NON-COMMUTATIVE EDMONDS' PROBLEM AND MATRIX SEMI-INVARIANTS

GÁBOR IVANYOS, YOUMING QIAO, AND K. V. SUBRAHMANYAM

Abstract. In 1967, Edmonds introduced the problem of computing the rank over the rational function field of an $n \times n$ matrix T with integral homogeneous linear polynomials. In this paper, we consider the *non-commutative version of Edmonds' problem*: compute the rank of T over the free skew field. This problem has been proposed, sometimes in disguise, from several different perspectives, e.g. in the study of matrix spaces of low rank (Fortin-Reutenauer, 2004), in the study of the original Edmonds' problem (Gurvits, 2004), and more recently, in the study of non-commutative arithmetic circuits with divisions (Hrubeš and Wigderson, 2014).

It is known that this problem relates to the following invariant ring, which we call the $\mathbb{F}$-*algebra of matrix semi-invariants*, denoted as $R(n, m)$. For a field $\mathbb{F}$, it is the ring of invariant polynomials for the action of $\mathrm{SL}(n, \mathbb{F}) \times \mathrm{SL}(n, \mathbb{F})$ on tuples of matrices – $(A, C) \in \mathrm{SL}(n, \mathbb{F}) \times \mathrm{SL}(n, \mathbb{F})$ sends $(B_1, \ldots, B_m) \in M(n, \mathbb{F})^{\oplus m}$ to $(AB_1C^{\mathrm{T}}, \ldots, AB_mC^{\mathrm{T}})$. Then those T with non-commutative rank $< n$ correspond to those points in the nullcone of $R(n, m)$. In particular, if the nullcone of $R(n, m)$ is defined by elements of degree $\leq \sigma$, then there follows a $\mathrm{poly}(n, \sigma)$-time randomized algorithm to decide whether the non-commutative rank of T is full. To our knowledge, previously the best bound for σ was $O(n^2 \cdot 4^{n^2})$ over algebraically closed fields of characteristic 0 (Derksen, 2001).

In this article we prove the following results:

- We observe that by using an algorithm of Gurvits, and assuming the above bound σ for $R(n, m)$ over $\mathbb{Q}$, deciding whether or not T has non-commutative rank $< n$ over $\mathbb{Q}$ can be done *deterministically* in time polynomial in the input size and σ.
- When $\mathbb{F}$ is large enough, we devise a deterministic algorithm for non-commutative Edmonds' problem in time polynomial in $(n+1)!$, with the following consequences.
  - If the commutative rank and the non-commutative rank of T differ by a constant, then there exists a randomized efficient algorithm that computes the non-commutative rank of T. This improves a result of Fortin and Reutenauer, who gave a randomized efficient algorithm to decide whether the commutative and non-commutative ranks are equal.
  - We prove that $\sigma \leq (n+1)!$. This not only improves the bound obtained from Derksen's work over algebraically closed field of characteristic 0 but, more importantly, also provides for the first time an explicit bound on σ for matrix semi-invariants over fields of positive characteristics. Furthermore, this does not require $\mathbb{F}$ to be algebraically closed.

Date: August 10, 2015.

2010 *Mathematics Subject Classification.* Primary 13A50, 68W30.

Key words and phrases. Edmonds' problem, symbolic determinant identity test, semi-invariants of quivers, non-commutative rank.

Research was supported in part by the Hungarian Research Fund (OTKA).

Research was supported by the Australian Research Council DECRA DE150100720.

1. Introduction

1.1. **Non-commutative Edmonds' problem.** In 1967, Edmonds introduced the following problem [24]: let $X = \{x_1, \ldots, x_m\}$ be a set of variables. Given an $n \times n$ matrix T whose entries are homogeneous linear polynomials from $\mathbb{Z}[X]$, determine the rank of T over the rational function field $\mathbb{Q}(X)$, denoted as $\mathrm{rk}(T)$. The decision version of Edmonds' problem is to decide whether T is of full rank or not; this decision version is better known now as the symbolic determinant identity testing (SDIT) problem. It is natural to consider this problem over any field $\mathbb{F}$. If $|\mathbb{F}|$ is constant, this problem is NP-hard [7]. This is not the setting we are concerned with - we will always assume $|\mathbb{F}|$ to be at least $\Omega(n)$.

When $|\mathbb{F}| \geq 2n$, the Schwartz-Zippel lemma provides a randomized efficient algorithm. To devise a deterministic efficient algorithm has a long history, and is of fundamental importance in complexity theory. Originally, the main motivation was its applications to certain combinatorial problems, most notably the maximum matching problem on graphs, as exploited by Tutte [56], Edmonds [24], Lovász [43], among others.[1] Since 2003, a major incentive to study SDIT arises from its implications to circuit lower bounds, as shown in the wonderful work by Kabanets and Impagliazzo [41]: they showed that such an algorithm implies either permanent does not have polynomial-size arithmetic formulas, or NEXP $\not\subseteq$ P/poly.

One interesting instance of Edmonds' problem is module isomorphism. To be more specific, assume that we are given two n-dimensional modules U and U' for the free algebra $\mathcal{A}$ over $\mathbb{F}$ with k generators as k-tuples $G_1, \ldots, G_k$ and $G'_1, \ldots, G'_k$ of n by n matrices. Then $\mathrm{Hom}_{\mathcal{A}}(U, U')$ is the $\mathbb{F}$-linear subspace of $\mathrm{Hom}_{\mathbb{F}}(U, U)'$, identified with $M(n, \mathbb{F})$, consising of matrices X with $XG_i = G'_i X$ $(i = 1, \ldots, k)$. As these conditions are linear in the entries of X, the space $\mathrm{Hom}_{\mathcal{A}}(U, U')$ can be obtained by solving a system of homogeneous linear equations in n^2 elements. Furthermore, U' is isomorphic to U if and only if there exists a nonsingular matrix in $\mathrm{Hom}_{\mathcal{A}}(U, U')$. In turn, any such nonsingular matrix will be an isomorphism, and, by the Schwartz-Zippel lemma, for sufficiently large base field a random homomorphism will be an isomorphism. Due to the special algebraic structure behind this problem, it can be solved even by deterministic polynomial-time methods, see the method of Chistov, Ivanyos and Karpinski [8] working over many fields, or a different approach of Brooksbank and Luks [4] which works over arbitrary fields, and an extension of the first method to arbitrary fields given by Ivanyos, Karpinski and Saxena in [39]. Interestingly, the general case of finding a surjective or injective homomorphism between non-isomorphic modules deterministically turns out to be as hard as the constructive version of the general Edmonds' problem [39].

In this paper, we propose to study Edmonds' problem in the non-commutative setting. In other words, we view the entries of T as elements of $\mathbb{F}\langle X \rangle$, the algebra of non-commutative polynomials over $\mathbb{F}$. To state this, we need a non-commutative counterpart of the rational function field. Note that, due to non-commutativity, the best we can hope for is a skew field. The *free skew field* introduced by P. M. Cohn [10] is the non-commutative analogue of the rational function field. We refer readers

[1] In these applications, T is usually of certain specific forms, for example, as a mixed matrix: each entry is a single variable or a field element, and each variable appears only once.

to Cohn's books [10, 11] for a comprehensive introduction to the free skew field.[2] For our purpose here, we only point out that for T, a matrix with homogeneous linear polynomials, the rank of T over the skew free field is the minimum $s \in \mathbb{Z}^+$ s.t. T can be written as PQ, where P (resp. Q) is a $n \times s$ matrix (resp. $s \times n$) with entries from $\mathbb{F}\langle X\rangle$ ([10], see also [27]).

Fortin and Reutenauer [27] defined the *non-commutative rank* of T, denoted as $\mathrm{ncrk}(T)$, as its rank over the free skew field. By the *non-commutative Edmonds' problem* we mean the problem of computing $\mathrm{ncrk}(M)$, and by the *non-commutative symbolic determinant identity testing problem* (non-commutative SDIT) we mean the problem of deciding whether $\mathrm{ncrk}(M)$ is full or not.

It will be clear soon that $\mathrm{rk}(T) \leq \mathrm{ncrk}(T)$, and Fortin and Reutenauer showed that $\mathrm{ncrk}(T) \leq 2\mathrm{rk}(T)$, and exhibited an example T for which $\mathrm{ncrk}(T) = 3/2 \cdot \mathrm{rk}(T)$ [27]. In [12], Cohn and Reutenauer presented an algorithm to decide whether $\mathrm{ncrk}(T)$ is full or not[3], but its time complexity can only be bounded as double exponential in n, due to the difficulty of testing the solvability of a system of multivariate polynomial equations. Unlike its commutative counterpart, it is not even clear that non-commutative Edmonds' problem has a randomized efficient algorithm. In Section 1.3, we will discuss a natural randomized algorithm for the non-commutative SDIT, but its efficiency will depend on an invariant-theoretic quantity.

1.2. **Equivalent formulations of non-commutative Edmonds' problem.** Recently, Hrubeš and Wigderson proposed the non-commutative SDIT problem, in their study of non-commutative arithmetic circuits with divisions [36]. It appears to be less known however, that in 2003, Gurvits had already posed this problem – but in a different form, see below – in his remarkable work on Edmonds' problem [33]. Indeed, a very intriguing feature of non-commutative Edmonds' problem is the existence of several interesting equivalent formulations. Instead of relying on the free skew field, these formulations use either just linear algebra, or concepts from invariant theory, or quantum information theory. They are scattered in the literature, so we collect them here, to illustrate the various facets of this problem, introduce some previous works, and motivate the study of non-commutative Edmonds' problem.

To state these formulations we need some notations. $M(n, \mathbb{F})$ denotes the linear space of $n \times n$ matrices over $\mathbb{F}$. A linear subspace of $M(n, \mathbb{F})$ is called a *matrix space*. Given T, a matrix of linear forms in variables $X = \{x_1, \ldots, x_m\}$, write $T = x_1B_1 + x_2B_2 + \cdots + x_mB_m$ where $B_i \in M(n, \mathbb{F})$. Let $\mathcal{B} := \langle B_1, \ldots, B_m\rangle$, where $\langle\cdot\rangle$ denotes linear span. The rank of $\mathcal{B}$, denoted as $\mathrm{rk}(\mathcal{B})$, is defined as $\max(\mathrm{rk}(B) \mid B \in \mathcal{B})$. We call $\mathcal{B}$ *singular*, if $\mathrm{rk}(\mathcal{B}) < n$. When $|\mathbb{F}| > n$, as we will assume throughout, $\mathrm{rk}(T) = \mathrm{rk}(\mathcal{B})$.[4] We shall soon see that $\mathrm{ncrk}(T)$ corresponds to some property of $\mathcal{B}$ as well, so that we can translate the study of commutative and non-commutative ranks of T entirely to the study of $\mathcal{B}$.

Some of these formulations make sense only subject to certain conditions. In such cases we indicate the conditions needed before that formulation.

---

[2] Hrubeš and Wigderson's paper ([36], full version) contains a nice introduction to non-commutative rational functions and the free skew field from the perspective of algebraic computations.

[3] As remarked in [27], this algorithm can be generalized to compute $\mathrm{ncrk}(T)$.

[4] As when the field size is large enough, the zero set of a nonzero polynomial is non-empty.

(1) Given $\mathcal{B} = \langle B_1, \dots, B_m \rangle \leq M(n, \mathbb{F})$, a subspace $U \leq \mathbb{F}^n$ is called a *c-shrunk subspace* of $\mathcal{B}$, for $c \in \mathbb{N}$, if there exists $W \leq \mathbb{F}^n$, s.t. $\dim(W) \leq \dim(U) - c$, and for every $B \in \mathcal{B}$, $B(U) \leq W$. U is called a shrunk subspace of $\mathcal{B}$, if it is a c-shrunk subspace for some $c \in \mathbb{Z}^+$.

Question: compute the maximum c s.t. there exists a c-shrunk subspace.

Remark: The relation between the existence of c-shrunk subspaces and the non-commutative rank was shown by Fortin and Reutenauer [27, Theorem 1]. Their motivation to consider this problem was to connect matrices over linear forms on one hand, and matrix spaces of low rank on the other. The latter topic was studied in e.g. [3, 25].

Using this formulation we can define the *non-commutative rank of* $\mathcal{B}$ as

$$n - \max(c \in \{0, 1, \dots, n\} \mid \exists c\text{-shrunk subspace of } \mathcal{B}).$$

Then $\mathrm{ncrk}(\mathcal{B}) = \mathrm{ncrk}(T)$. So we shall identify T with $\mathcal{B}$ in the following.

(2) ($\mathbb{F}$ is large enough) Given $\mathcal{B} = \langle B_1, \dots, B_m \rangle \leq M(n, \mathbb{F})$, the dth tensor blow-up of $\mathcal{B}$, is $\mathcal{B}^{[d]} := M(d, \mathbb{F}) \otimes \mathcal{B} \leq M(dn, \mathbb{F})$. It is clear that $\mathrm{rk}(\mathcal{B}^{[d]}) \geq d \cdot \mathrm{rk}(\mathcal{B})$. We shall prove that when $\mathbb{F}$ is large enough, then d always divides $\mathrm{rk}(\mathcal{B}^{[d]})$. Furthermore, when $d > n$, then $\mathrm{rk}(\mathcal{B}^{[d+1]})/(d+1) \geq \mathrm{rk}(\mathcal{B}^{[d]})/d$. See Lemma 4.5, Corollary 4.6, and Remark 4.4.

Question: compute $\lim_{d \to \infty} \mathrm{rk}(\mathcal{B}^{[d]})/d$.

Remark: the non-commutative SDIT is equivalent to decide whether $\mathrm{rk}(\mathcal{B}^{[d]}) = nd$ for some d, already shown by Hrubeš and Wigderson [36]. Our formulation here is a straightforward quantitative generalization of their statement. Hrubeš and Wigderson's motivation was to study non-commutative arithmetic formulas *with divisions*. In particular, they show that identity testing in this model can be reduced deterministically to the non-commutative SDIT.

(3) ($\mathbb{F} = \mathbb{C}$) Given $B_1, \dots, B_m \in M(n, \mathbb{Q})$, construct a completely positive operator $P : M(n, \mathbb{C}) \to M(n, \mathbb{C})$, sending $A \to \sum_{i \in [m]} B_i A B_i^\dagger$. For $c \in \mathbb{N}$, P is called rank c-decreasing, if there exists a positive semidefinite A, s.t. $\mathrm{rk}(A) - \mathrm{rk}(P(A)) = c$.

Question: compute the maximum c s.t. P is rank c-decreasing.

Remark: Gurvits stated the problem of deciding whether P is rank non-decreasing or not [33]. His original motivation was to study the original (commutative) Edmonds' problem, and the main result in [33] solves the case when the commutative and non-commutative SDIT coincide (see Theorem 3.1).

(4) (Non-commutative SDIT) Consider the action of $(A, C) \in \mathrm{SL}(n, \mathbb{F}) \times \mathrm{SL}(n, \mathbb{F})$ on a tuple of matrices $(B_1, \dots, B_m) \in M(n, \mathbb{F})^{\oplus m}$ by sending it to $(AB_1C^{\mathrm{T}}, \dots, AB_mC^{\mathrm{T}})$.[5] Let $R(n, m)$ be the ring ($\mathbb{F}$-algebra) of invariant polynomials w.r.t. this action. The *nullcone* of $R(n, m)$ is the common zero of all positive-degree polynomials in $R(n, m)$.

Question: decide whether or not $(B_1, \dots, B_m)$ is in the nullcone of R.

[5] This action can also be written as: (A, C) sending $(B_1, \dots, B_m)$ to $(AB_1C^{-1}, \dots, AB_mC^{-1})$. We adopt the transpose rather than the inverse, as the transpose yields a polynomial representation rather than a rational representation. Furthermore when Derksen's result is applied to the transpose, it gives a somewhat better bound (Fact 1.1).

That the original formulation is equivalent to (1) comes from [27]. While not difficult to prove directly, we refer to Theorem 4.8 for the equivalence between (1) and (2). The equivalence between (1) and (3) is straightforward. The equivalence among decision versions of (1) and (2), and (4) can be obtained via the ring of matrix semi-invariants, as described in Section 1.3.

To summarize, non-commutative Edmonds' problem can be derived naturally from the perspectives of quantum information theory,[6] and invariant theory. It is of great interests in non-commutative algebraic computation with divisions, and in the study of matrix spaces of low rank. Our motivation to study this is because a solution to the non-commutative Edmonds' problem will throw light on its commutative counterpart. Shrunk subspaces form a natural and important witness for the singularity of a matrix space. Therefore, if the non-commutative Edmonds' problem can be solved deterministically in polynomial time, it means that, for the commutative SDIT problem, the bottleneck lies in recognizing those singular matrix spaces without such witnesses. This connection will be detailed in Section 3.

1.3. **Matrix semi-invariants.** Formulation (1), (2) and (4) have a common origin, namely the invariant ring $R(n,m)$ described in (4). We shall call $R(n,m)$ *the ring of matrix semi-invariants*, as (1) it is closely related to the classical ring of matrix invariants [52] (see below for the definition, and [1,17] for the precise relationship between these two rings); and (2) it is the ring of semi-invariants of the representation of the m-Kronecker quiver with dimension vector (n,n). Here, the m-Kronecker quiver is the quiver with two vertices s and t, and m arrows pointing from s to t. When $m=2$, it is the classical Kronecker quiver. The reader is referred to [15,21,55] for a description of the semi-invariants for arbitrary quivers.

The equivalence between (1) and (4) comes from the observation that the $(B_1,\dots,B_m)$ with a shrunk subspace are exactly the points in the nullcone of $R(n,m)$ [1,5]. The equivalence between (2) and (4) can be seen from the first fundamental theorem (FFT) of matrix semi-invariants [1,15,21,55]. For this we need some notations: for $n\in\mathbb{N}$, $[n]:=\{1,\dots,n\}$. Note that $R(n,m)\subseteq\mathbb{F}[x_{i,j}^{(k)}]$ where $i,j\in[n]$, $k\in[m]$, and $x_{i,j}^{(k)}$ are independent variables. Let $X_k=(x_{i,j}^{(k)})_{i,j\in[n]}$. Then for $A_1,\dots,A_m\in M(d,\mathbb{F})$, $\det(A_1\otimes X_1+\cdots+A_m\otimes X_m)$ is a matrix semi-invariant, and every matrix semi-invariant is a linear combination of such polynomials. Therefore, $(B_1,\dots,B_m)$ is in the nullcone, if and only if for all $d\in\mathbb{Z}^+$ and all $(A_1,\dots,A_m)\in M(d,\mathbb{F})^{\oplus m}$, $A_1\otimes B_1+\cdots+A_m\otimes B_m$ is singular.

The second fundamental theorem (SFT) over fields of characteristic 0 is described in [40]. An explicit upper bound on $\beta(R(n,m))$ turns out to be particularly interesting for the purpose of the non-commutative SDIT problem. As already suggested by Hrubeš and Wigderson [36], if $R(n,m)$ has a degree bound $\beta=\beta(R(n,m))$, one can do the following: take m $d\times d$ variable matrices $Y_1,\dots,Y_m$, $Y_k=(y_{i,j}^{(k)})$ and form the polynomial

$$\det(Y_1\otimes B_1+\cdots+Y_m\otimes B_m)\in\mathbb{F}[y_{i,j}^{(k)}]_{k\in[m],i,j\in[d]}. \tag{1.1}$$

Letting d go from 1 to β, this system of polynomials characterizes $\mathrm{ncrk}(T)<n$: $\mathrm{ncrk}(T)<n$ if and only if all these polynomials are the zero polynomial. This

[6]It remains to investigate the physical meaning for a super-operator to be rank non-decreasing though.

immediately gives a randomized algorithm for the non-commutative SDIT in time $\mathrm{poly}(n, \beta)$ over large enough fields.

In fact, for the above application, what really matters is another important bound $\sigma = \sigma(R(n, m))$. This is defined as the minimum d with the property that $(B_1, \dots, B_m) \in M(n, \mathbb{F})^{\oplus m}$ is in the nullcone iff all polynomials of degree $\leq d$ in $R(n, m)$ vanish on $\{B_1, \dots, B_m\}$. The above reasoning goes obviously through even when β is replaced by σ. Clearly $\sigma \leq \beta$.

Over algebraically closed fields of characteristic 0, by directly employing Derksen's bounds for invariant rings satisfying certain general conditions [14], the following bound can be derived. For completeness we include a proof in Appendix A.

Fact 1.1 ([14]). *Over algebraically closed fields of characteristic* 0*, for* $R(n, m)$*,* $\beta \leq \max(2, 3/8 \cdot n^4 \cdot \sigma^2)$*, and* $\sigma \leq 1/4 \cdot n^2 \cdot 4^{n^2}$*.*

In particular, if σ is polynomial in n, then β is polynomial in n as well.

It is generally believed that over fields of characteristic 0, the bounds we get for $R(n, m)$ using Derksen's theorem is far from optimal. One reason to believe so is that $R(n, m)$ is closely related to another ring of invariants: let $A \in \mathrm{SL}(n, \mathbb{F})$ act on $(B_1, \dots, B_m) \in M(n, \mathbb{F})^{\oplus m}$ by conjugation – i.e A sends the tuple to $(AB_1A^{-1}, \dots, AB_mA^{-1})$. Denoted by $S(n, m)$, this is just the classical ring of matrix invariants [52]. The structure of $S(n, m)$ is well-understood. Over fields of characteristic 0, the FFT, the SFT, and an n^2 upper bound for $\beta(S(n, m))$ were established in 1970's, by the works of Procesi, Razmysolov, and Formanek [26,52,53]. See [1, 17] for the precise relationship between the rings $R(n, m)$ and $S(n, m)$. Note that when applied to $S(n, m)$ over characteristic 0, Derksen's bound yields $\beta(S(n, m)) \leq \max(2, 3/8 \cdot n^2 \cdot \sigma^2)$ and $\sigma(S(n, m)) = n^{O(n^2)}$, far from the n^2 bound given above. Another reason is that, for certain small m or n, explicit generating sets of $R(n, m)$ have been computed in e.g. [16, 17, 19, 40]. In these cases, elements of degree $\leq n^2$ generate the ring.[7]

If we turn to positive characteristic fields then, to our best knowledge, no explicit bounds for $\beta(R(n, m))$ nor $\sigma(R(n, m))$ have been derived. Note here that the relation between β and σ as in Fact 1.1 is not known to hold, due to the assumption on field characteristics there. This case is important, for example, in the application to identity testing, and division elimination for non-commutative arithmetic formulas with divisions over fields of positive characteristics [36]. For matrix invariants over fields of positive characteristics, FFT was established by Donkin in [22, 23], an n^3 upper bound for σ can be derived from [9, Proposition 9], and Domokos in [18, 20] proved an upper bound $O(n^7 m^n)$ on β.

1.4. **Our results.** In the previous sections, we defined non-commutative Edmonds' problem and non-commutative SDIT problem, and illustrated their connections to matrix semi-invariants. Indeed, our results suggest that progress on one topic helps to advance the other as well.

The first result shows that an upper bound for $\sigma(R(n, m))$ actually implies a *deterministic* algorithm for non-commutative SDIT over $\mathbb{Q}$, rather than just a randomized one as in Section 1.3.

---

[7] We thank M. Domokos for pointing out this fact to us.

Proposition 1.2. *Over $\mathbb{Q}$, if the nullcone of $R(n,m)$ is defined by elements of degree $\leq \sigma = \sigma(n,m)$, then there exists a deterministic algorithm that solves the non-commutative SDIT with bit complexity polynomial in σ and the input size.*

In particular, if σ is a polynomial in n and m, then the non-commutative SDIT can be solved deterministically in polynomial time over $\mathbb{Q}$. The key ingredient here is Gurvits' algorithm for the Edmonds' problem, although that algorithm works only under a promise [33]. The distinction between deterministic and probabilistic is important: as illustrated at the end of Section 1.2, our original motivation of studying the non-commutative SDIT is to gain understanding of the commutative SDIT, for which the question is to devise deterministic efficient algorithms.

Our main result is an algorithm that solves non-commutative Edmonds' problem using formulation (2). To ease the presentation, we give an informal statement of the main theorem in Section 4 (Theorem 4.8) here, and discuss its two consequences.

Theorem 1.3 (Theorem 4.8, informal)**.** *Given a matrix space $\mathcal{B} \leq M(n,\mathbb{F})$ over a large enough field, there exists a deterministic algorithm that computes* $\mathrm{rk}(\mathcal{B})$ *using* $\mathrm{poly}((n+1)!)$ *many arithmetic operations. Over $\mathbb{Q}$ the algorithm runs in time polynomial in the bit size of the input and $(n+1)!$.*[8]

In [27], Fortin and Reutenauer asked for "an algorithm which uses only linear-algebraic techniques." Indeed, the algorithm for Theorem 1.3 may be viewed as one, though it relies on certain routines dealing with objects from associative algebras.

Two interesting consequences now follow. Firstly, we have a randomized efficient algorithm to compute the non-commutative rank if it differs from the commutative rank by a constant. (Recall that $\mathrm{rk}(\mathcal{B}) \leq \mathrm{ncrk}(\mathcal{B}) \leq 2\mathrm{rk}(\mathcal{B})$.) Its easy proof is put after the statement of Theorem 4.8.

Corollary 1.4. *For $\mathcal{B} \leq M(n,\mathbb{F})$, let $c = \mathrm{ncrk}(\mathcal{B}) - \mathrm{rk}(\mathcal{B})$, and assume $\mathbb{F}$ is of size $\Omega(n\cdot(n+1)!)$. Then the non-commutative rank of $\mathcal{B}$ can be computed probabilistically in time polynomial in $(n+1)^{c+1}$.*

Secondly, we immediately obtain an explicit bound for $\sigma(R(n,m))$ as a consequence of Theorem 4.8. By Fact 1.1, we also get a bound on $\beta(R(n,m))$, over an algebraically closed field of characteristic 0. Its proof is also put after Theorem 4.8.

Corollary 1.5. *Over any field $\mathbb{F}$ of size $\Omega(n \cdot (n+1)!)$, $\sigma(R(n,m)) \leq (n+1)!$. If furthermore $\mathbb{F}$ is of characteristic 0 and algebraically closed, then $\beta(R(n,m)) \leq \max(2, 3/8 \cdot n^4 \cdot ((n+1)!)^2)$.*

This improves the bounds in Fact 1.1 over algebraically closed fields of characteristic 0. More importantly, to the best of our knowledge, this provides an explicit bound for $\sigma(R(n,m))$ over fields of positive characteristic for the first time. Furthermore to get this bound we only assume the field size to be large enough, whereas Fact 1.1 requires our field to be algebraically closed. In fact, we can improve this bound slightly to $n!/\lceil n/2 \rceil!$; see Remark 4.4.

While the improvement from $2^{O(n^2)}$ to $2^{O(n \log n)}$ is modest, we believe it is nonetheless an interesting improvement from the technical point of view: note that the dimension of $\mathrm{SL}(n,\mathbb{F})$ is n^2-1. In the line of research for bounds of an

[8] All algorithms presented in this paper, when working over $\mathbb{Q}$, have bit complexity polynomial in the input size, and some additional parameters. Sometimes, we may omit the input size but focus on those more important parameters.

invariant ring R w.r.t. a group G (cf. [14, 51]), the dimension of G has to stand on the exponent for $\sigma(R)$, and to get a bound as $2^{o(\dim(G))}$ seems difficult there. Furthermore, the idea of using correctness of algorithms to get bounds on quantities of interest in invariant theory seems new, and may deserve to be explored further.

1.5. **More previous works.**

Connections between invariant theory and complexity theory. The results in this paper suggest a new link between invariant theory and complexity theory. Connections between the two fields have been emerging in recent years. We have already alluded to the direct connection with noncommutative arithmetic circuits, in the work of Hrubeš and Wigderson [36] above. In the first of a series of papers titled geometric complexity theory (GCT), Mulmuley and Sohoni [47, 48] (see also [6, 45] pointed out possible deep connections between problems in invariant theory and complexity theory. GCT addresses the fundamental lower bound problems in complexity theory, e.g. the permanent versus determinant problem, by linking them to problems in representation theory and algebraic geometry. In particular, in [46], Mulmuley established a tight connection between derandomizing the Noether normalization lemma, and black-box derandomizing the polynomial identity test. The degree bounds of various invariant rings are of central importance in that work. We briefly remark that a polynomial bound for $\beta(R(n, m))$, if proven, will yield similar results as what the n^2 degree bound for $S(n, m)$ has yielded in [46].

More previous works on Edmonds' problem. Some earlier work on this problem was cited at the beginning of this article. Here we mention more related work.

Recall that one motivation to study Edmonds' problem is due to its implications to certain combinatorial problems. This line of research mostly focuses in the case when the given matrices are of particular form, e.g. rank-1 and certain generalizations [30, 35, 39, 49] as used in bipartite graph matchings, or skew symmetric rank-2 and certain generalizations [29, 31, 32] as used in general graph matchings.

Another line of research deals with matrix spaces that satisfy certain properties. Note that properties of matrix spaces should not depend on a particular basis. For example, we can define a property of matrix spaces as, "having a basis consisting of rank-1 matrices." So if $\mathcal{B}$ has a basis consisting of rank-1 matrices, $\mathcal{B}$ may not necessarily be presented using this rank-1 basis. We are not aware of any result on the complexity of finding rank-1 generators for rank-1 spanned matrix spaces, if it is given by a basis consisting of not necessarily rank one matrices. We believe that the problem is hard. Thus the results in [30, 35, 39, 49], which assume that the input is given by a rank-1 basis, do not translate to algorithms for rank-1 spanned matrix spaces.

As far as we are aware, there are two references for SDIT which assume only properties of matrix spaces. The first one is Gurvits' algorithm in [33]; this algorithm works over $\mathbb{C}$, and assumes a property which Gurvits called "Edmonds-Rado." His algorithm, put in the context of this paper, is rephrased as Theorem 3.1. Gurvits left open the problem of developing deterministic efficient algorithm for rank-1 spanned matrix spaces over finite fields. This was settled in affirmative in [38], the other reference that assumes properties of matrix spaces.

Recall that the other major incentive to study Edmonds' problem is to understand circuit lower bounds via [41]. We believe that for this goal, a better indication of progress is to use properties of matrix spaces, rather than properties of the given matrices. One reason is that, whether a matrix space contains a nonsingular matrix, is a property of matrix spaces. Another reason is that many properties of matrix spaces seem difficult to test algorithmically. Furthermore, note that in this paper we heavily rely on algorithmic techniques developed in [33] and [38]. This may be viewed as another evidence of the importance of working with properties of matrix spaces.

Connections to Kronecker coefficients. Recently, there was an interest in studying the semi-invariants of the m-Kronecker quivers due to its connection with the Kronecker coefficients [1,2,44], namely the multiplicities in the direct sum decompositions of the tensor products of two irreducible representations of symmetric groups. Giving a positive combinatorial description of these coefficients is considered to be one of the most important problem in the combinatorial representation theory of symmetric groups. This problem recently received much attention in complexity theory thanks to the geometric complexity theory (GCT). In GCT, these coefficients are expected to play a role in separating the determinant from the permanent. In [1,2,44], the authors exploit the fact that the $\mathrm{SL}(m,\mathbb{F})$-decomposition of matrix semi-invariants (acting on the copies of the matrices) gives Kronecker coefficients w.r.t. the rectangular shapes. These are just the cases of concern in GCT [37].

Organization. In Section 2 we present certain preliminaries. In Section 3 we give an exposition of the natural connection between commutative and non-commutative Edmonds' problem, and prove Proposition 1.2. In Section 4 we prove the formal version of Theorem 1.3 (Theorem 4.8) and deduce Corollary 1.4 and 1.5.

Update on a recent progress. In a recent advance [28], Garg et al. present a deterministic polynomial-time algorithm for computing the non-commutative rank over $\mathbb{Q}$. This is achieved via a closer analysis of Gurvits' algorithm [33]. This analysis relies crucially on the bound on σ in Corollary 1.5, though Derksen's bound suffices for their purpose as well.

2. Preliminaries

2.1. **Notation.** For the reader's convenience, we collect the main notation in this section. Some of this notation was already introduced in the introduction.

For $n \in \mathbb{N}$, $[n] := \{1, \dots, n\}$. Given two vector spaces U, V, $U \leq V$ denotes that U is a subspace of V. $\mathbf{0}$ denotes the zero vector.

$M(n,\mathbb{F})$ denotes the linear space of $n \times n$ matrices over $\mathbb{F}$. The rank of $A \in M(n,\mathbb{F})$ is denoted $\mathrm{rk}(A)$. The notation $\mathrm{cork}(A)$ is the corank of A, which is equal to $n - \mathrm{rk}(A)$.

A linear subspace of $M(n,\mathbb{F})$ is called a *matrix space*. For $B_i \in M(n,\mathbb{F})$, $i \in [m]$, the notation $\langle B_1, \dots, B_m \rangle$ denotes the matrix space spanned by the B_i's. For a matrix space $\mathcal{B}$, $\mathrm{rk}(\mathcal{B})$, is defined as $\max(\mathrm{rk}(B) \mid B \in \mathcal{B})$. We call $\mathcal{B}$ *singular*, if $\mathrm{rk}(\mathcal{B}) < n$. For $\mathcal{B} \in M(n,\mathbb{F})$ and $U \leq \mathbb{F}^n$ we use the notation $\mathcal{B}(U)$ to denote $\langle \cup_{B \in \mathcal{B}} B(U) \rangle$. The non-commutative rank, $\mathrm{ncrk}(\mathcal{B})$, is defined as the maximum $c \in \mathbb{N}$ s.t. there exists a c-shrunk subspace of $\mathcal{B}$.

For two matrices $A \in M(d,\mathbb{F})$, $B \in M(n,\mathbb{F})$, $A \otimes B$ is the element of $M(dn,\mathbb{F})$ given by the tensor product of matrices A, B. This is an $d \times d$ block matrix, with the size of each block (i,j), $1 \leq i,j \leq n$, being $n \times n$. For $A = (a_{i,j})_{i,j \in [d]}$, the entries in the (i,j)-th block of $A \otimes B$ are $a_{i,j}B$. For $\mathcal{B} = \langle B_1, \dots, B_m \rangle \leq M(n,\mathbb{F})$, the *dth tensor blow-up*, $\mathcal{B}^{[d]} := M(d, \mathbb{F}) \otimes \mathcal{B} \leq M(dn, \mathbb{F})$. This is the matrix space in $M(dn, \mathbb{F})$, spanned by $\{E_{i,j} \otimes B_k \mid i, j \in [d], k \in [m]\}$, where $E_{i,j}$ is the matrix with 1 at the (i, j)th position, and 0 otherwise. In Section 4 it will be easier to work with $\mathcal{B}^{\{d\}} := \mathcal{B} \otimes M(d, \mathbb{F})$. As $\mathcal{B}^{[d]} \cong \mathcal{B}^{\{d\}}$, the latter will also referred to as the dth tensor blow-up of $\mathcal{B}$.

2.2. **The second Wong sequences.** Let us introduce a key tool to be used in Section 4, called the second[9] (generalized) Wong sequence. This was used by Fortin and Reutenauer [27], and rediscovered by the first two authors with Karpinski and Santha in [38] to solve Edmonds' problem for rank-1 spanned matrix spaces over arbitrary fields.

Given $(A, \mathcal{B})$, $A \in M(n, \mathbb{F})$ and $\mathcal{B} \leq M(n, \mathbb{F})$, the second Wong sequence of $(A, \mathcal{B})$ is the following sequence of subspaces in $\mathbb{F}^n$: $W_0 = \mathbf{0}$, $W_1 = \mathcal{B}(A^{-1}(W_0))$, $\dots$, $W_i = \mathcal{B}(A^{-1}(W_{i-1}))$, $\dots$. It can be proved that $W_0 < W_1 < W_2 < \dots < W_\ell = W_{\ell+1} = \dots$ for some $\ell \in \{0, 1, \dots, n\}$. W_ℓ is then called the limit of this sequence, denoted as W^*.

A useful way to understand the second Wong sequence is to view it as a linear algebraic analogue of the augmenting path on bipartite graphs. While not precise, we find this intuition helpful. That is, we view matrices as linear maps from V to W, $V \cong W \cong \mathbb{F}^n$. Vectors in V and W may be thought of as the "vertices" on the left and right part, respectively. Then for $A \in \mathcal{B}$, thinking of A as a given matching, $A^{-1}(\mathbf{0})$ can be understood as identifying those "vertices unmatched by A on the left part." Then $\mathcal{B}(A^{-1}(\mathbf{0}))$ is understood as taking those "edges" outside A, and $A^{-1}(\mathcal{B}(A^{-1}(0)))$ is understood as taking a further step with those "edges" in A. And so on.

The key fact is that, when $A \in \mathcal{B}$, $W^* \leq \mathrm{im}(A)$ if and only if there exists a cork(A)-shrunk subspace [38, Lemma 9]. If this is the case, A is of maximum rank and $A^{-1}(W^*)$ is a cork(A)-shrunk subspace. It is clear that the second Wong sequence can be computed using polynomially many arithmetic operations. The direct way to compute the second Wong sequences over $\mathbb{Q}$ may cause the bit lengths to explode. If testing whether $W^* \leq \mathrm{im}(A)$ is the only concern (as in our application here), by replacing A^{-1} with some appropriate "pseudo-inverse" of A, the bit lengths of the intermediate numbers up to the first W_k, $W_k \not\leq \mathrm{im}(A)$, can be bounded by a polynomial of the input size. We refer the reader to [38, Lemma 10] for this trick.

When $|\mathbb{F}|$ is $\Omega(n)$, this immediately gives a method to decide whether $\mathrm{ncrk}(\mathcal{B}) = \mathrm{rk}(\mathcal{B})$ as in [27]: randomly choose a matrix $A \in \mathcal{B}$, which will be of maximal rank with high probability. Then compute the second Wong sequence of $(A, \mathcal{B})$ and check whether the limit $W^* \subseteq \mathrm{im}(A)$.

3. Gurvits' algorithm and Proposition 1.2

Commutative and non-commutative Edmonds' problem: a natural pair. Viewing matrices as linear maps between two vector spaces, one may suspect Edmonds problem to be a linear algebraic analogue of the maximum matching problem on bipartite graphs, with elements of the underlying vector spaces as being the left and right side vertices, and the matrices as giving us edges – mapping a vector on

[9] The first Wong sequence is the dual of the second one; this naming convention is due to Wong who in [58] defined the two sequences for the special case $\mathcal{B} = \langle B \rangle$.

the left side to one on the right side. Given such a correspondence, one may ask whether an analogue of Hall's theorem holds in this setting, i.e, is it true that a matrix space either has a matrix of rank s, or has a $n-s+1$-shrunk subspace; or put differently, whether $\mathrm{rk}(\mathcal{B}) = \mathrm{ncrk}(\mathcal{B})$ holds for all $\mathcal{B}$. This is far from the truth! For example, for the space of skew-symmetric matrices ($A = -A^{\mathrm{T}}$) of size 3, we have $\mathrm{ncrk} = 3$ and $\mathrm{rk} = 2$.

That is, while for the bipartite maximum matching problem, matchings and shrunk subsets are two sides of the same coin, in the linear algebraic setting, this coin splits into two problems: the commutative Edmonds' problem asks to compute the maximum rank, and the non-commutative Edmonds' problem asks to compute the maximum c for the existence of a c-shrunk subspace.

Rank-1 spanned matrix spaces. Now we point out that several results on the commutative Edmonds' problem can be viewed, and should be understood as, resolving the non-commutative counterpart. For this, note that shrunk subspaces are a natural witness for the singularity of matrix spaces: this construction can be dated back to 1930's in T. G. Room's book [54], and plays a key role in several results which solve special cases of the Edmonds' problem including [33, 38, 43].

A particular case of interest is rank-1 spanned matrix spaces: those matrix spaces that have a basis consisting of rank-1 matrices. For rank-1 spanned spaces, the analogue of Hall's theorem holds [43], so commutative and non-commutative Edmonds' problems conincide. Therefore, the known results for rank-1 spanned spaces [33, 38] can be viewed as solving non-commutative Edmonds' problem or SDIT. In retrospect, the results for rank-1 spanned spaces rely on shrunk subspaces in such a critical way that they should be understood as solving non-commutative Edmonds' problem, or SDIT, rather than the commutative version, for this special case:

- The core of Gurvits' algorithm [33] is an iterative procedure called the operator Sinkhorn's scaling procedure. When applied to a matrix space $\mathcal{B}$, this procedure converges, if and only if $\mathcal{B}$ has a shrunk subspace.
- The key tool in [38] is the second Wong sequence as described in Section 2. When applied to $\mathcal{B}$ and $A \in \mathcal{B}$, this sequence stablizes in a small number of steps, and the limit subspace is contained in $\mathrm{im}(A)$ if and only if $\mathcal{B}$ has an $n - \mathrm{rk}(A)$-shrunk subspace.

Gurvits' algorithm; Proof of Proposition 1.2. In fact, Gurvits' algorithm works by assuming that an analogue of Hall's theorem for perfect matchings holds.

Theorem 3.1 ([33]). *Over $\mathbb{Q}$, given a matrix space $\mathcal{B}$ s.t. either* $\mathrm{rk}(\mathcal{B}) = n$ *or* $\mathrm{ncrk}(\mathcal{B}) < n$*, there exists a deterministic polynomial-time algorithm that solves the commutative SDIT, and therefore the non-commutative SDIT.*

Gurvits' algorithm almost solves the non-commutative SDIT over $\mathbb{Q}$. The only problem is that for a matrix space $\mathcal{B}$ with $n = \mathrm{ncrk}(\mathcal{B}) > \mathrm{rk}(\mathcal{B})$ it may give a wrong answer. (When his algorithm takes such a matrix space, it still terminates in polynomially many steps.) We observe that this can be rectified by considering matrix semi-invariants up to the upper bound for $\sigma(R(n,m))$.

Proof of Proposition 1.2. Recall that, by assumption, the nullcone of $R(n,m)$ is defined by elements of degree $\leq \sigma(R(n,m))$ over $\mathbb{Q}$. Also, given a matrix space

$\mathcal{B} \in M(s, \mathbb{Q})$, Gurvits' algorithm either reports that $\mathrm{rk}(\mathcal{B}) = s$, or $\mathrm{ncrk}(\mathcal{B}) < s$. When $\mathrm{rk}(\mathcal{B}) = s$ or $\mathrm{ncrk}(\mathcal{B}) < s$, it is always correct.

The algorithm is easy to describe: for $d = 1, \ldots, \sigma = \sigma(R(n, m))$ run Gurvits' algorithm with input $\mathcal{B}^{[d]}$. If for some d, Gurvits' algorithm reports $\mathrm{rk}(\mathcal{B}^{[d]}) = dn$, then output $\mathrm{ncrk}(\mathcal{B}) = n$ and halt. Otherwise, return $\mathrm{ncrk}(\mathcal{B}) < n$.

It is clear that this algorithm runs in time polynomial in the input size and σ. Note that a linear basis of $\mathcal{B}^{[d]}$ can be constructed easily in time polynomial in the input size of $\mathcal{B}$ and d.

From the discussion in Section 1.3, the correctness is also easy to see. Specifically, assuming the bound on σ, the simultaneous vanishing of $\det(Y_1 \otimes B_1 + \cdots + Y_m \otimes B_m) \in \mathbb{Q}[y_{i,j}^{(k)}]$ for $d = 1, \ldots, \sigma$, characterizes whether $\mathrm{ncrk}(\mathcal{B}) < n$ or not. Therefore, if $\mathrm{ncrk}(\mathcal{B}) = n$, then for some $d \leq \sigma$, $\mathrm{rk}(\mathcal{B}^{[d]})$ is full. On the other hand if $\mathrm{ncrk}(\mathcal{B}) < n$, then there is a shrunk subspace U. For each d, $\mathbb{Q}^d \otimes U$ is a shrunk subspace of $\mathcal{B}^{[d]}$ and so $\mathrm{ncrk}(\mathcal{B}^{[d]}) < dn$ for any d. □

Implications to Gurvits' algorithm. The invariant-theoretic viewpoint also connects to a question of Gurvits in [33]. In [33], given a basis $\{B_1, \ldots, B_m\}$ of $\mathcal{B} \leq M(n, \mathbb{C})$, Gurvits associates with it a completely positive operator i.e., a linear map $F : M(n, \mathbb{C}) \to M(n, \mathbb{C})$. The main algorithmic technique is the so-called operator Sinkhorn's iterative scaling procedure, which is applied to F. This procedure is a quantum generalization of the classical Sinkhorn's iterative scaling procedure, which is applied to nonnegative matrices, and can be used to approximate the permanent, and to decide the existence of perfect matchings [34, 42]. Gurvits proved that this procedure, when applied to the operator T derived from a matrix space $\mathcal{B}$, converges if and only if $\mathcal{B}$ has a shrunk subspace. He proved this using a continuous but non-differentiable function, called the capacity of an operator, denoted as $\mathrm{Cap}(F)$. Specifically, he showed that $\mathcal{B}$ has a shrunk subspace if and only if $\mathrm{Cap}(F) = 0$. Gurvits asked whether there exists a "nice" function, like a polynomial with integer coefficients, that characterizes $\mathcal{B}$ with shrunk subspaces. Our previous argument suggests that there exists a set of polynomial functions with integer coefficients, whose simultaneous vanishing characterizes those $\mathcal{B}$ with shrunk subspaces, and therefore a "nice" substitute for Gurvits' capacity. However, the number of these polynomial functions depends on the degree bound for matrix semi-invariants.

4. Finding a nonsingular matrix in blow-ups

In this section we describe, given $\mathcal{B} \leq M(n, \mathbb{F})$, how to compute a nonsingular matrix in $\mathcal{B}^{[d]} = M(d, \mathbb{F}) \otimes \mathcal{B}$ for some $d \leq (n + 1)!$, or certify that this is not possible.

One important note is due here: as per our notation, matrices in $M(d, \mathbb{F}) \otimes M(n, \mathbb{F})$ are viewed as block matrices, where each block is of size $n \times n$. This is more convenient when describing semi-invariants. In this section, it will be more convenient to work with $M(n, \mathbb{F}) \otimes M(d, \mathbb{F})$, namely each block is of size $d \times d$. This is consistent with other parts simply because $M(d, \mathbb{F}) \otimes M(n, \mathbb{F}) \cong M(n, \mathbb{F}) \otimes M(d, \mathbb{F})$. Therefore, we identify $M(nd, \mathbb{F}) \cong M(n, \mathbb{F}) \otimes M(d, \mathbb{F})$, and fix such a decomposition. Recall the notation $\mathcal{B}(U)$, and $\mathcal{B}^{\{d\}}$ from Section 2.1.

After some preparation, we prove the main technical lemma called the regularity lemma for blow-ups. Then we prove Theorem 4.8 (the formal version of Theorem 1.3), and Corollary 1.4 and 1.5 follow easily.

4.1. **Preparations.**

A characterization of blow-ups.

Proposition 4.1. *For $\mathcal{A} \leq M(dn, \mathbb{F})$, $\mathcal{A} = \mathcal{B}^{\{d\}}$ for some $\mathcal{B} \leq M(n, \mathbb{F})$ if and only if $(I \otimes M(d, \mathbb{F}))\mathcal{A}(I \otimes M(d, \mathbb{F})) = \mathcal{A}$.*

Proof. The only if part is obvious. To see the reverse implication, notice that the bimodule action of $M(d, \mathbb{F})$ on itself gives an $M(d^2, \mathbb{F})$-module structure on $M(d, \mathbb{F})$. Let $C_1, \dots, C_{d^2}$ be an $\mathbb{F}$-basis for $M(d, \mathbb{F})$. For any index $j \in [d^2]$ and any matrix $C \neq 0 \in M(d, \mathbb{F})$, there exist matrices $U_1 \dots, U_{d^2}$ and $V_1, \dots, V_{d^2}$ from $M(d, \mathbb{F})$ such that $\sum_{k=1}^{d^2} U_k C_i V_k = \delta_{ij} C$, where δ_{ij} is the Kronecker delta (naturally U_i, V_i depend upon C). Any element A of $\mathcal{A}$ can be written as $B_1 \otimes C_1 + \dots + B_{d^2} \otimes C_{d^2}$. Then $\sum_{k=1}^{d^2} (I \otimes U_k) A (I \otimes V_k) = B_j \otimes C$, which implies that $B_j \otimes C$ is in $\mathcal{A}$. So define $\mathcal{B}$ as $\{B \in M(n, \mathbb{F}) : \exists C \neq 0 \in M(d, \mathbb{F}), \text{ s.t. } B \otimes C \in \mathcal{A}\}$, and we see that $\mathcal{A} = \mathcal{B}^{\{d\}}$. □

Note that $(I \otimes M(d, \mathbb{F}))\mathcal{A}(I \otimes M(d, \mathbb{F})) = \mathcal{A}$ is equivalent to saying that $\mathcal{A}$ is an $M(d, \mathbb{F})$ sub-bimodule of $M(n, \mathbb{F}) \otimes M(d, \mathbb{F})$, where we identify $M(d, \mathbb{F})$ with $I \otimes M(d, \mathbb{F})$. Similarly, for a subspace $W \leq \mathbb{F}^n \otimes \mathbb{F}^d$, one can see that W is of the form $W_0 \otimes \mathbb{F}^d$ if and only if $(I \otimes M(d, \mathbb{F}))W_0 = W_0$, that is, W_0 is an $M(d, \mathbb{F})$-submodule of $\mathbb{F}^n \otimes \mathbb{F}^d$.

Shrunk subspaces in the blow-up situation.

Proposition 4.2. *If $\mathcal{A} = \mathcal{B}^{\{d\}}$ has an s-shrunk subspace, then $\mathcal{A}$ has an s'-shrunk subspace where $s' \geq s$ s.t. d divides s', and $\mathcal{B}$ has an s'/d-shrunk subspace.*

Proof. As $\mathcal{A} = \mathcal{B}^{\{d\}}$, $(I \otimes M(d, \mathbb{F}))\mathcal{A}(I \otimes M(d, \mathbb{F})) = \mathcal{A}$. Assume that U is an s-shrunk subspace of $\mathcal{A}$: with $W = \mathcal{A}(U)$ we have $\dim_{\mathbb{F}} U - \dim_{\mathbb{F}} W = s$. Then $(I \otimes M(d, \mathbb{F}))W = (I \otimes M(d, \mathbb{F}))\mathcal{A}U = \mathcal{A}U = W$, thus $W = W_0 \otimes \mathbb{F}^d$ for some $W_0 \leq \mathbb{F}^n$. Similarly, as $\mathcal{A}(I \otimes M(d, \mathbb{F})) = \mathcal{A}$, we have $\mathcal{A}(I \otimes M(d, \mathbb{F}))U = W$, whence $U' = (I \otimes M(d, \mathbb{F}))U$ is an s'-shrunk subspaces with $s' \geq s$, and $U' = U_0 \otimes \mathbb{F}^d$ with some $U_0 \leq \mathbb{F}^n$. Note that $\dim(W)$, $\dim(U')$, and therefore s', are all divisible by d. Noting $\mathcal{A} = \mathcal{B} \otimes M(d, \mathbb{F})$, we have $W_0 \leq \mathcal{B}(U_0)$ and so U_0 is an s'/d-shrunk subspace. □

Using extension fields. Assume that for some extension field $\mathbb{K}$ of $\mathbb{F}$ we are given a matrix $A' \in \mathcal{B} \otimes_{\mathbb{F}} \mathbb{K} \leq M(n, \mathbb{K})$ of rank r. Then, if $|\mathbb{F}| > r$, using the method of [13], we can efficiently find a matrix $A \in \mathcal{B}$ of rank at least r. This procedure is also useful to keep sizes of the occurring field elements small. For completeness we include a brief description. Let $S \subseteq \mathbb{F}$ with $|S| = r + 1$ and let $B_1, \dots, B_\ell$ be an $\mathbb{F}$-basis for $\mathcal{B}$. Then $A' = a'_1 B_1 + \dots + a'_\ell B_\ell$, where $a'_i \in \mathbb{K}$. As A' is of rank r, there exists an $r \times r$ sub-matrix of A with nonzero determinant. Assume that $a'_1 \notin S$. Then we consider the determinant of the corresponding sub-matrix of the polynomial matrix $xB_1 + a'_2 B_2 + \dots a'_\ell B_\ell$. This determinant is a nonzero polynomial of degree at most r in x. Therefore there exists an element $a_1 \in S$ such that $a_1 B_1 + a'_2 B_2 + \dots a'_\ell B_\ell$ has rank at least r. Continuing with $a'_2, \dots, a'_\ell$, we can ensure that all the a_i's are from S. Since the B_i's span $\mathcal{B}$, the resulting matrix of rank at least r is in $\mathcal{B}$.

4.2. **Regularity of blow-ups.** Our goal in this subsection is to prove that when the field size is large enough, the maximum rank over $\mathcal{A} = \mathcal{B}^{\{d\}} \leq M(dn, \mathbb{F})$ is always divisible by d. The proof is constructive when $\mathrm{char}(\mathbb{F}) = 0$, or when d and $\mathrm{char}(\mathbb{F})$ are coprime: if we get a matrix in $\mathcal{A}$ of rank at least $rd + 1$, we will be able to construct a matrix in $\mathcal{A}$ of rank at least $(r+1)d$. This is the main technical lemma to be used in the proof of Theorem 4.8.

4.2.1. *Central division algebras.* We aim to prove that $\mathrm{rk}(\mathcal{B}^{\{d\}})$ is divisible by d. Therefore, it is desirable to devise an algebra D with a linear representation $\rho : D \to M(d, \mathbb{F})$ s.t. every element in $M(n, \mathbb{F}) \otimes \rho(D)$ has rank divisible by d. Central division algebras of appropriate parameters almost satisfy this criterion, with a little overhead of using extension fields.

Claim 4.3. *Let $\mathbb{F}'$ be an extension field of $\mathbb{F}$, and $\mathbb{K}$ be an extension field over $\mathbb{F}'$ with extension degree d. Let D be a central division algebra over $\mathbb{F}'$ of dimension d^2, and $\rho : D \to M(d, \mathbb{K})$ be a representation of D over $\mathbb{K}$. Then every matrix in $M(n, \mathbb{F}) \otimes_{\mathbb{F}} \rho(D)$ has rank divisible by d over $\mathbb{K}$.*

Proof. Consider the $\mathbb{F}'$-linear space $\mathbb{K}^{dn} \cong \mathbb{F}'^{d^2 n}$ as a module over $M(n, \mathbb{F}) \otimes_{\mathbb{F}} \rho(D)$. The centralizer of this action is the opposite algebra D^{op}. Therefore the image of $A'\mathbb{K}^{dn}$ of any $A' \in M(n, \mathbb{F}) \otimes_{\mathbb{F}} \rho(D)$ is a D^{op}-submodule, whence its dimension over $\mathbb{F}'$ is divisible by d^2. It follows that the dimension over $\mathbb{K}$ is divisible by d. $\square$

Due to this nice property, central division algebras play a key role in the proof of the regularity lemma. Furthermore, since we require this proof to be constructive, it is necessary to have a recipe of constructing such an algebra D with these parameters, together with this representation $\rho(D)$.

If $\mathbb{F}$ has characteristic 0 or prime to d, this can be achieved by using the following standard cyclic algebra construction based on Kummer extensions. (We refer the reader to [50, Chapter 15] for cyclic algebras.) First, we replace $\mathbb{F}$ with $\mathbb{F}' = \mathbb{F}[\zeta](z_1, z_2)$ where ζ is a primitive dth root of unity. Then the (noncommutative) associative algebra D over $\mathbb{F}'$ generated by X and Y with relations $X^d = z_1$, $Y^d = z_2$ and $XY = \zeta Y X$ is a central division algebra of dimension d^2 over $\mathbb{F}'$. Also, the subalgebra of D generated by X is isomorphic to the field $\mathbb{K} = \mathbb{F}'[\sqrt[d]{z_1}]$. (Note that $\mathbb{K} \cong \mathbb{F}[\zeta](t, z_1)$ and $\mathbb{F}'$ is embedded in $\mathbb{K}$ by identifying z_1 with t^d.) The minimum polynomial of X is $X^d - z_1$, whose roots are $\zeta^j t$ $(j = 0, \ldots, d-1)$. It follows that in $D \otimes \mathbb{K}$, the element $X - t$ generates a left ideal L of dimension d over K. The left action of D on L gives the desired representation ρ.

To compute with this algebra, we note the following. Firstly, if a nonsingular submatrix over $\mathbb{F}'$ is found, we can use the method of [13] described in Subsection 4.1 for going back to the original $\mathbb{F}$. Secondly, note that $\mathbb{K}$ is a purely transcendental extension of $\mathbb{F}[\zeta]$ of degree two. Therefore we can do the basic linear algebra computations over $\mathbb{K}$ once we can do the same over the finite extension $\mathbb{F}[\zeta]$. Thirdly, constructing $\mathbb{F}[\zeta]$ would require factoring the polynomial $x^d - 1$ over $\mathbb{F}$, a task which cannot be accomplished using basic arithmetic operations. To get around this issue, one can perform the required computations over an appropriate ideal R of the algebra $\mathbb{F}[x]/(x^d - 1)$ as if R were a field. The procedure may fail when a zero divisor in R is found. Then one can replace R with a proper ideal and restart the computation.

Remark 4.4. When the characteristic is a prime p divisor of d, say $d = p^e d'$ where d' is prime to p, the Kummer extension $\mathbb{K}$ above should be replaced by a cyclic extension which is a product of a Kummer extension of degree d' and a cyclic extension of degree p^e described by Artin, Schreier and Witt [57]. However, further work would be needed to investigate the complexity of computing over such extensions. As a consequence, in the following Lemma 4.5, Corollary 4.6, and Theorem 4.7, if the efficiency of the construction is not a concern, then the restrictions on the field characteristics can be removed.

In particular, this means that we can improve the upper bound of $\sigma(R(n,m))$ in Corollary 1.5 to $n!$. Recall that $\mathrm{ncrk}(\mathcal{B}) \le 2\mathrm{rk}(\mathcal{B})$. So if $\mathrm{rk}(\mathcal{B}) < \lceil n/2 \rceil$, then $\mathrm{ncrk}(\mathcal{B}) < n$ for sure. Otherwise, start with a matrix $B \in \mathcal{B}$ of rank $\lceil n/2 \rceil$, and if $\mathrm{ncrk}(\mathcal{B}) = n$, we can find a non-singular matrix in $M(d', \mathbb{F}) \otimes \mathcal{B}$ for some $d' \le n!/\lceil n/2 \rceil!$. This improves the upper bound of $\sigma(R(n,m))$ to $n!/\lceil n/2 \rceil!$.

4.2.2. *The regularity lemma.* Suppose $\mathbf{i} = (i_1, \dots, i_r)$, $\mathbf{j} = (j_1, \dots, j_r)$ are two sequences of integers, where $1 \le i_1 \le \dots \le i_r \le n$ and $1 \le j_1 \le \dots \le j_r \le n$. For a matrix $A \in M(n, \mathbb{F}) \otimes M(d, \mathbb{F})$, we call the sub-matrix of A consisting of the blocks indexed by (i_k, j_ℓ), $k, \ell \in [r]$, the $r \times r$ *window* indexed by $\mathbf{i}$, $\mathbf{j}$.

Lemma 4.5 (Regularity of blow-ups)**.** *For $\mathcal{B} \le M(n, \mathbb{F})$ and $\mathcal{A} = \mathcal{B}^{\{d\}}$, assume that* $\mathrm{char}(\mathbb{F}) = 0$ *or* $\mathrm{char}(\mathbb{F})$ *and d are coprime, and $|\mathbb{F}| > 2rd$. Given a matrix $A \in \mathcal{A}$ with* $\mathrm{rk}A > (r-1)d$, *there exists a deterministic algorithm that returns $\widetilde{A} \in \mathcal{A}$ and an $r \times r$ window W in $\widetilde{A}$ s.t. W is nonsingular (of rank rd). This algorithm uses* $\mathrm{poly}(nd)$ *arithmetic operations and, over $\mathbb{Q}$, all intermediate numbers have bit lengths polynomial in the input size.*

Proof. Fix $S \subseteq \mathbb{F}$ of size $2rd + 1$. The proof goes by induction on r. To see the initial case $r = 1$, let B be any nonzero matrix from $\mathcal{B}$. Assume that the (i,j)th entry of B is nonzero. Then the (i,j)th block of $B \otimes I$ is a nonzero $d \times d$ scalar matrix.

For the inductive step, assume $r > 1$. By the induction hypothesis, we can find a matrix $A' \in \mathcal{A}$ and an $(r-1) \times (r-1)$ nonsingular window in A'. We assume w.l.o.g. that the window corresponds to row and column indices $1, \dots, r-1$, that is, the nonsingular sub-matrix of A' consists of its upper left $(r-1)^2$ blocks. It is easy to see that for some $(\lambda, \mu) \in S \times S$, we have that the upper left $(r-1)d \times (r-1)d$ block of $\lambda A' + \mu A$ is still nonsingular and at the same time $\lambda A' + \mu A$ has rank larger than $(r-1)d$. We replace A with such a $\lambda A' + \mu A$. We can again w.l.o.g. assume that already the upper left r^2 blocks of A form a sub-matrix of rank larger that $(r-1)d$.

By the argument so far we can and do assume, in the sequel, that $n = r$ and the upper left $(n-1)d \times (n-1)d$ block of A is nonsingular. Our aim is to find a full rank matrix in $\mathcal{A}$.

By the construction in Section 4.2.1, we have in our disposal a central division algebra D of dimension d^2 over $\mathbb{F}'$, together with an extension field $\mathbb{K}$ of degree d over $\mathbb{F}'$, and a representation $\rho : D \to M(d, \mathbb{K})$. We consider $\mathcal{A}' = \mathcal{A} \otimes \mathbb{K}$ as a $\mathbb{K}$-linear subspace of $M(n, \mathbb{K}) \otimes_{\mathbb{K}} M(d, \mathbb{K})$. Notice that $\mathcal{A}' = \mathcal{B} \otimes_{\mathbb{F}} M(d, \mathbb{K})$. We consider the $\mathbb{F}'$-linear subspace $\mathcal{A}'_D = \mathcal{B} \otimes_{\mathbb{F}} \rho(D)$ of $\mathcal{A}'$. Then the $\mathbb{K}$-linear span of $\mathcal{A}'_D$ is $\mathcal{A}'$. Starting with A (more precisely, with its image in $\mathcal{A}'$), using the method of [13] we find an element A' of $\mathcal{A}'_D$ whose rank (as an $nd \times nd$ matrix over $\mathbb{K}$) is greater than $(r-1)d$. But the rank of every element from $\mathcal{A}'_D$ is divisible by d, by

Claim 4.3. Thus A' has full rank. Again using the procedure of [13], we can replace $A' \in \mathcal{A}'$ with a matrix $\widetilde{A}$ from $\mathcal{A}$ having full rank as well. $\square$

We mention an immediate consequence.

Corollary 4.6. *Let $\mathcal{B} \leq M(n, \mathbb{F})$. Assume that the characteristic of $\mathbb{F}$ is zero, $d \geq n$, and we are given a matrix $A \in \mathcal{B}^{\{d\}}$ with $\mathrm{rk} A = rd$. Then, for any $d' > d$ we can efficiently find $\widetilde{A} \in \mathcal{B}^{\{d'\}}$ with $\mathrm{rk}\widetilde{A} \geq rd'$.*

Proof. Induction on d' and d. Assume first that $d' = d+1$. Then $\frac{d+1}{d} \leq \frac{n+1}{n} < \frac{r}{r-1}$, whence $(d+1)(r-1) < dr$. Therefore, if we embed $\mathcal{B}^{\{d\}}$ into $\mathcal{B}^{\{d+1\}}$ then A has rank $rd > (r-1)(d+1)$ and Lemma 4.5 applies. Proceed with $(d+2, d+1)$ in place of $(d+1, d)$. $\square$

The statement probably remains true for $d < n$. However, we do not see how to prove it.

4.3. **Incrementing rank via blow-up.** We have introduced in Section 2.2 a key technique here, namely the second Wong sequences. Recall that given $A \in \mathcal{B} \leq M(n, \mathbb{F})$, the second Wong sequence can be used to detect whether there exists a $\mathrm{cork}(A)$-shrunk subspace. If such a space exists, then $\mathrm{ncrk}(\mathcal{B}) = n - \mathrm{cork}(A)$. The difficulty is the case when such $\mathrm{cork}(A)$-shrunk subspace does not exist. One natural idea to proceed is to find $A' \in \mathcal{B}$ of rank $> \mathrm{rk}(A)$, and use A' to test whether a $\mathrm{cork}(A')$-shrunk exists or not. If $\mathcal{B}$ is rank-1 spanned, such an A' does exist, and another ingredient in [38] is an update procedure that finds A' of higher rank in this case. However, this in general is not possible, since $\mathrm{rk}(\mathcal{B})$ and $\mathrm{ncrk}(\mathcal{B})$ may differ. Fortunately, the concept of blow-ups helps: instead of looking for $A' \in \mathcal{B}$ of rank $> \mathrm{rk}(A)$, we shall look for $A' \in \mathcal{B}^{\{d\}}$ of rank $> \mathrm{rk}(A)d$, where $d = \mathrm{rk}(A) + 1$. It turns out that this is achievable, and an application of the regularity lemma even yields $A'' \in \mathcal{B}^{\{d\}}$ of rank $\geq (\mathrm{rk}(A) + 1)d$.

Theorem 4.7. *Let $\mathcal{B} \leq M(n, \mathbb{F})$ and let $\mathcal{A} = \mathcal{B}^{\{d\}}$. Assume that we are given a matrix $A \in \mathcal{A}$ with $\mathrm{rk}(A) = rd$, and $|\mathbb{F}|$ is $\Omega(ndd')$, where $d' > r$ is any positive integer coprime with $\mathrm{char}(\mathbb{F})$. There exists a deterministic algorithm that returns either an $(n-r)d$-shrunk subspace for $\mathcal{A}$ (equivalently, an $(n-r)$-shrunk subspace for $\mathcal{B}$), or a matrix $B \in \mathcal{A} \otimes M(d', \mathbb{F})$ of rank at least $(r+1)dd'$. This algorithm uses $\mathrm{poly}(ndd')$ arithmetic operations and, over $\mathbb{Q}$, all intermediate numbers have bit lengths polynomial in the input size.*

Proof. Let $W_0 \leq W_1 \leq \ldots \leq \mathbb{F}^{nd}$ be the second Wong sequence for to the pair $(A, \mathcal{A})$: $W_0 = (0)$, $W_{t+1} = \mathcal{A}A^{-1}(W_t)$. As $(I \otimes M(d, \mathbb{F}))\mathcal{A} = \mathcal{A}$, we have $(I \otimes M(d, \mathbb{F}))\mathcal{A}W_t = \mathcal{A}W_t$, whence the dimension of $\mathcal{A}W_t$ is divisible by d for every t. It follows that, until stabilization, the dimension of $\mathcal{A}W_t$ increases by at least d.

If the sequence $\mathcal{A}W_t$ does not run out of $\mathrm{im}(A)$, then the limit is an $(n-r)d$-shrunk subspace.

Otherwise we have an index $t \leq r+1$, vectors $v_1, \ldots, v_t$ and matrices $B_1, \ldots, B_t$ such that $v_1 \in A^{-1}(0) = \ker A$ and $Av_j = B_{j-1}v_{j-1}$ for $j = 2, \ldots, t$ and $B_t v_t \notin \mathrm{im}(A)$. Let $u_1, \ldots, u_{d'}$ be a basis for $\mathbb{F}^{d'}$, and for $j = 1, \ldots, t$ let Z_j stand for the matrix from $M(d', \mathbb{F})$ that maps u_j to u_{j+1} and $u_{j'}$ to zero for $j' \neq j$. (If $j = r+1 = d'$ then send u_{r+1} to u_1, and $u_{j'}$ to zero for $j' \neq r+1$.) Put $A' = A \otimes I \in M(ndd', \mathbb{F})$, $B' = B_1 \otimes Z_1 + \ldots + B_t \otimes Z_t$, $w_1 = v_1 \otimes u_1$, $\ldots$, $w_t = v_t \otimes u_t$. Then $B'w_t = B_t v_t \otimes u_{t+1} \notin (A\mathbb{F}^{nd}) \otimes \mathbb{F}^{d'} \supseteq A'\mathbb{F}^{ndd'}$. Also, $w_1 \in \ker A'$

and $A'w_j = B'w_{j-1}$ for $j = 2, \ldots, t$. This means that the image under B' of the second Wong sequence for the pair $(A', \langle B' \rangle)$ runs out of the image of A'. If this happens, A' is not of maximum rank in $\langle A', B' \rangle$; see [38, Fact 11]. Therefore there exists λ (from some fixed $S \subseteq \mathbb{F}$ of size $\Omega(ndd')$) such that $A' + \lambda B'$ has rank greater than rdd'. Given such $A' + \lambda B'$, using Lemma 4.5 we can compute a matrix from $\mathcal{A} \otimes M(d', \mathbb{F})$ of rank at least even $(r+1)dd'$.

In the above procedure we only need to compute the second Wong sequences, form matrices in $M(\mathbb{F}, ndd')$, and apply the regularity lemma. The running time as stated then follows. $\square$

An iteration based on Theorem 4.7 proves the following Theorem 4.8. Note that in Theorem 4.7, d' can be chosen as either $r+1$ or $r+2$, depending on which is not divisible by $\mathrm{char}(\mathbb{F})$.

Theorem 4.8. *Suppose we are given $\mathcal{B} := \langle B_1, \ldots, B_m \rangle \leq M(n, \mathbb{F})$, and $A \in \mathcal{B}$ with $\mathrm{rk}(A) = s < n$. Let $d = (n+1)!/(s+1)!$, and assume that $|\mathbb{F}| = \Omega(nd)$. Then there exists a deterministic algorithm, that computes a matrix $B \in M(d', \mathbb{F}) \otimes \mathcal{B}$ of rank rd' for some $d' \leq d$ and, if $r < n$, an $(n-r)$-shrunk subspace for $\mathcal{B}$. The algorithm uses $\mathrm{poly}(n, d)$ arithmetic operations, and when working over $\mathbb{Q}$, has bit complexity polynomial in n, d and the input size.*

Now Corollary 1.4 and 1.5 follow easily. To see Corollary 1.4, note that if we choose a matrix in $\mathcal{B}$ randomly, it will be of maximum rank. Using that matrix as A in Theorem 4.8, Corollary 1.4 is proved. For Corollary 1.5, as if $\mathcal{B}$ has no shrunk subspaces, a full-rank matrix will be certainly present in $M(d', \mathbb{F}) \otimes \mathcal{B}$ for some $d' \leq (n+1)!$, giving us the upper bound on $\sigma(R(n, m))$.

Acknowledgement. We would like to thank Mátyás Domokos, Bharat Adsul, Ketan Mulmuley, Partha Mukhopadhyay and K. N. Raghavan for discussions related to this work. Part of the work was done when Youming was visiting the Simons Institute for the program Algorithms and Complexity in Algebraic Geometry.

References

1. B. Adsul, S. Nayak, and K. V. Subrahmanyam, *A geometric approach to the Kronecker problem II: rectangular shapes, invariants of matrices and the Artin–Procesi theorem*, preprint, 2007.
2. Bharat Adsul and K. V. Subrahmanyam, *A geometric approach to the Kronecker problem I: the two row case*, Proceedings Mathematical Sciences **118** (2008), no. 2, 213–226.
3. MD Atkinson and S Lloyd, *Primitive spaces of matrices of bounded rank*, Journal of the Australian Mathematical Society (Series A) **30** (1981), no. 04, 473–482.
4. Peter A. Brooksbank and Eugene M. Luks, *Testing isomorphism of modules*, Journal of Algebra **320** (2008), no. 11, 4020–4029.
5. M. Bürgin and J. Draisma, *The Hilbert null-cone on tuples of matrices and bilinear forms*, Mathematische Zeitschrift **254** (2006), no. 4, 785–809.
6. Peter Bürgisser, J. M. Landsberg, Laurent Manivel, and Jerzy Weyman, *An overview of mathematical issues arising in the geometric complexity theory approach to VP≠VNP*, SIAM J. Comput. **40** (2011), no. 4, 1179–1209.
7. Jonathan F. Buss, Gudmund S. Frandsen, and Jeffrey O. Shallit, *The computational complexity of some problems of linear algebra*, J. Comput. Syst. Sci. **58** (1999), no. 3, 572–596.
8. Alexander L. Chistov, Gábor Ivanyos, and Marek Karpinski, *Polynomial time algorithms for modules over finite dimensional algebras*, ISSAC, 1997, pp. 68–74.
9. Ajeh M Cohen, Gábor Ivanyos, and David B Wales, *Finding the radical of an algebra of linear transformations*, Journal of Pure and Applied Algebra **117** (1997), 177–193.

10. P. M. Cohn, *Free rings and their relations*, L.M.S. Monographs, Acad. Press, 1985, First edition 1971.
11. ———, *Skew fields: Theory of general division rings*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1995.
12. P. M. Cohn and C. Reutenauer, *On the construction of the free field*, International Journal of Algebra and Computation **9** (1999), no. 3-4, 307–323.
13. Willem A. de Graaf, Gábor Ivanyos, and Lajos Rónyai, *Computing Cartan subalgebras of Lie algebras*, Applicable Algebra in Engineering, Communication and Computing **7** (1996), no. 5, 339–349.
14. Harm Derksen, *Polynomial bounds for rings of invariants*, Proceedings of the American Mathematical Society **129** (2001), no. 4, 955–964.
15. Harm Derksen and Jerzy Weyman, *Semi-invariants of quivers and saturation for littlewood-richardson coefficients*, Journal of the American Mathematical Society **13** (2000), no. 3, 467–479.
16. M. Domokos, *Poincaré series of semi-invariants of 2× 2 matrices*, Linear Algebra and its Applications **310** (2000), no. 1, 183–194.
17. ———, *Relative invariants of 3× 3 matrix triples*, Linear and Multilinear Algebra **47** (2000), no. 2, 175–190.
18. ———, *Finite generating system of matrix invariants*, Math. Pannon **13** (2002), no. 2, 175–181.
19. M. Domokos and V. Drensky, *Defining relation for semi-invariants of three by three matrix triples*, Journal of Pure and Applied Algebra **216** (2012), no. 10, 2098–2105.
20. M. Domokos, S. G. Kuzmin, and A. N. Zubkov, *Rings of matrix invariants in positive characteristic*, Journal of Pure and Applied Algebra **176** (2002), no. 1, 61–80.
21. M. Domokos and A. N. Zubkov, *Semi-invariants of quivers as determinants*, Transformation groups **6** (2001), no. 1, 9–24.
22. Stephen Donkin, *Invariants of several matrices*, Inventiones mathematicae **110** (1992), no. 1, 389–401.
23. Stephen Donkin, *Invariant functions on matrices*, Mathematical Proceedings of the Cambridge Philosophical Society **113** (1993), 23–43.
24. Jack Edmonds, *Systems of distinct representatives and linear algebra*, J. Res. Nat. Bur. Standards Sect. B **71** (1967), 241–245.
25. David Eisenbud and Joe Harris, *Vector spaces of matrices of low rank*, Advances in Mathematics **70** (1988), no. 2, 135 – 155.
26. Edward Formanek, *Generating the ring of matrix invariants*, Ring Theory (Freddy M. J. van Oystaeyen, ed.), Lecture Notes in Mathematics, vol. 1197, Springer Berlin Heidelberg, 1986, pp. 73–82 (English).
27. M. Fortin and C. Reutenauer, *Commutative/noncommutative rank of linear matrices and subspaces of matrices of low rank*, Séminaire Lotharingien de Combinatoire **52** (2004), B52f.
28. Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson, *A deterministic polynomial time algorithm for non-commutative rational identity testing*, CoRR **abs/1511.03730** (2015).
29. James Geelen and Satoru Iwata, *Matroid matching via mixed skew-symmetric matrices*, Combinatorica **25** (2005), no. 2, 187–215.
30. James F. Geelen, *Maximum rank matrix completion*, Linear Algebra and its Applications **288** (1999), 211–217.
31. James F Geelen, *An algebraic matching algorithm*, Combinatorica **20** (2000), no. 1, 61–70.
32. James F. Geelen, Satoru Iwata, and Kazuo Murota, *The linear delta-matroid parity problem*, Journal of Combinatorial Theory, Series B **88** (2003), no. 2, 377–398.
33. Leonid Gurvits, *Classical complexity and quantum entanglement*, J. Comput. Syst. Sci. **69** (2004), no. 3, 448–484.
34. Leonid Gurvits and Peter N. Yianilos, *The deflation-inflation method for certain semidefinite programming and maximum determinant completion problems (extended abstract)*, Tech. report, NECI, 1998.
35. Nicholas J. A. Harvey, David R. Karger, and Kazuo Murota, *Deterministic network coding by matrix completion*, Proceedings of SODA, ACM-SIAM, 2005, pp. 489–498.
36. Pavel Hrubeš and Avi Wigderson, *Non-commutative arithmetic circuits with division*, Innovations in Theoretical Computer Science, ITCS'14, Princeton, NJ, USA, January 12-14, 2014,

2014, Full version at `http://www.math.ias.edu/~avi/PUBLICATIONS/HrubesWiDec2013.pdf`, pp. 49–66.
37. Christian Ikenmeyer, Ketan D. Mulmuley, and Michael Walter, *On vanishing of Kronecker coefficients*, CoRR **abs/1507.02955** (2015).
38. Gábor Ivanyos, Marek Karpinski, Youming Qiao, and Miklos Santha, *Generalized Wong sequences and their applications to Edmonds' problems*, STACS, 2014, pp. 397–408.
39. Gábor Ivanyos, Marek Karpinski, and Nitin Saxena, *Deterministic polynomial time algorithms for matrix completion problems*, SIAM J. Comput. **39** (2010), no. 8, 3736–3751.
40. Gábor Ivanyos, Youming Qiao, and K. V. Subrahmanyam, *On generating the ring of matrix semi-invariants*, CoRR **abs/1508.01554** (2015).
41. Valentine Kabanets and Russell Impagliazzo, *Derandomizing polynomial identity tests means proving circuit lower bounds*, Computational Complexity **13** (2004), no. 1-2, 1–46.
42. Nathan Linial, Alex Samorodnitsky, and Avi Wigderson, *A deterministic strongly polynomial algorithm for matrix scaling and approximate permanents*, Combinatorica **20** (2000), no. 4, 545–568.
43. László Lovász, *Singular spaces of matrices and their application in combinatorics*, Boletim da Sociedade Brasileira de Matemática-Bulletin/Brazilian Mathematical Society **20** (1989), no. 1, 87–99.
44. Laurent Manivel, *A note on certain Kronecker coefficients*, Proceedings of the American Mathematical Society **138** (2010), no. 1, 1–7.
45. Ketan Mulmuley, *On P vs. NP and geometric complexity theory: Dedicated to Sri Ramakrishna*, J. ACM **58** (2011), no. 2, 5.
46. ———, *Geometric complexity theory V: equivalence between blackbox derandomization of polynomial identity testing and derandomization of noether's normalization lemma*, 53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012, 2012, pp. 629–638.
47. Ketan Mulmuley and Milind A. Sohoni, *Geometric complexity theory I: an approach to the P vs. NP and related problems*, SIAM J. Comput. **31** (2001), no. 2, 496–526.
48. ———, *Geometric complexity theory II: towards explicit obstructions for embeddings among class varieties*, SIAM J. Comput. **38** (2008), no. 3, 1175–1206.
49. Kazuo Murota, *Matrices and matroids for systems analysis*, Springer, 2000.
50. R.S. Pierce, *Associative algebras*, http://link.springer.com/bookseries/136, Springer, 1982.
51. Vladimir L Popov, *The constructive theory of invariants*, Izvestiya: Mathematics **19** (1982), no. 2, 359–376.
52. C. Procesi, *The invariant theory of $n \times n$ matrices*, Advances in Mathematics **19** (1976), no. 3, 306–381.
53. Ju. P. Razmyslov, *Trace identities of full matrix algebras over a field of characteristic zero*, Mathematics of the USSR-Izvestiya **8** (1974), no. 4, 727, English translation available at `http://iopscience.iop.org/0025-5726/8/4/A01`.
54. T. G. Room, *The geometry of determinantal loci*, The Cambridge University Press, 1938.
55. Aidan Schofield and Michel Van den Bergh, *Semi-invariants of quivers for arbitrary dimension vectors*, Indagationes Mathematicae **12** (2001), no. 1, 125–138.
56. W. T. Tutte, *The factorization of linear graphs*, Journal of the London Mathematical Society **1** (1947), no. 2, 107–111.
57. E. Witt, *Zyklische körper und algebren der charakteristik p vom grad pn. struktur diskret bewerteter perfekter körper mit vollkommenem restklassenkörper der charakteristik p*, J. Reine Angew. Math, **176** (1937), no. 01, 126–140.
58. Kai-Tak Wong, *The eigenvalue problem $\lambda Tx+Sx$*, Journal of Differential Equations **16** (1974), no. 2, 270 – 280.

Appendix A. Derksen's bound applied to $R(n, m)$

Proof. We just need to indicate certain parameters for the matrix semi-invariants that are used in Derksen's bound.

Suppose a group G acts on a vector space V rationally, and let R be the resulting invariant ring. Theorem 1.1 in [14] shows that the degree bound is upper bounded by $\max(2, 3/8 \cdot s \cdot \sigma(R)^2)$, where σ is the degree bound for defining the nullcone, and s is the dimension of R.

s is upper bounded by the number of variables. Therefore for $R(m, n)$, $s \leq mn^2 \leq n^4$.

To bound σ, we use Proposition 1.2 in [14]. Recall that G as an algebraic group, is defined by a system of polynomial equations in $z_1, \ldots, z_t$. For example, $\mathrm{SL}(n, \mathbb{F}) \times \mathrm{SL}(n, \mathbb{F})$ is defined by $\det(X) = 1$ and $\det(Y) = 1$, where X and Y are $n \times n$ variable matrices. The action of G is rational, so it can be recorded as $\rho : G \to \mathrm{GL}(V)$ by $g \to (a_{i,j}(g))_{i,j \in \dim(V)}$, where each $a_{i,j}$ is a polynomial in $z_1, \ldots, z_t$.

$\sigma(R)$ is then upper bounded by $H^{t-m}A^m$, where t is the number of variables used to define G as above, $m = \dim(G)$, H is the maximum degree over polynomials defining G, and A is the maximum degree over polynomials defining the action. So for $R(n, m)$, $t = 2n^2$, $m = 2n^2 - 2$, $H = n$, and $A = 2$. It follows that $\sigma(R(n, m)) \leq n^2 \cdot 2^{2n^2-2}$.

Therefore $R(n, m)$ is generated by elements of degree $\leq 3/128 \cdot n^8 \cdot 16^{n^2}$. $\square$

Institute for Computer Science and Control, Hungarian Academy of Sciences, Budapest, Hungary

E-mail address: Gabor.Ivanyos@sztaki.mta.hu

Centre for Quantum Computation and Intelligent Systems, University of Technology, Sydney, Australia

E-mail address: jimmyqiao86@gmail.com

Chennai Mathematical Institute, Chennai, India

E-mail address: kv@cmi.ac.in