathbf{S}) \leqslant \frac{k^2}{2}$*, then it holds that*

$$\mathrm{vol}^2(\mathbf{S}) \leqslant 4^{k^2} \cdot \left(\frac{3}{4}\right)^{\mathrm{dup}(\mathbf{S})}.$$

We then present another upper bound on the volume of $\mathbf{S}$ in terms of the *inconsistency* of a partial solution of Grid Tiling constructed from the selected vectors. For a set $\mathbf{S}$ of k^2 vectors from $\mathbf{V}$, a *partial* assignment $\sigma_{\mathbf{S}} : [k]^2 \to [n]^2 \cup \{\bigstar\}$ for Grid Tiling is defined as

$$\sigma_{\mathbf{S}}(i,j) \triangleq \begin{cases} \text{any } (x,y) \text{ such that } \mathbf{v}_{x,y}^{(i,j)} \in \mathbf{S}^{(i,j)} & \text{if such } (x,y) \text{ exists,} \\ \bigstar & \text{otherwise (i.e., } \mathbf{S}^{(i,j)} = \emptyset), \end{cases}$$

where the symbol "$\bigstar$" means *undefined* and the choice of (x,y) is arbitrary. The inconsistency of a partial assignment $\sigma_{\mathbf{S}}$ is defined as

$$\sum_{(i_1,j_1,i_2,j_2)\in\mathcal{I}} \left[\!\left[\sigma(i_1,j_1) \neq \bigstar;\ \sigma(i_2,j_2) \neq \bigstar;\ \sigma(i_1,j_1) \text{ and } \sigma(i_2,j_2) \text{ are inconsistent} \right]\!\right].$$

Note that the sum of the consistency and inconsistency of $\sigma_{\mathbf{S}}$ is no longer necessarily $2k^2$. Using $\sigma_{\mathbf{S}}$, we define a partition $(\mathbf{P}, \mathbf{Q})$ of $\mathbf{S}$ as $\mathbf{P} \triangleq \{\mathbf{v}_{x,y}^{(i,j)} \in \mathbf{S} : i,j \in [k], \sigma_{\mathbf{S}}(i,j) = (x,y)\}$ and $\mathbf{Q} \triangleq \mathbf{S} \setminus \mathbf{P}$. We further prepare an arbitrary *ordering* $\prec$ over $[k]^2$; e.g., $(i,j) \prec (i',j')$ if $i < i'$, or $i = i'$ and $j < j'$. We abuse the notation by writing $\mathbf{v}_{x,y}^{(i,j)} \prec \mathbf{v}_{x',y'}^{(i',j')}$ for any two vectors of $\mathbf{V}$ whenever $(i,j) \prec (i',j')$. Define now $\mathbf{P}_{\prec \mathbf{v}} \triangleq \{\mathbf{u} \in \mathbf{P} : \mathbf{u} \prec \mathbf{v}\}$. The following lemma states that the squared volume of k^2 vectors exponentially decays in the minimum possible inconsistency among all assignments of $\mathcal{S}$.

▶ **Lemma 5.9** ($*$)**.** *Suppose* $\mathrm{opt}(\mathcal{S}) \leqslant 2k^2 - \delta k$ *for some* $\delta > 0$ *and* $\mathrm{cov}(\mathbf{S}) \geqslant k^2 - \gamma k$ *for some* $\gamma > 0$*. If* $\delta k - 4\gamma k$ *is positive, then it holds that*

$$\mathrm{vol}^2(\mathbf{S}) \leqslant 4^{k^2} \cdot \left(\frac{63}{64}\right)^{\frac{\delta - 4\gamma}{4} k}.$$

The proof of Lemma 5.9 involves the following claim.

▷ Claim 5.10 ($*$). Suppose the same conditions as in Lemma 5.9 are satisfied. Then, the inconsistency of $\sigma_{\mathbf{S}}$ is at least $\delta k - 4\gamma k$. Moreover, the number of vectors $\mathbf{v}$ in $\mathbf{P}$ such that $\mathbf{v}$ is inconsistent with some adjacent vector of $\mathbf{P}_{\prec \mathbf{v}}$ is at least $\frac{\delta k - 4\gamma k}{4}$.

Using Lemmas 5.8 and 5.9, we can easily conclude Lemma 5.6 as follows.

Proof of Lemma 5.6. Observe that the reduction described in Section 5.2 is a parameterized reduction as it requires polynomial time and an instance $\mathcal{S} = (S_{i,j})_{i,j\in[k]}$ of Grid Tiling is transformed into an instance $(\mathbf{V}, k^2)$ of Determinant Maximization. In addition, the construction of $\mathbf{B}^{(2\lceil \log n \rceil)}$ completes in time $\mathcal{O}(4^{2\lceil \log n \rceil}) = \mathcal{O}(n^4)$ by Lemma 5.7.

We now prove the correctness of the reduction. Let us begin with the completeness argument. Suppose $\mathrm{opt}(\mathcal{S}) = 2k^2$; i.e., there is an assignment σ of consistency $2k^2$. Then, k^2 vectors in the set $\mathbf{S} \triangleq \{\mathbf{v}^{(i,j)}_{\sigma(i,j)} : i,j \in [k]\}$ are orthogonal to each other, implying that $\mathrm{vol}^2(\mathbf{S}) = 4^{k^2}$. On the other hand, because every vector of $\mathbf{V}$ is of squared norm 4, the maximum possible squared volume among k^2 vectors in $\mathbf{V}$ is 4^{k^2}; namely, $\mathrm{maxdet}(\mathbf{V}, k^2) = 4^{k^2}$.

We then prove the soundness argument. Suppose $\mathrm{opt}(\mathcal{S}) \leqslant 2k^2 - \delta k$ for some constant $\delta > 0$. Then, for any set $\mathcal{S}$ of k^2 vectors from $\mathbf{V}$ such that $\mathrm{dup}(\mathbf{S}) > \frac{\log 0.999^{-1}}{\log(\frac{3}{4})^{-1}} \cdot \delta k$, we have that by Lemma 5.8, $\mathrm{vol}^2(\mathbf{S}) < 4^{k^2} \cdot 0.999^{\delta k}$. It is thus sufficient to consider the case that

$$\mathrm{dup}(\mathbf{S}) \leqslant \frac{\log 0.999^{-1}}{\log(\frac{3}{4})^{-1}} \cdot \delta k \approx 0.0035 \cdot \delta.$$

In particular, it suffices to assume that $\mathrm{dup}(\mathbf{S}) \leqslant \gamma k$ for some $\gamma \in (0, \frac{\delta}{4})$. Simple calculation using Lemmas 5.8 and 5.9 derives that

$$\begin{aligned}\mathrm{vol}^2(\mathbf{S}) &\leqslant \min\left\{4^{k^2} \cdot \left(\frac{3}{4}\right)^{\gamma k}, 4^{k^2} \cdot \left(\frac{63}{64}\right)^{\frac{\delta-4\gamma}{4}k}\right\} \leqslant 4^{k^2} \cdot \min\left\{\left(\frac{63}{64}\right)^{\gamma k}, \left(\frac{63}{64}\right)^{\frac{\delta-4\gamma}{4}k}\right\} \\ &\leqslant 4^{k^2} \cdot \left(\frac{63}{64}\right)^{\overbrace{\min\left\{\gamma, \frac{\delta-4\gamma}{4}\right\}}^{\heartsuit} \cdot k} \leqslant 4^{k^2} \cdot \left(\frac{63}{64}\right)^{\frac{\delta}{8}k} \leqslant 4^{k^2} \cdot 0.999^{\delta k},\end{aligned}$$

where the second-to-last inequality is due to the fact that $\heartsuit$ is maximized when $\gamma = \frac{\delta - 4\gamma}{4}$; i.e., $\gamma = \frac{\delta}{8} > 0$.

Because the diagonal entries of the Gram matrix $\mathbf{A}$ defined by the vectors of $\mathbf{V}$ are 4, we can construct another instance of Determinant Maximization as $(\widetilde{\mathbf{A}}, k^2)$, where $\widetilde{\mathbf{A}} \triangleq \frac{1}{4}\mathbf{A}$. Observe finally that the diagonal entries of $\widetilde{\mathbf{A}}$ are 1 and $\det(\widetilde{\mathbf{A}}_S) = 4^{-|S|} \cdot \det(\mathbf{A}_S)$ for any S, which completes the proof. ◀

5.3 ε-Additive FPT-Approximation Parameterized by Rank

Here, we develop an ε-additive FPT-approximation algorithm parameterized by the *rank* of an input matrix $\mathbf{A}$, provided that $\mathbf{A}$ is the Gram matrix of vectors of *infinity norm at most* 1. Our algorithm complements Lemma 5.6 in a sense that we can solve the promise problem in FPT time with respect to $\mathrm{rank}(\mathbf{A})$. The proof uses the standard rounding technique.

▶ **Observation 5.11** ($*$). *Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be n d-dimensional vectors in $\mathbb{Q}^d$ such that $\|\mathbf{v}_i\|_\infty \leqslant 1$ for all $i \in [n]$, $\mathbf{A}$ the Gram matrix defined by the vectors, $k \in [d]$ a positive integer, and $\varepsilon > 0$ an error tolerance parameter. Then, we can compute an approximate solution $S \in \binom{[n]}{k}$ to* Determinant Maximization *in $\varepsilon^{-d^2} \cdot d^{\mathcal{O}(d^3)} \cdot n^{\mathcal{O}(1)}$ time such that $\det(\mathbf{A}_S) \geqslant \mathrm{maxdet}(\mathbf{A}, k) - \varepsilon$.*

6 Open Problems

We investigated the W[1]-hardness of Determinant Maximization in the three restricted cases, improving upon the result due to Koutis [26]. Our parameterized hardness results leave a few natural open problems: For what kinds of sparse matrices is Determinant Maximization FPT? Is there a $(1-\varepsilon)$-factor (rather than "additive") FPT-approximation algorithm with respect to the matrix rank? Quantitative lower bounds can be also proved; e.g., due to the lower bound of k-Sum [41], Determinant Maximization on tridiagonal matrices cannot be solved in $n^{o(k)}$ time, unless *Exponential Time Hypothesis* [23, 24] fails.

References

1 Amir Abboud, Kevin Lewi, and Ryan Williams. Losing weight by gaining edges. In *ESA*, pages 1–12, 2014.
2 Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998.
3 Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998.
4 Rajendra Bhatia. *Perturbation Bounds for Matrix Eigenvalues*. SIAM, 2007.
5 Erdem Bıyık, Kenneth Wang, Nima Anari, and Dorsa Sadigh. Batch active learning using determinantal point processes. *CoRR*, abs/1906.07975, 2019. arXiv:1906.07975.
6 Alexei Borodin and Eric M. Rains. Eynard-Mehta theorem, Schur process, and their Pfaffian analogs. *J. Stat. Phys.*, 121(3–4):291–317, 2005.
7 L. Elisa Celis, Vijay Keswani, Damian Straszak, Amit Deshpande, Tarun Kathuria, and Nisheeth K. Vishnoi. Fair and diverse DPP-based data summarization. In *ICML*, pages 715–724, 2018.
8 Wei-Lun Chao, Boqing Gong, Kristen Grauman, and Fei Sha. Large-margin determinantal point processes. In *UAI*, pages 191–200, 2015.
9 Diego Cifuentes and Pablo A. Parrilo. An efficient tree decomposition method for permanents and mixed discriminants. *Linear Algebra Appl.*, 493:45–81, 2016.
10 Ali Çivril and Malik Magdon-Ismail. On selecting a maximum volume sub-matrix of a matrix and related problems. *Theor. Comput. Sci.*, 410(47-49):4801–4811, 2009.
11 Ali Çivril and Malik Magdon-Ismail. Exponential inapproximability of selecting a maximum volume sub-matrix. *Algorithmica*, 65(1):159–176, 2013.
12 Bruno Courcelle, Johann A. Makowsky, and Udi Rotics. On the fixed parameter complexity of graph enumeration problems definable in monadic second-order logic. *Discrete Appl. Math.*, 108(1-2):23–52, 2001.
13 Marek Cygan, Fedor V. Fomin, Łukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michał Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015.
14 Marco Di Summa, Friedrich Eisenbrand, Yuri Faenza, and Carsten Moldenhauer. On largest volume simplices and sub-determinants. In *SODA*, pages 315–323, 2014.
15 Rod G. Downey and Michael R. Fellows. Fixed-parameter tractability and completeness II: On completeness for $W[1]$. *Theor. Comput. Sci.*, 141(1-2):109–131, 1995.
16 Rodney G. Downey and Michael R. Fellows. *Parameterized Complexity*. Springer, 2012.
17 Andreas Emil Feldmann, C. S. Karthik, Euiwoong Lee, and Pasin Manurangsi. A survey on approximation in parameterized complexity: Hardness and algorithms. *Algorithms*, 13(6):146, 2020.
18 Leonardo Pisano Fibonacci. The book of squares, 1225.
19 Jörg Flum and Martin Grohe. *Parameterized Complexity Theory*. Springer, 2006.
20 Fedor V. Fomin and Dieter Kratsch. *Exact Exponential Algorithms*. An EATCS Series: Texts in Theoretical Computer Science. Springer, 2010.

21 Peter Gritzmann, Victor Klee, and David G. Larman. Largest j-simplices in n-polytopes. *Discrete Comput. Geom.*, 13:477–515, 1995.

22 Sariel Har-Peled and Soham Mazumdar. On coresets for k-means and k-median clustering. In *STOC*, pages 291–300, 2004.

23 Russell Impagliazzo and Ramamohan Paturi. On the complexity of k-SAT. *J. Comput. Syst. Sci.*, 62(2):367–375, 2001.

24 Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *J. Comput. Syst. Sci.*, 63(4):512–530, 2001.

25 Chun-Wa Ko, Jon Lee, and Maurice Queyranne. An exact algorithm for maximum entropy sampling. *Oper. Res.*, 43(4):684–691, 1995.

26 Ioannis Koutis. Parameterized complexity and improved inapproximability for computing the largest j-simplex in a V-polytope. *Inf. Process. Lett.*, 100(1):8–13, 2006.

27 Alex Kulesza and Ben Taskar. k-DPPs: Fixed-size determinantal point processes. In *ICML*, pages 1193–1200, 2011.

28 Alex Kulesza and Ben Taskar. Determinantal point processes for machine learning. *Found. Trends Mach. Learn.*, 5(2–3):123–286, 2012.

29 Donghoon Lee, Geonho Cha, Ming-Hsuan Yang, and Songhwai Oh. Individualness and determinantal point processes for pedestrian detection. In *ECCV*, pages 330–346, 2016.

30 Daniel Lokshtanov, M. S. Ramanujan, Saket Saurab, and Meirav Zehavi. Parameterized complexity and approximability of directed odd cycle transversal. In *SODA*, pages 2181–2200, 2020.

31 Odile Macchi. The coincidence approach to stochastic point processes. *Adv. Appl. Probab.*, 7(1):83–122, 1975.

32 Vivek Madan, Aleksandar Nikolov, Mohit Singh, and Uthaipon Tantipongpipat. Maximizing determinants under matroid constraints. In *FOCS*, pages 565–576, 2020.

33 Dániel Marx. On the optimality of planar and geometric approximation schemes. In *FOCS*, pages 338–348, 2007.

34 Dániel Marx. Parameterized complexity and approximation algorithms. *Comput. J.*, 51(1):60–78, 2008.

35 Dániel Marx. A tight lower bound for planar multiway cut with fixed number of terminals. In *ICALP*, pages 677–688, 2012.

36 Aleksandar Nikolov. Randomized rounding for the largest simplex problem. In *STOC*, pages 861–870, 2015.

37 Aleksandar Nikolov and Mohit Singh. Maximizing determinants under partition constraints. In *STOC*, pages 192–201, 2016.

38 Naoto Ohsaka. On the parameterized intractability of determinant maximization. *CoRR*, abs/2209.12519, 2022. `arXiv:2209.12519`.

39 Naoto Ohsaka. Some inapproximability results of MAP inference and exponentiated determinantal point processes. *J. Artif. Intell. Res.*, 73:709–735, 2022.

40 Naoto Ohsaka and Tatsuya Matsuoka. On the (in)tractability of computing normalizing constants for the product of determinantal point processes. In *ICML*, pages 7414–7423, 2020.

41 Mihai Pătraşcu and Ryan Williams. On the possibility of faster SAT algorithms. In *SODA*, pages 1065–1075, 2010.

42 Mark Wilhelm, Ajith Ramanathan, Alexander Bonomo, Sagar Jain, Ed H. Chi, and Jennifer Gillenwater. Practical diversified recommendations on YouTube with determinantal point processes. In *CIKM*, pages 2165–2173, 2018.

43 Jin-ge Yao, Feifan Fan, Wayne Xin Zhao, Xiaojun Wan, Edward Y. Chang, and Jianguo Xiao. Tweet timeline generation with determinantal point processes. In *AAAI*, pages 3080–3086, 2016.

44 Martin J. Zhang and Zhijian Ou. Block-wise MAP inference for determinantal point processes with application to change-point detection. In *SSP*, pages 1–5, 2016.