Deterministic Polynomial Time Algorithms for Matrix Completion Problems

Gábor Ivanyos * Marek Karpinski † 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 as to maximize the resulting matrix *rank*. Matrix completion belongs to the fundamental problems in computational complexity with numerous important algorithmic applications, among others, in computing dynamic transitive closures or multicast network codings [HKM05, HKY06]. We design efficient deterministic algorithms for common generalizations of the results of Lovász and Geelen on this problem by allowing *linear functions* in the entries of the input matrix such that the submatrices corresponding to each variable have rank one. We present also a deterministic polynomial time algorithm for finding the *minimal* number of generators of a given module structure given by matrices. We establish further several hardness results related to matrix algebras and modules. As a result we connect the classical problem of polynomial identity testing with checking surjectivity (or injectivity) between two given modules. One of the elements of our algorithm is a construction of a greedy algorithm for finding a maximum rank element in the more general setting of the problem. The proof methods used in this paper could be also of independent interest.

1 Introduction

A *linear matrix* is a matrix having linear functions as its entries, say the linear functions 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). 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 (actually, more or less a special case) is *nonsingular matrix completion*, where we have a square linear matrix and we are interested in an assignment resulting in a nonsingular matrix. If the ground field is sufficiently large then the maximum

*Computer and Automation Research Institute of the Hungarian Academy of Sciences (MTA SZTAKI), Lágymányosi u. 11, 1111 Budapest. E-mail: Gabor.Ivanyos@sztaki.hu

†Department of Computer Science and Hausdorff Center for Mathematics, University of Bonn, 53117 Bonn. E-mail: marek@cs.uni-bonn.de

‡Hausdorff Center for Mathematics, Endenicher Allee 62, University of Bonn, 53115 Bonn. E-mail: ns@hcm.uni-bonn.de

§The authors thank the Hausdorff Research Institute for Mathematics, Bonn for the kind support.

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 (resp. determining the maximum rank) is in deterministic polynomial time reducible to instances of finding (resp. 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 contains a special case of PIT: any arithmetic formula can be written as the determinant of a linear matrix [Val79], hence the formula would be zero iff the corresponding matrix could attain full rank (assuming a large enough field). Over large fields, this makes matrix completion an important problem in coRP as its derandomization would imply circuit lower bounds (see Kabanets & 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 were first studied by Buss et al. [BFS99] and more recently by Harvey et al. [HKY06]. In the former paper nonsingular matrix completion is proved to be NP-hard over fields of constant size, while the latter showed that matrix completion over the field $\mathbb{F}_2$ is NP-hard even if we restrict 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 et al. [HKM05], building on the works of Geelen [Gee99] and Murota [Mur00], gave an efficient deterministic algorithm for matrix completion over *any* field if the mixed matrix has each variable appearing at most once. While Geelen at al. [GIM03, GI05] gave an efficient deterministic algorithm when the mixed matrix is *skew-symmetric* and has each variable appearing at most twice.

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 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 and showing that a greedy approach can be utilized to gradually increase the rank of an element in L. In fact our method is robust enough to check the rank maximality of a given matrix in L *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 (an *algebra* is 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.

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 algebras, however the main concepts and computational tasks we are concerned with can be understood without any knowledge from the theory of algebras.) Let $\mathcal{S}$ be a finite set. A vector space V over the field $\mathbb{F}$ equipped 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 $\mathcal{S}'$.

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 $\nu(B)$ corresponding to $B \in \mathcal{S}$. 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 $V = \mathcal{A}U$.

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 $\{\mu_v | v \in V\}$ is a linear space of linear maps from $\mathcal{A}$ to V. The rank of μ_v is the dimension of the submodule $\mathcal{A}v$ generated by v. 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. We also give analogous hardness results for existence of injective resp. surjective homomorphisms between modules (a $\mathcal{S}$-module *homomorphism* 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 Brookbanks & Luks [BL08] for module isomorphism problem (see also Chistov et al. [CIK97] over special base fields).

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 latter:

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

Note that the above result includes testing cyclicity of modules efficiently over any field. This problem was considered in [CIK97] over special fields. The algorithm is based on a greedy approach analogous to the method for Theorem 1.1 implicitly using certain submodule dimension optimization technique for a special class of (so called *semisimple*) modules.

2 Matrix Completion with Rank One Matrices

Let V be a finite dimensional vector space over the field $\mathbb{F}$ and $L \leq \mathrm{Lin}(V)$ be a $\mathbb{F}$-linear space of linear transformations. Recall that $\mathrm{Env}(L)$, the *enveloping algebra* of L is the linear span of products $h_1h_2\cdots h_s$ ($s \geq 1,\ h_1,\ldots,h_s \in L$). Obviously, $\mathrm{Env}(L)$ is also spanned by products of elements from an arbitrary basis of L. We will finally 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.

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 $h_s \cdots h_2h_1v \notin eV$ 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 $h_{j-1}\cdots h_1v \in eV$, therefore $h_s\cdots h_{j+1}h_jh_{j-1}\cdots h_1v = h_s\cdots h_{j+1}eh_{j-1}\cdots h_1v = h_s\cdots h_{j+1}h_{j-1}\cdots h_1v$, contradicting the minimality of s. Thus all the matrices $h_1,\ldots,h_s$ are of rank one. Set $v_0 = v$ and for $1 \leq i \leq s$, $v_i = h_iv_{i-1}$. The minimality of s implies that for every $1 \leq i \leq s$ we have $v_i \in L^iv \setminus \sum_{j=0}^{i-1} L^jv$. In particular, the vectors $v_0,\ldots,v_s$ are linearly independent. Since h_i is a rank one transformation on V, $h_iV = \mathbb{F}v_i$ ($i = 1,\ldots,s$). From this, and from the minimality of s we infer $h_jv_{i-1} = 0$ for every $1 \leq i < j \leq s$ (otherwise $v_j \in h_jL^{i-1}\ker e \subseteq L^i\ker e$). We show below that $a := e+h_1+\ldots+h_s$ 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 $av_{i-1} = ev_{i-1} + \sum_{j<i} h_jv_{i-1} + v_i + \sum_{j>i} h_jv_{i-1} = v_i + ev_{i-1} + \sum_{j<i} h_jv_{i-1} \in v_i + ev_{i-1} + \sum_{j<i}\mathbb{F}v_j \subseteq v_i + \langle v_1,\ldots,v_{i-1}\rangle$ ($i = 1,\ldots,s$). (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 $w_1,\ldots,w_t$ be a basis of W ($t = \mathrm{rk}\ e - s + 1$). Then $aw_i = ew_i + \sum_{j=1}^{s} h_jw_i = w_i + \sum_{j=1}^{s} h_jw_i \in w_i + \langle v_1,\ldots,v_s\rangle$. 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 $\mathrm{rk}\ e+1$ and hence $\mathrm{rk}\ a \geq \mathrm{rk}\ e+1$, as claimed. $\square$

In the proof above, 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 $h, h'' \in \mathrm{Lin}(V, U)$ are transformations such that $h'' \ker h \not\subseteq hV$ then $h' := h + \alpha h''$ will be of a higher rank than h except for at most* rk $h + 1$ *many* $\alpha \in \mathbb{F}$.

Proof. Let $k = \mathrm{rk}\ h$, and let V_0 be a subspace of V complementary to $\ker h$. Let $v_1, \ldots, v_k$ be a basis of V_0 and let $u_1, \ldots, u_k$ be a basis of the image hV. Choose a vector v_{k+1} from $\ker h$ such that $u_{k+1} := h'' v_{k+1} \notin hV$. Consider the matrix of the restriction of $h + xh''$ to $V_0 + \mathbb{F}v_{k+1}$ in the bases $v_1, \ldots, v_k, v_{k+1}$ and $u_1, \ldots, u_k, u_{k+1}$. The last row and column 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 $k+1$ by $k+1$ 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 $k+1$ by $k+1$ 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 $L \leq \mathrm{Lin}(V, U)$, where U and V are finite dimensional spaces over the field $\mathbb{F}$. Then for every $h \in L$ we have* $\mathrm{rk}\ h \leq \dim V - max\{\dim W - \dim LW | W \leq V\}$.

Proof. For any subspace $W \leq V$ pick a direct complement W' of W in V. Now $\dim V - \mathrm{rk}\ h = \dim V - \dim hV \geq (\dim W - \dim hW) + (\dim W' - \dim hW') \geq \dim W - \dim LW$. □

Using Edmonds' Matroid Intersection Theorem, Lovász (Section 3, [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 $L \leq \mathrm{Lin}(V, U)$ be given by a basis and let an $h \in L$ be also 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 if h is an element of L of maximum rank. If h is of maximum rank then a subspace W of V is constructed such that $\mathrm{rk}\ h = \dim V - max\{\dim W - \dim LW | W \leq V\}$.
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 $\mathrm{rk}\ h' > \mathrm{rk}\ h$ *in deterministic polynomial time.*

Proof. We may assume wlog that $\dim V = \dim U$, for otherwise we can pad transformations from L with zeros to obtain a space $L' \leq \mathrm{Lin}(V \oplus V', U \oplus U')$, where $\dim U \oplus U' = \dim V \oplus V'$ with some (possibly zero) spaces U', V'. By padding a transformation $b \in \mathrm{Lin}(V, U)$ we mean the map $b' \in \mathrm{Lin}(V \oplus V', U \oplus U')$ which is the direct sum of b and the zero map: $b'(v, v') = (bv, 0)$.

Let $g : U \to V$ be an arbitrary nonsingular linear map such that $gh : V \to V$ 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 iff 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, $\mathrm{Env}(gL) \ker gh \leq ghV$. Conversely if $\mathrm{Env}(gL) \ker gh \leq ghV$ then, with $W_0 := \mathrm{Env}(gL) \ker gh$ and $W_1 := \ker gh$, we have $gLW_0, gLW_1 \leq W_0 \leq ghV$, and $W_0 \cap W_1 = 0$ (if $v \in W_0 \cap W_1$ then $v = ghu$ for some $u \in V$ and $ghv = 0$, implying $0 = ghghu = ghu = v$). Therefore with $W = W_0 + W_1 \leq V$ we have $gLW \leq W_0$ and $\dim V - \mathrm{rk}\, gh = \dim W_1 = \dim W - \dim W_0 \leq \dim W - \dim gLW$. Now g being invertible also implies that $\dim V - \mathrm{rk}\, h \leq \dim W - \dim LW$, which together with Fact 2.3 implies that h has maximal rank. Thus if $\mathrm{Env}(gL) \ker gh \leq ghV$ then we can efficiently construct W with the requested property, it is a witness of the maximality of the rank of gh (resp. h) in gL (resp. L). Thus, h and hence gh is not of maximum rank if and only if $\mathrm{Env}(gL) \ker gh$ is not contained in ghV. This can be decided in an obvious way.

Furthermore, if L is spanned by h and (known) rank one matrices $h_1, \ldots, h_\ell$ then the proof of Lemma 2.1 gives a linear combination of gh and $gh_1, \ldots, gh_\ell$ of higher rank. Multiplying by g^{-1} we obtain an element of L of rank larger than $\mathrm{rk}\, h$. □

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

We remark that the shortest product $\Pi = gh_1 \cdots gh_\ell$ with $\Pi \ker h \subsetneq hV$ in the proof of Lemma 2.1 can be interpreted as a generalization of the notion of *augmenting paths* in the classical bipartite matching algorithms.

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 e_1 and e_2 be the projections corresponding to the decomposition of W: $e_1(w_1, w_2) = (w_1, 0)$ and $e_2(w_1, w_2) = (0, w_2)$. Let $\mathcal{S}$ be the set $\{e_1, e_2, r_1, \ldots, r_\ell\}$ and the map $\nu : \mathcal{S} \to \mathrm{Lin}(W)$ defining the $\mathcal{S}$-module structure be just the identity map. This $\mathcal{S}$-module W is a bipartite module.

3.2 Universality of cyclic submodule optimization

Assume that we are given a linear space L of $\mathbb{F}$-linear maps from U to V. It would be straightforward to consider the bipartite module W for $W_1 = U$, $W_2 = V$ and $R = L$. Instead of this we consider another view: put $W_1 = L$, $W_2 = V$, $R = \{\mu_u | u \in U\}$, where $\mu_u(h) = h \cdot u$. Then, if U is spanned by $u_1, \ldots, u_\ell$ then $\mathcal{S}$ is $\{e_1, e_2, \mu_{u_1}, \ldots, \mu_{u_\ell}\}$. Let $(h, 0) \in W = L \oplus V$. Then the $\mathcal{S}$-submodule of W generated by $(h, 0)$ is $\mathbb{F}(h, 0) + (0, hU)$ and its dimension is $1 + \mathrm{rk}\, h$ if h is not the zero map. Therefore this 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 form 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 that the cyclic submodules of a bipartite module W arise as homomorphic images of an $\mathcal{S}$-module W_0 in W, where $\mathcal{S} = \{e_1, e_2, r_1, \ldots, r_\ell\}$ and W_0 has basis $b_0, b_1, \ldots, b_\ell$ satisfying, $e_1 b_0 = b_0$, $e_2 b_0 = 0$, $e_1 b_i = 0$, $e_2 b_i = b_i$, $r_i b_0 = b_i$, $r_i b_j = 0$ $(i, j = 1, \ldots, \ell)$.

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] over 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, 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 for showing hardness of cyclic submodule optimization: let L be a space of linear maps from U to V, put $W_1 = L$ and $W_2 = U$. Then the image of W_0 at the map ψ under which the image of b_0 is $(h, 0)$ is the subspace spanned by $(h, 0), (0, hu_1), \ldots, (0, hu_\ell)$. 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 special case) to deciding and finding an injective homomorphism from W_0 to W.

Module Surjection: Existence of and 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 $\mathbb{F}$, 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 $\psi(b_0) = (h, 0)$. 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 ϕ is injective, where $\psi(b_0) = (h, 0)$. This completes the proof of Theorem 1.2.

4 Minimizing Number of Generators in Modules

In this section we give an algorithm to minimize 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 that in Section 2). The *radical* of a module V is the intersection of its maximal (more accurately, maximal proper) submodules. A finite dimensional module is *semisimple* if its radical is the zero submodule. Submodules, direct sums and factor modules of semisimple modules as well as the factor of a module by its radical are semisimple. (Recall that submodules, direct sums of spaces which are $\mathcal{S}$-modules inherit the $\mathcal{S}$-module structure in a natural way and so do factor spaces by submodules.)

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 $\mathcal{A}v$.

4.1 A Greedy Optimization of the Submodule Dimension in Semisimple Modules

The annihilator $\mathrm{Ann}_{\mathcal{A}}(U)$ of $U \subseteq V$ is $\{a \in \mathcal{A} | au = 0 \text{ for every } u \in U\}$. Note that the annihilator $\mathrm{Ann}_{\mathcal{A}}(v)$ of the single element $v \in V$ is just the kernel of the 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 4.1. *Assume that V is semisimple. Then, for an arbitrary $u \in V$,* $\dim \mathcal{A}u = \max\{\dim \mathcal{A}u' | u' \in V\}$ *iff* $\mathrm{Ann}_{\mathcal{A}}(u)V \subseteq \mathcal{A}u$. *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* $\mathbb{F}$.

The proof can be found in [CIK97]. For completeness, we also include a proof in the Appendix. The second part of the above lemma is especially interesting for small base fields where Lemma 2.2 does not apply. The next lemma can be used to give a generalization for submodules generated by larger systems.

Lemma 4.2. *Assume that V is semisimple. Then, for arbitrary positive integer ℓ and for elements $u_1, \ldots, u_\ell \in V$,* $\dim \mathcal{A}\{u_1, \ldots, u_\ell\} = \max\{\mathcal{A}U | U \subseteq V, \#U \leq \ell\}$ *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 $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_\ell$.*

We relegate the proof of the lemma to the Appendix. The two lemmas above together with Lemma 2.2 immediately give the following.

Proposition 4.3. *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 $\mathbb{F}$ is of size larger than 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 4.4. *Let $v_1, \ldots, v_n$ be a basis of the $\mathcal{S}$-module V 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 $V_0 = U_0 \oplus V_0$. Then the factor module $V/V_0 \cong U/U_0 \oplus W/W_0$ is semisimple and we can apply the preceding Proposition 4.3 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 λ's, for which either dim $U' \leq$ dim U or $V \neq U' + W$, is at most dim $V/V_0 +$ dim $V \leq 2n$. □

4.2 Algorithm for Finding ℓ Generators

Using the previous Lemma, now we describe an iterative algorithm to find 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 $\{u_1, \ldots, u_\ell, u_{\ell+1}, \ldots\}$, 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 $W' = W$, we are done and output U.

   `Inner loop:`

3) apply Lemma 4.4 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"; exit
   Else set $U := U'$.
5) If $W \not\leq U + W'$ then
   continue *inner* loop with $W' = (U + W') \cap W$.
6) 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 6 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 easily 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 4.3 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' containing U and hence Γ. But $U' \geq V_0$ by the definition of V_0, therefore $U' \supseteq \Gamma \cup V_0$, a contradiction with that this union generates the whole V. 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 4.1). 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 $h + \alpha_1 h_{i_1} + \ldots + \alpha_c h_{i_c}$ has higher rank for some $\alpha_1, \ldots, \alpha_c \in \mathbb{F}$ ($\mathbb{F}$ is large enough.)

References

[BFS99] J. F. Buss, G. S. Frandsen, and J. Shallit. The computational complexity of some problems of linear algebra. *J. Comput. Syst. Sci.*, 58(3):572–596, 1999.

[BL08] P.A. Brooksbank and E.M. Luks. Testing isomorphism of modules. *Journal of Algebra*, 320(11):4020–4029, 2008.

[CIK97] A. Chistov, G. Ivanyos, and M. Karpinski. Polynomial time algorithms for modules over finite dimensional algebras. In *ISSAC '97: Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation*, pages 68–74, 1997.

[CIW97] A. M. Cohen, G. Ivanyos, and D. B. Wales. Finding the radical of an algebra of linear transformations. *Journal of Pure and Applied Algebra*, 117-118:177–193, 1997. (Proc. MEGA'96).

[Edm67] J. Edmonds. Systems of distinct representatives and linear algebra. *Journal of Research of the National Bureau of Standards*, 71B:241–245, 1967.

[FR85] K. Friedl and L. Rónyai. Polynomial time solutions of some problems of computational algebra. In *STOC '85: Proceedings of the seventeenth annual ACM Symposium on Theory of Computing*, pages 153–162, 1985.

[Gee99] J. F. Geelen. Maximum rank matrix completion. *Linear Algebra and its Applications*, 288:211–217, 1999.

[GI05] J. F. Geelen and S. Iwata. Matroid matching via mixed skew-symmetric matrices. *Combinatorica*, 25(2):187–215, 2005.

[GIM03] J. F. Geelen, S. Iwata, and K. Murota. The linear delta-matroid parity problem. *Journal of Combinatorial Theory, Series B*, 88(2):377–398, 2003.

[Gur03] L. Gurvits. Classical deterministic complexity of edmonds' problem and quantum entanglement. In *STOC '03: Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 10–19, 2003.

[Gur04] L. Gurvits. Classical complexity and quantum entanglement. *Journal of Computer and System Sciences*, 69(3):448–484, 2004.

[HKM05] N. J. A. Harvey, D. R. Karger, and K. Murota. Deterministic network coding by matrix completion. In *SODA '05: Proceedings of the sixteenth annual ACM-SIAM Symposium on Discrete Algorithms*, pages 489–498, 2005.

[HKY06] N. J. A. Harvey, D. R. Karger, and S. Yekhanin. The complexity of matrix completion. In *SODA '06: Proceedings of the seventeenth annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1103–1111, 2006.

[IM83] O. Ibarra and S. Moran. Probabilistic algorithms for deciding equivalence of straight-line programs. *JACM*, 30(1):217–228, 1983.

[KI03] V. Kabanets and R. Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. In *STOC '03: Proceedings of the thirty-fifth annual ACM Symposium on Theory of Computing*, pages 355–364, 2003.

[Lov79] L. Lovász. On determinants, matchings and random algorithms. In *FCT '79: Fundamentals of Computation Theory*, pages 565–574, 1979.

[Lov89] L. Lovász. Singular spaces of matrices and their applications in combinatorics. *Bol. Soc. Braz. Mat*, 20:87–99, 1989.

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

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

[Sch80] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *JACM*, 27(4):701–717, 1980.

[Val79] L. G. Valiant. Completeness classes in algebra. In *STOC '79: Proceedings of the eleventh annual ACM Symposium on Theory of Computing*, pages 249–261, 1979.

[Zip79] R. Zippel. Probabilistic algorithms for sparse polynomials. *Symbolic and Algebraic Computation*, pages 216–226, 1979.

A Appendix

Algebras and modules

In this part we summarize some basic notions and facts from the representation theory of algebras needed in the proof of Lemmas 4.1 and 4.2. For details, we refer the reader to the first few chapters of the textbook [Pie82].

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 Appendix 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 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 Appendix 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. (Recall that in Section 4 we used this property as the definition for semisimplicity.) 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 and, conversely, the isomorphism class of a direct sum of simple modules depend only on the isomorphism classes and multiplicities of the constituents. 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 $\mathrm{Lin}(V)$ with the same identity generated by $\nu(\mathcal{S})$). 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. In this appendix 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]. (Recall that we consider only finite dimensional and unital modules.)

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:

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

where V_i are pairwise non-isomorphic $\mathcal{A}$-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

$$V \cong \bigoplus_{i=1}^{t} V_i^{s_i}, \tag{2}$$

where the multiplicities s_i are non-negative integers.

Lemma A.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 that the submodule dimension in V is maximized iff 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 cyclic 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 $W/W_0 \cong V_0$. □

Proof of Lemma 4.1

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 (1) resp. (2). Let $u \in V$. Assume that the dimension of the submodule $\mathcal{A}u$ is not maximal. Then, by Lemma A.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 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 $\{a \in \mathcal{A} | ax = 0 \text{ for every } x \in X\}$. Assume that $\mathrm{Ann}_{\mathcal{A}}(u)V \subseteq \mathcal{A}u$. Then every element of $\mathrm{Ann}_{\mathcal{A}}(u)$ act 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: 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 of $\mathcal{A}$ (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 it 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 4.1.

Proof of Lemma 4.2

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 (1) resp. (2). Let $u_1, \ldots, u_\ell \in V$ and let $W_i = \mathcal{A}(\{u_1, \ldots, u_\ell\} \setminus \{i\})$. 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 A.1, for every i, the multiplicity of V_j in W/W_i equals either the multiplicity of V_j in V/W_i or it 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 $U_{ij} \cap W_i = 0$ 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 4.2.

□