SIAM J. COMPUT.
Vol. 39, No. 8, pp. 3736–3751

DETERMINISTIC POLYNOMIAL TIME ALGORITHMS FOR MATRIX COMPLETION PROBLEMS*

GÁBOR IVANYOS†, MAREK KARPINSKI‡, AND NITIN SAXENA§

Abstract. We present new deterministic algorithms for several cases of the *maximum rank matrix completion* problem (for short *matrix completion*), i.e., the problem of assigning values to the variables in a given symbolic matrix to maximize the resulting matrix *rank*. Matrix completion is one of the fundamental problems in computational complexity. It has numerous important algorithmic applications, among others, in computing dynamic transitive closures or multicast network codings [N. J. A. Harvey, D. R. Karger, and K. Murota, *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2005, pp. 489–498; N. J. A. Harvey, D. R. Karger, and S. Yekhanin, *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2006, pp. 1103–1111]. We design efficient deterministic algorithms for common generalizations of the results of Lovász and Geelen on this problem by allowing *linear polynomials* in the entries of the input matrix such that the submatrices corresponding to each variable have rank one. Our methods are algebraic and quite different from those of Lovász and Geelen. We look at the problem of matrix completion in the more general setting of linear spaces of linear transformations and find a maximum rank element there using a greedy method. Matrix algebras and modules play a crucial role in the algorithm. We show (hardness) results for special instances of matrix completion naturally related to matrix algebras; i.e., in contrast to computing isomorphisms of modules (for which there is a known deterministic polynomial time algorithm), finding a surjective or an injective homomorphism between two given modules is as hard as the general matrix completion problem. The same hardness holds for finding a maximum dimension *cyclic* submodule (i.e., generated by a single element). For the "dual" task, i.e., finding the minimal number of generators of a given module, we present a deterministic polynomial time algorithm. The proof methods developed in this paper apply to fairly general modules and could also be of independent interest.

Key words. matrix completion, identity testing, modules, generators, morphisms

AMS subject classifications. 68Q17, 68W30, 16D99

DOI. 10.1137/090781231

1. Introduction.

A *linear matrix* is a matrix having linear polynomials as its entries, say the linear polynomials are over a field $\mathbb{F}$ and in $\mathbb{F}[x_1, \ldots, x_n]$. The problem of *maximum rank matrix completion*, or just *matrix completion* for short, is the problem of assigning values from the field $\mathbb{F}$ to the variables $x_1, \ldots, x_n$ such that the rank of a given linear matrix is maximized (over all possible assignments). (Throughout this paper for a matrix M we denote its rank by rk M.) The notion of linear matrices appears in several places including both theory and applications; see [HKM05, HKY06] for several references. The problem of matrix completion is a well-studied problem, dating back to the work of Edmonds [Edm67] and Lovász [Lov79]. A similar problem (basically an equivalent one) is *nonsingular matrix completion*, where we have a square linear matrix and are interested in an assignment resulting in a nonsingular

*Received by the editors December 28, 2009; accepted for publication (in revised form) September 16, 2010; published electronically December 8, 2010.
http://www.siam.org/journals/sicomp/39-8/78123.html

†Computer and Automation Research Institute, Hungarian Academy of Sciences (MTA SZTAKI), Lágymányosi u. 11, 1111 Budapest, Hungary (Gabor.Ivanyos@sztaki.hu). This author's research was supported by Hungarian Scientific Research Fund (OTKA) grants NK72845 and T77476.

‡Department of Computer Science and Hausdorff Center for Mathematics, University of Bonn, D-53117 Bonn, Germany (marek@cs.uni-bonn.de).

§Hausdorff Center for Mathematics, University of Bonn, Endenicher Allee 62, D-53115 Bonn, Germany (ns@hcm.uni-bonn.de).

matrix. If the ground field is sufficiently large, then the maximum rank achieved by completion coincides with the rank of the linear matrix considered as a matrix over the function field $\mathbb{F}(x_1, \ldots, x_n)$ and, hence, by standard linear algebra, finding a maximum rank completion (equivalently, determining the maximum rank) is in deterministic polynomial time reducible to instances of finding (equivalently, deciding the existence of) nonsingular completion of certain minors. Lovász gave an efficient randomized algorithm to find a matrix completion using the Schwartz–Zippel lemma [Sch80, Zip79], deducing that a random assignment of the variables will maximize the rank if the field is large enough (see also [IM83]). This is a method also useful in the fundamental problem of polynomial identity testing (PIT). Indeed matrix completion is equivalent to a special case of PIT: any arithmetic formula can be written as the determinant of a linear matrix [Val79]; hence the formula would be nonzero if and only if the corresponding matrix could attain full rank (assuming a large enough field). Over large fields, this makes matrix completion an important problem in ZPP (zero-error probabilistic polynomial time), as its derandomization would imply circuit lower bounds (see Kabanets and Impagliazzo [KI03]).

Over small fields, matrix completion soon becomes a hard problem. This version has some important practical applications, for example, in constructing multicast network codes [HKM05], and hence there are several results in the literature specifying the exact parameters for which the problem becomes NP-hard. The hardness of matrix completion and various related problems was first studied by Buss, Frandsen, and Shallit [BFS99] and more recently by Harvey, Karger, and Yekhanin [HKY06]. In the former paper nonsingular matrix completion is proved to be NP-hard over fields of constant size, while the latter shows that matrix completion over the field $\mathbb{F}_2$ is NP-hard even if we restrict the input to a matrix where each variable occurs at most twice in its entries. This naturally raises the question, Can we solve matrix completion by restricting the way the variables appear in the input matrix?

Few such cases are already known, and they all look at *mixed matrices*, i.e., linear matrices where each entry is either a variable or a constant. Harvey, Karger, and Murota [HKM05], building on the works of Geelen [Gee99] and Murota [Mur00], give an efficient deterministic algorithm for matrix completion over *any* field if the mixed matrix has each variable appearing at most once, while Geelen, Iwata, and Murota [GIM03] and Geelen and Iwata [GI05] give an efficient deterministic algorithm when the mixed matrix is *skew-symmetric* and has each variable appearing at most twice.

Completion by rank one matrices. In this paper we are interested in cases that are more general than the first case [HKM05]. Consider a linear matrix $A \in \mathbb{F}[x_1, \ldots, x_n]^{m \times m}$, where the submatrix "induced" by each variable is of rank one, i.e., $A = B_0 + x_1 B_1 + \cdots + x_n B_n$, where $B_1, \ldots, B_n$ are constant matrices of rank one (note that B_0 is also a constant matrix but of arbitrary rank). The case $B_0 = 0$ was first considered by Lovász in [Lov89], where it is shown how Edmonds' matroid intersection algorithm can be applied to solve this special case in deterministic polynomial time. The first main result in this paper is a common generalization of the results of Lovász [Lov89] and Geelen [Gee99]: we show that the matrix completion problem for an *arbitrary* B_0 can be solved in deterministic polynomial time over any field.

THEOREM 1.1. *Let $\mathbb{F}$ be a field and let $B_0, \ldots, B_n$ be $m \times m$ matrices over $\mathbb{F}$. If $B_1, \ldots, B_n$ are of rank one, then matrix completion for the matrix $(B_0 + x_1 B_1 + \cdots + x_n B_n)$ can be done deterministically in poly(m, n) field operations.*

The proof of this theorem basically involves looking at the linear space $L := \langle B_0, B_1, \ldots, B_n \rangle$ of matrices (i.e., we consider all linear combinations of these $n + 1$

matrices over $\mathbb{F}$) and showing that a greedy approach can be utilized to gradually increase the rank of an element in L. Our methods are more algebraic and quite different from those of Lovász and Geelen. In particular our method is robust enough to check whether a given matrix in L has the largest possible rank *without* needing the rank one generators of L; they are needed only if we want to increase the rank (see section 2).

Matrix algebras or algebras of linear transformations (in this paper by an *algebra* we mean a linear space of matrices or linear transformations that is also closed under multiplication) play a crucial role in the algorithm for Theorem 1.1. We consider special instances of matrix completion problems where algebras of linear transformations arise naturally. These are certain *module* problems.

Preliminaries about modules. If U and V are vector spaces over the field $\mathbb{F}$, then we denote the vector space of linear maps from U to V by $\mathrm{Lin}(U,V)$. For $\mathrm{Lin}(U,U)$ we use the notation $\mathrm{Lin}(U)$. For simplicity, in this paper we consider modules over finite sets. (Actually, we work with modules over free associative algebras; however, the main concepts and computational tasks we are concerned with can be understood without any knowledge from the theory of abstract associative algebras.) Let $\mathcal{S}$ be a finite set. A vector space V over the field $\mathbb{F}$ together with a map ν from $\mathcal{S}$ into $\mathrm{Lin}(V)$ is called an $\mathbb{F}\{\mathcal{S}\}$-*module* (or an $\mathcal{S}$-module for short if $\mathbb{F}$ is clear from the context). We assume that the data for an $\mathcal{S}$-module is input by an $|\mathcal{S}|$-tuple of $\dim V$ by $\dim V$ matrices. In cases when the map ν is clear from the context—most typically when $\mathcal{S}$ is itself a set of linear transformations—we omit ν and denote the result $\nu(B)v$ of the action of $B \in \mathcal{S}$ on $v \in V$ by Bv. For a set $\mathcal{S}' \subseteq \mathrm{Lin}(V)$ of linear transformations the *enveloping algebra* $\mathrm{Env}(\mathcal{S}')$ is the smallest algebra containing $\mathcal{S}'$. It is the linear span of finite products of transformations from $\mathcal{S}'$ (and may be noncommutative).

In the context of $\mathcal{S}$-modules the algebra $\mathcal{A} = \mathrm{Env}(\nu(\mathcal{S}) \cup I)$ is of special interest (I is the identity in $\mathrm{Lin}(V)$). An $\mathcal{S}$-*submodule* of V is a linear subspace closed under the action of all the transformations in $\nu(\mathcal{S})$. Obviously, the intersection of a family of submodules is again a submodule. In particular, if T is a subset of V, then there is a smallest submodule of V containing T: the submodule generated by T. It is $\mathcal{A}T$, the linear span of vectors obtained by application of transformations from $\mathcal{A}$ to vectors from T. The set $T \subseteq V$ is a system of *generators* for the $\mathcal{S}$-module V if $V = \mathcal{A}T$.

Cyclic submodules, i.e., those generated by a single element, are of particular interest. For $v \in V$ we consider the map $\mu_v : \mathcal{A} \to V$ given by $\mu_v(B) = Bv$. Obviously, μ_v is a linear map from $\mathcal{A}$ into V, and the set $\{\mu_v \mid v \in V\}$ is a linear space of linear maps from $\mathcal{A}$ to V. The *rank* of μ_v is the rank of its image, i.e., the dimension of the submodule $\mathcal{A}v$ generated by v.

A "universal" module problem. The matrix completion problem in this context is finding an element v which generates a submodule of maximum dimension. It turns out that this problem, which we call *cyclic submodule optimization*, is universal in matrix completion: there is a deterministic polynomial time reduction from maximum rank matrix completion to cyclic submodule optimization (over an arbitrary base field). We show this universality in section 3. Universality implies two hardness results. First, existence of a deterministic polynomial time algorithm for cyclic submodule optimization would imply deterministic solvability of the matrix completion problem over sufficiently large fields. Also, over small fields, cyclic submodule optimization is NP-hard. Second, we get analogous hardness results for the existence of injective (resp., surjective) homomorphisms between modules (an $\mathcal{S}$-module *homo-*

morphism from V to V' is a linear map in $\mathrm{Lin}(V, V')$ that commutes with the action of $\mathcal{S}$).

THEOREM 1.2. *There is a deterministic polynomial time reduction from the existence of (resp., finding) a nonsingular matrix completion to the problem of checking for the existence of (resp., finding) a surjective (or injective) homomorphism between two modules.*

This result is remarkable in view of the recent deterministic polynomial time algorithm of Brooksbank and Luks [BL08] for the module isomorphism problem (see also Chistov, Ivanyos, and Karpinski [CIK97] regarding special base fields).

A "dual" problem. In section 4 we consider a problem which is in some sense "dual" to the cyclic submodule optimization. This is finding a system of generators of smallest size for a module. In contrast to hardness of the former problem, we have an efficient solution to the generator problem.

THEOREM 1.3. *Given a module structure on the n-dimensional vector space V over the field $\mathbb{F}$ in terms of m $n \times n$ matrices $\mathcal{S}$, one can find the minimum number of generators of V deterministically using $\mathrm{poly}(m, n)$ field operations.*

Note that the above result includes efficiently testing the cyclicity of modules over any field, a special instance of cyclic submodule optimization. This problem was considered in [CIK97] as a tool for constructing isomorphisms between modules and was efficiently solved over special fields. Together with the reduction given in [CIK97], Theorem 1.3 gives a method, completely different from that of [BL08], for constructing isomorphisms between modules in polynomial time over arbitrary fields. The algorithm of Theorem 1.3 is based on a greedy approach analogous to the method for Theorem 1.1, and it implicitly uses certain submodule dimension optimization techniques for a special class of (so-called *semisimple*) modules.

2. Matrix completion with rank one matrices. In this section we prove Theorem 1.1. The iterative step (formulated in Theorem 2.4 below) provides a tool for increasing the rank of a matrix by adding some matrices from a collection of rank one matrices, if possible. The proof is based on the special case (see Lemma 2.1) where we have square matrices and the matrix whose rank is to be increased is idempotent (equals its square). This assumption enables us to use matrix multiplication and matrix algebras. In the general case we will achieve this situation by padding and multiplying all of our matrices by an appropriate matrix.

Let V be a finite dimensional vector space over the field $\mathbb{F}$ and let $L \leq \mathrm{Lin}(V)$ be an $\mathbb{F}$-linear subspace of linear transformations. Recall that $\mathrm{Env}(L)$, the *enveloping algebra* of L, is the linear span of products $h_1 h_2 \cdots h_s$ ($s \geq 1$, $h_1, \ldots, h_s \in L$). Obviously, $\mathrm{Env}(L)$ is also spanned by products of elements from an arbitrary basis of L. We will use the action of the enveloping algebra on the kernel of an idempotent transformation to optimize rank in a linear space; to that effect we present the following lemma. Its proof will also suggest how to greedily increment the rank. Broadly speaking, Lemma 2.1 and Fact 2.3 are used to show in Theorem 2.4 that an $h \in L$ is of maximum rank if and only if $\mathrm{Env}(L)$ maps the kernel of h in the image set of h, a condition which is then easy to check algorithmically.

LEMMA 2.1. *Let V be a finite dimensional vector space over the field $\mathbb{F}$, let $L \leq \mathrm{Lin}(V)$, and assume that $e \in L$ is an idempotent ($e^2 = e$) such that* $\mathrm{rk}\, e \geq \mathrm{rk}\, h$ *for every $h \in L$. If L is spanned by e and certain rank one transformations, then* $\mathrm{Env}(L) \ker e \subseteq eV$.

Proof. Assume, for contradiction, that $\mathrm{Env}(L) \ker e$ is not contained in eV. Then there exists a vector $v \in \ker e$ such that $L^s v \not\subseteq eV$ for some integer s. Let $s \geq 1$

be the smallest among such integers. Then there are matrices $h_1, \ldots, h_s \in L$ with $h_s \cdots h_2 h_1 v \notin eV$ such that, for every i, the matrix h_i is either e or has rank one. Assume that $h_j = e$ for some $j \leq s$. Then $j > 1$ as $ev = 0$. Furthermore, the minimality of s implies

$$h_{j-1} \cdots h_1 v \in eV; \tag{2.1}$$

therefore, as $ew = w$ for every $w \in eV$, we have $h_s \cdots h_{j+1} h_j h_{j-1} \cdots h_1 v = h_s \cdots h_{j+1} e h_{j-1} \cdots h_1 v = h_s \cdots h_{j+1} h_{j-1} \cdots h_1 v$, contradicting the minimality of s. Thus all the matrices $h_1, \ldots, h_s$ are of rank one. Set $v_0 = v$ and for $1 \leq i \leq s$, $v_i = h_i v_{i-1}$. The minimality of s implies that for every $1 \leq i \leq s$ we have $v_i \in L^i v \setminus \sum_{j=0}^{i-1} L^j v$. In particular, the vectors $v_0, \ldots, v_s$ are linearly independent. Since h_i is a rank one transformation on V,

$$h_i V = \mathbb{F} v_i \quad \text{for all } 1 \leq i \leq s. \tag{2.2}$$

From this, and from the minimality of s, we infer $h_j v_{i-1} = 0$ for every $1 \leq i < j \leq s$ (otherwise $v_j \in h_j L^{i-1} v \subseteq L^i v$). We show below that $a := e + h_1 + \cdots + h_s$ is of a rank higher than e, leading to the desired contradiction.

Informally, we build a basis in which the matrix of a is upper triangular. First, we see that (keep in mind (2.1) and (2.2)) $av_{i-1} = ev_{i-1} + \sum_{j<i} h_j v_{i-1} + v_i + \sum_{j>i} h_j v_{i-1} = v_i + ev_{i-1} + \sum_{j<i} h_j v_{i-1} \in v_i + ev_{i-1} + \sum_{j<i} \mathbb{F} v_j \subseteq v_i + \langle v_1, \ldots, v_{i-1} \rangle$ $(i = 1, \ldots, s)$. (With some abuse of notation, for $i = 1$, by $\langle v_1, \ldots, v_{i-1} \rangle$ we mean the zero subspace.) Hence the vectors $av_0, \ldots, av_{s-1}$ span the subspace $\langle v_1, \ldots, v_s \rangle$. Let W be a direct complement to the subspace $\langle v_1, \ldots, v_{s-1} \rangle$ in eV and let $w_1, \ldots, w_t$ be a basis of W ($t = \operatorname{rk} e - s + 1$). Then $aw_i = ew_i + \sum_{j=1}^{s} h_j w_i = w_i + \sum_{j=1}^{s} h_j w_i \in w_i + \langle v_1, \ldots, v_s \rangle$ (by (2.2)). Therefore the vectors $aw_1, \ldots, aw_t$ are linearly independent even modulo the subspace $\langle v_1, \ldots, v_s \rangle$. Together with the fact that $\langle av_0, \ldots, av_{s-1} \rangle = \langle v_1, \ldots, v_s \rangle$, this implies that $\langle av_0, \ldots, av_{s-1}, aw_1, \ldots, aw_t \rangle = \langle v_1, \ldots, v_s, w_1, \ldots, w_t \rangle = \langle eV, v_s \rangle$. Thus the image of a contains a subspace of dimension $\operatorname{rk} e + 1$, and hence $\operatorname{rk} a \geq \operatorname{rk} e + 1$, as claimed. □

In the above proof, the special case $s = 1$ deserves special attention. In that case we have a simple method for increasing the rank over sufficiently large fields which works even *without* any assumption on the presence of rank one matrices. We will use this simple observation later in section 4.

LEMMA 2.2. *If $h, h'' \in \operatorname{Lin}(U, V)$ are transformations such that $h'' \ker h \not\subseteq hU$, then $h' := h + \alpha h''$ will be of a higher rank than h except for at most* $(\operatorname{rk} h + 1)$ *elements $\alpha \in \mathbb{F}$.*

Proof. Let $k = \operatorname{rk} h$, and let U_0 be a subspace of U complementary to $\ker h$. Let $u_1, \ldots, u_k$ be a basis of U_0, and let $v_1, \ldots, v_k$ be a basis of the image hU. Choose a vector $u_{k+1} \in \ker h$ such that $v_{k+1} := h'' u_{k+1} \notin hU$. Consider the matrix of the restriction of $h + xh''$ to $U_0 + \mathbb{F} u_{k+1}$ in the bases $u_1, \ldots, u_k, u_{k+1}$ and $v_1, \ldots, v_k, v_{k+1}$. The last row of the constant term (the matrix of h) is zero, while the lower right entry of the linear term (the matrix of xh'') is x. Expanding by the last row, we obtain that the linear term of the determinant of this $(k+1) \times (k+1)$ matrix is dx, where $d \neq 0$ is the determinant of the upper left $k \times k$ block of h. Thus the determinant is a nonzero polynomial in x of degree at most $(k+1)$, and hence the corresponding $(k+1) \times (k+1)$ block of $h' = h + \alpha h''$ is nonsingular, showing that h' has rank higher than k unless α is a root of this polynomial. □

We state below a simple fact about the linear spaces of matrices that is useful in providing a certificate for the rank maximality of a given matrix.

FACT 2.3. *Let $L \le \mathrm{Lin}(U,V)$, where U and V are finite dimensional spaces over the field $\mathbb{F}$. Then for every $h \in L$ we have* $\operatorname{rk} h \le \dim U - \max\{\dim W - \dim LW \mid W \le U\}$.

Proof. For any subspace $W \le U$ pick a direct complement W' of W in U. Now $\dim U - \operatorname{rk} h = \dim U - \dim hU \ge (\dim W - \dim hW) + (\dim W' - \dim hW') \ge (\dim W - \dim LW) + 0$. □

Using Edmonds' matroid intersection theorem, Lovász (section 3 of [Lov89]) has shown that equality holds provided that h is of maximum rank and if L is spanned by rank one matrices. We give the following algorithmic generalization to the case when L is spanned by rank one matrices and an *arbitrary* rank matrix.

THEOREM 2.4. *Let U and V be two finite dimensional vector spaces over the field $\mathbb{F}$, let $L \le \mathrm{Lin}(U,V)$ be given by a basis, and let an $h \in L$ also be given. Suppose that L is spanned by h and certain (unknown) transformations of rank one.*

1. *Then there exists a deterministic polynomial time algorithm which decides whether h is an element of L of maximum rank. If h is of maximum rank, then a subspace W of U is constructed such that* $\operatorname{rk} h = \dim U - (\dim W - \dim LW)$.

2. *If h is not of maximum rank, then, given rank one transformations that together with h span L, we can compute an element $h' \in L$ with* $\operatorname{rk} h' > \operatorname{rk} h$ *in deterministic polynomial time.*

Proof. We may assume without loss of generality that $\dim U = \dim V$, for otherwise we can pad transformations from L with zeros to obtain a space $L' \le \mathrm{Lin}(U \oplus U', V \oplus V')$, where $\dim U \oplus U' = \dim V \oplus V'$ with some (possibly zero) spaces U', V'. By padding a transformation $b \in \mathrm{Lin}(U,V)$ we mean the map $b' \in \mathrm{Lin}(U \oplus U', V \oplus V')$ which is the direct sum of b and the zero map: $b'(u,u') = (bu, 0)$.

Let $g : V \to U$ be an arbitrary nonsingular linear map such that $gh : U \to U$ is an idempotent. (The matrix of such a map g can be obtained as the product of the matrices corresponding to the pivoting steps in Gaussian elimination for the matrix of h.) As g is invertible, h is of maximum rank within L if and only if gh is of maximum rank within gL. Also, rank one generators of L are mapped to rank one generators of gL. If gh is of maximum rank, then by Lemma 2.1, $\mathrm{Env}(gL) \ker gh \le ghU$. Conversely, if $\mathrm{Env}(gL) \ker gh \le ghU$, then, with $W_0 := \mathrm{Env}(gL) \ker gh$ and $W_1 := \ker gh$, we have $gLW_0, gLW_1 \le W_0 \le ghU$ and $W_0 \cap W_1 = 0$ (if $v \in W_0 \cap W_1$, then $v = ghu$ for some $u \in U$ and $ghv = 0$, implying $0 = ghghu = ghu = v$). Therefore, with $W := W_0 + W_1 \le U$ we have $gLW \le W_0$ and $\dim U - \operatorname{rk} gh = \ker gh = \dim W_1 = \dim W - \dim W_0 \le \dim W - \dim gLW$. Now g being invertible also implies that $\dim U - \operatorname{rk} h \le \dim W - \dim LW$, which together with Fact 2.3 implies that h has maximal rank. Thus if $\mathrm{Env}(gL) \ker gh \le ghU$, then we can efficiently construct W with the required property that it is a witness of the maximality of the rank of gh (resp., h) in gL (resp., L). Thus, h and hence gh are not of maximum rank if and only if $\mathrm{Env}(gL) \ker gh$ is not contained in ghU. This can be easily decided, e.g., by the following breadth-first-search–like algorithm.

Algorithm:

(0) Let A, B, and C be bases for gL, $\ker gh$, and ghU, respectively, and set $D := \emptyset$. Also initialize Q to be an empty FIFO list.

(1) Enqueue all elements of B into Q.
`Outer loop, repeat until` Q `becomes empty`

(2) Dequeue front element u from Q.
`Inner loop, do for each` $a \in A$`:`

(3) If au is linearly independent of C, then output "not contained" and exit.

(4) If au is linearly independent of D, then add au to D and enqueue au into Q.
`End of inner loop`
`End of outer loop`
(5) Output "contained" and exit.

Furthermore, if L is spanned by h and (known) rank one matrices $h_1, \dots, h_\ell$, we use $\{gh_1, \dots, gh_\ell\}$ as basis A in the above algorithm, and supplement step (4) with keeping track of a and u along au when it is enqueued, then, by the proof of Lemma 2.1, we can easily find a linear combination of gh and $gh_1, \dots, gh_\ell$ of higher rank if the algorithm exits in step (3). Multiplying by g^{-1} we obtain an element of L of rank larger than $\operatorname{rk} h$.

Note that initially $\dim Q = |B|$, and it is always dequeued in the outer loop, while sometimes it is enqueued in the inner loop. However, whenever enqueuing happens, $\dim D$ increases; thus Q can be enqueued at most $\dim ghU = \operatorname{rk} h$ times. Thus the number of iterations in the algorithm is at most $(|B| + \operatorname{rk} h) \cdot |A| = (\dim U) \cdot (\dim L)$. □

It is obvious that repeated applications of Theorem 2.4 complete the proof of Theorem 1.1.

We remark that the shortest product $\Pi = gh_1 \cdots gh_\ell$ with $\Pi \ker h \not\subseteq hU$ found by the above algorithm can be interpreted as a generalization of the notion of *augmenting paths* in the classical bipartite matching algorithms. Indeed, let $G = (V_1, V_2, E)$ be a bipartite graph. For simplicity we assume $|V_1| = |V_2|$. We take vector spaces U and V over $\mathbb{F}$ with bases V_1 and V_2, respectively. For an edge $(v_1, v_2) \in E$ we define the linear map $a_{(v_1,v_2)}$, which maps v_1 to v_2 and the other basis elements from V_1 to zero, and let L be the space of linear maps spanned by $a_{(v_1,v_2)}$, where $(v_1, v_2) \in E$. Let H be a (partial) matching in G, and let h be its adjacency matrix, i.e., $h = \sum_{(v_1,v_2)\in H} a_{(v_1,v_2)}$. We define a bijection $\Gamma : V_2 \to V_1$ by reversing the edges in H and supplementing this system to a perfect matching by some independent pairs from $V_2 \times V_1$. Let g be the linear map $V \to U$ extending Γ. We run the above algorithm, where $A = \{a_{(v_1,v_2)} \mid (v_1, v_2) \in E\}$, B consists of the elements of V_1 which are unmatched by H, and C is the set of matched elements of V_2. In this setting the algorithm literally behaves like the classical method, which builds alternating forests, and the edges corresponding to the maps h_i taking part in the product Π form (together with the appropriate edges from H) an augmenting path.

Skew-symmetric matrices of odd size are perhaps the best known examples showing that the inequality in Fact 2.3 cannot be replaced by equality in general; see [Lov79]. These examples can be used to demonstrate that the assertion of Lemma 2.1 fails as well if we omit the assumption that L is generated by e and rank one matrices. To see a small counterexample, let L' be the set of skew-symmetric 3×3 matrices and let $V = \mathbb{F}^3$. Then, as skew-symmetric matrices have even rank, all the nonzero matrices in L' have rank two. Put

$$h = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad g = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{and} \quad e := gh = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then e is an idempotent of rank two, and in $L = gL'$ all the matrices have rank at most two. On the other hand, it is straightforward to check that $\operatorname{Env}(L)$ is the whole algebra of the 3×3 matrices. Therefore $\operatorname{Env}(L) \ker e$ is the whole three-dimensional space V and is not contained in eV.

3. Module morphism problems and matrix completion. In this section we present hardness results of certain problems concerning modules. The key constructions are modules that we call *bipartite modules* as they resemble bipartite graphs.

3.1. Bipartite modules. Let W_1 and W_2 be two linear spaces over $\mathbb{F}$, and assume that we are given a linear subspace $R \leq \mathrm{Lin}(W_1, W_2)$ of linear maps from W_1 to W_2. We assume that R is spanned by ℓ maps: $r_1, \ldots, r_\ell$. We consider the direct sum $W = W_1 \oplus W_2$. We extend transformations $r \in R$ to linear transformations of W by letting r act on W_2 as the zero map. (That is, the extension maps (w_1, w_2) to $(0, rw_1)$.) With some abuse of notation, we denote the extended map also by r and consider R as a subspace of $\mathrm{Lin}(W)$. Let $\mathcal{S}$ be the set $\{r_1, \ldots, r_\ell\}$, and let the (now identity) map $\nu : \mathcal{S} \to \mathrm{Lin}(W)$ define the $\mathcal{S}$-module structure on W (more precisely, W is an $\mathbb{F}\{\mathcal{S}\}$-module). This $\mathcal{S}$-module W is a *bipartite module*.

3.2. Universality of cyclic submodule optimization. In this section we show the following result.

THEOREM 3.1. *There is a deterministic polynomial time reduction from maximum rank matrix completion to cyclic submodule optimization.*

Proof. Let L be a linear space of $\mathbb{F}$-linear maps from U to V. As a first idea, it would be straightforward to consider the bipartite module W for $W_1 = U$, $W_2 = V$, and $R = L$. However, this module does not turn out to be useful for our purposes. The main idea is based on the insight that the data for L can be considered as an element of the tensor product space $U \otimes V \otimes L'$, where L' is a vector space isomorphic to L. Such a tensor can be viewed as a linear space of linear maps in six ways: a space of linear maps from U to V, a space of linear maps from V to U, a space of linear maps from U to L', a space of linear maps from L' to U, a space of linear maps from V to L', and finally a space of linear maps from L' to V. We consider here the last view; that is (identifying L' with L), we give linear maps from L to V.

For every vector u the map $\mu_u : L \to V$ given as $\mu_u(h) := h(u)$ is a linear map, and $\{\mu_u \mid u \in U\}$ is a linear space of linear maps from L to V. We use the bipartite module construction with $W_1 = L$, $W_2 = V$, and $R = \{\mu_u \mid u \in U\}$. Assume that U is spanned by $u_1, \ldots, u_\ell$. Then we put $\mathcal{S} := \{\mu_{u_1}, \ldots, \mu_{u_\ell}\}$. There is a relation between the rank of $h \in L$ and the dimension of the $\mathcal{S}$-submodule of the bipartite module $W = L \oplus V$ generated by (h, v). This submodule is the subspace $\mathbb{F}(h, v) + (0, hU)$. (Indeed, (h, v) must be included in the submodule generated by itself, and application of μ_{u_i} to (h, v) gives $(0, \mu_{u_i}(h)) = (0, hu_i)$ $(i = 1, \ldots, \ell)$. These vectors span the subspace $\mathbb{F}(h, v) + (0, hU)$, and this space remains invariant under the action of elements of $\mathbb{F}\{\mathcal{S}\}$.) The dimension of $\mathbb{F}(h, v) + (0, hU)$ is clearly $(1 + \mathrm{rk}\ h)$ if h is not the zero map. Therefore the maximum dimension of a cyclic submodule of W is one plus the maximum rank in L, and this optimum is taken at generators of the form (h, v), where h is of maximum rank. Thus our construction transforms matrix completion in L to cyclic submodule optimization in W. ∎

3.3. Module morphisms. Let U and V be two $\mathbb{F}\{\mathcal{S}\}$-modules. An $\mathbb{F}$-linear map $\phi \in \mathrm{Lin}(U, V)$ is an $\mathcal{S}$-module *homomorphism* if for every $s \in \mathcal{S}$ and $u \in U$ we have $\phi(su) = s\phi(u)$. The module homomorphism from U to V forms a linear subspace $\mathrm{Hom}_{\mathbb{F}\{\mathcal{S}\}}(U, V)$ of $\mathrm{Lin}(U, V)$. Given the $\mathcal{S}$-module structure on U and V in terms of matrices over bases, a basis for the matrix space representing $\mathrm{Hom}_{\mathbb{F}\{\mathcal{S}\}}(U, V)$ can be computed with $poly(\dim U + \dim V + |\mathcal{S}|)$ field operations by solving a system of homogeneous linear equations.

It is not difficult to construct subspaces of $\mathrm{Lin}(U, V)$ which do not arise as spaces

of module homomorphisms. Thus it is natural to ask how difficult are the matrix completion problems in spaces of module morphisms. It turns out (as shown below) that the cyclic submodules of a bipartite module W (defined as $L \oplus V$ in the last subsection) arise as homomorphic images of another $\mathcal{S}$-module W_0, where $\mathcal{S} = \{r_1, \ldots, r_\ell\}$ and W_0 has basis $b_0, b_1, \ldots, b_\ell$ that by definition satisfy $r_i b_0 = b_i$, $r_i b_j = 0$ $(i, j = 1, \ldots, \ell)$.

This shows that hard matrix completion problems do arise in module morphism spaces. However, curiously enough, deciding existence and construction of module *isomorphisms*, i.e., module homomorphisms which are bijective linear maps, can be accomplished in polynomial time (see [CIK97] with some restriction for the base field and [BL08] for arbitrary fields). We show that this is not the case for testing existence of injective or surjective module morphisms.

Module injection. For the injective case, consider the bipartite modules W and W_0 discussed above. The module W_0 is cyclic; i.e., it is generated by b_0. Therefore a module homomorphism is determined by the image of b_0. In this case for every pair (w_1, w_2) there is indeed a homomorphism with $\psi(b_0) = (w_1, w_2)$. (For $i > 0$ set $\psi(b_i) = (0, r_i w_1)$.) Consider the special case of the bipartite module W used in the proof of Theorem 3.1: let L be a space of linear maps from U to V, and put $W_1 = L$ and $W_2 = V$. Then the image of W_0 at the map ψ, under which the image of b_0 is (h, v), is the subspace spanned by $(h, v), (0, hu_1), \ldots, (0, hu_\ell)$. This ψ is injective if and only if h is. This construction reduces both deciding and finding an injective transformation in L (and also nonsingular matrix completion as a special case) to deciding and finding an injective homomorphism from W_0 to W.

Module surjection. Existence of (resp., finding) injective module morphisms can be transformed to the existence of (resp., finding) surjective morphisms between modules by standard *dualization*. If M is a vector space over $\mathbb{F}$, then by M^* we denote the space of (homogeneous) linear functions from M to $\mathbb{F}$ (that is, $M^* = \mathrm{Lin}(M, \mathbb{F})$). If ϕ is an $\mathbb{F}$-linear map from the space M_1 to M_2, then the map $\phi^* : M_2^* \to M_1^*$ given as $(\phi^* f)v = \phi(fv)$ is again a linear map. (Note that if ϕ is interpreted as multiplication of column vectors by a matrix from the left, then ϕ^* can be interpreted as multiplication of row vectors by the transposed matrix from the right.) Furthermore, if both M_1 and M_2 are finite dimensional, then ϕ is injective (resp., surjective) if and only if ϕ^* is surjective (resp., injective). If M_1 and M_2 are $\mathcal{S}$-modules given by the maps ν_1 and ν_2, then ν_1^* and ν_2^* given as $\nu_i^*(s) = \nu_i(s)^*$ make M_1^* and M_2^* $\mathcal{S}$-modules. Furthermore, the linear map $\phi \in \mathrm{Lin}(M_1, M_2)$ is a module homomorphism from M_1 to M_2 if and only if ϕ^* is a module homomorphism from M_2^* to M_1^*.

So when given vector spaces U, V over $\mathbb{F}$, and a linear subspace L of $\mathrm{Lin}(U, V)$ with $\ell := \dim U$, we first construct modules W and W_0 as in the previous reduction so that the module homomorphism ψ from W_0 to W is injective if and only if $h \in L$ is an injective map where $\psi(b_0) = (h, v)$. Therefore $\Psi \in \mathrm{Hom}_{\mathbb{F}\{\mathcal{S}\}}(W^*, W_0^*)$ is surjective if and only if for the unique $\mathbb{F}$-linear map $\psi : W_0 \to W$ such that $\Psi = \psi^*$ we have that h is injective, where $\psi(b_0) = (h, v)$. This completes the proof of Theorem 1.2.

4. Minimizing number of generators in modules. We saw that cyclic submodule optimization is matrix completion hard. Now we will study the "dual" problem of finding the minimal number of generators of a given module. In this section we give an *efficient* algorithm for minimizing the number of generators in a given $\mathbb{F}\{\mathcal{S}\}$-module. It depends on a greedy property of the dimension of submodules in so-called semisimple modules (which will be vaguely similar to the property in section 2). But first we need to summarize some basic notions and facts from the representation

theory of algebras needed in the proof. For details, we refer the reader to the first few chapters of the textbook [Pie82].

4.1. Preliminaries: Algebras, modules, and their decompositions. Let $\mathbb{F}$ be an arbitrary field. An associative algebra with identity, or *algebra* for short, is a vector space $\mathcal{A}$ over $\mathbb{F}$ equipped with an associative $\mathbb{F}$-bilinear multiplication having a two-sided identity element $1_{\mathcal{A}}$ with respect to the multiplicative structure. If V is a finite dimensional vector space of $\mathbb{F}$, then the $\mathbb{F}$-linear transformations of V form a finite dimensional algebra $\mathrm{Lin}(V)$. Subalgebras of $\mathrm{Lin}(V)$, that is, subspaces closed under multiplication, containing the identity matrix are further examples. (In contrast to section 2, where we considered algebras of linear transformations not necessarily having an identity, in this section it will be convenient to consider algebras with identity only.) An algebra *homomorphism* from $\mathcal{A}$ to $\mathcal{B}$ is an $\mathbb{F}$-linear map $\phi : \mathcal{A} \to \mathcal{B}$ also satisfying $\phi(a_1 \cdot a_2) = \phi(a_1) \cdot \phi(a_2)$ and $\phi(1_{\mathcal{A}}) = 1_{\mathcal{B}}$.

A left $\mathcal{A}$-module, or an *$\mathcal{A}$-module* for short, is an $\mathbb{F}$-linear space V equipped with a bilinear multiplication $\cdot : \mathcal{A} \times V \to V$ which commutes with the multiplication within $\mathcal{A}$ (that is, $a_1 \cdot (a_2 \cdot v) = (a_1 \cdot a_2) \cdot v$). (In [Pie82], right modules are used. Here we use left modules, which are somewhat more common in the literature.) A module V is *unital* if $1_{\mathcal{A}} v = v$ for every $v \in V$. All modules in this work are assumed to be unital and *finite dimensional* over $\mathbb{F}$.

If V is an $\mathcal{A}$-module, then the map $\nu : \mathcal{A} \to \mathrm{Lin}(V)$ defined as $\nu(a)v = a \cdot v$ is a homomorphism from $\mathcal{A}$ into $\mathrm{Lin}(V)$. We say that V is a *faithful* $\mathcal{A}$-module if the kernel of ν is zero; that is, if $a \in \mathcal{A}$ such that $av = 0$ for every $v \in V$, then $a = 0$. If $\mathcal{S}$ is a finite set, then $\mathbb{F}\{\mathcal{S}\}$, the algebra of noncommutative polynomials over $\mathbb{F}$ with indeterminates from $\mathcal{S}$, is an example of an infinite dimensional $\mathbb{F}$-algebra. It is the *free* algebra generated by $\mathcal{S}$: if $\mathcal{A}$ is an algebra and ν is a map from $\mathcal{S}$ into $\mathcal{A}$, then ν can be extended to a unique algebra homomorphism from $\mathbb{F}\{\mathcal{S}\}$ to $\mathcal{A}$. In view of this, an $\mathbb{F}\{\mathcal{S}\}$-module structure on V can be given by an arbitrary map $\nu : \mathcal{S} \to \mathrm{Lin}(V)$. Thus the notion of $\mathcal{S}$-module used in this paper is consistent with the notion of modules over free algebras.

A *submodule* of an $\mathcal{A}$-module is a linear subspace also closed under multiplication by elements of $\mathcal{A}$. The *factor space* of a submodule inherits the $\mathcal{A}$-module structure in a natural way and so do *direct sums* of linear spaces which are $\mathcal{A}$-modules. An $\mathcal{A}$-module V is called *simple* if it has exactly two submodules: the whole V and the zero submodule. The *radical* of a module is the intersection of its maximal (more precisely, maximal proper) submodules. A module V is called *semisimple* if it is isomorphic to a direct sum of simple modules. By section 2.7 of [Pie82], V is semisimple if and only if its radical is the zero submodule. Furthermore, the factor module of V by its radical is always semisimple. By section 2.5 of [Pie82], the isomorphism classes of the constituents and their multiplicities in a decomposition of a semisimple module into a direct sum of simple modules are uniquely determined. Direct sums and homomorphic images of semisimple modules are semisimple.

Let V be a finite dimensional $\mathbb{F}\{\mathcal{S}\}$-module, and let $\mathcal{A}$ be the enveloping algebra $\mathrm{Env}(I \cup \nu(\mathcal{S}))$ (the subalgebra of $\mathrm{Lin}(V)$ generated by the identity and $\nu(\mathcal{S})$). Then $\mathcal{A}$ is the image of $\mathbb{F}\{\mathcal{S}\}$ under the unique algebra homomorphism from $\mathbb{F}\{\mathcal{S}\}$ to $\mathrm{Lin}(V)$ extending ν, and V is a faithful $\mathcal{A}$-module in the natural way. We work with the $\mathcal{A}$-module structures of V, its submodules, and factors as they coincide with the $\mathcal{S}$-module structures of the same objects. Assume that V is semisimple. Then by section 4.1 of [Pie82], $\mathcal{A}$ considered as a left module over itself by the algebra multiplication is also semisimple. Such algebras are called *semisimple*. Modules over

semisimple algebras are semisimple, again by section 4.1 of [Pie82].

Let $\mathcal{A}$ be a semisimple algebra over $\mathbb{F}$, and let $\mathcal{A}$ as a left module over itself be isomorphic to the direct sum:

$$\bigoplus_{i=1}^{t} V_i^{m_i}, \tag{4.1}$$

where V_i are pairwise nonisomorphic $\mathcal{A}$-modules. Let V be an $\mathcal{A}$-module. As V is a homomorphic image of at most $\dim V$ copies of the module $\mathcal{A}$, we have

$$V \cong \bigoplus_{i=1}^{t} V_i^{s_i}, \tag{4.2}$$

where the multiplicities s_i are nonnegative integers.

LEMMA 4.1. *Let $\mathcal{A}$ and V be as above, and let ℓ be a positive integer. Let U be a submodule of V generated by ℓ elements. Then U is of maximum dimension among the ℓ-generated submodules of V if and only if $U \cong \bigoplus_{i=1}^{t} V_i^{d_i}$, where $d_i := \min(s_i, \ell m_i)$.*

Proof. Let W_i be the sum of all simple submodules of V not isomorphic to V_i. Then the $\mathcal{A}$-module V/W_i is isomorphic to $V_i^{s_i}$ and the submodule dimension in V is maximized if and only if it is maximized in V/W_i for all $i \in [t]$. As a single generator in $V_i^{s_i}$ can generate a submodule of dimension at most that of $V_i^{\min(s_i, m_i)}$, we get that ℓ generators in $V_i^{s_i}$ can generate a submodule of dimension at most that of $V_i^{d_i}$. Repeating this for every $i \in \{1, \ldots, t\}$, we obtain that the maximum dimension is at most the dimension of the direct sum in the statement.

To see that this module occurs in fact as a submodule of V, let W be the direct sum of ℓ copies of $\mathcal{A}$ (as a left $\mathcal{A}$-module), and let $w_1 = (1_{\mathcal{A}}, 0, \ldots, 0), \ldots, w_\ell = (0, \ldots, 0, 1_{\mathcal{A}})$. Let W_0 be a submodule of W isomorphic to $\bigoplus_{i=1}^{t} V_i^{\ell m_i - d_i}$, and let V_0 be a submodule of V isomorphic to $\bigoplus_{i=1}^{t} V_i^{d_i}$. Then $V_0 \cong W/W_0$ and W/W_0 is generated by ℓ elements: the images of $w_1, \ldots, w_\ell$ under the projection $W \to W/W_0$. Thus V_0 can be generated by the images of the latter ℓ elements under any isomorphism $W/W_0 \cong V_0$. ∎

4.2. A greedy optimization of the submodule dimension in semisimple modules. In this section V denotes a finite dimensional $\mathbb{F}\{\mathcal{S}\}$-module, and $\mathcal{A}$ stands for the enveloping algebra $\mathrm{Env}(\nu(\mathcal{S}) \cup I)$. For subsets $\mathcal{B} \subseteq \mathcal{A}$ and $U \subseteq V$ by $\mathcal{B}U$ we denote the linear span of the products bu, where $b \in \mathcal{B}$ and $u \in U$. In this context we omit braces around one-element sets. In particular, for $v \in V$, the submodule generated by v is $\mathcal{A}v$.

The *annihilator* $\mathrm{Ann}_{\mathcal{A}}(U)$ of $U \subseteq V$ is $\{a \in \mathcal{A} \mid au = 0 \text{ for every } u \in U\}$. Note that the annihilator $\mathrm{Ann}_{\mathcal{A}}(v)$ of the single element $v \in V$ is just the kernel of the linear map $\mu_v : \mathcal{A} \to V$ given as $\mu_v(a) = av$. The following lemma states that if the rank of μ_v is not maximal, then we are in the situation of Lemma 2.2.

LEMMA 4.2. *Assume that V is semisimple. Then, for an arbitrary $u \in V$, $\dim \mathcal{A}u = \max\{\dim \mathcal{A}u' \mid u' \in V\}$ if and only if $\mathrm{Ann}_{\mathcal{A}}(u)V \subseteq \mathcal{A}u$.*

Furthermore, if $\mathrm{Ann}_{\mathcal{A}}(u)V \not\subseteq \mathcal{A}u$, then an element u' with $\dim \mathcal{A}u' > \dim \mathcal{A}u$ can be constructed using poly$(|\mathcal{S}| + \dim V)$ operations in $\mathbb{F}$.

Remark 4.3. The lemma generalizes a result of Babai and Rónyai which was used in [BR90] for solving the cyclic submodule optimization in modules over simple algebras. The proof can be found in [CIK97]. For completeness, we discuss it here as well. The second part of the lemma is especially interesting for small base fields where Lemma 2.2 does not apply.

Proof. Let V be a semisimple $\mathcal{S}$-module, and let $\mathcal{A} = \mathrm{Env}(I \cup \nu(\mathcal{S}))$. Let $\mathcal{A}$ (resp., V) be decomposed as in (4.1) (resp., (4.2)). Let $u \in V$. Assume that the dimension of the submodule $\mathcal{A}u$ is not maximal. Then, by Lemma 4.1, there exists an index i such that the multiplicity of V_i in $\mathcal{A}u$ is less than both s_i and m_i. Let W be the submodule of V which is the direct sum of the constituents of V not isomorphic to V_i. Then $V/W \cong V_i^{s_i}$ and $V/(W + \mathcal{A}u)$ is isomorphic to V_i^h with some $h > 0$. Recall that for a subset X of V the annihilator of X in $\mathcal{A}$, denoted by $\mathrm{Ann}_{\mathcal{A}}(X)$, is $\{a \in \mathcal{A} \mid ax = 0 \text{ for every } x \in X\}$. Assume that $\mathrm{Ann}_{\mathcal{A}}(u)V \subseteq \mathcal{A}u$. Then every element of $\mathrm{Ann}_{\mathcal{A}}(u)$ acts as zero on the factor module $V/\mathcal{A}u$ and hence also on the factor $V/(W + \mathcal{A}u)$. As the latter module is isomorphic to V_i^h, we obtain that $\mathrm{Ann}_{\mathcal{A}}(u) \subseteq \mathrm{Ann}_{\mathcal{A}}(V_i)$. Recall that the map $\mu_u : \mathcal{A} \to V$ is given as $\mu_u(a) = au$. It is an $\mathcal{A}$-module homomorphism from the left module $\mathcal{A}$ to V. Its kernel is $\mathrm{Ann}_{\mathcal{A}}(u)$, and its image is $\mathcal{A}u$. Therefore $\mathcal{A}u \cong \mathcal{A}/\mathrm{Ann}_{\mathcal{A}}(u)$. Now $\mathrm{Ann}_{\mathcal{A}}(V_i)$ is also an $\mathcal{A}$-submodule of $\mathcal{A}$. Let L be a submodule of $\mathcal{A}$ isomorphic to V_i. We claim that $LV_i \neq 0$. Indeed, if $LV_i = 0$, then, by the assumed isomorphism, $LL = 0$ as well, which is impossible by section 3.2 of [Pie82]. The claim implies that the multiplicity of V_i in $\mathrm{Ann}_{\mathcal{A}}(V_i)$ is zero and the same holds in $\mathrm{Ann}_{\mathcal{A}}(u) \subseteq \mathrm{Ann}_{\mathcal{A}}(V_i)$. But then the multiplicity of V_i in the factor module $\mathcal{A}/\mathrm{Ann}_{\mathcal{A}}(u) \cong \mathcal{A}u$ is m_i. This contradiction finishes the proof of the statement that if $\mathcal{A}u$ is not of maximum dimension, then in fact $\mathrm{Ann}_{\mathcal{A}}(u)V \not\subseteq \mathcal{A}u$.

To see the reverse implication, assume that $\mathrm{Ann}_{\mathcal{A}}(u)V \not\subseteq \mathcal{A}u$, and let $w \in V$ and $b \in \mathrm{Ann}_{\mathcal{A}}(u)$ such that $bw \notin \mathcal{A}u$. By section 2.4 of [Pie82], there exists a submodule W' of V such that $W' \cap \mathcal{A}u = 0$ and $W' + \mathcal{A}u = V$. Write $w = au + w'$, where $a \in \mathcal{A}$ and $w' \in W'$. Put $u' = u + w'$. As $\mathcal{A}w' \in W'$, we have $\mathcal{A}u' + W' = \mathcal{A}u + W'$. On the other hand, from $bw \notin \mathcal{A}u$ but $bau \in \mathcal{A}u$, we infer that bw' is a nonzero element of W', and by the equality $bu' = bu + bw' = bw'$, it is also an element of $\mathcal{A}u'$. Therefore $\dim \mathcal{A}u' > \dim V - \dim W' = \dim \mathcal{A}u$, as required.

For a polynomial time implementation of the construction above, notice that a basis for $\mathrm{Ann}_{\mathcal{A}}(u)$ can be found by solving a system of linear equations. Then b and w can be found by testing membership of products of pairs of basis elements for $\mathrm{Ann}_{\mathcal{A}}(u)$ and those for V. To compute a direct complement of $\mathcal{A}u$, we first compute a projection π of V onto $\mathcal{A}u$ such that $\pi a = a\pi$ for every element $a \in \mathcal{A}$ (equivalently, for every element of a system of generators for $\mathcal{A}$, say $\nu(\mathcal{S})$). (Recall that a projection π onto a subspace V' of V is a map whose image is V' and which acts as the identity on V'. If W' is submodule complementary to $\mathcal{A}u$, then the unique linear map which is the identity on $\mathcal{A}u$ and zero on W' is a projection onto $\mathcal{A}u$ which commutes with the action of $\mathcal{A}$ on V.) Once π is constructed we take $\pi' = I - \pi$. It is straightforward to see that the image $W' = \pi' V$ is in fact a direct complement of $\mathcal{A}u$. The element w' in the argument above is then just $\pi' w$ and $u' = u + \pi' w$. This finishes the proof of Lemma 4.2. ∎

The next lemma can be used to give a generalization for submodules generated by larger systems (e.g., noncyclic modules).

LEMMA 4.4. *Assume that V is semisimple. Then, for arbitrary positive integer ℓ and for elements $u_1, \ldots, u_\ell \in V$, $\dim \mathcal{A}\{u_1, \ldots, u_\ell\} = \max\{\dim \mathcal{A}U \mid U \subseteq V, \#U \leq \ell\}$ if and only if for every $i \in [\ell]$, the $\mathcal{S}$-submodule generated by $u_i + W_i$ in the factor module V/W_i is of maximum dimension, where W_i denotes the submodule generated by $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_\ell$.*

Proof. Let V be a semisimple $\mathcal{S}$-module, and let $\mathcal{A} = \mathrm{Env}(I \cup \nu(\mathcal{S}))$. Let $\mathcal{A}$ (resp., V) be decomposed as in (4.1) (resp., (4.2)). Let $u_1, \ldots, u_\ell \in V$, and let $W_i = \mathcal{A}(\{u_1, \ldots, u_\ell\} \setminus \{u_i\})$. As $\mathcal{A}\{u_1, \ldots, u_\ell\} = \mathcal{A}u_i + W_i$, it is obvious that if, for

some index i there is an element u_i' such that modulo W_i, $\mathcal{A}u_i'$ has a larger dimension than $\mathcal{A}u_i$, then replacing u_i with u_i' results in a system generating a submodule of larger dimension.

To see the reverse implication let $W = \mathcal{A}\{u_1, \ldots, u_\ell\}$ and assume that for every i, the submodule of V/W_i generated by $u_i + W_i$, that is, W/W_i, is a maximal dimensional cyclic submodule of V/W_i. Let $j \in \{1, \ldots, t\}$. By Lemma 4.1, for every i, the multiplicity of V_j in W/W_i either equals the multiplicity of V_j in V/W_i or is just m_j. If for some index i the former is the case, then the multiplicity of V_j in V/W is zero. Otherwise the multiplicity of V_j in W/W_i is m_j for every index i. In the former case the multiplicity of V_j in W is the maximum possible among all submodules. Assume the latter case and let U_{ij} denote the direct sum of the constituents of $\mathcal{A}u_i$ isomorphic to V_j. Then, for every index i, we have that $U_{ij} \cap W_i = 0$ and U_{ij} is isomorphic to a direct sum of m_j copies of V_j, as otherwise the multiplicity of V_j in W/W_i would be less than m_j. Thus U_{ij} intersects $\sum_{i' \neq i} U_{i'j}$ trivially; therefore they form an independent system and hence $\sum_{i=1}^{\ell} U_{ij} \cong V_j^{\ell m_j}$, showing that the multiplicity of V_j is optimal in this case as well. Repeating this for every irreducible module V_j, we obtain that the dimension of W is indeed the maximum possible. This finishes the proof of Lemma 4.4. ∎

The two lemmas above together with Lemma 2.2 immediately give the following.

PROPOSITION 4.5. *Let $v_1, \ldots, v_n$ be a basis of the semisimple $\mathbb{F}\{\mathcal{S}\}$-module V. Assume that $u_1, \ldots, u_\ell$ are elements of V such that the submodule generated by $u_1, \ldots, u_\ell$ is not of maximum dimension among the submodules of V generated by at most ℓ elements. If the $\mathbb{F}\{\mathcal{S}\}$-module structure on V is given by an array of matrices, then we can find an index i and construct $u_i' \in V$ using poly$(|\mathcal{S}| + n)$ operations such that replacing u_i with u_i' results in a submodule of larger dimension.*

Furthermore, if $|\mathbb{F}| > n$, then there exist indices $i \in [\ell]$, $j \in [n]$ such that replacing u_i with $(u_i + \omega v_j)$ results in a submodule of larger dimension except for at most n elements ω from $\mathbb{F}$.

The above greedy property for the submodules of a semisimple module gives us the following technical lemma for *general* modules. It will be useful in the subsequent algorithm for optimizing the number of generators in any module without computing the radical explicitly.

LEMMA 4.6. *Let $v_1, \ldots, v_n$ be a basis of the $\mathcal{S}$-module V which can be generated by ℓ elements, and let $u_1, \ldots, u_\ell$ be elements of V such that $U = \mathcal{A}\{u_1, \ldots, u_\ell\} < V$. If W is a nonzero submodule such that $V = U \oplus W$, then there exist $i \in [\ell]$, $j \in [n]$ such that for $U' := \mathcal{A}\{u_1, \ldots, u_i + \lambda v_j, \ldots, u_\ell\}$, $V = U' + W$, but $U' \cap W \neq \{0\}$ except for at most $2n$ elements $\lambda \in \mathbb{F}$.*

Proof. Let U_0, W_0 be the radicals of U, W, respectively. Let $V_0 = U_0 \oplus W_0$. Then the factor module $V/V_0 \cong U/U_0 \oplus W/W_0$ is semisimple, and we can apply Proposition 4.5 to choose $i \in [\ell]$, $j \in [n]$ such that the number of λ's, for which the dimension of $(U' + V_0)/V_0$ is not larger than the dimension of $(U + V_0)/V_0$, is at most $\dim V/V_0$. Also for the same i, j the λ's, for which $\{u_1, \ldots, u_i + \lambda v_j, \ldots, u_\ell\} \cup W$ do not span the whole of V, are the roots of a nonzero $\mathbb{F}$-polynomial of degree at most $\dim V$. Thus for this i, j the number of λ's, for which either $\dim U' \leq \dim U$ or $V \neq U' + W$, is at most $\dim V/V_0 + \dim V \leq 2n$. ∎

4.3. Algorithm for finding ℓ generators. Using the previous lemma, now we describe an iterative algorithm for finding a minimal set of generators of a given module over a sufficiently large ground field.

Input: An $\mathcal{A}$-module V given in terms of a set of generators. We assume that $\mathcal{A}$ is an $\mathbb{F}$-algebra where $|\mathbb{F}| > 2 \dim V$.
Output: A set of at most ℓ elements generating V over $\mathcal{A}$.
Algorithm:

(0) Initially pick any irredundant generating set $\{u_1, \ldots, u_\ell, u_{\ell+1}, \ldots\}$, and set $U := \mathcal{A}\{u_1, \ldots, u_\ell\}$ and $W := \mathcal{A}\{u_{\ell+1}, \ldots\}$. Then $V = U + W$.
`Outer loop:`
(1) Set $W' := U \cap W$.
(2) If $W' = W$, then output U and exit.
`Inner loop:`
(3) Apply Lemma 4.6 in V/W' to obtain U' generated by ℓ elements and satisfying $U' + W = V$ and $(U' + W') \cap W > W'$.
(4) If such a U' cannot be found, then report "ℓ generators are insufficient for V" and exit.
Else set $U := U'$.
(5) If $W \not\leq U + W'$, then continue *inner* loop with $W' = (U + W') \cap W$.
Else continue *outer* loop with $W = W'$.

Analysis of the algorithm. At each step of the algorithm there is a pair (U, W) of $\mathcal{S}$-modules such that $V = U + W$ and U is known in terms of ℓ generators. At every repetition of the inner loop, W' becomes a larger submodule of W, since at step (5) we know (from step (3)) that $(U + W') \cap W$ is strictly larger than W'. At every repetition of the outer loop, W becomes a smaller submodule of V, since at step (5) we know (again from step (3)) that W' is strictly smaller than W. Thus, the number of times the algorithm can loop is bounded by $(\dim V)^2$, which makes the algorithm polynomial time. This gives a proof of Theorem 1.3 over large base fields.

Over small base fields we use the algorithm of [FR85] or [CIW97] to compute the radical of $\mathcal{A}$ and the radical V_0 of V therefrom and compute a minimal generating set Γ_0 of the factor module V/V_0 using Proposition 4.5 directly. For each $u_0 \in \Gamma_0$ we pick a representative $u \in u_0 + V_0$ and obtain a subset $\Gamma \subseteq V$ such that $|\Gamma| = |\Gamma_0|$ and $\Gamma \cup V_0$ generates V. By a standard property of the radical, we show that Γ itself generates V. Indeed, let U be the submodule generated by Γ. If $U \neq V$, then there is a *maximal* (proper) submodule $U' \supseteq U \supseteq \Gamma$. But $U' \geq V_0$ by the definition of V_0, and therefore $U' \supseteq \Gamma \cup V_0$, implying $U' \supseteq V$, which is a contradiction to U' being proper. This ends the proof of Theorem 1.3.

5. Concluding remarks. We have shown that the maximum rank matrix in a linear space generated by rank one matrices and a further matrix of arbitrary rank can be found in deterministic polynomial time if the rank one generators are given. It would be interesting to know if there is an efficient deterministic method in the case where the rank one generators are not known. In this direction we have a deterministic polynomial time algorithm, which, given a matrix of maximum rank, constructs a certificate that the rank is in fact maximal (see Theorem 2.4) without knowing the rank one generators. This implies that, over sufficiently large base fields, the maximum rank matrix can be constructed in *Las Vegas* polynomial time. The best result of this flavor is the deterministic polynomial time algorithm of Gurvits [Gur03, Gur04] which decides whether there exists a nonsingular matrix in the space generated by rational matrices under the assumption that the span over the complex numbers can be generated by unknown rank one matrices (with not necessarily rational entries). Unfortunately, this algorithm decides the mere existence of a nonsingular matrix without explicitly constructing one.

The space of the maps $\mu_v : \mathcal{A} \to V$, where V is a *semisimple* $\mathcal{S}$-module and $\mathcal{A}$ is the corresponding enveloping algebra, has a curious property that if μ_v is not of maximum rank, there is a $v'' \in V$ such that Lemma 2.2 applies for $h = \mu_v$ and $h'' = \mu_{v''}$ (see Lemma 4.2). In particular, over a sufficiently large field $\mathbb{F}$ the rank of $\mu_v + \alpha\mu_{v''}$ will be higher for some v'' chosen from an arbitrary basis of V and a "generic" $\alpha \in \mathbb{F}$.

It would be interesting to find more classes $\mathcal{L}$ of spaces of linear maps with such a "local rank incrementing" property: there is a constant c such that, for every $L \in \mathcal{L}$, if $h \in L$ is not of maximum rank, then from an arbitrary basis $h_1, \ldots, h_\ell$ of L one can choose maps $h_{i_1}, \ldots, h_{i_c}$ such that $h + \alpha_1 h_{i_1} + \cdots + \alpha_c h_{i_c}$ has higher rank for some $\alpha_1, \ldots, \alpha_c \in \mathbb{F}$ ($\mathbb{F}$ is large enough).

Acknowledgments. We would like to thank the anonymous referees for several suggestions. We are grateful to the Hausdorff Research Institute for Mathematics, Bonn, for its hospitality and kind support.

REFERENCES

[BR90] L. Babai and L. Rónyai, *Computing irreducible representations of finite groups*, Math. Comp., 55 (1990), pp. 705–722.

[BL08] P. A. Brooksbank and E. M. Luks, *Testing isomorphism of modules*, J. Algebra, 320 (2008), pp. 4020–4029.

[BFS99] J. F. Buss, G. S. Frandsen, and J. Shallit, *The computational complexity of some problems of linear algebra*, J. Comput. System Sci., 58 (1999), pp. 572–596.

[CIK97] A. Chistov, G. Ivanyos, and M. Karpinski, *Polynomial time algorithms for modules over finite dimensional algebras*, in Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation (ISSAC '97), ACM, New York, 1997, pp. 68–74.

[CIW97] A. M. Cohen, G. Ivanyos, and D. B. Wales, *Finding the radical of an algebra of linear transformations*, J. Pure Appl. Algebra, 117/118 (1997), pp. 177–193.

[Edm67] J. Edmonds, *Systems of distinct representatives and linear algebra*, J. Res. Nat. Bur. Standards Sect. B, 71B (1967), pp. 241–245.

[FR85] K. Friedl and L. Rónyai, *Polynomial time solutions of some problems of computational algebra*, in Proceedings of the Seventeenth Annual ACM Symposium on Theory of Computing (STOC '85), 1985, pp. 153–162.

[Gee99] J. F. Geelen, *Maximum rank matrix completion*, Linear Algebra Appl., 288 (1999), pp. 211–217.

[GI05] J. F. Geelen and S. Iwata, *Matroid matching via mixed skew-symmetric matrices*, Combinatorica, 25 (2005), pp. 187–215.

[GIM03] J. F. Geelen, S. Iwata, and K. Murota, *The linear delta-matroid parity problem*, J. Combin. Theory Ser. B, 88 (2003), pp. 377–398.

[Gur03] L. Gurvits, *Classical deterministic complexity of Edmonds' problem and quantum entanglement*, in Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing (STOC '03), 2003, pp. 10–19.

[Gur04] L. Gurvits, *Classical complexity and quantum entanglement*, J. Comput. System Sci., 69 (2004), pp. 448–484.

[HKM05] N. J. A. Harvey, D. R. Karger, and K. Murota, *Deterministic network coding by matrix completion*, in Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '05), 2005, pp. 489–498.

[HKY06] N. J. A. Harvey, D. R. Karger, and S. Yekhanin, *The complexity of matrix completion*, in Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '06), 2006, pp. 1103–1111.

[IM83] O. Ibarra and S. Moran, *Probabilistic algorithms for deciding equivalence of straight-line programs*, J. ACM, 30 (1983), pp. 217–228.

[KI03] V. Kabanets and R. Impagliazzo, *Derandomizing polynomial identity tests means proving circuit lower bounds*, in Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing (STOC '03), 2003, pp. 355–364.

[Lov79] L. Lovász, *On determinants, matchings, and random algorithms*, in Fundamentals of Computation Theory, Akademie-Verlag, Berlin, 1979, pp. 565–574.

[Lov89] L. Lovász, *Singular spaces of matrices and their applications in combinatorics*, Bol. Soc. Brasil. Mat. (N.S.), 20 (1989), pp. 87–99.

[Mur00] K. Murota, *Matrices and Matroids for Systems Analysis*, Springer-Verlag, Berlin, 2000.

[Pie82] R. S. Pierce, *Associative Algebras*, Springer-Verlag, New York, 1982.

[Sch80] J. T. Schwartz, *Fast probabilistic algorithms for verification of polynomial identities*, J. ACM, 27 (1980), pp. 701–717.

[Val79] L. G. Valiant, *Completeness classes in algebra*, in Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing (STOC '79), 1979, pp. 249–261.

[Zip79] R. Zippel, *Probabilistic algorithms for sparse polynomials*, in Symbolic and Algebraic Computation, Springer-Verlag, Berlin, New York, 1979, pp. 216–226.