RC24610 (W0807-126) July 22, 2008
Mathematics

IBM Research Report

Nonlinear Optimization for Matroid Intersection and Extensions

Yael Berstein[1], Jon Lee[2], Shmuel Onn[1], Robert Weismantel[3]

[1]Technion - Israel Institute of Technology
32000 Haifa
Israel

[2]IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598
USA

[3]Otto-von-Guericke Universität Magdeburg
D-39106 Magdeburg
Germany

Research Division
Almaden - Austin - Beijing - Cambridge - Haifa - India - T. J. Watson - Tokyo - Zurich

Y. Berstein · J. Lee · S. Onn · R. Weismantel

Nonlinear optimization for matroid intersection and extensions

For Tom Liebling, on the occasion of his retirement

Abstract We address optimization of nonlinear functions of the form $f(Wx)$, where $f : \mathbb{R}^d \to \mathbb{R}$ is a nonlinear function, W is a $d \times n$ matrix, and feasible x are in some large finite set $\mathcal{F}$ of integer points in $\mathbb{R}^n$. Generally, such problems are intractable, so we obtain positive algorithmic results by looking at broad natural classes of f , W and $\mathcal{F}$.

One of our main motivations is multi-objective discrete optimization, where f trades off the linear functions given by the rows of W . Another motivation is that we want to extend as much as possible the known results about polynomial-time linear optimization over trees, assignments, matroids, polymatroids, etc. to nonlinear optimization over such structures.

We assume that the convex hull of $\mathcal{F}$ is well-described by linear inequalities (i.e., we have an efficient separation oracle). For example, the set of characteristic vectors of common bases of a pair of matroids on a common ground set satisfies this property for $\mathcal{F}$. In this setting, the problem is already known to be intractable (even for a single matroid), for general f (given by a comparison oracle), for (i) $d = 1$ and binary-encoded W , and for (ii) $d = n$ and $W = I$.

Our main results (a few technicalities suppressed):

1- When $\mathcal{F}$ is well described, f is convex (or even quasiconvex), and W has a fixed number of rows and is unary encoded or with entries in a fixed set, we give an efficient deterministic algorithm for maximization.

2- When $\mathcal{F}$ is well described, f is a norm, and binary-encoded W is nonnegative, we give an efficient deterministic constant-approximation algorithm for maximization.

3- When $\mathcal{F}$ is well described, f is "ray concave" and non-decreasing, and W has a fixed number of rows and is unary encoded or with entries in a fixed set, we give an efficient deterministic constant-approximation algorithm for minimization.

4- When $\mathcal{F}$ is the set of characteristic vectors of common bases of a pair of vectorial matroids on a common ground set, f is arbitrary, and W has a fixed number of rows and is unary encoded, we give an efficient randomized algorithm for optimization.

Yael Berstein
Technion - Israel Institute of Technology, 32000 Haifa, Israel
E-mail: yaelber@tx.technion.ac.il

Jon Lee
IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA
E-mail: jonlee@us.ibm.com

Shmuel Onn
Technion - Israel Institute of Technology, 32000 Haifa, Israel
E-mail: onn@ie.technion.ac.il

Robert Weismantel
Otto-von-Guericke Universität Magdeburg, D-39106 Magdeburg, Germany
E-mail: weismantel@imo.math.uni-magdeburg.de

Introduction

Generally, we are considering nonlinear discrete optimization problems of the form

$$P(\mathcal{F}, \text{opt}, f, W): \qquad \text{opt } \{f(Wx) \ : \ x \in \mathcal{F}\} \ ,$$

where $\text{opt} \in \{\min, \max\}$, $W \in \mathbb{Z}^{d \times n}$ is is a matrix of integers, function $f : \mathbb{R}^d \to \mathbb{R}$ is given by a comparison oracle, and $\mathcal{F}$ is a large finite subset of $\mathbb{Z}^n_+$.

One motivation for this setting is multiobjective optimization, where we seek an optimal "balancing", via a nonlinear function f , of d competing linear objectives $W_i.x$. Another motivation is the desire to extend known results about polynomial-time linear optimization over combinatorial structures to nonlinear optimization over such structures.

We assume that $\mathcal{F}$ is "well described" — that is, $\mathcal{F} \subset \mathcal{S}(n, \beta) := \{x \in \mathbb{R}^n_+ \ : \ \mathbf{1}^\top x \leq \beta\}$, for some unary encoded $\beta \in \mathbb{Z}_+$, and that we have a separation oracle for $\mathcal{P} := \text{conv}(\mathcal{F})$. For example, the set of characteristic vectors of common bases of a pair of matroids on the common ground set $\{1, 2, \ldots, n\}$ satisfies this property for $\mathcal{F}$ (with $\beta = n$) .

In this setting, the problem is already known to be intractable (even for a single matroid), for general f (given by a comparison oracle), for (i) $d = 1$ and binary-encoded W , and for (ii) $d = n$ and $W = I$. Therefore, in much of our work, we assume that the objective vectors $W_i.$ have a coarse encoding. This can take the form of (i) unary encoding, (ii) entries in a fixed set of binary-encoded numbers, or (iii) "generalized unary encoding" (a common generalization of (i) and (ii)). Furthermore, to obtain efficient algorithms, we often require that the number d of rows of W is fixed.

We have four main results which below we describe and relate to the literature. Precise statements of the theorems are given in the later sections.

(**Theorem 1**) *When $\mathcal{F}$ is well described, f is convex (or even quasiconvex), and W has a fixed number of rows and is unary encoded or with entries in a fixed set, we give an efficient deterministic algorithm for maximization.*

- As a very special case ($\mathcal{F} = \{n' \times n' \text{ permutation matrices}\}$, $W \in \mathbb{Z}^{d \times n' \times n'}$ unary encoded, f convex), we obtain a result of [3]. The special case in which $\mathcal{F}$ is the set of characteristic vectors of common independent sets or common bases of a pair of matroids on $N := \{1, 2, \ldots, n\}$ appears to be new (note that in such a case we may take $\beta = n$), even for f convex. A fortiori, also new is the special case in which $\mathcal{F}$ is the set of integer points that are in a pair of integral polymatroids or associated base polytopes. Furthermore, the special case in which $\mathcal{F}$ is the set of characteristic vectors of matchings of a *nonbipartite* graph appears to be new, even for f convex.

(**Theorem 2**) *When $\mathcal{F}$ is well described, f is a norm, and binary-encoded W is nonnegative, we give an efficient deterministic constant-approximation algorithm for maximization.*

- As a very special case ($\mathcal{F} = \{n' \times n' \text{ permutation matrices}\}$, $W \in \mathbb{Z}^{d \times n' \times n'}$, f a p-norm), we obtain a result of [3].

(**Theorem 3**) *When $\mathcal{F}$ is well described, f is "ray concave" and non-decreasing, and W has a fixed number of rows and is unary encoded or with entries in a fixed set, we give an efficient deterministic constant-approximation algorithm for minimization.*

- This theorem generalizes a result of [3] which looked at the very special case of: $f \in \{\text{p}{-}norms\}$, $\mathcal{F} := \{n' \times n' \text{ permutation matrices}\}$, $W \in \mathbb{Z}^{d \times n' \times n'}$ unary encoded. In addition, our more detailed analysis of p-norms further generalizes the result of [3] pertaining to the 2-norm.

(**Theorem 4**) *When $\mathcal{F}$ is the set of characteristic vectors of common bases of a pair of vectorial matroids on a common ground set, f is arbitrary, and W has a fixed number of rows and is unary encoded, we give an efficient randomized algorithm for optimization.*

- This theorem can be contrasted with a result in [2] which established a *deterministic* algorithm for the case of a *single* vectorial matroid.

Note that all of our results extend to the generalization where we have a binary-encoded $c \in \mathbb{Z}^n$ giving a "primary" linear objective function $c^\top x$ that we maximize, and, subject to that maximization, we seek an optimal balancing, via a nonlinear function f , of d competing linear objectives $W_i.x$:

$$P_c(\mathcal{F}, \text{opt}, f, W): \qquad \text{opt } \left\{ \ f(Wx) \ : \ x \in \text{argmax}\left\{c^\top \tilde{x} \ : \ \tilde{x} \in \mathcal{F}\right\}\right\} \ .$$

The reason is that we can find the optimal value z^* of $c^\top x$ over $\mathcal{F}$ using the ellipsoid method and the separation oracle for $\mathcal{P} = \text{conv}(\mathcal{F})$, and then the equation $c^\top x = z^*$ together with the separation oracle for $\mathcal{P}$ yields a separation oracle for the face $\mathcal{P}_c := \mathcal{P} \cap \left\{x \in \mathbb{R}^n : c^\top x = z^*\right\}$ of $\mathcal{P}$.

Relevant background material (including polytopes, matroids, polymatroids and linear programming) can be found in [4,9,11]. Additional relevant material on nonlinear discrete optimization can be found in [6,7,8].

In §1, we develop an algorithm for quasiconvex combinatorial maximization. Some of the elements that we develop for this are used in later sections as well. In §1.1, we introduce our "generalized unary encoding" scheme which we use for weight matrices. In §1.2, we investigate relevant properties of the fibers (i.e., inverse images) of points in the linear image of a polytope. In §1.3, we apply properties of fibers to give an efficient algorithm for quasiconvex combinatorial maximization. In §2, we consider approximation algorithms as a means of relaxing the assumptions that gave us an efficient algorithm in the previous section. We develop one algorithm for norms and another for a very broad generalization of norms. In §2.1, relaxing a requirement on the input weight matrix W, we give an efficient approximation algorithm for combinatorial maximization of norms f . In §2.2, we further exploit the geometry of fibers to give an efficient approximation algorithm for nonlinear combinatorial minimization, when W is nonnegative and f is non-decreasing and "ray-concave" on the nonnegative orthant. In §3, for unary-encoded weight matrices W , we give an efficient algorithm for optimizing an arbitrary nonlinear function f , over the set $\mathcal{F}$ of characteristic vectors of common bases of a pair of vectorial matroids on a common ground set.

1 Quasiconvex combinatorial maximization

1.1 Generalized unary encoding of weights

General binary encoding of weights is typically too parsimonious to allow for the construction of theoretically efficient algorithms for nonlinear discrete optimization. So we turn our attention to less parsimonious encodings. We consider weights $W_{i,j}$ of the form

$$W_{i,j} = \sum_{k=1}^{p} \delta_{i,j}^k a_k \ ,$$

with integer $p \geq 1$ fixed, the distinct positive integers a_k being binary encoded, but the integers $\delta_{i,j}^k$ (unrestricted in sign) being unary encoded. It is convenient to think of unary-encoded matrices $\delta^k = ((\delta_{i,j}^k)) \in \mathbb{Z}^{d\times n}$, for $1 \leq k \leq p$, and binary encoded $a \in \mathbb{Z}^p$, with $a_k > 0$, for $1 \leq k \leq p$. Then we have $W := ((W_{i,j})) = \sum_{k=1}^{p} a_k \delta^k$.

We have the following special cases:

1. **Unary-encoded weights:** With $p = 1$ and $a_1 = 1$, we get the ordinary model of unary-encoded $W = \delta^1$.
2. **Binary-encoded $\{0, a_1, a_2, ..., a_p\}$-valued weights:** With $\delta^k \geq 0$ for all k , and $\sum_{k=1}^{p} \delta_{i,j}^k \leq 1$, for all i, j , we get the case of all $W_{i,j}$ in the set $\{0, a_1, a_2, ..., a_p\}$ having binary-encoded elements .
3. **0/1-valued weights:** This is, of course, the important common special case of 1 and 2.

Because 1 is a much more common generalization of 3 than is 2, we refer to our general setting as *generalized unary encoding*. We do note that there are cases where, with respect to a particular nonlinear combinatorial optimization problem, one algorithm is efficient for case 2 but not for case 1, and for another algorithm vice versa (see [2,5] for example). Furthermore, there are broad cases where with binary encoding a problem is provably intractable, but we will see that we can construct an efficient algorithm when the weights have a generalized unary encoding.

1.2 Fibers

We begin with some definitions. For $W \in \mathbb{Z}^{d\times n}$ and polytope $\mathcal{P} \subset \mathbb{R}^n$, we let $W\mathcal{P}$ denote the image of $\mathcal{P}$ under the linear map $x \mapsto Wx$. Clearly $W\mathcal{P} \subset \mathbb{R}^d$ is also a polytope. For $u \in \mathbb{R}^d$, we define

$$W^{-1}u := \{x \in \mathbb{R}^n : Wx = u\} \ .$$

For $u \in \mathbb{R}^d$, we refer to $(W^{-1}u) \cap \mathcal{P}$ as the $W_{\mathcal{P}}$-fiber of u . If W and $\mathcal{P}$ are clear from context, then we may just say "the fiber of u ." For $u \in \mathbb{R}^d$, we note that $u \in W\mathcal{P}$ if and only if the $W_{\mathcal{P}}$-fiber of u is nonempty.

First, we point out how we can optimize efficiently a linear function on the fiber of any $u \in \mathbb{R}^d$. In particular, we can also determine nonemptiness of such a fiber.

Lemma 1 *Assume that we have a separation oracle for the polytope $\mathcal{P} \subset \mathbb{R}^n$. Let $c \in \mathbb{Z}^n$, $W \in \mathbb{Z}^{d \times n}$ and $u \in \mathbb{R}^d$ be binary encoded. Then we can solve the linear-objective optimization problem*

$$\max\left\{c^\top x \ : x \in (W^{-1}u) \cap \mathcal{P}\right\}$$

in polynomial time. In particular, we can test nonemptiness of the $W_{\mathcal{P}}$-fiber of u in polynomial time.

Proof The problem is just the linear program

$$\max\left\{c^\top x \ : Wx = u \ , \ x \in \mathcal{P}\right\} \ ,$$

which we can solve in polynomial time via the ellipsoid method, as we assume that we have available a separation oracle for $\mathcal{P}$. □

Unfortunately, even though $\mathcal{P}$ is assumed to have integer vertices (it is the convex hull of the finite set $\mathcal{F}$ of integer points), the $W_{\mathcal{P}}$-fiber of u may have some integer vertices and some fractional ones (after all, we are just intersecting $\mathcal{P}$ with a linear subspace). Therefore, an optimal extreme point solution of the linear program of Lemma 1 may be at a fractional vertex of the $W_{\mathcal{P}}$-fiber of u . In short, fibers of arbitrary $u \in W\mathcal{P}$ are not well behaved. However, we will see that fibers of $u \in \mathrm{vert}(W\mathcal{P})$ are better behaved. Specifically, we will now demonstrate that if $u \in \mathrm{vert}(W\mathcal{P})$, then we can find efficiently an integer optimizer of any linear function on the $W_{\mathcal{P}}$-fiber of u .

Lemma 2 *Assume that we have a separation oracle for the polytope $\mathcal{P} \subset \mathbb{R}^n$. Let $c \in \mathbb{Z}^n$, $W \in \mathbb{Z}^{d \times n}$ and $u \in \mathrm{vert}(W\mathcal{P})$ be binary encoded. Then we can solve the optimization problem*

$$\max\left\{c^\top x \ : x \in (W^{-1}u) \cap \mathcal{P} \cap \mathbb{Z}^n\right\}$$

in polynomial time.

Proof It is well known that (i) the image $\mathcal{Q}$ of a polytope $\mathcal{P}$ under a linear (even affine) map is a polytope, and (ii) the preimage (in $\mathcal{P}$) of every face of $\mathcal{Q}$ is a face of $\mathcal{P}$. In particular, the preimage (in $\mathcal{P}$) of every vertex of $\mathcal{Q}$ is a face of $\mathcal{P}$.

Therefore, if $u \in \mathrm{vert}(W\mathcal{P})$, then the $W_{\mathcal{P}}$-fiber of u is a face of $\mathcal{P}$, and so such a fiber has itself integer vertices (because $\mathcal{P}$ does). Therefore, an extreme-point solution of the linear program

$$\max\left\{c^\top x \ : Wx = u \ , \ x \in \mathcal{P}\right\}$$

will be integer. □

The following result is not needed for our purposes, but it is interesting in its own right.

Proposition 1 *Assume that we have $\mathcal{P} = \{x \in \mathbb{R}^n : Ax \leq b\}$, for some binary encoded $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$. Let $W \in \mathbb{Z}^{d \times n}$ and $u \in \mathrm{vert}(W\mathcal{P})$ be binary encoded, Then we can compute an inequality description of the $W_{\mathcal{P}}$-fiber of u as a face of $\mathcal{P}$ in polynomial time.*

Proof For each row $1 \leq i \leq m$, we simply solve the linear program

$$z_i := \min_{x \in \mathbb{R}^n} \left\{A_i.x \ : Wx = u \ , \ Ax \leq b\right\} \ .$$

Then

$$(W^{-1}u) \cap \mathcal{P} = \left\{x \in \mathbb{R}^n \ : \ A_i.x \left\{\begin{matrix}\leq\\=\end{matrix}\right\} b_i \text{ if } z_i \left\{\begin{matrix}<\\=\end{matrix}\right\} b_i \ , \text{ for } 1 \leq i \leq m\right\} \ .$$

□

Next, we demonstrate that under an appropriate encoding of the data, the set of vertices of $W\mathcal{P}$ can be computed efficiently.

Lemma 3 *Let $W \in \mathbb{Z}^{d\times n}$ be given. We assume that $\mathcal{F}$ is a finite subset of Z_+^n , and further that $\mathcal{F} \subset \mathcal{S}(n,\beta) := \{x \in \mathbb{R}_+^n \ : \ \mathbf{1}^\top x \leq \beta\}$, for some $\beta \in \mathbb{Z}_+$. Then, for fixed d , we can compute the vertices of $W\mathcal{P}$ in time polynomial in n , β , and the length of the generalized unary encoding of W .*

Proof For $x \in \mathcal{F}$, consider $u := Wx$. For $1 \leq i \leq d$, we have

$$u_i \ := \ W_{i\cdot}x \ = \ \sum_{j=1}^n W_{i,j}x_j \ = \ \sum_{j=1}^n \left(\sum_{k=1}^p \delta_{i,j}^k a_k\right) x_j \ = \ \sum_{k=1}^p \left(\sum_{j=1}^n \delta_{i,j}^k x_j\right) a_k \ .$$

Observe that the nonnegative integer coefficients $\sum_{j=1}^n \delta_{i,j}^k x_j$ of the a_k are bounded by a polynomial in the unary β and the unary $\delta_{i,j}^k$: Letting $\omega := \max\left|\delta_{i,j}^k\right|$, we observe that

$$\left|\sum_{j=1}^n \delta_{i,j}^k x_j\right| \leq \left\|\delta_{i,j}^k\right\|_\infty \|x\|_1 \leq \omega\beta \ .$$

This means that there are only a polynomial number of possible values for each u_i (p is fixed), and then (with d fixed) only a polynomial number of images $u = Wx$ of $x \in \mathcal{F}$.

Concretely, for $k = 1,\ldots,p$, we let $\hat{u}^k$ range over

$$\{0, \pm 1, \ldots, \pm\omega\beta\}^d \ \subset \mathbb{Z}^d \ .$$

For each of the $(2\omega\beta+1)^{pd}$ choices of $(\hat{u}^1,\ldots,\hat{u}^p)$, we let $u := \sum_{k=1}^p a_k\hat{u}^k$. For each such u we check if $u \in W\mathcal{P}$ using the algorithm of Lemma 1, and so we obtain a set $U \subset \mathbb{Z}^d$, comprising these points u having a nonempty fiber, having a polynomial number of elements.

Now, because U is contained in $W\mathcal{P}$ and contains $\mathrm{vert}(W\mathcal{P})$, it is clear that $W\mathcal{P} = \mathrm{conv}(U)$.

Finally, because the dimension d is fixed, we can compute the vertices of $\mathrm{conv}(U)$ in polynomial time. □

1.3 Exact algorithm for quasiconvex maximization

In this subsection, we take the facts that we have already built up to give an efficient algorithm to maximize a function that is quasiconvex (that is, its "lower level sets" $\{z \in \mathbb{R}^d \ : \ f(z) \leq \tilde{f}\}$ are convex subsets of $\mathbb{R}^d$, for all $\tilde{f} \in \mathbb{R}$; see [1], for example). Of course, the algorithm also applies to ordinary convex functions.

Theorem 1 *Suppose that we are given $W \in \mathbb{Z}^{d\times n}$, quasiconvex function $f : \mathbb{R}^d \to \mathbb{R}$ specified by a comparison oracle, and a finite subset $\mathcal{F}$ of $\mathbb{Z}_+^n$. We further assume that $\mathcal{F} \subset \mathcal{S}(n,\beta) := \{x \in \mathbb{R}_+^n \ : \ \mathbf{1}^\top x \leq \beta\}$, for some $\beta \in \mathbb{Z}_+$, and that we have a separation oracle for $\mathcal{P} := conv(\mathcal{F})$. Then, if d is fixed, we can solve $P(\mathcal{F},\max,f,W)$ in time polynomial in n , β , and the size of the generalized unary encoding of W .*

Proof Because $f(u)$ is quasiconvex on the polytope $W\mathcal{P}$ and $f_W(x) := f(Wx)$ is quasiconvex on the polytope $\mathcal{P}$, and the maximum of a quasiconvex function on a polytope is attained at a vertex of the polytope, we have

$$\begin{array}{rcl} \max\{f(Wx) \ : \ x \in \mathrm{vert}(\mathcal{P})\} & = & \max\{f(Wx) \ : \ x \in \mathcal{P}\} \\ = \max\{f(u) \ : \ u \in W\mathcal{P}\} & = & \max\{f(u) \ : \ u \in \mathrm{vert}(W\mathcal{P})\} \ . \end{array}$$

First, via Lemma 3, we compute efficiently $\mathrm{vert}(W\mathcal{P})$. Next, using the comparison oracle for f , we can compare values of $f(u)$, for all $u \in \mathrm{vert}(W\mathcal{P})$, to obtain the best u . Then, for the best $u \in \mathrm{vert}(W\mathcal{P})$, via Lemma 1, using an arbitrary linear objective vector $c \in \mathbb{Z}^n$ in general position, we efficiently find an extreme point x^u of the $W_{\mathcal{P}}$-fiber of u . By Lemma 2, such a point x^u will be integer and will thus solve $P(\mathcal{F},\max,f,W)$. □

2 Approximation algorithms for norms and related functions

In this section, we consider approximation algorithms as a means of relaxing the assumptions that gave us an efficient algorithm in the previous section. In §2.1, we provide an efficient approximation algorithm for combinatorial norm maximization. In §2.2, we provide an efficient approximation algorithm for combinatorial minimization. This latter algorithm applies to a very broad generalization of norms.

2.1 Approximation algorithm for norm maximization

The following theorem provides an approximation algorithm for maximizing a norm f , that runs in time that is polynomial even in the bit size of the weights $W^k_{i,j}$ and even if d is variable. The approximation ratio depends on the (positive) constants for the standard equivalence of the norm f with the infinity norm:

$$C_f \ \|u\|_\infty \ \leq \ f(u) \ \leq \ C^f \ \|u\|_\infty \ .$$

Note that appropriate constants C_f and C^f can be concretely described, and in some cases efficiently calculated. Letting $\{e_1, \ldots, e_d\}$ denote the standard basis of $\mathbb{R}^d$, a valid choice of C^f is the sum of f evaluated on each of the basis elements (i.e., $C_f := \sum_{i=1}^d f(e_i)$). Furthermore, the best C_f is the minimum of the convex function f on the unit sphere in the ∞-norm (i.e., $C_f := \min\{f(u) \ : \ \|u\|_\infty = 1\}$. This is not a convex optimization problem because we are restricted to the *boundary* of the (unit) ball (of the ∞-norm). If however we take d to be constant or even say $O(\log \ n)$, then we can solve 2^d convex minimization problems (via say an ellipsoid method), one for reach facet of the unit ball of the ∞-norm, in order to efficiently calculate a valid (in fact, best) value of C_f . Finally, for the special case of $f(u) := \|u\|_p$, our algorithm finds a $d^{1/p}$-approximate solution.

Theorem 2 *Suppose that we are given nonnegative $W \in \mathbb{Z}^{d\times n}$, and $\mathcal{F}$ a finite subset of $\mathbb{Z}^n_+$. We further assume that $\mathcal{F} \subset \mathcal{S}(n,\beta) := \{x \in \mathbb{R}^n_+ \ : \ \mathbf{1}^\top x \leq \beta\}$, for some $\beta \in \mathbb{Z}_+$, and that we have a separation oracle for $\mathcal{P} := conv(\mathcal{F})$. Then, for any norm f given by a comparison oracle, there is an algorithm that determines a $\left(C^f/C_f\right)$-approximate solution to $P(\mathcal{F}, \max, f, W)$, in time that is polynomial in d ,n , β , and $\max \left\lceil \log w^k_{i,j} \right\rceil$. Moreover, for $f(u) := \|u\|_p$, for any p satisfying $1 \leq p \leq \infty$, our algorithm determines a $d^{1/p}$-approximate solution.*

Proof First, let f be an arbitrary norm. The algorithm is the following: For $i = 1, \ldots, d$, solve the linear-programming problem

$$\max\{W_i.x \ : \ x \in \mathcal{P}\} \ ,$$

obtaining an optimal vertex x^i of $\mathcal{P}$, and let u^i be its W-image. Then output x^r such that $\|u^r\|_p = \max^d_{i=1} \|u^i\|_p$.

We now show that this provides the claimed approximation. Let s satisfy $\|u^s\|_\infty = \max^d_{i=1} \|u^i\|_\infty$. First, we claim that any $\tilde{u} \in \text{vert}(W\mathcal{P})$ satisfies $\|\tilde{u}\|_\infty \leq \|u^s\|_\infty$. To see this, recall that W and hence the u^i are nonnegative; choose any point $\tilde{x} \in (W^{-1}\tilde{u}) \cap \mathcal{F}$ (hence $\tilde{u} = W\tilde{x}$), and let t satisfy $\tilde{u}_t = \|\tilde{u}\|_\infty = \max^d_{i=1} \tilde{u}_i$. Then, as claimed, we get

$$\|\tilde{u}\|_\infty \ = \ \tilde{u}_t \ = \ W_t.\tilde{x} \ \leq \ \max\{W_t.x : x \in \mathcal{F}\} \ = \ W_t.x^t \ = \ u^t_t \ \leq \ \|u^t\|_\infty \ \leq \ \|u^s\|_\infty \ .$$

Let x^* be an optimal solution, and let u^* be its W-image. As norms are (quasi)convex, we can without loss of generality take x^* to be a vertex of $\mathcal{P}$ and u^* to be a vertex of $W\mathcal{P}$.

Then we have the following inequalities:

$$\begin{aligned} f(Wx^*) = f(u^*) \ &\leq \ C^f \ \|u^*\|_\infty \ \leq \ C^f \ \|u^s\|_\infty \\ &\leq \left(C^f/C_f\right) \ f(u^s) \ \leq \ \left(C^f/C_f\right) \ f(u^r) \ = \ \left(C^f/C_f\right) f(Wx^r) \ . \end{aligned}$$

Finally, observing that for p satisfying $1 \leq p \leq \infty$, we have $C_f = 1$ and $C^f = d^{1/p}$, we conclude that for p-norms we get a $d^{1/p}$-approximate solution. □

2.2 Approximation algorithm for ray-concave minimization

In this subsection, using some of the methodology that we already developed, we obtain an efficient approximation algorithm for nonlinear combinatorial minimization. Our algorithm applies to certain functions that generalize norms via, what will at first seem paradoxical, a type of concavity.

A function $f : \mathbb{R}^d_+ \to \mathbb{R}$ is *ray-concave* if

(i) $\lambda f(u) \le f(\lambda u)$ for $u \in \mathbb{R}^d_+$, $0 \le \lambda \le 1$;

Analogously $f : \mathbb{R}^d_+ \to \mathbb{R}$ is *ray-convex* if

(i') $\lambda f(u) \ge f(\lambda u)$ for $u \in \mathbb{R}^d_+$, $0 \le \lambda \le 1$;

Naturally, we say that $f : \mathbb{R}^d_+ \to \mathbb{R}$ is *homogeneous* if it is both ray-concave and ray-convex.

A function $f : \mathbb{R}^d_+ \to \mathbb{R}$ is *non-decreasing* if

(ii) $f(u) \le f(\tilde{u})$, for $u, \tilde{u} \in \mathbb{R}^d_+$, $u \le \tilde{u}$.

We are mainly interested in ray-concave non-decreasing functions $f : \mathbb{R}^d_+ \to \mathbb{R}$. In particular, by (i) with $\lambda = 0$, we have $f(\mathbf{0}) \ge 0$, and hence by (ii) , we have $f(u) \ge 0$ for all $u \in \mathbb{R}^d_+$.

Notice that ordinary concavity of a function f has the special case:

$$\lambda f(u) + (1-\lambda) f(\mathbf{0}) \le f(\lambda u + (1-\lambda)\mathbf{0}) \text{ , for } u \in \mathbb{R}^d_+ \text{ , } 0 \le \lambda \le 1 \text{ ,}$$

and so if f is concave with $f(\mathbf{0}) = 0$, then it is ray-concave — thus justifying our terminology.

Notice further that if f is a norm on $\mathbb{R}^d$, then it of course non-decreasing on $\mathbb{R}^d_+$; additionally, it is not only ray-concave, but it is more strongly homogeneous (and more of course). In general, homogeneity already implies $f(\mathbf{0}) = 0$. A concrete example of a ray-convex function that is not generally homogeneous is $h(u) := \prod_{i=1}^d u_i$.

There are other interesting and useful functions that are ray-concave. Suppose that ray-concave $g : \mathbb{R}^d_+ \to \mathbb{R}$ and ray-convex $h : \mathbb{R}^d_+ \to \mathbb{R}$ satisfy

(iii) $g(u+v) - g(u) \ge h(u+v) - h(u)$, for all $u, v \in \mathbb{R}^d_+$.

That is, g grows no slower than h on $\mathbb{R}^d_+$. Then $f(u) := g(u) - h(u)$ is of course non-decreasing. Moreover, if g is ray-concave on $\mathbb{R}^d_+$ and h is ray-convex on $\mathbb{R}^d_+$, then f is also ray-concave on $\mathbb{R}^d_+$.

As a concrete and natural example, consider $g(u) := \|u\|_1$ and $h(u) := \|u\|_p$ for any integer $p \ge 1$ or infinity. Then $f(u) := g(u) - h(u) = \|u\|_1 - \|u\|_p$ is ray-concave and non-decreasing on $\mathbb{R}^d_+$. Notice that already for $d = 2$ and $p = \infty$, $f(u)$ is not a norm — indeed, for this case $f(u) = \min(u_1, u_2)$.

Theorem 3 *Suppose that nonnegative $W \in \mathbb{Z}^{d \times n}$ is given, d is fixed, $f : \mathbb{R}^d_+ \to \mathbb{R}$ is a ray-concave non-decreasing function given by a comparison oracle, and $\mathcal{F}$ is a finite subset of $\mathbb{Z}^n_+$. We further assume that $\mathcal{F} \subset \mathcal{S}(n, \beta) := \{x \in \mathbb{R}^n_+ \ : \ \mathbf{1}^\top x \le \beta\}$, for some $\beta \in \mathbb{Z}_+$, and that we have a separation oracle for $\mathcal{P} := conv(\mathcal{F})$. Then we can compute a d-approximate solution to $P(\mathcal{F}, \min, f, W)$ in time polynomial in n , β , and the size of the generalized unary encoding of W . Moreover, for $f(u) := \|u\|_p$ (the p-norm), $1 \le p \le \infty$, the algorithm actually determines a $d^{1/q}$-approximate solution, where $1/p + 1/q = 1$ (with obvious interpretations when a denominator is ∞) .*

Proof Because we are minimizing a function that need not be concave, it may well be the case that the optimal solution of $P(\mathcal{F}, \min, f, W)$ is *not* realized at an extreme point of $\mathcal{P}$ and consequently, it may be that $\min\{f(u) \ : \ u \in W\mathcal{P}\}$ is *not* realized at an extreme point of $W\mathcal{P}$. Nonetheless, we will focus on extreme points of $W\mathcal{P}$.

Apply the algorithm of Lemma 3 so as to construct the set of vertices of $W\mathcal{P}$. Using the comparison oracle of f , identify a vertex $\hat{u} \in \text{vert}(W\mathcal{P})$ attaining minimum value $f(u)$. Now apply the algorithm of Lemma 1 to $\hat{u}$ and, as guaranteed by Lemma 2, obtain an integer $\hat{x}$ in the fiber of $\hat{u}$, so that $\hat{u} = W\hat{x}$. Output the point $\hat{x}$.

We now show that this provides the claimed approximation. Let x^* be an optimal integer point, and let $u^* := Wx^*$ be its image. Let u' be a point on the boundary of $W\mathcal{P}$ satisfying $u' \le u^*$. By Carathéodory's theorem (on a facet of $W\mathcal{P}$ containing u'), u' is a convex combination $u' = \sum_{i=1}^r \lambda_i u^i$

of some $r \leq d$ vertices of $W\mathcal{P}$ for some coefficients $\lambda_i \geq 0$ with $\sum_{i=1}^r \lambda_i = 1$. Let t be an index for which $\lambda_t = \max_i\{\lambda_i\}$. Then $\lambda_t \geq \frac{1}{r}\sum_{i=1}^r \lambda_i = 1/r \geq 1/d$.

Because W is nonnegative, we find that so are u' and the u^i , and hence we obtain

$$\begin{array}{rcccccccl} f(W\hat{x}) & = & f(\hat{u}) & \leq & f(u^t) & \leq & d\lambda_t \cdot f(u^t) & \leq & d \cdot f(\lambda_t u^t) \\ & \leq & d \cdot f(\sum_{i=1}^r \lambda_i u^i) & = & d \cdot f(u') & \leq & d \cdot f(u^*) & = & d \cdot f(Wx^*) \ . \end{array}$$

This proves that $\hat{x}$ provides a d-approximate solution.

Now consider the case of the p-norm $f(u) = \|u\|_p$. First, we note that the first part already covers the case of $p = \infty$. So we confine our attention to $p < \infty$. We complete now the proof for $1 < p < \infty$, and then it is an easy observation that the proof still goes through (in a simplified form) establishing also that for $p = 1$ we get a 1-approximate (i.e., optimal) solution.

By the Hölder inequality ($|a^\top b| \leq \|a\|_p \|b\|_q$, for $1/p + 1/q = 1$), we have

$$1 \ = \ \textstyle\sum_{i=1}^r 1 \cdot \lambda_i \leq (\sum_{i=1}^r 1^q)^{1/q} \, (\sum_{i=1}^r \lambda_i^p)^{1/p} \ = \ r^{1/q} \, (\sum_{i=1}^r \lambda_i^p)^{1/p} \ \leq \ d^{1/q} \, (\sum_{i=1}^r \lambda_i^p)^{1/p} \ .$$

Find s with $\|u^s\|_p = \min_i\{\|u^i\|_p\}$ and recall that the u^i are nonnegative. We then have

$$\begin{array}{rcccccccl} f^p(W\hat{x}) & = & \|\hat{u}\|_p^p & \leq & \|u^s\|_p^p & \leq & d^{p/q} \, (\sum_{i=1}^r \lambda_i^p) \, \|u^s\|_p^p & \leq & d^{p/q} \sum_{i=1}^r \left(\lambda_i^p \|u^i\|_p^p\right) \\ & \leq & d^{p/q} \left\|\sum_{i=1}^r \lambda_i u^i\right\|_p^p & = & d^{p/q} \|u'\|_p^p & \leq & d^{p/q} \|u^*\|_p^p & = & d^{p/q} f^p(Wx^*) \ , \end{array}$$

which proves that in this case, as claimed, $\hat{x}$ provides moreover a $d^{1/q}$-approximate solution. □

3 Randomized algorithm for vectorial matroid intersection

The algorithm of §1.3 can be used to solve efficiently the problem $P(\mathcal{F}, \max, f, W)$, when $\mathcal{F}$ is the set of characteristic vectors of sets that are common independent sets or bases of a pair of matroids on a common ground set, and f is convex. Such an algorithm only needs an independence oracle for each of the matroids. Even better, we can do likewise when $\mathcal{F}$ is the set of integer points that are in a pair of integral polymatroids or associated base polytopes. For polymatroids, we only need an oracle that evaluates the associated submodular functions.

In this section, we relax the assumption of f being convex. Indeed, we give a *randomized* algorithm for arbitrary f presented by a comparison oracle. By this we mean an algorithm that has access to a random bit generator, and on any input, it outputs the optimal solution with probability at least half. We restrict our attention to unary encoded W and $\mathcal{F}$ being the set of characteristic vectors of common bases of a pair of vectorial matroids M_1 and M_2 on the common ground set $\{1, 2, \ldots, n\}$. Our broad generalization of the standard linear-objective matroid intersection problem is defined as follows.

Nonlinear Matroid Intersection. Given two matroids M_1 and M_2 on the common ground set $N = \{1, \ldots, n\}$, integer weight matrix $W \in \mathbb{Z}^{d \times n}$ and an arbitrary function $f : \mathbb{R}^d \to \mathbb{R}$ specified by a comparison oracle . Let $\mathcal{F}$ be the set of characteristic vectors of common bases of M_1 and M_2 . Find an $x \in \mathcal{F}$ maximizing (or minimizing) $f(Wx)$.

When M_1 and M_2 are vectorial matroids, they can be represented by a pair of $r \times n$ matrices. We assume, without loss of generality, that r is the common rank of M_1 and M_2. Note that in this case $\mathcal{F}$ satisfies $\mathcal{F} \subseteq I_{n,r}$, where $I_{n,r}$ is the set of vectors in $\mathbb{R}^n$ with exactly r nonzero entries which are equal to 1. Thus our problem consists on finding an x that maximizes (minimizes) $f(Wx)$, where x is the characteristic vector corresponding to the set of column indices defining a common base. We assume the matroids have at least one common base which can be efficiently checked by the standard matroid intersection algorithm.

By adding to each $W_{i,j}$ a suitable positive integer v and replacing the function f by the function that maps each $u \in \mathbb{R}^d$ to $f(u_1 - rv, \ldots, u_d - rv)$ if necessary, we may and will assume without loss of generality throughout this section that the given weight matrix is nonnegative, $W \in \mathbb{Z}_+^{d \times n}$.

In this section we will be working with polynomials with integer coefficients in the $n + d$ variables a_j , $j = 1, \ldots, n$ and b_i , $i = 1, \ldots, d$. Define a vector γ in $\mathbb{R}^n$ whose entries are monomials by

$$\gamma_j \;:=\; a_j \prod_{i=1}^{d} b_i^{W_{i,j}},\ j=1,\ldots,n\ .$$

For each vectors $x \in \mathbb{Z}_+^n$ and $u \in \mathbb{Z}_+^d$, the corresponding monomials are

$$a^x := \prod_{j=1}^{n} a_j^{x_j} \quad , \qquad b^u := \prod_{i=1}^{d} b_i^{u_i}\ .$$

Let M_1^x and M_2^x be the $r \times r$ matrices comprising the r columns whose indices correspond to the nonzero entries in x . Finally, for each $u \in \mathbb{Z}_+^d$ define the following polynomial in the variables $a = (a_j)$ only.

$$g_u(a) \quad := \quad \sum \{\det(M_1^x)\cdot\det(M_2^x)a^x \,:\, x\in\mathcal{F}\ ,\ \ Wx=u\}\ .$$

Define $U := W\mathcal{F} = \{Wx : x \in \mathcal{F}\}$ as the set of points that are the projection of $\mathcal{F}$ under the weight matrix W . Let $M_1(\gamma)$ be a matrix whose j-th column is $M_1^j\gamma_j$. We then have the following identity in terms of the $g_u(a)$.

$$\begin{aligned}
\det(M_1(\gamma)M_2^\top) \;&=\; \sum_{x\in I_{n,r}} \det(M_1^x(\gamma))\cdot\det(M_2^x) \;=\; \sum_{x\in I_{n,r}} \det(M_1^x)\cdot\det(M_2^x)\prod_{j=1}^{n}\gamma_j^{x_j}\\
&= \sum_{x\in\mathcal{F}} \det(M_1^x)\cdot\det(M_2^x)\prod_{j=1}^{n}\gamma_j^{x_j} \;=\; \sum_{x\in\mathcal{F}} \det(M_1^x)\cdot\det(M_2^x)\prod_{j=1}^{n}(a_j\prod_{i=1}^{d} b_i^{W_{i,j}})^{x_j}\\
&= \sum_{x\in\mathcal{F}} \det(M_1^x)\cdot\det(M_2^x)\prod_{j=1}^{n}a_j^{x_j}\prod_{j=1}^{n}\prod_{i=1}^{d} b_i^{W_{i,j}x_j} \;=\; \sum_{x\in\mathcal{F}} \det(M_1^x)\cdot\det(M_2^x)\prod_{j=1}^{n}a_j^{x_j}\prod_{i=1}^{d} b_i^{W_i\cdot x}\\
&= \sum_{x\in\mathcal{F}} \det(M_1^x)\cdot\det(M_2^x)a^x b^{W\cdot x} \;=\; \sum_{u\in U}\sum_{\substack{x\in\mathcal{F}\ :\\ Wx=u}} \det(M_1^x)\cdot\det(M_2^x)a^x b^{W\cdot x} \;=\; \sum_{u\in U} g_u(a)b^u\ .
\end{aligned}$$

Next we consider integer substitutions for the variables a_j . Under such substitutions, each $g_u(a)$ becomes an integer, and $\det(M_1(\gamma)M_2^\top) = \sum_{u\in U} g_u(a)b^u$ becomes a polynomial in the variables $b = (b_i)$ only. Given such a substitution, let $U(a) := \{u \in U \,:\, g_u(a) \neq 0\}$ be the *support* of $\det(M_1(\gamma)M_2^\top)$, that is, the set of exponents of monomial b^u appearing with nonzero coefficient in $\det(M_1(\gamma)M_2^\top)$.

The next proposition concerns substitutions of independent identical random variables uniformly distributed on $\{1,2,\ldots,s\}$, under which $U(a)$ becomes a random subset $\widehat{U} \subseteq U$.

Proposition 2 *Suppose that independent uniformly-distributed random variables on* $\{1,2,\ldots,s\}$ *are substituted for the* a_j *, and let* $\widehat{U} := \{u \in U \,:\, g_u(a) \neq 0\}$ *be the random support of* $\det(M_1(\gamma)M_2^\top)$ *. Then, for every* $u \in U = \{Wx : x \in \mathcal{F}\}$ *, the probability that* $u \notin \widehat{U}$ *is at most* r/s *.*

Proof Consider any $u \in U$, and consider $g_u(a)$ as a polynomial in the variables $a = (a_j)$. Because $u = Wx$ for some $x \in \mathcal{F}$, there is as least one term $\det(M_1^x)\cdot\det(M_2^x)a^x$ in $g_u(a)$. Because distinct $x \in \mathcal{F}$ give distinct monomials a^x , no cancelations occur among the terms $\det(M_1^x)\cdot\det(M_2^x)a^x$ in $g_u(a)$. Thus, $g_u(a)$ is a nonzero polynomial of degree r . The claim now follows from a lemma of Schwartz [10] stating that the substitution of independent uniformly-distributed random variable on $\{1,2,\ldots,s\}$ into a nonzero multivariate polynomial of degree r is zero with probability at most r/s . □

The next lemma establishes that, given a_j , the support $U(a)$ of $\det(M_1(\gamma)M_2^\top)$ is polynomial-time computable.

Lemma 4 *For every fixed* d *, there is an algorithm that, given* $W \in \mathbb{Z}_+^{d\times n}$ *, and substitutions* $a_j \in \{1,2,\ldots,s\}$ *, computes* $U(a) = \{u \in U \,:\, g_u(a) \neq 0\}$ *in time polynomial in* n *,* $\max W_{i,j}$ *and* $\lceil \log s \rceil$ *.*

Proof The proof is based on interpolation. For each u , let $g_u := g_u(a)$ be the fixed integer obtained by substituting the given integers a_j . Let $z := r \max W_{i,j}$ and $Z := \{0,1,\ldots,z\}^d$. Note that z and $|Z| = (1+z)^d$ are polynomial in n and the unary encoding of W . Then $U(a) \subseteq U \subseteq Z$, and hence $\det(M_1(\gamma)M_2^\top) = \sum_{u\in Z} g_u b^u$ is a polynomial in d variables $b = (b_i)$ involving at most $|Z| = (z+1)^d$ monomials. For $t = 1,2,\ldots,(z+1)^d$, consider the substitution $b_i := t^{(z+1)^{i-1}}$, $i = 1,\ldots,d$, of suitable points on the moment curve. Under this substitution of b and a given substitution of a , the matrix $M_1(\gamma(t))M_2^\top$ becomes an integer matrix and so its determinant $\det(M_1(\gamma(t))M_2^\top)$ becomes an integer number that can be computed in polynomial time by Gaussian elimination. So we obtain the following system of $(z+1)^d$ equations in $(z+1)^d$ variables g_u , $u \in Z = \{0,1,\ldots,z\}^d$,

$$\det(M_1(\gamma(t))M_2^\top) \quad = \quad \sum_{u\in Z} g_u \prod_{k=1}^{d} b_k^{u_k} \quad = \quad \sum_{u\in Z} t^{\sum_{k=1}^{d} u_k (z+1)^{k-1}} \cdot g_u \,, \qquad t = 1,2,\ldots,(z+1)^d \,.$$

As $u = (u_1,\ldots,u_d)$ runs through Z , the sum $\sum_{k=1}^{d} u_k (z+1)^{k-1}$ attains precisely all $|Z| = (z+1)^d$ distinct values $0,1,\ldots,(z+1)^d - 1$. This implies that, under the total order of the points u in Z by increasing value of $\sum_{k=1}^{d} u_k (z+1)^{k-1}$, the vector of coefficients of the g_u in the equation corresponding to t is precisely the point $(t^0, t^1, \ldots, t^{(z+1)^d-1})$ on the moment curve in $\mathbb{R}^Z \simeq \mathbb{R}^{(z+1)^d}$. Therefore, the equations are linearly independent, and hence the system can be solved for the $g_u = g_u(a)$ and the desired support $U(a) = \{u \in U \,:\, g_u(a) \neq 0\}$ of $\det(M_1(\gamma)M_2)$ can indeed be computed in polynomial time. □

Next, we demonstrate that finding an optimal common base for a nonlinear matroid intersection problem can be reduced to finding an optimal W-image for a small number of subproblems. Consider data for a nonlinear matroid intersection problem, consisting of two matroids M_1 and M_2 on the common ground set $N = \{1,\ldots,n\}$, weight matrix $W \in \mathbb{Z}^{d\times n}$, and function $f : \mathbb{R}^d \to \mathbb{R}$. Each subset $S \subseteq N$ gives a *subproblem* of nonlinear matroid intersection as follows. The matroids of the subproblem are the *restriction* of M_1 and M_2 to S ; that is, the matroid $M_i.S$, $i = 1,2$, on ground set S in which a subset $I \subseteq S$ is independent if and only if it is independent in M_i . Note that the *restriction* of the matrix representing M_i to the columns indexed by S is the matrix representation of $M_i.S$. Then, the subproblem is the nonlinear matroid intersection problem, consisting of two matroids $M_1.S$ and $M_2.S$ on the common ground set S , the new weight matrix is the restrictions of the original weight matrix to S , and the function $f : \mathbb{R}^d \to \mathbb{R}$ is the same as in the original problem. We have the following useful statement.

Lemma 5 *The nonlinear matroid intersection problem of finding an optimal common base of two matroids M_1 and M_2 is reducible in time polynomial in n to finding an optimal W-image for at most $n+1$ subproblems.*

Proof Denote by $u^*(S)$ the optimal W-image for the matroid intersection subproblem defined by the matroids $M_1.S$ and $M_2.S$ on the ground set $S \subseteq N$. Now compute an optimal common base for the original problem, by computing an optimal W-image for $n+1$ such subproblems as follows.

```
Start with S := N;
Compute an optimal W-image u* := u*(N) of the original problem;
for j=1,2,...,n do
    Let T := S \ {j};
    Compute rank(M1.T) and rank(M2.T);
    if rank(M1.T) = r and rank(M2.T) = r then
        Compute an optimal W-image u*(T);
        if f(u*(T)) ≥ f(u*) then let S := T;
    end
end
return B := S;
```

It is not hard to verify that the set B obtained is indeed an optimal common base for the original problem. □

We are now in position to state our main theorem of this section. By a *randomized algorithm* that solves the nonlinear matroid intersection problem we mean an algorithm that has access to a random bit generator and on any input to the problem outputs a common base that is optimal with probability at least a half. The running time of the algorithm includes a count of the number of random bits used. Note that by repeatedly applying such an algorithm many times and picking the best common base, the probability of failure can be decreased arbitrarily; in particular, repeating it n times decreases the failure probability to as negligible a fraction as $1/2^n$ while increasing the running time by a linear factor only.

Theorem 4 *For every fixed d , there is a randomized algorithm that, given an integer weight matrix $W \in \mathbb{Z}^{d\times n}$, and a function $f : \mathbb{R}^d \to \mathbb{R}$ presented by a comparison oracle, solves the nonlinear matroid intersection problem in time polynomial in n and $\max W_{i,j}$.*

Proof As explained at the beginning of this section, we may and will assume that W is nonnegative. We claim that the following adaptation of the algorithm of Lemma 5 provides a common base which is optimal with probability at least $1/2$. Apply the algorithm of Lemma 5, but at each iteration apply the algorithm of Lemma 4 to determine a random optimal W-image $\hat{u}^*(T)$ which we claim to be the optimal W-image $u^*(T)$ with probability at least $1 - 1/\left(2(n+1)\right)$. At each iteration, the algorithm either returns: a correct optimal W-image with probability at least $1 - 1/\left(2(n+1)\right)$; or a wrong one, or the computed $\hat{U}$ is empty with probability at most $1/\left(2(n+1)\right)$. In case that the $\hat{U}$ computed by the randomized algorithm is empty, we set $\hat{u}^*(T)$ to be a virtual point with objective value $f(\hat{u}^*(T)) = -\infty$.

Now we show that the probability of success in each iteration, that is, of computing an optimal W-image $u^*(T)$ for any $T \subseteq N$ in which $rank(M_1.T) = rank(M_2.T) = r$, is at least $1 - 1/\left(2(n+1)\right)$, as claimed before. Indeed, using polynomially many random bits, draw independent and uniformly-distributed integers from $\{1, 2, \ldots, 2r(n+1)\}$ and substitute them for the a_j . Next compute a point $\hat{u}^*(T) \in \widehat{U}$ attaining $\max\{f(u) : u \in \widehat{U}\}$ using the algorithm underlying Lemma 4 and a comparison oracle. By Proposition 2, with probability at least $1 - 1/\left(2(n+1)\right)$, we have $u^*(T) \in \widehat{U}$, in which event $\max\{f(u) : u \in \widehat{U}\} = \max\{f(u) : u \in U\}$ and then $u^*(T)$ is indeed an optimal W-image.

Therefore, the probability that the algorithm obtains a correct optimal W-image at all iterations, and consequently outputs an optimal common base B of M_1 and M_2 , is at least the product over all iterations of the probability of success in each iteration, that is $\left(1 - 1/\left(2(n+1)\right)\right)^{n+1}$. Hence at least $1/2$ as desired. This completes the proof.

□

Acknowledgements Yael Berstein was supported by the Israel Ministry of Science Scholarship for Women and by a scholarship from the Graduate School of the Technion. The research of Jon Lee, Shmuel Onn and Robert Weismantel was partially supported by the Mathematisches Forschungsinstitut Oberwolfach during a stay within the Research in Pairs Programme. Shmuel Onn was also supported by the ISF - Israel Science Foundation. Robert Weismantel was also supported by the European TMR Network ADONET 504438.

References

1. Avriel, M., Diewert, W.E., Schaible, S., Zang, I.: Generalized concavity, *Mathematical Concepts and Methods in Science and Engineering*, vol. 36. Plenum Press, New York (1988)
2. Berstein, Y., Lee, J., Maruri-Aguilar, H., Onn, S., Riccomagno, E., Weismantel, R., Wynn, H.: Nonlinear matroid optimization and experimental design. SIAM Journal on Discrete Mathematics **to appear** (2008)
3. Berstein, Y., Onn, S.: Nonlinear bipartite matching. Discrete Optimization **5**(1), 53–65 (2008)
4. Lee, J.: A first course in combinatorial optimization. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge (2004)
5. Lee, J., Onn, S., Weismantel, R.: Nonlinear optimization over a weighted independence system. IBM Research Report RC24513
6. Lee, J., Onn, S., Weismantel, R.: Nonlinear discrete optimization. draft monograph (2008)
7. Onn, S.: Convex matroid optimization. SIAM J. Discrete Math. **17**(2), 249–253 (2003)
8. Onn, S., Rothblum, U.: Convex combinatorial optimization. Discrete and Computational Geometry **32**, 549–566 (2004)
9. Schrijver, A.: Combinatorial optimization. Springer, Berlin, Germany (2002)
10. Schwartz, J.T.: Fast probabilistic algorithms for verification of polynomial identities. J. Assoc. Comput. Mach. **27**(4), 701–717 (1980)
11. Ziegler, G.M.: Lectures on polytopes, *Graduate Texts in Mathematics*, vol. 152. Springer-Verlag, New York (1995)