Quasipolynomial-Time Deterministic Kernelization and (Gammoid) Representation

Rohit Gurjar
Indian Institute of Technology Bombay, India

Daniel Lokshtanov
University of California, Santa Barbara, CA, USA

Pranabendu Misra
Chennai Mathematical Institute, India

Fahad Panolan
School of Computing, University of Leeds, UK

Saket Saurabh
Institute of Mathematical Sciences, HBNI, Chennai, India

Meirav Zehavi
Ben-Gurion University of the Negev, Beersheba, Israel

Abstract

In this paper, we suggest to extend the notion of a kernel to *permit the kernelization algorithm to be executed in quasi-polynomial time rather than polynomial time*. So far, we are only aware of one work that addressed this negatively, showing that some lower bounds on kernel sizes proved for kernelization also hold when quasi-polynomial time complexity is allowed. When we, anyway, deal with an NP-hard problem, sacrificing polynomial time in preprocessing for quasi-polynomial time may often not be a big deal, but, of course, the question is – does it give us more power? The only known work, mentioned above, seems to suggest that the answer is "no". In this paper, we show that this is not the case – in particular, we show that this notion is extremely powerful for derandomization. Some of the most basic kernelization algorithms in the field are based on inherently randomized tools whose derandomization is a huge problem that has remained (and may still remain) open for many decades. Still, some breakthrough advances for derandomization in quasi-polynomial time have been made. Can we harness these advancements to design quasi-polynomial deterministic kernelization algorithms for basic problems in the field? To this end, we revisit the question of deterministic polynomial-time computation of a linear representation of transversal matroids and gammoids, which is a longstanding open problem. We present a deterministic computation of a representation matrix of a transversal matroid in time quasipolynomial in the rank of the matroid, where each entry of the matrix can be represented in quasipolynomial (in the rank of the matroid) bits. As a corollary, we obtain a linear representation of a gammoid in deterministic quasipolynomial time and quasipolynomial bits in the size of the underlying ground set of the gammoid. In turn, as applications of our results, we present deterministic quasi-polynomial time kernels of polynomial size for several central problems in the field.

2012 ACM Subject Classification Mathematics of computing → Matroids and greedoids; Theory of computation → Pseudorandomness and derandomization; Mathematics of computing → Graph algorithms

Keywords and phrases Network Flows, Gammoids, Matchings, Transversal Matroids, Matroid Representation, Derandomization

Digital Object Identifier 10.4230/LIPIcs.MFCS.2025.54

50th International Symposium on Mathematical Foundations of Computer Science (MFCS 2025).
Editors: Paweł Gawrychowski, Filip Mazowiecki, and Michał Skrzypczak; Article No. 54; pp. 54:1–54:17
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

Preprocessing (or *data reduction*) is an integral part of almost any application: both systematic and intuitive approaches to tackle difficult problems often involve it. Even in our everyday lives, we often rely on preprocessing, sometimes without even noticing it. A natural question in this regard is how to measure the quality of preprocessing rules proposed for a specific problem, yet for a long time the mathematical analysis of polynomial time preprocessing algorithms was neglected. One central reason for this anomaly stems from the following observation: showing that in polynomial time an instance I of an NP-hard problem can be replaced by an equivalent instance whose size is *necessarily* smaller than the size of I, even by a single bit, implies that P=NP. The situation has changed drastically with the advent of mulivariate complexity theory, known as *Parameterized Complexity*. By combining tools from Parameterized Complexity and classical (i.e., univariate) complexity, it has become possible to derive upper and lower bounds on sizes of reduced instances, or so called kernels.

Formally, a problem Π is *parameterized* if each instance of Π is associated with a *parameter* k. It admits a *compression* if there exists a (not necessarily parameterized) problem Π', and a polynomial-time algorithm that, given an instance (I, k) of Π, outputs an equivalent instance I' of Π' (i.e., (I, k) is a Yes-instance of Π if and only if I' is a Yes-instance of Π') such that $|I'| \leq p(k)$ where p is any computational function that depends only on k. In particular, when $\Pi' = \Pi$, we say that Π admits a *kernel*. Further, when p is polynomial or quasi-polynomial, we say that Π admits a kernel of polynomial or quasi-polynomial size (or, simply, a polynomial or quasi-polynomial kernel), respectively. So, when the parameter is substantially smaller than the entire input size, we can, just in polynomial time, reduce the input instance to be significantly smaller than it was before – depending only on how small or large k is! Nowadays, Kernelization is a major sub-field of research within Parameterized Complexity, and we refer to books such as [7, 1, 3] for more information.

In this paper, we suggest to extend the notion of a kernel to *permit the kernelization algorithm to be executed in quasi-polynomial time rather than polynomial time*. So far, we are only aware of one work [13] that addressed this negatively, showing that the lower bounds on kernel sizes proved for kernelization in that paper also hold when quasi-polynomial time complexity is allowed. When we, anyway, deal with an NP-hard problem, sacrificing polynomial time in preprocessing for quasi-polynomial time may often not be a big deal, but, of course, the question is – does it give us more power? The only known work, mentioned above, seems to suggest that the answer is "no". In this paper, we show that, for obtaining deterministic kernelization algorithms, quasipolynomial time can be quite powerful. Some of the most basic kernelization algorithms in the field are based on inherently randomized tools whose derandomization is a huge problem that has remained (and may still remain) open for many decades. Still, some breakthrough advances for derandomization in quasi-polynomial time have been made. Can we harness these advancements to design quasi-polynomial deterministic kernelization algorithms for basic problems in the field?

We also note that some recent works [27, 28] have investigated (randomized) quasipolynomial size kernels (i.e. $k^{\mathcal{O} \log^{\mathcal{O}(1)} k)}$) size) for several problems, such as Edge Multiway Cut, Group Feedback Edge Set and others.

For this purpose, we revisit the question of deterministic polynomial-time computation of a linear representation of transversal matroids and gammoids, which is a longstanding open problem. We build upon the earlier work of Lokshtanov et.al. on union representation [19] to present a deterministic computation of a representation matrix of a transversal matroid (and, more generally, of a matching matroid) in time quasipolynomial in the rank of the matroid, where each entry of the matrix can be represented in quasipolynomial (in the rank of the matroid) bits. As a corollary, we obtain a linear representation of a gammoid in deterministic quasipolynomial time and quasipolynomial bits in the size of the underlying ground set of the gammoid. Although we were mainly interested in these results for the purpose of the design of quasipolynomial-time deterministic kernels, and we mainly think of them as the tools that we present for this purpose, they are of broad and significant independent interest for computer science researchers outside Parameterized Complexity – indeed, matroids have become, over the past decade, ubiquitous in many fields in theoretical computer science.

Specifically, we apply the aforementioned tools to derandomize (in quasipolynomial time) the celebrated kernelization algorithms of Kratsch and Wahlström [17], who gave randomized polynomial kernels for several problems such as Odd Cycle Transversal, Multiway Cut with deletable terminals and Almost 2-SAT, based on representation of gammoids (see Section 4). Notably, the sizes of our kernels are polynomial – the only compensation is done in terms of time complexity. We remark that the derandomization (in polynomial time) of these kernels is among the most well-known and central open problems in the field of Kernelization (see, e.g., [7]). Similarly, we derandomize other kernels based on representation of gammoids, such as those given in [14, 16]. We also obtain derandomizations of the *Cut Covering Lemma* in quasipolynomial time and of the computation of representative sets regarding matroids, which are of independent interest in (not only) Parameterized Complexity. So, in particular, given an n-vertex (di)graph G and $S, T \subseteq V(G)$, a set $Z \subseteq V$ of cardinality $\mathcal{O}(|S| \cdot |T| \cdot r)$ such that for every $A \subseteq S$ and $B \subseteq T$, Z contains a minimum (A, B)-vertex cut, can be found in time $n^{\mathcal{O}(\log n)}$. Here, r is the size of a minimum (S, T)-cut in G. For example, succinct representations of cuts have been studied in network design (see, e.g., [18]), where our result naturally fits in.

In what follows in this introduction, we present relevant background on matroid representation and discuss our contribution in this regard. After that, we discuss some related works, as well as the relation of this paper to the earlier work on union representation.

1.1 Matroid Representation

During the past few decades, matroids have gained particular interest in computer science. These mathematical objects have taken lead roles in algorithm design, combinatorial optimization and computational complexity. For example, in algorithm design, analysis of these objects can yield algorithmic meta theorems. Such theorems unify classical results such as polynomial-time solvability of a wide-variety of problems as central as Minimum Weight Spanning Tree and Perfect Matching. In fact, if a problem admits a greedy algorithm, then it can be embedded in a matroid so that solutions correspond to maximum independent sets in the matroid. Recently, matroids also stand in the forefront of studies of approximation algorithms, parameterized algorithms and kernels.

A matroid is a pair $M = (E, \mathcal{I})$, where $\mathcal{I}$ is a family of subsets of E (called *independent sets*), that satisfy three conditions called *matroid axioms* (see Section 2). As the size of $\mathcal{I}$ can be exponential in the size of E, explicit listing of all independent sets is often rendered prohibitive. Then, it is necessary to have an *independence oracle* that, given a subset $X \subseteq E$, determines (in polynomial time) whether X is present in $\mathcal{I}$. For a wide class of matroids, known as *linear matroids*, such oracle is given by a matrix called a *representation*. Roughly speaking, the columns of the matrix are in bijection with the elements in E, and a set of columns is linearly independent if and only if the set of corresponding elements is independent. Such representations are immensely useful in algorithmic settings, since a large class of tools and techniques based on linear algebra become available. Unfortunately, despite substantial efforts, some central forms of linear matroids are not yet known to admit efficient deterministic computations of their representations, where (arguably) the most well-known and studied such classes of matroids are gammoids and transversal matroids.

Among the most central classes of matroids are the classes of uniform (or, more generally, partition) matroids, graphic and cographic matroids, matching matroids, transversal matroids and gammoids. A common property of all of these classes is that all of them are contained in the wider class of linear matroids. However, for the last three classes in this list a polynomial-time deterministic computation of a representation is not known. Indeed, developing polynomial-time deterministic computation of a representation of transversal matroids and gammoids is a longstanding open problem. Formally, a *transversal matroid* is a matroid derived from a bipartite graph G with a fixed bipartition (A, B) as follows: the ground set E is simply A, and a subset $X \subseteq A$ is independent if and only if G has a matching that saturates it. Matching constraints are ubiquitous in computational problems, e.g. modelling scenarios where some given objects should be partitioned into pairs, or allocations of a set of objects to a set of agents. Transversal matroids and matching matroids are the translation of these constraints into the language of matroids. A *gammoid* is a matroid derived from a (di)graph G with subsets $S, T \subseteq V(D)$ as follows: the ground set E is simply T, and a subset $X \subseteq T$ is independent if and only if there exists a collection of vertex disjoint paths $\mathcal{P}$ from S to X in G where every vertex in X is the end-vertex of some path in $\mathcal{P}$. As the dual of a transversal matroid is a (strict) gammoid and vice versa, based on the work of [21], we know that a polynomial-time computation of a representation for one also yields such a computation for the other. Representation of gammoids is of particular importance in parameterized complexity as some of the most celebrated kernels in this field build upon it [17].

Our main contribution is the resolution of the questions of the representation of transversal matroids and gammoids under the relaxation of having quasipolynoimal rather than strictly polynomial time. Specifically, our main theorems are as follows.

▶ **Theorem 1** (Quasipolynomial Representation of Transversal Matroids). *Let G be an n-vertex bipartite graph with a fixed vertex bipartition (A, B), and let r be the size of a maximum matching in G. Then, in time $r^{\mathcal{O}(\log r)} n^{\mathcal{O}(1)}$ we can compute a representation (M, ϕ) of the transversal matroid of G, where each entry of the matrix is an element of $\mathbb{Q}$ that can be encoded in $r^{\mathcal{O}(\log r)}$ bits.*

▶ **Theorem 2** (Quasipolynomial Representation of Gammoids). *Let G be an n-vertex (di)graph and let $S, T \subseteq V(G)$. Then, in time $n^{\mathcal{O}(\log n)}$ we can compute a representation (M, ϕ) of the gammoid of G with respect to S on ground set T, where each entry of the matrix is an element of $\mathbb{Q}$ that can be represented in $n^{\mathcal{O}(\log n)}$ bits.*

Matching matroids are a generalization of Transversal matroids to general graphs. Using the well known Gallai-Edmonds decomposition and Theorem 1, we obtain the following.

▶ **Theorem 3** (Quasipolynomial Representation of Matching Matroids). *Let G be an n-vertex graph, and let r be the size of a maximum matching in G. Then, in time $r^{\mathcal{O}(\log r)} n^{\mathcal{O}(1)}$ we can compute a representation (M, ϕ) of the matching matroid of G, where each entry of the matrix is an element of $\mathbb{Q}$ that can be encoded in $r^{\mathcal{O}(\log r)}$ bits.*

We remark that Theorem 4.9 in [12] can also be applied to obtain a quasi-NC-computable representation of transversal matroids; although this not immediate.

1.2 Previous Works and the Relation to Union Representation

Prior to our work, the fastest (deterministic) computations of representations of transversal matroids and gammoids were only slightly better than trivial brute-force. More precisely, Misra et al. [22] showed that given a bipartite graph G with a fixed bipartition (A, B), a representation of the transversal matroid can be computed deterministically in (exponential) time $\binom{|A|}{r}|A|^{\mathcal{O}(1)}$ where r is the rank of the matroid, which equals the maximum size of a matching in G. In this context, it is important to note that a randomized polynomial-time algorithm to compute a representation of a transversal matroid is well known (see, e.g., [21, 25]). Here, randomization means that with some (low) probability, the algorithm may output a matrix that is not a representation of the matroid. This algorithm utilizes the Schwartz-Zippel lemma [2, 26, 29], and hence it is inherently randomized. The above mentioned trivial brute-force, which runs in time $2^{\mathcal{O}(|A|^2|B|)}$ (see [22]), refers to a loop through all choices made by the randomized algorithm. Building upon their representation for transversal matroids, Misra et al. [22] presented a similar result for gammoids.

The notion of *union representation* was introduced in [19]. Roughly speaking, a union representation of a matroid $M = (E, \mathcal{I})$ is a collection of matrices such that a subset X of E is independent in M if and only if for at least one of the matrices, the set of columns corresponding to X is linearly independent. Standard representation is precisely union representation where the size of the collection is one. While only linear matroids admit standard representations, note that all matroids admit union representations: to see this, for every base of the matroid, create one matrix with a set of linearly independent columns corresponding to the base, and vectors having only 0 entries as the rest of the columns. However, this procedure can create a *huge* number of matrices. In [19], a quasipolynomial-time union representation of transversal matroids with quasipolynomially many matrices and with entries represented in quasipolynomially many bits was presented. Unfortunately, a union representation of transversal matroids, apart from not being a proper representation in itself, also does *not* yield even a union representation for gammoids (the duality relation can only be exploited if a standard representation is provided).

Nevertheless, we use the earlier work as a starting point of the proof of Theorem 1. In Section 3, we essentially show that a union representation of a transversal matroid can be *merged* into a single matrix, using polynomial interpolation. It is pertinent to ask if this can achieved for any arbitrary matroid. The answer to this question is in the negative, since there exist well known matroids, such as the Vamos Matroid [25], which are not representable over any field, while all matroids admit union representations. The key reason this breaks down is in the reverse direction of Lemma 21, where given a collection of linearly independent columns of the interpolation matrix, we must show that the corresponding elements form an independent set in the matroid. Here, we rely on the fact that we deal with a transversal matroid. However, it is an interesting problem to determine for a class of non-representable matroids of interest, if a union representation with $|\mathcal{B}|$ elements (where $\mathcal{B}$ is the set of bases), which could be of exponential size, can be compressed to a substantially smaller collection of matrices.

2 Preliminaries

Basic Definitions

Given $t \in \mathbb{N}$, we use $[t]$ as a shorthand for $\{1, 2, \ldots, t\}$. Given a function $f : A \to B$ and a subset $A' \subseteq A$, we denote $f(A') = \{f(a) : a \in A'\}$, and we define $f|_{A'}$ as the restriction of f to A'. We slightly abuse notation, and given a function $g : A \to \mathbb{N}$, called a *weight function*, and a subset $A' \subseteq A$, we denote $g(A') = \sum_{a \in A'} g(a)$. Whenever we refer to a function that is a weight function, we use the second notation.

Linear Algebra. Given a matrix M, a set R of rows of M and a set C of columns of M, the submatrix of M formed by choosing the rows in R and the columns in C is denoted by $M[R, C]$. Moreover, we denote the determinant of M by $\mathsf{det}(M)$. We remind that a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t\}$ is *linearly dependent* over some field $\mathbb{F}$ if there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_t \in \mathbb{F}$, not all equal 0, such that $\sum_{i=1}^{t} \lambda_i \mathbf{v}_i = 0$. Moreover, a set of vectors is *linearly independent* if it is not linearly dependent.

Graphs. Given a (di)graph G, we denote the vertex set of G and edge(arc) set of G by $V(G)$ and $E(G)$, respectively. Given a subset $U \subseteq V(G)$, we let $G[U]$ denote the subgraph of G induced by U. A subset $U \subseteq V(G)$ is an *independent set* if there does not exist any edge in $E(G)$ with both endpoints in U. We say that (A, B) is a *vertex bipartition* of G if it is a partition of $V(G)$ such that A and B are independent sets. Moreover, we say that G is a *bipartite graph* if it has a vertex bipartition. A *matching* μ is a family of pairwise-disjoint edges of $E(G)$. If an edge (u, v) is in a matching μ, then the vertices u and v are said to be *matching partners* in μ. The maximum size of a matching is denoted by $\kappa(G)$. A subset $X \subseteq V(G)$ is said to be *saturated* by μ if every vertex in X belongs to some pair in μ. Moreover, μ is called *perfect* if it saturates $V(G)$. The set of all vertices saturated by a matching μ is denoted by $V(\mu)$. A graph G is called *factor-critical* if for every $v \in V(G)$ the subgraph graph $G[V(G) \setminus \{v\}]$ has a perfect matching. A *path* P in G is a subgraph defined by sequence of distinct vertices $v_1, v_2, \ldots, v_t$, where for each $i \in [t-1]$, we have an edge(arc) $(v_i, v_{i+1}) \in E(G)$. The vertices v_1 and v_t are called the *start-vertex* and *end-vertex* of P, respectively, and collectively they are called the *endpoints* of P. Further, P is said to be a path from v_1 to v_t. Two paths P and P' are *vertex disjoint* if they have no common vertices. A *path system* in G from $S \subseteq V(G)$ to $T \subseteq V(G)$ is a collection of pairwise vertex disjoint paths with their start-vertices in S and end-vertices in T.

Two graphs G and G' are said to be *isomorphic* if $|V(G)| = |V(G')|$ and there is a bijection $\phi : V(G) \to V(G')$ such that $\{u, v\} \in E(G)$ if and only if $\{\phi(u), \phi(v)\} \in E(G')$. A *graph class* $\mathcal{G}$ is a collection of graphs closed under isomorphism, i.e., it contains a graph G if and only if it contains all graphs G' that are isomorphic to G.

Matroids

Let us begin by presenting the definition of a matroid.

▶ **Definition 4** (Matroid, [25]). *A pair $\mathcal{M} = (E, \mathcal{I})$, where E is a ground set and $\mathcal{I}$ is a family of subsets of E (called* independent sets*), is a* matroid *if it satisfies the following conditions:*

(I1) $\emptyset \in \mathcal{I}$.

(I2) *If $X \subseteq Y$ and $Y \in \mathcal{I}$, then $X \in \mathcal{I}$.*

(I3) *If $X, Y \in \mathcal{I}$ and $|X| < |Y|$, then there exists $e \in (Y \setminus X)$ such that $X \cup \{e\} \in \mathcal{I}$.*

We remark that conditions (I1), (I2) and (I3) are called *matroid axioms*. An inclusion-wise maximal independent set in a matroid is called a *basis*. We say that two matroids $\mathcal{M} = (E, \mathcal{I})$ and $\mathcal{M}' = (E', \mathcal{I}')$ are *isomorphic* if there exists a bijection $\varphi : E \to E'$ such that for every $X \subseteq E$, $X \in \mathcal{I}$ if and only if $\varphi(X) \in \mathcal{I}'$. Using (I3) it is easy to show that all the bases of a matroid have the same size, which is called the *rank* of the matroid. Using (I2) it is easy to see that if two matroids have the same set of bases, then they are the same matroid.

Given two matroids $\mathcal{M} = (E, \mathcal{I})$ and $\mathcal{M}' = (E, \mathcal{I}')$ over the same ground set, we say that $\mathcal{M}$ and $\mathcal{M}'$ are *dual matroids*, if they satisfy the following: $B \subseteq E$ is a basis of $\mathcal{M}$ if and only if $(E \setminus B)$ is a basis of $\mathcal{M}'$. The dual of a matroid $\mathcal{M}$ is denoted by $\mathcal{M}^*$, and clearly $(\mathcal{M}^*)^* = \mathcal{M}$. Given a matroid $\mathcal{M} = (E, \mathcal{I})$ and $F \subseteq E$, the *deletion* of F in $\mathcal{M}$ is the matroid $\mathcal{M} \setminus F = (E \setminus F, \mathcal{I} \setminus F)$ where $\mathcal{I} \setminus F$ denotes the collection of sets in $\mathcal{I}$ that are disjoint from F. In this paper, we are specifically interested in transversal matroids and gammoids.

▶ **Definition 5** (Transversal Matroid, [25]). *Let G be a bipartite graph with a fixed vertex bipartition (A, B). The* transversal matroid *of G is a matroid $(A, \mathcal{I})$ where $\mathcal{I}$ is the family that consists of every subset $X \subseteq A$ such that there exists a matching that saturates X.*

▶ **Definition 6** (Gammoid, [25]). *Let G be a (di)graph, and let $S, T \subseteq V(D)$. The* gammoid *of G with respect to S on ground set T is a matroid $(T, \mathcal{I})$ where $\mathcal{I}$ is the family that consists of every subset $X \subseteq T$ that satisfies the following condition: there exists a path system $\mathcal{P}$ from S to X in G where every vertex in X is the end-vertex of some path in $\mathcal{P}$. When $T = V(G)$, the gammoid is called a* strict gammoid.

It is well known that transversal matroids and strict gammoids are duals [25]. Moreover, it is easy to see that if we consider the strict gammoid of a graph G with respect to some subset $S \subseteq V(G)$, then for any subset $T \subseteq V(G)$, the gammoid of G with respect to S on the ground set T can be obtained by deleting $V(G) \setminus T$ from the strict gammoid.

Having a *representation* of a matroid, given by a matrix that compactly encodes the matroid, is a central component in many algorithmic applications [17]. Matroids having a representation are called *linear*, as formally defined below.

▶ **Definition 7.** *Let A be a matrix over an arbitrary field $\mathbb{F}$, and let C be the set of columns of A. The* matroid represented by *A is the pair $(C, \mathcal{I})$ where a subset $X \subseteq C$ belongs to $\mathcal{I}$ if and only if the columns in X are linearly independent over $\mathbb{F}$.*

It is well known that the pair $(C, \mathcal{I})$ in Definition 7 indeed defines a matroid [25].

▶ **Definition 8** (Linear Matroid, Representation). *A matroid $\mathcal{M} = (E, \mathcal{I})$ is a* linear matroid *if there exists a matrix A, called a* representation *of $\mathcal{M}$, such that $\mathcal{M}$ and the matroid represented by A are isomorphic. Furthermore, $\mathcal{M}$ is* representable over a field $\mathbb{F}$ *if it has a representation A over $\mathbb{F}$.*

The phrasing that $\mathcal{M}$ is represented by (A, φ) is an abbreviation to the statement that A is a representation of $\mathcal{M}$ and φ is an isomorphism that witnesses this. It is easy to see that that if $\mathcal{M} = (E, \mathcal{I})$ is a linear matroid with representation M, then for any $F \subseteq E$, the matroid $\mathcal{M} \setminus F$ is also a linear matroid whose representation can be obtained from M by deleting the columns that correspond to elements in $E \setminus F$. The following lemma relates the representations of a matroid and its dual.

▶ **Proposition 9** ([25]). *Let $\mathcal{M} = (E, \mathcal{I})$ be a linear matroid that is representable over a field $\mathbb{F}$, and let $|E| = n$. Then, $\mathcal{M}^*$ is also a linear matroid. Further, given a representation M of $\mathcal{M}$, a representation M' of $\mathcal{M}^*$ can computed in polynomially many (in n) field operations.*

Isolation of Perfect Matchings

For the sake of clarity, let us first introduce the following notation. Given $n \in \mathbb{N}$, let G be a bipartite graph with a *fixed* bipartition (A, B) such that $|A|, |B| \leq n$, and *fixed* injective functions $\gamma_A : A \to [n]$ and $\gamma_B : B \to [n]$. Given a weight function $w : [n] \times [n] \to \mathbb{N}$, we define the weight of an edge $\{a,b\} \in E(G)$, where $a \in A$ and $b \in B$, by $\widetilde{w}(\{a,b\}) = w(\gamma_A(a), \gamma_B(b))$. Thus, $\widetilde{w}$ can be thought of as a function from $E(G)$ to $\mathbb{N}$. Let us remind that for a subset $U \subseteq E(G)$, $\widetilde{w}(U) = \sum_{e \in U} \widetilde{w}(e)$.

We remark that we need to define a weight function via injective functions of the form γ_A and γ_B as above (rather than letting the domain directly be an edge set) in order to prove the correctness of our main result, particularly in its general form. Now, for perfect matchings, isolating weight functions are defined as follows.

▶ **Definition 10** (Isolating Weight Function). *Let G be a bipartite graph with a fixed bipartition (A,B) such that $|A|,|B| \leq n$, and fixed injective functions $\gamma_A : A \to [n]$ and $\gamma_B : B \to [n]$. A weight function $w : [n] \times [n] \to \mathbb{N}$ is* isolating *(for G) if it satisfies the following condition: If G has a perfect matching, then G also has a unique perfect matching μ of minimum weight (i.e. for every perfect matching $\mu' \neq \mu$, $\widetilde{w}(\mu) < \widetilde{w}(\mu')$).*

Such isolating weight functions are particularly relevant to the detection of a perfect matching. To see this, we first need to define the matrix associated with an isolating weight function.

▶ **Definition 11.** *Let G be a bipartite graph with a fixed bipartition (A,B) such that $|A|,|B| \leq n$, and fixed injective functions $\gamma_A : A \to [n]$ and $\gamma_B : B \to [n]$. In addition, let $w : [n] \times [n] \to \mathbb{N}$ be a weight function. Then, $W_{(G,w)}$ is the matrix on $|A|$ columns indexed by the vertices in A and $|B|$ rows indexed by the vertices in B, where*

$$W_{(G,w)}[b,a] = \begin{cases} 2^{\widetilde{w}(\{b,a\})} & \text{if } \{b,a\} \in E(G) \\ 0 & \text{otherwise} \end{cases}$$

for all $a \in A$ and $b \in B$.

The following well-known result, due to Mulmuley et al. [23], reveals a connection between isolating weight functions, determinants and perfect matchings.

▶ **Proposition 12** ([23]). *Let G be a bipartite graph with a fixed bipartition (A,B) such that $|A| = |B| \leq n$, and fixed injective functions $\gamma_A : A \to [n]$ and $\gamma_B : B \to [n]$. In addition, let $w : [n] \times [n] \to \mathbb{N}$ be a weight function. If $\mathsf{det}(W_{(G,w)}) \neq 0$, then G has a perfect matching. Moreover, if w is isolating and G has a perfect matching, then $\mathsf{det}(W_{(G,w)}) \neq 0$.*

▶ **Definition 13** (Isolating Collection). *Let $\mathcal{G}$ be a graph class and let $n \in \mathbb{N}$. An n-*isolating collection *for $\mathcal{G}$ is a set $\mathcal{W}_n$ of weight functions $w : [n] \times [n] \to \mathbb{N}$ with the following property. For any bipartite graph $G \in \mathcal{G}$ with a fixed bipartition (A,B) such that $|A|,|B| \leq n$, and fixed bijective functions $\gamma_A : A \to [n]$ and $\gamma_B : B \to [n]$, there exists a weight function $w \in \mathcal{W}_n$ such that w is isolating.*

Let us remark that the above definition can be generalized to graph classes that do not contain only bipartite graphs. However, as we are primarily interested in transversal matroids that are defined using bipartite graphs, the above definition is sufficient (and necessary). Fenner et al. [5] presented a (deterministic) computation of a collection of weight functions that, for any bipartite graph, has at least one isolating weight function. Formally,

▶ **Proposition 14** ([5]). *Let $n \in \mathbb{N}$. An n-isolating collection for the class of bipartite graphs, denoted by $\mathcal{W}_n$, containing $n^{\mathcal{O}(\log n)}$ weight functions with the following property can be obtained in time $n^{\mathcal{O}(\log n)}$: for any weight function $w \in \mathcal{W}_n$, every weight assigned by w can be represented (in binary) using $\mathcal{O}(\log^2 n)$ bits.*

Splitters, Representative Families

Splitters are well-known tools in derandomization, formally defined as follows.

▶ **Definition 15** (Splitter). *Let $n, k, \ell \in \mathbb{N}$ where $k \leq \ell$. An (n, k, ℓ)-*splitter *is a family $\mathcal{F}$ of functions from $[n]$ to $[\ell]$ such that for every $S \subseteq [n]$ of size k, there is a function $f \in \mathcal{F}$ that satisfies $f(i) \neq f(j)$ for all distinct $i, j \in S$.*

We are specifically interested in an (n, k, k^2)-splitter. The following lemma asserts that such a small splitter can be computed efficiently.

▶ **Proposition 16** ([24]). *Given $n, k \in \mathbb{N}$, an (n, k, k^2)-splitter of size $k^{\mathcal{O}(1)} \log n$ can be constructed in time $k^{\mathcal{O}(1)} n \log n$.*

The notion of a representative family (implicitly linked to that of a splitter), introduced by Fomin et al. [6], plays a central role in the design of fast deterministic parameterized algorithms.

▶ **Definition 17** (Representative Family). *Given a matroid $M = (E, \mathcal{I})$ and a family $\mathcal{S}$ of subsets of E, a subfamily $\widehat{\mathcal{S}} \subseteq \mathcal{S}$ is q-*representative *for $\mathcal{S}$, denoted by $\widehat{\mathcal{S}} \subseteq^q_{rep} \mathcal{S}$, if the following holds: for every set $Y \subseteq E$ of size at most q, if there is a set $X \in \mathcal{S}$ disjoint from Y with $X \cup Y \in \mathcal{I}$, then there is a set $\widehat{X} \in \widehat{\mathcal{S}}$ disjoint from Y with $\widehat{X} \cup Y \in \mathcal{I}$.*

Isolation of Maximum Matchings

We begin with the definition of the matrix $W_{(G,w,f)}$ that combines isolating weight functions with splitter functions, which are functions from $[2n]$ to $[(2r)^2]$ where $n, r \in \mathbb{N}$ will be clear from context.

▶ **Definition 18.** *Let G be a bipartite graph with a fixed bipartition (A, B) such that $|A|, |B| \leq n$, and fixed injective functions $\gamma_A : A \to [n]$ and $\gamma_B : B \to [n]$. In addition, let $w : [(2r)^2] \times [(2r)^2] \to \mathbb{N}$ be a weight function and $f : [2n] \to [(2r)^2]$ be a splitter function for some $r \in \mathbb{N}$. Then, $W_{(G,w,f)}$ is the matrix on $|A|$ columns indexed by the vertices in A and $|B|$ rows indexed by the vertices in B, where*

$$W_{(G,w,f)}[b, a] = \begin{cases} 2^{w(f(\gamma_A(a)), f(n+\gamma_B(b)))} & \text{if } \{b, a\} \in E(G) \\ 0 & \text{otherwise} \end{cases}$$

for all $a \in A$ and $b \in B$.

Next, we define good pairs of a weight function and a splitter function with respect to a subset of vertices, which "isolates" a maximum matching saturating this subset.

▶ **Definition 19.** *Let G be a bipartite graph with a fixed vertex bipartition (A, B) such that $|A|, |B| \leq n$, and fixed injective functions $\gamma_A : A \to [n]$ and $\gamma_B : B \to [n]$. In addition, let $r \in \mathbb{N}$. For a weight function $w : [(2r)^2] \times [(2r)^2] \to \mathbb{N}$ and a splitter function $f_A : [2n] \to [(2r)^2]$, the pair (w, f) is* good *for a subset $X \subseteq A$ if $\mathsf{det}(W_{(G,w,f)}[Y, X]) \neq 0$ for some $Y \subseteq B$.*

The following result asserts that by combining splitters and isolating weight functions for perfect matchings, we can obtain isolating weight functions for maximum matchings.

▶ **Proposition 20** ([19]). *Let $\mathcal{G}$ be a graph class. Let $G \in \mathcal{G}$ be a bipartite graph with a fixed vertex bipartition (A, B) such that $|A|, |B| \leq n$, and fixed injective functions $\gamma_A : A \to [n]$ and $\gamma_B : B \to [n]$. In addition, let $\mathcal{W}$ be a $(2r)^2$-isolating collection for $\mathcal{G}$, and $\mathcal{F}$ be a $(2n, 2r, (2r)^2)$-splitter for some $r \in \mathbb{N}$. For every subset $X \subseteq A$ of size at most r, if X is independent in the transversal matroid of G, then there exist $w \in \mathcal{W}$ and $f \in \mathcal{F}$ such that (w, f) is good for X.*

3 Representation via Polynomial Interpolation

In this section, we show that a collection of isolating weight functions for maximum matchings can be converted to a representation matrix for a transversal matroid by using *polynomial interpolation*. Our first step is to obtain a representation over $\mathbb{Q}[Z]$, the ring of univariate polynomials with coefficients from $\mathbb{Q}$ (in the formal variable Z).

▶ **Lemma 21.** *Let $\mathcal{G}$ be a graph class. Let $G \in \mathcal{G}$ be an n-vertex bipartite graph with a fixed vertex bipartition (A, B), and let r be the size of a maximum matching in G. Let $\mathcal{W}$ be an r-isolating collection for $\mathcal{G}$. Let $b \in \mathbb{N}$ be the maximum weight assigned by any function in $\mathcal{W}$. Then, we can compute a representation $(\widehat{M}, \widehat{\phi})$ of the transversal matroid of G in time $(nb|\mathcal{W}|)^{\mathcal{O}(1)}$, where each entry $\widehat{M}[\beta, \alpha]$ is a polynomial in $\mathbb{Q}[Z]$ with integer coefficients in the form*

$$\widehat{M}[\beta, \alpha] = \sum_{\ell=1}^{t} \big(q_\ell \cdot \prod_{j \in [t] \setminus \{\ell\}} (Z - j)\big),$$

where $t = |\mathcal{W}| n^{\mathcal{O}(1)}$ and for any $\ell \in [t]$, $q_\ell \in \mathbb{Z}$ is encoded in $b + t \log t$ bits.

Proof. Observe that $r \leq n$, and let us begin by constructing a family of matrices over $\mathbb{Q}$ with integer entries. First, we apply Proposition 15 to obtain a $(2n, 2r, (2r)^2)$-splitter $\mathcal{F}$ of size $r^{\mathcal{O}(1)} \log n$ in time $r^{\mathcal{O}(1)} \cdot n \log n$. Then, we select arbitrary bijective functions $\gamma_A : A \to [|A|]$ and $\gamma_B : B \to [|B|]$. Note that the collection $\{W_{(G,w,f)}\}|_{w \in \mathcal{W}, f \in \mathcal{F}}$ contains $t = |\mathcal{W}| \cdot |\mathcal{F}| = |\mathcal{W}| r^{\mathcal{O}(1)} \log n$ matrices. By Definition 18, every entry in $W_{(G,w,f)}$, $w \in \mathcal{W}$ and $f \in \mathcal{F}$, is an integer of bit-length b. Thus, the time to construct $\{W_{(G,w,f)}\}|_{w \in \mathcal{W}, f \in \mathcal{F}}$ is bounded by $b|\mathcal{W}|n^{\mathcal{O}(1)}$. For convenience, let us define $\tau : [t] \to \mathcal{W} \times \mathcal{F}$ as an arbitrary bijection, and for $i \in [t]$ let M_i denote the matrix $W_{(G,w_i,f_i)}$ where $\tau(i) = (w_i, f_i)$. Then, each entry of M_i is an integer that can be represented using b bits. Observe that the rows and columns of each M_i are indexed by B and A, respectively, where the indexing "agree" for any two matrices M_i and M_j.

Let us describe the construction of the matrix $\widehat{M}$ using the collection $\{M_i\}|_{i \in [t]}$. Let $\widehat{M}[\beta, \alpha]$ denote the entry of the matrix in row β and column α. Then, we will define $\widehat{M}[\beta, \alpha]$ as the multiplication of some scalar s with a (univariate) polynomial $p_{\beta,\alpha}(Z) \in \mathbb{Q}[Z]$ in the formal variable Z of degree $t-1$ that will have the following property: for each $i \in [t]$, we have $p_{\beta,\alpha}(Z{=}i) = M_i[\beta, \alpha]$. It follows that $\widehat{M} = \widehat{M}(Z)$ is a matrix over $\mathbb{Q}[Z]$, and for each $i \in [t]$, we have $\widehat{M}(Z{=}i) = s \cdot M_i$. For this purpose, we choose the polynomial $p_{\beta,\alpha}(Z)$ to be the Lagrange polynomial of the points $\{(i, M_i[\beta, \alpha])\}|_{i \in [t]}$, which is a degree $t-1$ polynomial in $\mathbb{Q}[Z]$. Specifically, it is described by the following formula.

$$p_{\beta,\alpha}(Z) = \sum\nolimits_{\ell=1}^{t} \Big(\prod_{j \in [t] \setminus \{\ell\}} \frac{Z - j}{\ell - j} \Big) M_\ell[\beta, \alpha]$$

Indeed, for any $i \in [t]$, if we substitute Z by i above, then for all $\ell \in [t] \setminus \{i\}$, the expression $(\prod_{j\in[t]\setminus\{\ell\}} \frac{Z-j}{\ell-j})M_\ell[\beta,\alpha]$ evaluates to 0, while for $\ell = i$ it evaluates to $M_i[\beta,\alpha]$.

Now, we choose $s = (t-1)\cdot(t-2)\cdot\dots\cdot 1\cdot(-1)\cdot(-2)\cdot\dots\cdot(1-t)$. Accordingly, we define

$$\widehat{M}[\beta,\alpha] = \sum_{\ell=1}^{t} \big(q_\ell \cdot \prod_{j\in[t]\setminus\{\ell\}} (Z-j)\big),$$

where for any $\ell \in [t]$, let $q_\ell = \dfrac{s\cdot M_\ell[\beta,\alpha]}{\prod_{j\in[t]\setminus\{\ell\}}(\ell-j)}$. Then, for any $\ell \in [t]$, the definition of s directly yields that $q_\ell \in \mathbb{Z}$ and $q_\ell \le t^t \cdot M_\ell[\beta,\alpha]$, and because $M_\ell[\beta,\alpha]$ can be encoded in b bits, we deduce that q_ℓ can be encoded in $t\cdot\log t + b$ bits.

Clearly, the computation of $\widehat{M}$ is done in time $(nb|\mathcal{W}|)^{\mathcal{O}(1)}$. Let us note that the rows and columns of $\widehat{M}$ are indexed by the vertex sets B and A, respectively, as is directly obtained from the same indexing of the matrices $\{M_i\}|_{i\in[t]}$. Hence the mapping $\widehat{\phi}$ from A to columns of $\widehat{M}$ is well defined. Furthermore, by Definition 18, we have the following claim.

▷ Claim 22. For any $\alpha \in A$ and $\beta \in B$, if $\widehat{M}(Z)[\beta,\alpha] \neq 0$, then $\{\alpha,\beta\} \in E(G)$.

To see this, consider some $\alpha \in A$ and $\beta \in B$. If $\widehat{M}(Z)[\beta,\alpha] \neq 0$, then the polynomial $p_{\beta,\alpha}(Z)$ is not identically zero, therefore in particular there must exist $i \in [t]$ such that $M_i[\beta,\alpha] \neq 0$. However, Definition 18 then directly yields that $\{\alpha,\beta\} \in E(G)$. (We note that the reverse direction of the claim also holds, as can be easily seen to follow as a special case from the arguments in the following paragraph.)

Let us argue that $\widehat{M}(Z)$ is indeed a representation of the transversal matroid of G. In the forward direction, consider a subset $X \subseteq A$ such that there is a matching in G saturating X, i.e., X is an independent set (of size at most r) in the transversal matroid of G. Then, we must argue that the columns of $\widehat{M}(Z)$ corresponding to X are linearly independent. Consider a subset $Y \subseteq B$ such that $|Y| = |X|$, and there is a perfect matching in $G[X \cup Y]$. Then, consider the submatrix $\widehat{M}(Z)[Y,X]$ indexed by the rows corresponding to Y and columns corresponding to X. We claim that $\mathsf{det}(\widehat{M}(Z)[Y,X]) \neq 0$, which implies that this submatrix has full rank, and hence the columns of $\widehat{M}(Z)$ corresponding to X are linearly independent. Suppose not, that is, $\mathsf{det}(\widehat{M}(Z)[Y,X]) = 0$. In particular, this means that $\mathsf{det}(\widehat{M}(Z{=}i)[Y,X]) = 0$ for every $i \in [t]$. However, by Proposition 20, there exist a pair of $w \in \mathcal{W}$ and $f \in \mathcal{F}$ that is good for X, and thus $\mathsf{det}(W_{(G,w,f)}[Y,X]) \neq 0$. Now, consider $i \in [t]$ such that $\tau(i) = (w,f)$ and observe that we have reached a contradiction since $\widehat{M}(Z{=}i)[Y,X] = s\cdot M_i[Y,X] = s\cdot W_{(G,w,f)}[Y,X]$.

In the reverse direction, consider $X \subseteq A$ such that the column vectors of $\widehat{M}$ corresponding to X are linearly independent. Therefore, $\mathsf{det}(\widehat{M}[Y,X]) \neq 0$ for some $Y \subseteq B$ with $|Y| = |X|$. By the definition of determinant, we have that

$$\sum_{\sigma\in\mathcal{S}} \left(\mathsf{sign}(\sigma) \prod_{y\in Y} (\widehat{M}[Y,X])[y,\sigma(y)]\right) \neq 0,$$

where $\mathcal{S}$ is the set of all bijective functions from Y to X. Thus, there exists $\sigma \in \mathcal{S}$ such that $(\widehat{M}[Y,X])[y,\sigma(y)] \neq 0$ for all $y \in Y$. By Claim 22, this means that there exists a bijective function $g : Y \to X$ (that is, σ above) such that $\{g(y), y\} \in E(G)$ for all $y \in Y$. Therefore, $G[X \cup Y]$ has a perfect matching. We thus conclude that X is independent in the transversal matroid of G. ◀

The above lemma gives us a representation matrix $\widehat{M}$ over the ring of (univariate) polynomials in $\mathbb{Q}[Z]$. Our next step is to obtain a representation over $\mathbb{Q}$.

▶ **Lemma 23.** *Let $\mathcal{G}$ be a graph class. Let $G \in \mathcal{G}$ be an n-vertex bipartite graph with a fixed vertex bipartition (A, B), and let r be the size of a maximum matching in G. Let $\mathcal{W}$ be an r-isolating collection for $\mathcal{G}$. Let $b \in \mathbb{N}$ be the maximum weight assigned by any function in $\mathcal{W}$. Then, we can compute a representation (M, ϕ) of the transversal matroid of G in time $(nb|\mathcal{W}|)^{\mathcal{O}(1)}$, where each entry of M is an integer that can be encoded in $bn^{\mathcal{O}(1)}|\mathcal{W}|^2 \log |\mathcal{W}|$ bits.*

Proof. We begin by applying Lemma 21 and obtaining the representation $\widehat{M}$ over $\mathbb{Q}[Z]$. Note that each entry $\widehat{M}[\beta, \alpha]$ is a univariate polynomial in the formal variable Z with integer coefficients, which is of the form

$$\widehat{M}[\beta, \alpha] = \sum_{\ell=1}^{t} \big(q_\ell \cdot \prod_{j \in [t] \setminus \{\ell\}} (Z - j)\big),$$

where $t = |\mathcal{W}|n^{\mathcal{O}(1)}$ and for any $\ell \in [t]$, $q_\ell \in \mathbb{Z}$ is encoded in $b + t \log t$ bits. Let $\gamma = 2^{b+4t \log t}$, and $\Gamma = (2\gamma)^n + 2$. Then, we have the following immediate claim.

▷ Claim 24. The absolute value of any coefficient of $\widehat{M}[\beta, \alpha]$ is upper bounded by γ.

To see this, consider some $\ell \in [t]$ and $r \in [t-1] \cup \{0\}$, and observe that the absolute value of the coefficient of Z^r in $\prod_{j \in [t] \setminus \{\ell\}} (Z - j)$ is upper bounded by $\binom{t-1}{r} t^{t-1-r} \leq (2t)^t$. Because $q_\ell \leq 2^{b+t \log t} = 2^b t^t$, and we sum over ℓ from 1 to t, we derive that the absolute value of any coefficient of $\widehat{M}[\beta, \alpha]$ is upper bounded by $t \cdot 2^b t^t \cdot (2t)^t \leq \gamma$.

Now, observe that $\Gamma \leq 2^{\mathcal{O}(n(b+4t \log t))}$, and hence it can be encoded using $\mathcal{O}(n(b + 4t \log t)) \leq bn^{\mathcal{O}(1)}|\mathcal{W}| \log |\mathcal{W}|$ bits. Consider the matrix $M = \widehat{M}(Z = \Gamma)$. Observe that M is a matrix whose entries are integers. Each entry of M, say $M[\beta, \alpha] = \widehat{M}(Z = \Gamma)[\beta, \alpha]$ at row β and column α, is the evaluation at Γ of a degree $t-1$ polynomial with coefficients of absolute value at most γ (by Claim 24). Therefore, the absolute value of $M[\beta, \alpha]$ is upper bounded by $t \cdot \gamma \cdot \Gamma^{t-1} \leq \gamma^{\mathcal{O}(nt)} \leq 2^{\mathcal{O}(ntb+nt^2 \log t)}$, and hence it can be encoded in $ntb + nt^2 \log t = bn^{\mathcal{O}(1)}|\mathcal{W}|^2 \log |\mathcal{W}|$ bits.

Next, let us argue that M is a correct representation matrix. Consider a square submatrix M' of M, and let $\widehat{M}'$ be the corresponding submatrix of $\widehat{M}$. It is sufficient to argue that $\mathsf{det}(M') \neq 0$ if and only if $\mathsf{det}(\widehat{M}') \neq 0$. Then, as $\widehat{M}'$ is a correct representation, so is M. Observe that $\mathsf{det}(\widehat{M}')$ is a univariate polynomial in $\mathbb{Q}[Z]$, and $\mathsf{det}(M')$ is an evaluation of this polynomial at $Z = \Gamma$. Hence, if $\mathsf{det}(M') \neq 0$, then $\mathsf{det}(\widehat{M}') \neq 0$ as well. In the reverse direction, let us argue that Γ is not a root of the polynomial $\mathsf{det}(\widehat{M}')$. Then, it follows that if $\mathsf{det}(\widehat{M}') \neq 0$, then $\mathsf{det}(M') \neq 0$ as well. To obtain that Γ is not a root of $\mathsf{det}(\widehat{M}')$, we will apply Cauchy's upper bound on the absolute value of any root of a univariate polynomial with coefficients in $\mathbb{Q}$. It states that if $a_0 + a_1 Z + a_2 Z^2 + \ldots + a_r Z^r$ is a degree r polynomial with coefficients $a_i \in \mathbb{Q}$ for all $i \in [r]$, then the absolute value of any root of this polynomial is upper bounded by $1 + \max\{\frac{a_{r-1}}{a_r}, \frac{a_{r-2}}{a_r}, \ldots, \frac{a_0}{a_r}\}$. Now, consider the coefficients of $\mathsf{det}(\widehat{M}')$. Observe that $\mathsf{det}(\widehat{M}')$ is a polynomial that is obtained by summing over at most 2^n terms, each of which is a product of n entries of $\widehat{M}'$. Recall that the entries in $\widehat{M}'$ are polynomials with integer coefficients of absolute value at most γ (by Claim 24). Thus, every coefficient of $\mathsf{det}(\widehat{M}')$ is an integer that has absolute value at most $(2\gamma)^n$. Thus, any root of $\mathsf{det}(\widehat{M}')$ must have absolute value at most $(2\gamma)^n + 1$. Since Γ is larger than this upper bound, it cannot be a root of $\mathsf{det}(\widehat{M}')$. This concludes the proof of this lemma. ◀

Now we can obtain a representation as a corollary of Proposition 14 and Lemma 23.

▶ **Theorem 1** (Quasipolynomial Representation of Transversal Matroids). *Let G be an n-vertex bipartite graph with a fixed vertex bipartition (A, B), and let r be the size of a maximum matching in G. Then, in time $r^{\mathcal{O}(\log r)} n^{\mathcal{O}(1)}$ we can compute a representation (M, ϕ) of the transversal matroid of G, where each entry of the matrix is an element of $\mathbb{Q}$ that can be encoded in $r^{\mathcal{O}(\log r)}$ bits.*

Proof. By Proposition 14, we can compute in time $r^{\mathcal{O}(\log r)}$ an r-isolating family $\mathcal{W}$ for the class of all bipartite graphs such that $|\mathcal{W}| = r^{\mathcal{O}(\log r)}$, and the maximum weight assigned by any function in $\mathcal{W}$ is $b = 2^{\mathcal{O}(\log^2 r)} = r^{\mathcal{O}(\log r)}$. Then, by Lemma 23, in time $(nb|\mathcal{W}|)^{\mathcal{O}(1)} = r^{\mathcal{O}(\log r)} n^{\mathcal{O}(1)}$ we can compute a representation (M, ϕ) of the transversal matroid of G, where each entry of M is a integer encoded in $bn^{\mathcal{O}(1)}|\mathcal{W}|^2 \log |\mathcal{W}| = r^{\mathcal{O}(\log r)} n^{\mathcal{O}(1)}$ bits. ◀

3.1 Representation of Gammoids in Deterministic Quasipolynomial Time

We use the representation of transversal matroids to obtain a representation of gammoids. First, we consider strict gammoids, and then obtain the result for (not necessarily strict) as an almost direct consequence. In this context, recall that strict gammoids are duals of transversal matroids, which is central to the proof.

▶ **Lemma 25** ($\star$). [1] *Let G be an n-vertex (di)graph and let $S \subseteq V(G)$. Then, in time $n^{\mathcal{O}(\log n)}$ we can compute a representation (M, ϕ) of the strict gammoid of G with respect to S, where each entry of the matrix is an element of $\mathbb{Q}$ that can be encoded in $n^{\mathcal{O}(\log n)}$ bits.*

▶ **Theorem 2** (Quasipolynomial Representation of Gammoids). *Let G be an n-vertex (di)graph and let $S, T \subseteq V(G)$. Then, in time $n^{\mathcal{O}(\log n)}$ we can compute a representation (M, ϕ) of the gammoid of G with respect to S on ground set T, where each entry of the matrix is an element of $\mathbb{Q}$ that can be represented in $n^{\mathcal{O}(\log n)}$ bits.*

Proof. Consider the strict gammoid of G with respect to S; then, as noted in Section 2, the gammoid of G with respect to S on ground set T can be obtained by deleting $V(G) \setminus T$ from the strict gammoid. Thus, given a representation matrix M of the strict gammoid of G with respect to S, by deleting the columns indexed by $V(G) \setminus T$ we obtain a representation matrix of the gammoid of G with respect to S on ground set T. Thus, the statement is a direct consequence of Lemma 25. ◀

3.2 Representative Sets

We now proceed to the computation of representative sets over gammoids in deterministic quasipolynomial time. We require the following result of Fomin et al. [6].

▶ **Proposition 26** ([6]). *Let $\mathcal{M} = (E, \mathcal{I})$ be a linear matroid of rank $p+q$, and let $\mathcal{S}$ be a family of ℓ independent sets, each of size p. Let A be an $n \times |E|$ matrix representing M over a field $\mathbb{F}$, and let $\omega < 2.373$ be the exponent of matrix multiplication [8]. Then, there is a deterministic algorithm computing $\widehat{\mathcal{S}} \subseteq^q_{rep} \mathcal{S}$ of size $\binom{p+q}{p}$ in $\mathcal{O}\left(\binom{p+q}{p}\ell p^\omega + \ell\binom{p+q}{q}^{\omega-1}\right) + (n + |E|)^{\mathcal{O}(1)}$ operations over $\mathbb{F}$.*

[1] The proof of results marked with $\star$ are omitted due to space constraints. They will appear in the full-version of the paper.

The algorithm of Proposition 26 requires the multiplication of at most $(p+q)^2$ entries of M for any computation during its execution. Hence, given a linear matroid with a representation where each entry requires $n^{\mathcal{O}(\log n)}$ bits, all intermediate results can be stored in $n^{\mathcal{O}(\log n)}$ bits as well. Hence, we obtain the following corollary.

▶ **Lemma 27.** *Let $p, q \in \mathbb{N}$. Let $\mathcal{M}$ be a linear matroid of rank $p+q$, and let $\mathcal{S}$ be a family of ℓ independent sets, each of size p. Let M be a representation of this matroid, where each entry of M is an element of $\mathbb{Q}$ that requires $n^{\mathcal{O}(\log n)}$ bits. Then, there is a deterministic algorithm computing $\widehat{\mathcal{S}} \subseteq^q_{rep} \mathcal{S}$ of size $\binom{p+q}{p}$ in $\mathcal{O}(\left(\binom{p+q}{p}\ell p^\omega + \ell\binom{p+q}{q}^{\omega-1}\right) \cdot n^{\mathcal{O}(\log n)})$ time, where ω is the exponent of matrix multiplication.*

4 Applications

In this section, we apply our results to derandomize several important results in the area of kernelization. Specifically, we apply Theorem 2 and Lemma 27 to derandomize the kernelization algorithms of [17, 16, 14], as well as the *cut covering lemma*. These algorithms, first compute an approximate solution for a given problem instance, and then construct a representation of an associated gammoid, where the construction deployed is a randomized one (also see [21]). To derandomize these results, we must ensure that there is a deterministic approximation algorithm for these problems, and then apply Theorem 2 and Lemma 27. For some of these problems, we have the following proposition from [15], that gives us deterministic polynomial time FPT-approximation algorithm. These results are obtained via parameter preserving reductions to a problem called d-Skew Symmetric Multicut. The other problems are known to admit standard approximation algorithms in deterministic polynomial time [11, 17].

▶ **Proposition 28** ([15]). Odd Cycle Transversal, Almost 2-SAT, Vertex Cover *above* $(2LP - MM)$, Vertex Cover *parameterized by* Konig Deletion Set, Konig Deletion Set *in graphs with a perfect matching,* RHorn-Backdoor Deletion Set *admit a deterministic polynomial time FPT-approximation, with approximation factor $\mathcal{O}(k)$.*

The subsequent step in these algorithms is to compute representative sets over a matroid of rank $(p+q)$ where p is a (polynomial) function of k and q is a constant. Applying our results, we directly obtain the following.

▶ **Lemma 29** (Cut Covering Lemma). *Let G be an n-vertex (di)graph and let $S, T \subseteq V(G)$. Let r be the size of a minimum (S,T)-cut in G. Then, a set $Z \subseteq V$ of cardinality $\mathcal{O}(|S| \cdot |T| \cdot r)$ such that for every $A \subseteq S$ and $B \subseteq T$, Z contains a minimum (A,B)-vertex cut, can be found in time $n^{\mathcal{O}(\log n)}$.*

▶ **Theorem 30.** *The following problems admit a polynomial kernel in deterministic quasi-polynomial (i.e., $n^{\mathcal{O}(\log n)}$) time:* Odd Cycle Transversal, MultiwayCut *with deletable terminals,* Group Feedback Vertex Set, Almost 2-SAT, Vertex Cover *above* $(2LP - MM)$, Vertex Cover *parameterized by* Konig Deletion Set, Konig Deletion Set *in graphs with a perfect matching,* RHorn-Backdoor Deletion Set *and* Subset Feedback Vertex Set.

5 Representation of Matching Matroids

In this section, we give a quasi-polynomial time representation for *Matching Matroid* which are a generalization of Transversal Matroid.

▶ **Definition 31** (Matching Matroid, [20]). *Let G be a graph. The* matching matroid *of G is a matroid $(V(G), \mathcal{I})$ where $\mathcal{I}$ is the family that consists of every subset $X \subseteq V(G)$ such that there exists a matching that saturates X.*

▶ **Observation 32.** *Let G be a graph and let $\mathcal{M}$ be the matching matroid of G. Then, $X \subseteq V(G)$ is a basis of $\mathcal{M}$ if and only if there is a maximum matching μ in G such that $V(\mu) = X$.*

Our algorithm for computing a representation of matching matroids, requires the *Gallai-Edmonds* decomposition of graphs from matching theory. It is defined as follows.

▶ **Definition 33** (Gallai-Edmonds Decomposition). *Let G be a graph. Let $V(G) = O \uplus N \uplus P$ be a partition of $V(G)$ defined as follows:*

- *$O \subseteq V(G)$ contains all vertices v such that there is a maximum matching μ in G and $v \notin V(\mu)$.*
- *$N = N_G(O)$ is the set of all neighbors of O in G.*
- *$P = V(G) \setminus (N \cup O)$, are the remaining vertices.*

▶ **Proposition 34** ([4, 9, 10, 20]). *Let G be a graph, and let $V(G) = O \uplus N \uplus P$ be the Gallai-Edmonds decomposition of G. Then the following hold:*

- *$G[O]$ is a collection of factor-critical graphs.*
- *$G[P]$ has a perfect matching.*
- *A matching μ in G is a maximum matching if and only if the following hold.*
  - *$P \cup N \subseteq V(\mu)$; furthermore $\mu \cap E(G[P])$ is a perfect matching in $G[P]$.*
  - *The matching partners of vertices in N are in O.*
  - *For each component O_i of $G[O]$, $\mu \cap E(O_i)$ matches all but one vertex $o_i \in V(O_i)$. The vertex o_i is either matched to a vertex in N or it remains unmatched.*

▶ **Proposition 35** ([4, 9, 10, 20]). *Given a graph G, a Gallai-Edmonds decomposition can be computed in polynomial time.*

▶ **Observation 36.** *Let G be a graph, $\mathcal{M}$ be it's matching matroid and let X be a basis of $\mathcal{M}$. Then $P \cup N \subseteq X$. And for any component O_i of $G[O]$, $|V(O_i) \cap X| \geq |V(O_i)| - 1$.*

The above observation implies that in any basis X of $\mathcal{M}$, the "non-trivial" part arises from the "bipartite graph" between N and O. Intuitively, we can capture this part via a transversal matroid. The following construction formalizes this intuition.

▶ **Lemma 37** (⋆). *Given a graph G, there exists a bipartite graph H with vertex bipartition $(U \uplus W)$ such that the matching matroid $\mathcal{M}$ of G is isomorphic to the transversal matroid of H with ground set U.*

From Lemma 37 and Theorem 1 we obtain the following.

▶ **Theorem 3** (Quasipolynomial Representation of Matching Matroids). *Let G be an n-vertex graph, and let r be the size of a maximum matching in G. Then, in time $r^{\mathcal{O}(\log r)} n^{\mathcal{O}(1)}$ we can compute a representation (M, ϕ) of the matching matroid of G, where each entry of the matrix is an element of $\mathbb{Q}$ that can be encoded in $r^{\mathcal{O}(\log r)}$ bits.*

References

1 Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015. doi:10.1007/978-3-319-21275-3.

2 R A DeMillo and R J Lipton. A probabilistic remark on algebraic program testing. *Inform. Process Lett.*, 7(4):193–195, 1978. doi:10.1016/0020-0190(78)90067-4.

3 Rodney G. Downey and Michael R. Fellows. *Fundamentals of Parameterized Complexity*. Texts in Computer Science. Springer, 2013. doi:10.1007/978-1-4471-5559-1.

4 Jack Edmonds. Paths, trees, and flowers. *Canadian Journal of mathematics*, 17:449–467, 1965.

5 Stephen A. Fenner, Rohit Gurjar, and Thomas Thierauf. Bipartite perfect matching is in quasi-nc. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 754–763, 2016. doi:10.1145/2897518.2897564.

6 Fedor V. Fomin, Daniel Lokshtanov, Fahad Panolan, and Saket Saurabh. Efficient computation of representative families with applications in parameterized and exact algorithms. *J. ACM*, 63(4):29:1–29:60, 2016. doi:10.1145/2886094.

7 F.V. Fomin, D. Lokshtanov, S. Saurabh, and M. Zehavi. *Kernelization: Theory of Parameterized Preprocessing*. Cambridge University Press, 2018.

8 François Le Gall. Powers of tensors and fast matrix multiplication. In *International Symposium on Symbolic and Algebraic Computation, ISSAC '14, Kobe, Japan, July 23-25, 2014*, pages 296–303, 2014. doi:10.1145/2608628.2608664.

9 Tibor Gallai. Kritische graphen ii. *Magyar Tud. Akad. Mat. Kutato Int. Kozl.*, 8:373–395, 1963.

10 Tibor Gallai. Maximale systeme unabhangiger kanten. *Magyar Tud. Akad. Mat. Kutato Int. Kozl.*, 9:401–413, 1964.

11 Naveen Garg, Vijay V Vazirani, and Mihalis Yannakakis. Multiway cuts in node weighted graphs. *Journal of Algorithms*, 50(1):49–61, 2004. doi:10.1016/S0196-6774(03)00111-1.

12 Rohit Gurjar and Thomas Thierauf. Linear matroid intersection is in quasi-nc. *Comput. Complex.*, 29(2):9, 2020. doi:10.1007/S00037-020-00200-Z.

13 Danny Hermelin and Xi Wu. Weak compositions and their applications to polynomial lower bounds for kernelization. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2012, Kyoto, Japan, January 17-19, 2012*, pages 104–113, 2012. doi:10.1137/1.9781611973099.9.

14 Eva-Maria C. Hols and Stefan Kratsch. A randomized polynomial kernel for subset feedback vertex set. In *33rd Symposium on Theoretical Aspects of Computer Science, STACS 2016, February 17-20, 2016, Orléans, France*, pages 43:1–43:14, 2016. doi:10.4230/LIPICS.STACS.2016.43.

15 Sudeshna Kolay, Pranabendu Misra, MS Ramanujan, and Saket Saurabh. Faster graph bipartization. *J. Comput. Syst. Sci.*, 109:45–55, 2020. doi:10.1016/j.jcss.2019.11.001.

16 Stefan Kratsch. A randomized polynomial kernelization for vertex cover with a smaller parameter. In *24th Annual European Symposium on Algorithms, ESA 2016, August 22-24, 2016, Aarhus, Denmark*, pages 59:1–59:17, 2016. doi:10.4230/LIPICS.ESA.2016.59.

17 Stefan Kratsch and Magnus Wahlström. Representative sets and irrelevant vertices: New tools for kernelization. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012*, pages 450–459, 2012. doi:10.1109/FOCS.2012.46.

18 Robert Krauthgamer and Inbal Rika. Mimicking networks and succinct representations of terminal cuts. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2013, New Orleans, Louisiana, USA, January 6-8, 2013*, pages 1789–1799, 2013. doi:10.1137/1.9781611973105.128.

19 Daniel Lokshtanov, Pranabendu Misra, Fahad Panolan, Saket Saurabh, and Meirav Zehavi. Quasipolynomial representation of transversal matroids with applications in parameterized complexity. In *9th Innovations in Theoretical Computer Science Conference, ITCS 2018, January 11-14, 2018, Cambridge, MA, USA*, pages 32:1–32:13, 2018. doi:10.4230/LIPICS.ITCS.2018.32.

20 László Lovász and Michael D Plummer. *Matching theory*, volume 367. American Mathematical Soc., 2009.

21 Dániel Marx. A parameterized view on matroid optimization problems. *Theor. Comput. Sci.*, 410(44):4471–4479, 2009. doi:10.1016/j.tcs.2009.07.027.

22 Pranabendu Misra, Fahad Panolan, M. S. Ramanujan, and Saket Saurabh. Linear representation of transversal matroids and gammoids parameterized by rank. In *Computing and Combinatorics - 23rd International Conference, COCOON 2017, Hong Kong, China, August 3-5, 2017, Proceedings*, pages 420–432, 2017. doi:10.1007/978-3-319-62389-4_35.

23 Ketan Mulmuley, Umesh V. Vazirani, and Vijay V. Vazirani. Matching is as easy as matrix inversion. *Combinatorica*, 7(1):105–113, 1987. doi:10.1007/BF02579206.

24 Moni Naor, Leonard J. Schulman, and Aravind Srinivasan. Splitters and near-optimal derandomization. In *36th Annual Symposium on Foundations of Computer Science, Milwaukee, Wisconsin, 23-25 October 1995*, pages 182–191, 1995. doi:10.1109/SFCS.1995.492475.

25 J.G. Oxley. *Matroid Theory*. Oxford graduate texts in mathematics. Oxford University Press, 2006.

26 J T Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. Assoc. Comput. Mach.*, 27(4):701–717, 1980. doi:10.1145/322217.322225.

27 Magnus Wahlström. Quasipolynomial multicut-mimicking networks and kernels for multiway cut problems. *ACM Transactions on Algorithms (TALG)*, 18(2):1–19, 2022. doi:10.1145/3501304.

28 Magnus Wahlström. Representative set statements for delta-matroids and the mader delta-matroid. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 780–810. SIAM, 2024. doi:10.1137/1.9781611977912.31.

29 R Zippel. Probabilistic algorithms for sparse polynomials. In *EUROSAM*, pages 216–226, 1979.