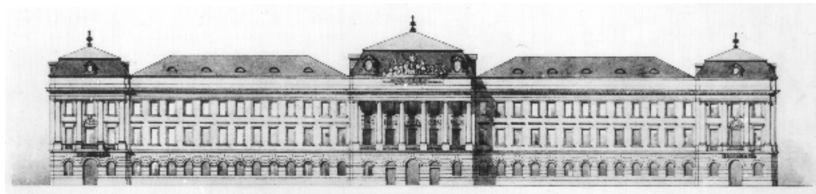
I N F S Y S
RESEARCH
R E P O R T



INSTITUT FÜR INFORMATIONSSYSTEME
ABTEILUNG WISSENSBASIERTE SYSTEME

A PROBABILISTIC APPROACH TO PROBLEMS PARAMETERIZED ABOVE OR BELOW TIGHT BOUNDS

Gregory Gutin Eun Jung Kim Stefan Szeider
Anders Yeo

INFSYS RESEARCH REPORT 1843-11-04
JULY 2011

Institut für Informationssysteme
Abtg. Wissensbasierte Systeme
Technische Universität Wien
Favoritenstraße 9-11
A-1040 Wien, Austria
Tel: +43-1-58801-18405
Fax: +43-1-58801-18493
sek@kr.tuwien.ac.at
www.kr.tuwien.ac.at

TECHNISCHE UNIVERSITÄT WIEN

INFSYS RESEARCH REPORT
INFSYS RESEARCH REPORT 1843-11-04, JULY 2011

A PROBABILISTIC APPROACH TO PROBLEMS PARAMETERIZED ABOVE OR BELOW TIGHT BOUNDS

Gregory Gutin[1] Eun Jung Kim[1] Stefan Szeider[2]
Anders Yeo[1]

Abstract. We introduce a new approach for establishing fixed-parameter tractability of problems parameterized above tight lower bounds or below tight upper bounds. To illustrate the approach we consider two problems of this type of unknown complexity that were introduced by Mahajan, Raman and Sikdar (J. Comput. Syst. Sci. 75, 2009). We show that a generalization of one of the problems and three nontrivial special cases of the other problem admit kernels of quadratic size. As a byproduct we obtain a new probabilistic inequality that could be of independent interest. Our new inequality is dual to the Hypercontractive Inequality.
Keywords: parameterized problems; above tight bounds; fixed-parameter tractable; kernel; Hypercontractive Inequality; probabilistic method.

[1] Department of Computer Science, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK
[2] Institute of Information Systems, Vienna University of Technology, A-1040 Vienna, Austria

Acknowledgements: Research of Gutin, Kim and Yeo was supported in part by an EPSRC grant. Research of Gutin was also supported in part by the IST Programme of the European Community, under the PASCAL 2 Network of Excellence. Research of Szeider was partially supported by the European Research Council (ERC), Grant 239962 (COMPLEX REASON).

Publication Information: This is the author's self-archived copy of a paper that appeared in the *Journal of Computer and Systems Sciences*, vol. 77, no. 2, pp. 422-429, 2011. doi:10.1016/j.jcss.2010.06.001. The final publication is available at www.sciencedirect.com

1 Introduction

A parameterized problem Π can be considered as a set of pairs (x, k) where x is the *main part* and k (usually an integer) is the *parameter*. Π is called *fixed-parameter tractable* (FPT) if membership of (x, k) in Π can be decided in time $O(f(k)|x|^c)$, where $|x|$ denotes the size of x, $f(k)$ is a computable function, and c is a constant independent of k and I (for further background and terminology on parameterized complexity we refer the reader to the monographs [9, 10, 19]). If the nonparameterized version of Π (where k is just a part of the input) is NP-hard, then the function $f(k)$ must be superpolynomial provided P $\neq$ NP. Often $f(k)$ is "moderately exponential," which makes the problem practically feasible for small values of k. Thus, it is important to parameterize a problem in such a way that the instances with small values of k are of real interest.

Consider the following well-known problem: given a digraph $D = (V, A)$, find an acyclic subdigraph of D with the maximum number of arcs. We can parameterize this problem "naturally" by asking whether D contains an acyclic subdigraph with at least k arcs. It is easy to prove that this parameterized problem is fixed-parameter tractable by observing that D always has an acyclic subdigraph with at least $|A|/2$ arcs. (Indeed, consider a bijection $\alpha : V \to \{1, \dots, |V|\}$ and the following subdigraphs of D: $(V, \{ xy \in A : \alpha(x) < \alpha(y) \})$ and $(V, \{ xy \in A : \alpha(x) > \alpha(y) \})$. Both subdigraphs are acyclic and at least one of them has at least $|A|/2$ arcs.) However, $k \leq |A|/2$ for every small value of k and almost every practical value of $|A|$ and, thus, our "natural" parameterization is of almost no practical or theoretical interest.

Instead, one should consider the following parameterized problem: decide whether $D = (V, A)$ contains an acyclic subdigraph with at least $|A|/2 + k$ arcs. We choose $|A|/2 + k$ because $|A|/2$ is a *tight lower bound* on the size of a largest acyclic subdigraph. Indeed, the size of a largest acyclic subdigraph of a symmetric digraph $D = (V, A)$ is precisely $|A|/2$. (A digraph $D = (V, A)$ is *symmetric* if $xy \in A$ implies $yx \in A$.)

In a recent paper [18] Mahajan, Raman and Sikdar provided several examples of problems of this type and argued that a natural parameterization is one above a tight lower bound for maximization problems, and below a tight upper bound for minimization problems. Furthermore, they observed that only a few non-trivial results are known for problems parameterized above a tight lower bound [14, 15, 17, 21], and they listed several problems parameterized above a tight lower bound whose complexity is unknown. The difficulty in showing whether such a problem is fixed-parameter tractable can be illustrated by the fact that often we even do not know whether the problem is in XP, i.e., can be solved in time $O(|I|^{g(k)})$ for a computable function $g(k)$. For example, it is non-trivial to see that the above-mentioned digraph problem is in XP when parameterized above the $|A|/2$ bound.

In this paper we introduce the *Strictly Above/Below Expectation Method (SABEM)*, a novel approach for establishing the fixed-parameter tractability of maximization problems parameterized above tight lower bounds and minimization problems parameterized below tight upper bounds. The new method is based on probabilistic arguments and utilizes certain probabilistic inequalities. We will state the equalities in the next section, and in the subsequent sections we will apply SABEM to two open problems posed in [18].

Now we give a very brief description of the new method with respect to a given problem Π

parameterized above a tight lower bound or below a tight upper bound. We first apply some reductions rules to reduce Π to its special case Π'. Then we introduce a random variable X such that the answer to Π is YES if X takes, with positive probability, a value greater or equal to the parameter k. Now using some probabilistic inequalities on X, we derive upper bounds on the size of NO-instances of Π' in terms of a function of the parameter k. If the size of a given instance exceeds this bound, then we know the answer is YES; otherwise, we produce a *kernel*; see the next paragraph.

Given a parameterized problem Π, a *kernelization of* Π is a polynomial-time algorithm that maps an instance (x, k) of Π to an instance (x', k') of Π, the *kernel*, such that (i) $(x, k) \in \Pi$ if and only if $(x', k') \in \Pi$, (ii) $k' \leq f(k)$, and (iii) $|x'| \leq g(k)$ for some functions f and g. The function $g(k)$ is called the *size* of the kernel. A parameterized problem is fixed-parameter tractable if and only if it is decidable and admits a kernelization [9, 10, 19]; however, the kernels obtained by this general result have impractically large size. Therefore, one tries to develop kernelizations that yield kernels of smaller size; polynomial size kernels are of great interest.

In Section 2, we describe probabilistic inequalities used in the new method. The inequalities include a recent inequality of Alon et al. [1], the well-known Hypercontractive Inequality and a new result, Lemma 1, which is an analog of the Hypercontractive Inequality and is dual to the Hypercontractive Inequality, in a sense.

In Section 3, we consider the LINEAR ORDERING problem, a generalization of the problem discussed above: Given a digraph $D = (V, A)$ in which each arc ij has a positive integral weight w_{ij}, find an acyclic subdigraph of D of maximum weight. Observe that $W/2$, where W is the sum of all arc weights, is a tight lower bound for LINEAR ORDERING. We prove that the problem parameterized above $W/2$ is fixed-parameter tractable and admits a quadratic kernel. Note that this parameterized problem generalizes the maximum acyclic subdigraph problem parameterized above a tight lower bound considered in [18]; thus, our result answers the corresponding open question of [18].

In Section 4, we consider the problem MAX LIN-2: Given a system of m linear equations $e_1, \ldots, e_m$ in n variables over GF(2), and for each equation e_j a positive integral weight w_j; find an assignment of values to the n variables that maximizes the total weight of the satisfied equations. We will see that $W/2$, where $W = w_1 + \cdots + w_m$, is a tight lower bound for MAX LIN-2. The complexity of the problem parameterized above $W/2$ is open [18]. We prove that the following three special cases of the parameterized problem are fixed-parameter tractable: (1) there is a set U of variables such that each equation has an odd number of variables from U, (2) there is a constant r such that each equation involves at most r variables, (3) there is a constant ρ such that any variable appears in at most ρ equations. For all three cases we obtain kernels with $O(k^2)$ variables and equations. We also show that if we allow the weights w_j to be positive rational numbers, the problem is NP-hard already if $k = 1$ and each equation involves two variables.

In Section 5, we briefly mention minimization problems parameterized below tight upper bounds, provide further discussions of problems considered in this paper and point out to recent results obtained using our new method.

2 Probabilistic Inequalities

In this paper all random variables are real. A random variable is *discrete* if its distribution function has a finite or countable number of positive increases. A random variable X is *symmetric* if $-X$ has the same distribution function as X. If X is discrete, then X is symmetric if and only if $\text{Prob}(X = a) = \text{Prob}(X = -a)$ for each real a. Let X be a symmetric variable for which the first moment $\mathbb{E}(X)$ exists. Then $\mathbb{E}(X) = \mathbb{E}(-X) = -\mathbb{E}(X)$ and, thus, $\mathbb{E}(X) = 0$. The following is easy to prove [22].

Lemma 1. *If X is a symmetric random variable and $\mathbb{E}(X^2)$ is finite, then*

$$\text{Prob}(\ X \geq \sqrt{\mathbb{E}(X^2)}\ ) > 0.$$

See Sections 3 and 4 for applications of Lemma 1. If X is not symmetric then the following lemma can be used instead (a similar result was already proved in [2]).

Lemma 2 (Alon et al. [1]). *Let X be a real random variable and suppose that its first, second and fourth moments satisfy $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = \sigma^2 > 0$ and $\mathbb{E}[X^4] \leq c\sigma^4$, respectively, for some constant c. Then $\text{Prob}(X > \frac{\sigma}{2\sqrt{c}}) > 0$.*

We combine this result with the following result from harmonic analysis.

Lemma 3 (Hypercontractive Inequality [5, 11]). *Let $f = f(x_1, \ldots, x_n)$ be a polynomial of degree r in n variables $x_1, \ldots, x_n$ each with domain $\{-1, 1\}$. Define a random variable X by choosing a vector $(\epsilon_1, \ldots, \epsilon_n) \in \{-1, 1\}^n$ uniformly at random and setting $X = f(\epsilon_1, \ldots, \epsilon_n)$. Then $\mathbb{E}[X^4] \leq 9^r \mathbb{E}[X^2]^2$.*

If $f = f(x_1, \ldots, x_n)$ is a polynomial in n variables $x_1, \ldots, x_n$ each with domain $\{-1, 1\}$, then it can be written as $f = \sum_{S \subseteq [n]} c_S \prod_{i \in S} x_i$, where $[n] = \{1, \ldots, n\}$ and c_S is a real for each $S \subseteq [n]$. The following dual, in a sense, form of the Hypercontractive Inequality is proved in Section 4 (see an explanation after Lemma 7).

Proposition 1. *Let $f = f(x_1, \ldots, x_n)$ be a polynomial in n variables $x_1, \ldots, x_n$ each with domain $\{-1, 1\}$ such that $f = \sum_{S \subseteq [n]} c_S \prod_{i \in S} x_i$. Suppose that no variable x_i appears in more than $\rho \geq 2$ monomials of f. Define a random variable X by choosing a vector $(\epsilon_1, \ldots, \epsilon_n) \in \{-1, 1\}^n$ uniformly at random and setting $X = f(\epsilon_1, \ldots, \epsilon_n)$. Then $\mathbb{E}[X^4] \leq 2\rho^2 \mathbb{E}[X^2]^2$.*

3 Linear Ordering

Let $D = (V, A)$ be a digraph with no loops or parallel arcs in which every arc ij has a positive weight w_{ij}. The problem of finding an acyclic subdigraph of D of maximum weight, known as LINEAR ORDERING, has applications in economics [3]. Let $n = |V|$ and consider a bijection $\alpha : V \to \{1, \ldots, n\}$. Observe that the subdigraphs $(V, \{\, ij \in A : \alpha(i) < \alpha(j)\,\})$ and $(V, \{\, ij \in A : \alpha(i) > \alpha(j)\,\})$ are acyclic. Since the two subdigraphs contain all arcs of D, at least one of them has weight at least $W/2$, where $W = \sum_{ij \in A} w_{ij}$, the *weight* of D. Thus, $W/2$ is a lower bound on the maximum weight of an acyclic subdigraph of D. Consider a digraph D where for every arc ij of D there is also an arc ji of the same weight. Each maximum weight subdigraph of D has weight exactly $W/2$. Hence the lower bound $W/2$ is tight.

LINEAR ORDERING ABOVE TIGHT LOWER BOUND (LOALB)

Instance: A digraph $D = (V, A)$, each arc ij has an integral positive weight w_{ij}, and a positive integer k.

Parameter: The integer k.

Question: Is there an acyclic subdigraph of D of weight at least $W/2 + k$, where $W = \sum_{ij \in A} w_{ij}$?

Mahajan, Raman, and Sikdar [18] asked whether LOALB is fixed-parameter tractable for the special case when all arcs are of weight 1 (i.e., D is unweighted). In this section we will prove that LOALB admits a kernel with $O(k^2)$ arcs; consequently the problem is fixed-parameter tractable. Note that if we allow weights to be positive reals, then we can show, similarly to the NP-completeness proof given in the next section, that LOALB is NP-complete already for $k = 1$.

Consider the following reduction rule:

Reduction Rule 1. *Assume D has a directed 2-cycle iji; if $w_{ij} = w_{ji}$ delete the cycle, if $w_{ij} > w_{ji}$ delete the arc ji and replace w_{ij} by $w_{ij} - w_{ji}$, and if $w_{ji} > w_{ij}$ delete the arc ij and replace w_{ji} by $w_{ji} - w_{ij}$.*

It is easy to check that the answer to LOALB for a digraph D is YES if and only if the answer to LOALB is YES for a digraph obtained from D using the reduction rule as long as possible. A digraph is called an *oriented graph* if it has no directed 2-cycle. Note that applying Rule 1 as long as possible results in an oriented graph.

Let $D = (V, A)$ be an oriented graph, let $n = |V|$ and $W = \sum_{ij \in A} w_{ij}$. Consider a random bijection: $\alpha : V \to \{1, \dots, n\}$ and a random variable $X(\alpha) = \frac{1}{2} \sum_{ij \in A} \epsilon_{ij}(\alpha)$, where $\epsilon_{ij}(\alpha) = w_{ij}$ if $\alpha(i) < \alpha(j)$ and $\epsilon_{ij}(\alpha) = -w_{ij}$, otherwise. It is easy to see that $X(\alpha) = \sum\{ w_{ij} : ij \in A, \alpha(i) < \alpha(j) \} - W/2$. Thus, the answer to LOALB is YES if and only if there is a bijection $\alpha : V \to \{1, \dots, n\}$ such that $X(\alpha) \geq k$. Since $\mathbb{E}(\epsilon_{ij}) = 0$, we have $\mathbb{E}(X) = 0$.

Let $W^{(2)} = \sum_{ij \in A} w_{ij}^2$. We will prove the following:

Lemma 4. $\mathbb{E}(X^2) \geq W^{(2)}/12$.

Proof. Let $N^+(i)$ and $N^-(i)$ denote the sets of out-neighbors and in-neighbors of a vertex i in D. By the definition of X,

$$4 \cdot \mathbb{E}(X^2) = \sum_{ij \in A} \mathbb{E}(\epsilon_{ij}^2) + \sum_{ij,pq \in A} \mathbb{E}(\epsilon_{ij}\epsilon_{pq}), \tag{1}$$

where the second sum is taken over ordered pairs of distinct arcs. Clearly, $\sum_{ij \in A} \mathbb{E}(\epsilon_{ij}^2) = W^{(2)}$. To compute $\sum_{ij,pq \in A} \mathbb{E}(\epsilon_{ij}\epsilon_{pq})$ we consider the following cases:

Case 1: $\{i,j\} \cap \{p,q\} = \emptyset$. Then ϵ_{ij} and ϵ_{pq} are independent and $\mathbb{E}(\epsilon_{ij}\epsilon_{pq}) = \mathbb{E}(\epsilon_{ij})\mathbb{E}(\epsilon_{pq}) = 0$.

Case 2a: $|\{i,j\} \cap \{p,q\}| = 1$ and $i = p$. Since the probability that $i < \min\{j,q\}$ or $i > \max\{j,q\}$ is $2/3$, $\epsilon_{ij}\epsilon_{iq} = w_{ij}w_{iq}$ with probability $\frac{2}{3}$ and $\epsilon_{ij}\epsilon_{iq} = -w_{ij}w_{iq}$ with probability $\frac{1}{3}$. Thus, for every $i \in V$ we have $\sum_{ij,iq\in A} \mathbb{E}(\epsilon_{ij}\epsilon_{iq}) = \frac{1}{3}\sum_{j\neq q\in N^+(i)} w_{ij}w_{iq} = \frac{1}{3}(\sum_{j\in N^+(i)} w_{ij})^2 - \frac{1}{3}\sum_{j\in N^+(i)} w_{ij}^2$.

Case 2b: $|\{i,j\} \cap \{p,q\}| = 1$ and $j = q$. Similarly to Case 2a, we obtain $\sum_{ij,pj\in A} \mathbb{E}(\epsilon_{ij}\epsilon_{pj}) = \frac{1}{3}(\sum_{i\in N^-(j)} w_{ij})^2 - \frac{1}{3}\sum_{i\in N^-(j)} w_{ij}^2$.

Case 3a: $|\{i,j\} \cap \{p,q\}| = 1$ and $i = q$. Since $\epsilon_{ij}\epsilon_{pi} = w_{ij}w_{pi}$ with probability $\frac{1}{3}$ and $\epsilon_{ij}\epsilon_{pi} = -w_{ij}w_{pi}$ with probability $\frac{2}{3}$, we obtain $\sum_{ij,pi\in A} \mathbb{E}(\epsilon_{ij}\epsilon_{pi}) = -\frac{1}{3}\sum_{j\in N^+(i),\ p\in N^-(i)} w_{ij}w_{pi} = -\frac{1}{3}\sum_{j\in N^+(i)} w_{ij}\sum_{p\in N^-(i)} w_{pi}$.

Case 3b: $|\{i,j\} \cap \{p,q\}| = 1$ and $j = p$. Similarly to Case 3a, we obtain $\sum_{ij,jq\in A} \mathbb{E}(\epsilon_{ij}\epsilon_{jq}) = -\frac{1}{3}\sum_{i\in N^-(j)} w_{ij}\sum_{q\in N^+(j)} w_{jq}$.

Equation (1) and the subsequent computations imply that $4\cdot\mathbb{E}(X^2) = W^{(2)} + \frac{1}{3}(Q-R)$, where

$$Q = \sum_{i\in V}\left(\Big(\sum_{j\in N^+(i)} w_{ij}\Big)^2 - \sum_{j\in N^+(i)} w_{ij}^2 + \Big(\sum_{j\in N^-(i)} w_{ji}\Big)^2 - \sum_{j\in N^-(i)} w_{ji}^2 \right),$$

and

$$R = 2\cdot\sum_{i\in V}\Big(\sum_{j\in N^+(i)} w_{ij}\Big)\Big(\sum_{j\in N^-(i)} w_{ji}\Big).$$

By the inequality of arithmetic and geometric means, for each $i \in V$, we have

$$\Big(\sum_{j\in N^+(i)} w_{ij}\Big)^2 + \Big(\sum_{j\in N^-(i)} w_{ji}\Big)^2 - 2\Big(\sum_{j\in N^+(i)} w_{ij}\Big)\Big(\sum_{j\in N^-(i)} w_{ji}\Big) \geq 0.$$

Therefore,

$$Q - R \geq -\sum_{i\in V}\sum_{j\in N^+(i)} w_{ij}^2 - \sum_{i\in V}\sum_{j\in N^-(i)} w_{ji}^2 = -2W^{(2)},$$

and $4\cdot\mathbb{E}(X^2) \geq W^{(2)} - 2W^{(2)}/3 = W^{(2)}/3$, implying $\mathbb{E}(X^2) \geq W^{(2)}/12$. $\square$

Now we can prove the main result of this section.

Theorem 1. *The problem* LOALB *admits a kernel with $O(k^2)$ arcs.*

Proof. Let H be a digraph. We know that the answer to LOALB for H is YES if and only if the answer to LOALB is YES for a digraph D obtained from H using Reduction Rule 1 as long as possible. Observe that D is an oriented graph. Let $\mathcal{B}$ be the set of bijections from V to $\{1, \dots, n\}$. Observe that $f : \mathcal{B} \to \mathcal{B}$ such that $f(\alpha(v)) = |V| + 1 - \alpha(v)$ for each $\alpha \in \mathcal{B}$ is a bijection. Note that $X(f(\alpha)) = -X(\alpha)$ for each $\alpha \in \mathcal{B}$. Therefore, Prob$(X = a) =$ Prob$(X = -a)$ for each real a and, thus, X is symmetric. Thus, by Lemmas 1 and 4, we have Prob$(\ X \geq \sqrt{W^{(2)}/12}\ ) > 0$. Hence, if $\sqrt{W^{(2)}/12} \geq k$, there is a bijection $\alpha : V \to \{1, \dots, n\}$ such that $X(\alpha) \geq k$ and, thus, the answer to LOALB (for both D and H) is YES. Otherwise, $|A| \leq W^{(2)} < 12 \cdot k^2$. □

We close this section by outlining how Theorem 1 can be used to actually find a solution to LOALB if one exists. Let (D, k) be an instance of LOALB where $D = (V, A)$ is a directed graph with integral positive arc-weights and $k \geq 1$ is an integer. Let W be the total weight of D. As discussed above, we may assume that D is an oriented graph. If $|A| < 12k^2$ then we can find a solution, if one exists, by trying all subsets $A' \subseteq A$, and testing whether (V, A') is acyclic and has weight at least $W/2 + k$; this search can be carried out in time $2^{O(k^2)}$. Next we assume $|A| \geq 12k^2$. We know by Theorem 1 that (D, k) is a YES-instance; it remains to find a solution.

For a vertex $i \in V$ let $d_D(i)$ denote its unweighted degree in D, i.e., the number of arcs (incoming or outgoing) that are incident with i. Consider the following reduction rule:

Reduction Rule 2. *If there is a vertex $i \in V$ with $|A| - 12k^2 \geq d_D(i)$, then delete i from D.*

Observe that by applying the rule we obtain again a YES-instance $(D - i, k)$ of LOALB since $D - i$ has still at least $12k^2$ arcs. Moreover, if we know a solution D'_i of $(D - i, k)$, then we can efficiently obtain a solution D' of (D, k): if $\sum_{j \in N^+(i)} w_{ij} \geq \sum_{j \in N^-(i)} w_{ij}$ then we add i and all outgoing arcs $ij \in A$ to D'_i; otherwise, we add i and all incoming arcs $ji \in A$ to D'_i. After multiple applications of Rule 2 we are left with an instance (D_0, k) to which Rule 2 cannot be applied. Let $D_0 = (V_0, A_0)$. We pick a vertex $i \in V_0$ arbitrarily. If i has a neighbor j with $d_{D_0}(j) = 1$, then $|A_0| \leq 12k^2$, since $|A_0| - d_{D_0}(j) < 12k^2$. On the other hand, if $d_{D_0}(j) \geq 2$ for all neighbors j of i, then i has less than $2 \cdot 12k^2$ neighbors, since $D_0 - i$ has less than $12k^2$ arcs; thus $|A_0| < 3 \cdot 12k^2$. Therefore, as above, time $2^{O(k^2)}$ is sufficient to try all subsets $A'_0 \subseteq A_0$ to find a solution to the instance (D_0, k). Let n denote the input size of instance (D, k). Rule 2 can certainly be applied in polynomial time $n^{O(1)}$, and we apply it less than n times. Hence, we can find a solution to (D, k), if one exists, in time $n^{O(1)} + 2^{O(k^2)}$.

Recall that a kernelization reduces in polynomial time an instance (I, k) of a parameterized problem to a *decision-equivalent* instance (I', k'), its *problem kernel*, where $k' \leq k$ and the size of I' is bounded by a function of k. Solutions for (I, k) and solutions for (I', k') are possibly unrelated to each other. We call (I', k') a *faithful problem kernel* if from a solution for (I', k') we can construct a solution for (I, k) in time polynomial in $|I|$ and k. Clearly the above (D_0, k) is a faithful kernel.

4 Max Lin-2

Consider a system of m linear equations $e_1, \ldots, e_m$ in n variables $z_1, \ldots, z_n$ over GF(2), and suppose that each equation e_j has a positive integral weight w_j, $j = 1, \ldots, m$. The problem MAX LIN-2 asks for an assignment of values to the variables that maximizes the total weight of the satisfied equations. Let $W = w_1 + \cdots + w_m$.

To see that the total weight of the equations that can be satisfied is at least $W/2$, we describe a simple procedure suggested in [16]. We assign values to the variables $z_1, \ldots, z_n$ one by one and simplify the system after each assignment. When we wish to assign 0 or 1 to z_i, we consider all equations reduced to the form $z_i = b$, for a constant b. Let W' be the total weight of all such equations. We set $z_i := 0$, if the total weight of such equations is at least $W'/2$, and set $z_i := 1$, otherwise. If there are no equations of the form $z_i = b$, we set $z_i := 0$. To see that the lower bound $W/2$ is tight, consider a system consisting of pairs of equations of the form $\sum_{i \in I} z_i = 1$ and $\sum_{i \in I} z_i = 0$ where both equations have the same weight.

The parameterized complexity of MAX LIN-2 parameterized above the tight lower bound $W/2$ was stated by Mahajan, Raman and Sikdar [18] as an open question:

MAX LIN-2 PARAMETERIZED ABOVE TIGHT LOWER BOUND (LINALB)

Instance: A system S of m linear equations $e_1, \ldots, e_m$ in n variables $z_1, \ldots, z_n$ over GF(2), each equation e_i with a positive integral weight w_i, $i = 1, 2, \ldots, m$, and a positive integer k. Each equation e_j can be written as $\sum_{i \in I_j} z_i = b_j$, where $\emptyset \neq I_j \subseteq \{1, \ldots, n\}$.

Parameter: The integer k.

Question: Is there an assignment of values to the variables $z_1, \ldots, z_n$ such that the total weight of the satisfied equations is at least $W/2 + k$, where $W = \sum_{i=1}^{m} w_i$?

Let r_j be the number of variables in equation e_j, and let $r(S) = \max_{i=1}^{m} r_j$. We are not able to determine whether LINALB is fixed-parameter tractable or not, but we can prove that the following three special cases are fixed-parameter tractable: (1) there is a set U of variables such that each equation contains an odd number of variables from U, (2) there is a constant r such that $r(S) \leq r$, (3) there is a constant ρ such that any variable appears in at most ρ equations.

Notice that in our formulation of LINALB it is required that each equation has a positive integral weight. In a relaxed setting in which an equation may have any positive rational number as its weight, the problem is NP-complete even for $k = 1$ and each $r_j = 2$. Indeed, let each linear equation be of the form $z_u + z_v = 1$. Then the problem is equivalent to MAXCUT, the problem of finding a cut of total weight at least L in an undirected graph G, where $V(G)$ is the set of variables, $E(G)$ contains an edge $\{z_u, z_v\}$ if and only if there is a linear equation $z_u + z_v = 1$, and the weight of an edge $\{z_u, z_v\}$ equals the weight of the corresponding linear equation. The problem MAXCUT is a well-known NP-complete problem. Let us transform an instance I of MAXCUT into an instance I' of the "relaxed" LINALB by replacing the weight w_i by $w'_i := w_i/(L - W/2)$. We may assume that $L - W/2 > 0$ since otherwise the instance is immediately seen as a YES-instance.

Observe that the new instance I' has an assignment of values with total weight at least $W'/2+1$ if and only if I has a cut with total weight at least L. We are done.

Let A be the matrix of the coefficients of the variables in S. It is well-known that the maximum number of linearly independent columns of A equals $\mathrm{rank}A$, and such a collection of columns can be found in time polynomial in n and m, using, e.g., the Gaussian elimination on columns [4]. We have the following reduction rule and supporting lemma.

Reduction Rule 3. *Let A be the matrix of the coefficients of the variables in S, let $t = \mathrm{rank}A$ and let columns $a^{i_1}, \dots, a^{i_t}$ of A be linearly independent. Then set all variables not in $\{z_{i_1}, \dots, z_{i_t}\}$ to 0 and simplify the equations of S.*

Lemma 5. *Let T be obtained from S by Rule 3. Then T is a* YES*-instance if and only if S is a* YES*-instance. Moreover, T can be obtained from S in time polynomial in n and m.*

Proof. The remark before the lemma immediately implies that T can be obtained from S in time polynomial in n and m. The rest of the proof is taken from [8]. Consider an independent set I of columns of A of cardinality $\mathrm{rank}A$ and a column $a^j \notin I$. Observe that $a^j = \sum_{i \in I'} a^i$, where $I' \subseteq I$. Consider an assignment $z = z^0$. If $z_j^0 = 1$ then for each $i \in I' \cup \{j\}$ replace z_i^0 by $z_i^0 + 1$. The new assignment satisfies exactly the same equations as the initial assignment. Thus, we may assume that $z_j = 0$ and remove z_j from the system. $\square$

Consider the following reduction rule for LINALB.

Reduction Rule 4. *If we have, for a subset I of $\{1, 2, \dots, n\}$, the equation $\sum_{i \in I} z_i = b'$ with weight w', and the equation $\sum_{i \in I} z_i = b''$ with weight w'', then we replace this pair by one of these equations with weight $w' + w''$ if $b' = b''$ and, otherwise, by the equation whose weight is bigger, modifying its new weight to be the difference of the two old ones. If the resulting weight is 0, we omit the equation from the system.*

If Rule 4 is not applicable to a system we call the system *reduced under Rule 4*. Note that the problem LINALB for S and the system obtained from S by applying Rule 4 as long as possible have the same answer.

Let $I_j \subseteq \{1, \dots, n\}$ be the set of indices of the variables participating in equation e_j, and let $b_j \in \{0, 1\}$ be the right hand side of e_j. Define a random variable $X = \sum_{j=1}^{m} X_j$, where $X_j = (-1)^{b_j} w_j \prod_{i \in I_j} \epsilon_i$ and all the ϵ_i are independent uniform random variables on $\{-1, 1\}$ (X was first introduced in [16]). We set $z_i = 0$ if $\epsilon_i = 1$ and $z_i = 1$, otherwise, for each i. In other words, $\epsilon_i = (-1)^{z_i}$. Then z_i are independent uniform random variables on $\{0, 1\}$ and observe that $X_j = w_j$ if e_j is satisfied and $X_j = -w_j$, otherwise. Note that the relation $\epsilon_i = (-1)^{z_i}$ is well-known for Fourier expansions of pseudo-boolean functions, i.e., functions $f : \{-1, +1\}^n \to \mathbb{R}$, see, e.g., [20, 23].

Lemma 6. *Let S be reduced under Rule 4. The weight of the satisfied equations is at least $W/2+k$ if and only if $X \geq 2k$. We have $\mathbb{E}(X) = 0$ and $\mathbb{E}(X^2) = \sum_{j=1}^{m} w_j^2$.*

Proof. Observe that X is the difference between the weights of satisfied and falsified equations. Therefore, the weight of the satisfied equations equals $(X+W)/2$, and it is at least $W/2+k$ if and only if $X \geq 2k$. Since ϵ_i are independent, $\mathbb{E}(\prod_{i\in I_j}\epsilon_i) = \prod_{i\in I_j}\mathbb{E}(\epsilon_i) = 0$. Thus, $\mathbb{E}(X_j) = 0$ and $\mathbb{E}(X) = 0$ by linearity of expectation. Moreover,

$$\mathbb{E}(X^2) = \sum_{j=1}^{m}\mathbb{E}(X_j^2) + \sum_{1\leq j\neq q\leq m}\mathbb{E}(X_jX_q) = \sum_{j=1}^{m} w_j^2 > 0$$

as $\mathbb{E}(\prod_{i\in I_j}\epsilon_i \cdot \prod_{i\in I_q}\epsilon_i) = \mathbb{E}(\prod_{i\in I_j\Delta I_q}\epsilon_i) = 0$ implies $\mathbb{E}(X_jX_q) = 0$, where $I_j\Delta I_q$ is the symmetric difference between I_j and I_q ($I_j\Delta I_q \neq \emptyset$ due to Rule 4). □

Lemma 7. *Let S be reduced under Rule 4 and suppose that no variable appears in more than $\rho \geq 2$ equations of S. Then $\mathbb{E}(X^4) \leq 2\rho^2(\mathbb{E}(X^2))^2$.*

Proof. Observe that

$$\mathbb{E}(X^4) = \sum_{(p,q,s,t)\in[m]^4}\mathbb{E}(X_pX_qX_sX_t), \tag{2}$$

where $[m] = \{1,\dots,m\}$. Note that if the product $X_pX_qX_sX_t$ contains a variable ϵ_i in only one or three of the factors, then $\mathbb{E}(X_pX_qX_sX_t) = A\cdot\mathbb{E}(\epsilon_i) = 0$, where A is a polynomial in random variables ϵ_l, $l \in \{1,\dots,n\}\setminus\{i\}$. Thus, the only nonzero terms in (2) are those for which either (Case 1) $p=q=s=t$, or (Case 2) there are two distinct integers j,l such that each of them coincides with two elements in the sequence p,q,s,t, or (Case 3) $|\{p,q,s,t\}| = 4$, but each variable ϵ_i appears in an even number of the factors in $X_pX_qX_sX_t$. In Cases 1 and 2, we have $\mathbb{E}(X_pX_qX_sX_t) = w_p^4$ and $\mathbb{E}(X_pX_qX_sX_t) = w_j^2w_l^2$, respectively. In Case 3,

$$\mathbb{E}(X_pX_qX_sX_t) \leq w_pw_qw_sw_t \leq (w_p^2w_q^2 + w_s^2w_t^2)/2.$$

Let $1 \leq j < l \leq m$. Observe that $\mathbb{E}(X_pX_qX_sX_t) = w_j^2w_l^2$ in Case 2 for $\binom{4}{2} = 6$ 4-tuples $(p,q,s,t)\in[m]^4$. In Case 3, we claim that there are at most $4\cdot(\rho-1)^2$ 4-tuples $(p,q,s,t)\in[m]^4$ with $j,l \in \{p,q,s,t\}$ which contribute $w_j^2w_l^2/2$ to the bound on $\mathbb{E}(X_pX_qX_sX_t)$. Indeed, there are only four possible ways for $w_j^2w_l^2/2$ to appear in our upper bound, namely the following: (i) $j=p, l=q$, (ii) $l=p, j=q$, (iii) $j=s, l=t$, and (iv) $l=s, j=t$. Now assume, without loss of generality, that $j=p$ and $l=q$. Since S is reduced under Rule 4, the product X_jX_l must have a variable ϵ_i of degree one. Thus, ϵ_i must be in X_s or X_t, but not in both (two choices). Assume that ϵ_i is in X_s. Observe that there are at most $\rho-1$ choices for s. Note that $X_jX_lX_s$ must contain a variable $\epsilon_{i'}$ of odd degree. Thus, $\epsilon_{i'}$ must be in X_t and, hence, there are at most $\rho-1$ choices for t.

Therefore, we have

$$\mathbb{E}(X^4) \leq \sum_{j=1}^{m} w_j^4 + (6+4(\rho-1)^2)\sum_{1\leq j<l\leq m} w_j^2w_l^2 < 2\rho^2\left(\sum_{j=1}^{m} w_j^2\right)^2.$$

Thus, by Lemma 6, $\mathbb{E}(X^4) \leq 2\rho^2(\mathbb{E}(X^2))^2$. □

Observe that Lemma 7 and the relation $\epsilon_i = (-1)^{z_i}$, described before Lemma 6 between weighted systems of linear equations on GF(2) and n-variate polynomials with domain $\{-1, 1\}^n$, imply immediately Proposition 1 (essentially Proposition 1 and Lemma 7 are equivalent via the relation).

Now we can prove the following:

Theorem 2. *Let S be reduced under Rule 4. The following three special cases of* LINALB *are fixed-parameter tractable: (1) there is a set U of variables such that each equation contains an odd number of variables from U, (2) there is a constant r such that $r(S) \leq r$, (3) there is a constant ρ, such that any variable appears in at most ρ equations. In each case, there exists a kernel with $O(k^2)$ equations and variables.*

Proof.

Case 1: Due to the relation $\epsilon_i = (-1)^{z_i}$ we may consider X as a random variable depending on random variables $z_1, \ldots, z_n$. Let $z^0 = (z_1^0, \ldots, z_n^0) \in \{0, 1\}^n$ be an assignment of values to the variables $z_1, \ldots, z_n$, and let $-z^0 = (z_1', \ldots, z_n')$, where $z_i' = 1 - z_i^0$ if $z_i \in U$ and $z_i' = z_i^0$, otherwise, $i = 1, \ldots, n$. Observe that $f : z^0 \mapsto -z^0$ is a bijection on the set of assignments and $X(-z^0) = -X(z^0)$. Thus, X is a symmetric random variable. Therefore, by Lemmas 1 and 6, Prob($X \geq \sqrt{m}$) $\geq$ Prob($X \geq \sqrt{\sum_{j=1}^m w_j^2}$) > 0. Hence, if $\sqrt{m} \geq 2k$, the answer to LINALB is YES. Otherwise, $m < 4k^2$ and after applying Rule 3, we obtain a kernel with $O(k^2)$ equations and variables.

Case 2: Since X is a polynomial of degree at most r, it follows by Lemma 3 that $\mathbb{E}(X^4) \leq 9^r \mathbb{E}(X^2)^2$. This inequality and Lemma 6 show that the conditions of Lemma 2 are satisfied and, thus,

$$\text{Prob}\left(X > \frac{\sqrt{\sum_{j=1}^m w_j^2}}{2 \cdot 3^r}\right) > 0, \quad \text{implying} \quad \text{Prob}\left(X > \frac{\sqrt{m}}{2 \cdot 3^r}\right) > 0.$$

Consequently, if $2k - 1 \leq \sqrt{m}/(2 \cdot 3^r)$, then there is an assignment of values to the variables $z_1, \ldots, z_n$ which satisfies equations of total weight at least $W/2 + k$. Otherwise, $2k - 1 > \sqrt{m}/(2 \cdot 3^r)$ and $m < 4(2k-1)^2 9^r$. After applying Rule 3, we obtain the required kernel.

Case 3: If $\rho = 1$, it is easy to find an assignment to the variables that satisfies all equations of S. Thus, we may assume that $\rho \geq 2$. To prove that there exists a kernel with $O(k^2)$ equations, we can proceed as in Case 2, but use Lemma 7 rather than Lemma 3. $\square$

Case 1 of Theorem 2 is of interest since its condition can be checked in polynomial time due to the following:

Proposition 2. *We can check, in polynomial time, whether there exists a set U of variables such that each equation of S contains an odd number of variables from U.*

Proof. Observe that such a set U exists if and only if the unweighted system S' of linear equations over GF(2) obtained from S by replacing each b_j with 1 has a solution. Indeed, if U exists, set $z_j = 1$ for each $z_j \in U$ and $z_j = 0$ for each $z_j \notin U$. This assignment is a solution to S'. If a solution to S' exists, form U by including in it all variables z_j which equal 1 in the solution. We can check whether S' has a solution using the Gaussian elimination or other polynomial-time algorithms, see, e.g., [6]. □

Remark 1 Note that even if S does not satisfy Case 2 of the theorem, T, the system obtained from S using Rule 3, may still satisfy Case 2. However, we have not formulated the theorem for S reduced under Rule 3 as the reduced system depends on the choice of a maximum linear independent collection of columns of A.

5 Discussion

We have shown that the new method allows us to prove that some maximization problems parameterized above tight lower bounds are fixed-parameter tractable. Our method can also be used for minimization problems parameterized below tight upper bounds. As a simple example, consider the feedback arc problem: given a digraph $D = (V, A)$ find a minimum set F of arcs such that $D - F$ is acyclic. Certainly, $|A|/2$ is a tight upper bound on a minimum feedback set and we can consider the parameterized problem which asks whether D has a feedback arc set with at most $|A|/2 - k$ arcs. Fixed-parameter tractability of this parameterized problem follows immediately from fixed-parameter tractability of LOALB, but we could prove this result directly using essentially the same approach as for LOALB.

It would be interesting to obtain applications of our method to other problems parameterized above tight lower bounds or below tight upper bounds. One such application is given by Gutin et al. [13], who solved an open problem due to Benny Chor described in [19]. Another recent application is given by Alon et al. [1], who obtained a quadratic kernel for MAX-r-SAT parameterized above a tight lower bound. This solved another open problem from [18]. The most recent application is given by Gutin et al. [12], who extended the main result of [13] to all ternary permutation constraint satisfaction problems.

Using different approaches, Crowston et al. [7, 8] obtained significant extensions of Cases 2 and 3 of Theorem 2.

References

[1] N. Alon, G. Gutin, E. J. Kim, S. Szeider, and A. Yeo. Solving MAX-r-SAT above a tight lower bound. In *Proc. SODA 2010*, pages 511–517, 2010.

[2] N. Alon, G. Gutin, and M. Krivelevich. Algorithms with large domination ratio. *J. Algorithms*, 50(1):118–131, 2004.

[3] J. Bang-Jensen and G. Gutin. *Digraphs: Theory, Algorithms and Applications*. Springer-Verlag, London, 2nd edition, 2009.

[4] T. S. Blyth and E. F. Robertson, *Basic Linear Algebra*, Springer, 2000.

[5] A. Bonami. Étude des coefficients de Fourier des fonctions de $L^p(G)$. *Ann. Inst. Fourier*, 20(2):335–402, 1970.

[6] D. Coppersmith. Solving linear systems over GF(2): block Lanczos algorithm. *Lin. Algebra Applic.*, 192:33–60, 1993.

[7] R. Crowston, G. Gutin, and M. Jones. Note on Max Lin-2 above average. *Inform. Proc. Lett.* 110:451–454, 2010.

[8] R. Crowston, G. Gutin, M. Jones, E. J. Kim, and I. Ruzsa. Systems of linear equations over $\mathbb{F}_2$ and problems parameterized above average. In *Proc. SWAT 2010*, to appear.

[9] R. G. Downey and M. R. Fellows, *Parameterized Complexity*, Springer, 1999.

[10] J. Flum and M. Grohe. *Parameterized Complexity Theory*, Springer Verlag, 2006.

[11] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97:1061–1083, 1975.

[12] G. Gutin, L. van Iersel, M. Mnich, and A. Yeo. All ternary permutation constraint satisfaction problems parameterized above average have kernels with quadratic numbers of variables. Tech. Report, arXiv:1004.1956.

[13] G. Gutin, E. J. Kim, M. Mnich, and A. Yeo. Betweenness Parameterized Above Tight Lower Bound. *J. Comput. Syst. Sci.*, to appear.

[14] G. Gutin, A. Rafiey, S. Szeider, and A. Yeo. The linear arrangement problem parameterized above guaranteed value. *Theory Comput. Syst.*, 41:521–538, 2007.

[15] G. Gutin, S. Szeider, and A. Yeo. Fixed-parameter complexity of minimum profile problems. *Algorithmica*, 52(2):133–152, 2008.

[16] J. Håstad and S. Venkatesh. On the advantage over a random assignment. *Random Structures Algorithms* 25(2):117–149, 2004.

[17] M. Mahajan and V. Raman. Parameterizing above guaranteed values: MaxSat and MaxCut. *J. Algorithms*, 31(2):335–354, 1999.

[18] M. Mahajan, V. Raman, and S. Sikdar. Parameterizing above or below guaranteed values. *J. Comput. Syst. Sci.*, 75(2):137–153, 2009.

[19] R. Niedermeier. *Invitation to Fixed-Parameter Algorithms*. Oxford Univ. Press, 2006.

[20] R. O'Donnell, Some topics in analysis of Boolean functions. Technical report, ECCC Report TR08-055, 2008. Paper for an invited talk at STOC'08, www.eccc.uni-trier.de/eccc-reports/2008/TR08-055/

[21] Y. Villanger, P. Heggernes, C. Paul, and J. A. Telle. Interval completion is fixed parameter tractable. *SIAM J. Comput.*, 38(5):2007–2020, 2008.

[22] V. Vovk, Private communication, August, 2009.

[23] R. de Wolf, A Brief Introduction to Fourier Analysis on the Boolean Cube, Theory Of Computing Library Graduate Surveys 1 (2008), 1–20, http://theoryofcomputing.org