Quadratizations of symmetric pseudo-Boolean functions: sub-linear bounds on the number of auxiliary variables

Endre Boros*[1], Yves Crama†[2], and Elisabeth Rodríguez-Heck‡[2]

[1]MSIS Department and RUTCOR, Rutgers University, NJ, USA
[2]QuantOM, HEC Management School, University of Liège, Belgium

Abstract

The problem of minimizing a pseudo-Boolean function with no additional constraints arises in a variety of applications. A quadratization is a quadratic reformulation of the nonlinear problem obtained by introducing a set of auxiliary binary variables which can be optimized using quadratic optimization techniques. Using the well-known result that a pseudo-Boolean function can be uniquely expressed as a multilinear polynomial, any pseudo-Boolean function can be quadratized by providing a quadratization for negative monomials and one for positive monomials. A desirable property for a quadratization is to introduce small number of auxiliary variables. For the case of negative monomials there exists a quadratization using a single auxiliary variable which is, in addition, submodular. However, much less is known for positive monomials. The best lower bound on the number of variables required by a quadratization in the literature is $\lfloor \frac{n-1}{2} \rfloor$. We present here new quadratizations of the positive monomial that significantly improve this lower bound decreasing it to a logarithmic bound. This lower bound is derived from a more general result, stating that a quadratization with a logarithmic number of variables can be defined for exact k-out-of-n functions, which is a more general class of symmetric functions. Moreover, for exact k-out-of-n functions we also prove that a logarithmic number of variables is necessary when requiring certain symmetry conditions to the quadratization. Finally, we provide quadratizations for general symmetric functions using $O(\sqrt{n})$ auxiliary variables. This upper bound nicely matches a lower bound of $\Omega(\sqrt{n})$ variables that was recently introduced.

Introduction

A pseudo-Boolean function is a mapping $f : \{0,1\}^n \to \mathbb{R}$ that assigns a real value to each tuple of n binary variables $(x_1, \ldots, x_n)$. Pseudo-Boolean functions have been extensively used and studied during the last century and especially in the last 50 years, given that they model problems in a wide range of areas such as reliability theory, computer science, statistics, economics, finance, operations research, management science, discrete mathematics, or computer vision (see (Boros and Hammer 2002) and (Crama and Hammer 2011) for a list of applications and references). In most of these applications f has to be optimized, therefore we are interested in the problem

$$\min_{x \in \{0,1\}^n} f(x), \tag{1}$$

which is $\mathcal{NP}$-hard even when f is a quadratic function.

Several techniques have been proposed to solve problem (1), such as enumerative methods, algebraic methods, linear reformulations and quadratic reformulations, which are then solved using a linear or quadratic solver, respectively. It is not clear whether one of the previous techniques is generally better than the others. In fact, the performance of the different approaches seems to depend on the underlying structure of the problem, among other factors.

In this paper, we focus on quadratic reformulations of problem (1), which are also called quadratizations. Interestingly, much progress in the understanding of quadratizations has been made in the field of computer vision, where these type of techniques perform especially well for problems such as image restoration. A systematic study of quadratizations and its properties has been initiated in (Anthony et al. 2017), where a quadratization is defined as follows.

Definition 1. *Given a pseudo-Boolean function $f(x)$ on $\{0,1\}^n$, we say that $g(x, y)$ is a* quadratization *of f if $g(x, y)$ is a quadratic polynomial depending on x and on m* auxiliary variables *$y_1, \ldots, y_m$, such that*

$$f(x) = \min_{y \in \{0,1\}^m} g(x, y) \quad \forall x \in \{0,1\}^n. \tag{2}$$

It is clear that given a pseudo-Boolean function f and a quadratization g, minimizing f over $x \in \{0,1\}^n$ is equivalent to minimizing g over the $(x, y) \in \{0,1\}^{n+m}$. Moreover, it is known that every pseudo-Boolean function f admits a quadratization (Rosenberg 1975).

Not all quadratizations will perform equally well when solving the resulting quadratic problem. A desirable property of a quadratization is to have a small set of auxiliary variables, so that the size of the reformulation does not increase too much with respect to the size of the original problem. Another interesting property is submodularity. Of course, if the original pseudo-Boolean function f is not submodular, then it does not admit a submodular quadratization. However, even if a quadratization is not submodular, a small number of positive quadratic terms might have a positive impact in resolution performance.

*endre.boros@rutgers.edu
†yves.crama@ulg.ac.be
‡elisabeth.rodriguezheck@ulg.ac.be

Literature review

Given a pseudo-Boolean function f, there are many different methods to define a quadratization. In particular, *termwise quadratizations* have attracted much interest in the literature. This type of procedure uses the well-known result that a pseudo-Boolean function f can be represented by a unique multilinear polynomial (Hammer, Rosenberg, and Rudeanu 1963), (Hammer and Rudeanu 1968), i.e.,

$$f(x_1,\dots,x_n)=\sum_{S\in 2^{[n]}} a_S \prod_{i\in S} x_i, \tag{3}$$

where $[n]=\{1,\dots,n\}$ and $2^{[n]}$ is the sets of subsets of $[n]$. A termwise quadratizations of f can be defined by providing a quadratization for each term, therefore it is necessary to understand quadratizations of positive monomials ($a_S>0$) and of negative monomials ($a_S<0$).

The case of negative monomials is well understood. A simple expression to quadratize cubic negative monomials was introduced in (Kolmogorov and Zabih 2004). This expression was later extended to higher degrees in (Freedman and Drineas 2005). Their quadratization for a degree n monomial $N_n(x)=-\prod_{i=1}^n x_i$ is

$$s_n(x,y)=(n-1)y-\sum_{i=1}^n x_i y. \tag{4}$$

This quadratization is submodular because all quadratic terms have negative coefficients.

Surprisingly, much less is known for positive monomials. Until now, the quadratization introducing the smallest number of artificial variables was defined in (Ishikawa 2011). Consider a positive monomial $P_n(x)=\prod_{i=1}^n x_i$. Then,

$$\prod_{i=1}^n x_i=\min_{y_1,\dots y_m\in\{0,1\}}\sum_{i=1}^m y_i(c_{i,n}(-S_1+2i)-1)+aS_2, \tag{5}$$

where S_1,S_2 are the elementary linear and quadratic symmetric polynomials in n variables,

$$S_1=\sum_{i=1}^n x_i,\quad S_2=\sum_{i=1}^{n-1}\sum_{j=i+1}^n x_i x_j=\frac{S_1(S_1-1)}{2},$$

$m=\lfloor\frac{n-1}{2}\rfloor$ and $c_{i,n}=\begin{cases}1, \text{if } n \text{ is odd and } i=m,\\ 2, \text{otherwise.}\end{cases}$ Quadratization (5) introduces $\binom{n}{2}$ positive quadratic terms. Despite this fact, this quadratization performs very well computationally (Ishikawa 2009).

In this paper we define several quadratizations that improve the number of artificial variables used in quadratization (5); one uses $\lceil\frac{n}{4}\rceil$ auxiliary variables, and the other two only use a logarithmic number on n.

A more general class of pseudo-Boolean functions are *symmetric* functions, which are defined as follows.

Definition 2. *A pseudo-Boolean function $f:\{0,1\}^n\to\mathbb{R}$ is* symmetric *if there is a discrete function $r:\{0,1,\dots,n\}\to\mathbb{R}$ such that $f(x)=r(X)$, where $X=\sum_{i=1}^n x_i$ is the Hamming weight (number of ones) in x.*

In the next section we define several quadratizations for a special class of symmetric functions, namely, exact k-out-of-n functions, which use a logarithmic number of auxiliary variables.

An upper bound of $n-2$ and a lower bound of $\Omega(\sqrt{n})$ on the minimum number of variables required to quadratize a symmetric function was recently established in (Anthony et al. 2016). In the last section we define a quadratization using $O(\sqrt{n})$ auxiliary variables, thus providing a matching upper bound.

Quadratizations of exact k-out-of-n functions

This section is concerned with quadratizations of exact k-out-of-n functions. We provide a quadratization using a logarithmic number of variables in n and we prove that, when requiring certain symmetry conditions to the quadratization, then it is also necessary to use a logarithmic number of auxiliary variables.

An exact k-out-of-n function is defined as follows. Consider n binary variables, x_i, $i=1,...,n$, set $X=\sum_{i=1}^n x_i$, and let $0\le k\le n$ be an integer. Let us consider a symmetric function of the form

$$f_k(x_1,...,x_n)=\begin{cases}1 & \text{if } X=k\\ 0 & \text{otherwise.}\end{cases}$$

We are concerned with the quadratizations of f_k. Notice that the positive monomial is an exact k-out-of-n function with $k=n$.

Observation 1. *Let $\bar{x}_i=1-x_i$, for all i. Then $f_k(x_1,\dots,x_n)=f_{n-k}(\bar{x}_1,\dots,\bar{x}_n)$.*

We define a quadratization G_k of f_k as follows. Assume first that $k\ge 2$. Let $l=\lceil\log k\rceil$, consider variables $z\in\{0,1\}$ and $y=(y_0,...,y_{l-1})\in\{0,1\}^l$ and define

$$A_k(X,y,z)=X-(k-2^l)z-(k+1)(1-z)-\sum_{i=0}^{l-1}2^i y_i.$$

With these notation define the quadratic function

$$G_k(x,y,z)=\begin{cases}A_k(X,y,z)^2 & \text{if } \frac{n}{2}\le k\le n,\\ A_{n-k}(n-X,y,z)^2 & \text{if } 0\le k\le\frac{n}{2}.\end{cases}$$

Theorem 1. *$G_k(x,y,z)$ is a quadratization of f_k with $K+1\le 1+\lceil\log n\rceil$ auxiliary variables, where $K=\max(\lceil\log(k)\rceil,\lceil\log(n-k)\rceil)$.*

Proof. Clearly, $G_k(x,y,z)\ge 0$ for all (x,y,z) and for all k.

Assume that $\frac{n}{2}\le k\le n$. We show that, for all x with $X\ne k$, there exists (y,z) such that $G_k(x,y,z)=f_k(x)=0$.

1. If $0\le X<k$, set $z=1$ so that $A_k(X,y,1)=X-k+2^l-\sum_{i=0}^{l-1}2^i y_i$. Note that $0\le X-k+2^l\le 2^l-1$ (the first inequality holds because $k\le 2^l$, and the second inequality follows from $X<k$). Hence, one can choose y_i, for $i=0,\dots,l-1$, in such a way that $X-k+2^l=\sum_{i=1}^l 2^{i-1}y_i$, and $G_k(x,y,1)=A_k(X,y,1)=0$.

2. If $k < X \le n$, set $z = 0$ so that $A_k(X, y, 0) = X - k - 1 - \sum_{i=0}^{l-1} 2^i y_i$. Now, $0 \le X - k - 1 \le 2^l - 1$ (since $k < X$, and since $X - k \le n - k \le k \le 2^l$ by definition of l and because $k \ge \frac{n}{2}$). Hence, one can choose y_i, for $i = 0, \ldots, l - 1$, in such a way that $X - k - 1 = \sum_{i=0}^{l-1} 2^i y_i$, and $G_k(x, y, 0) = A_k(X, y, 0) = 0$.

Consider now the case where $X = k$. When $z = 1$, we obtain $A_k(k, y, 1) = 2^l - \sum_{i=0}^{l-1} 2^i y_i \ge 1$, and hence $G_k(x, y, z) \ge 1$. When $z = 0$, $A_k(k, y, 0) = -1 - \sum_{i=0}^{l-1} 2^i y_i \le -1$, and hence again $G_k(x, y, z) \ge 1$. The minimum value $G_k(x, y, z) = 1$ is obtained by setting either $z = y_i = 1$ for $i = 0, \ldots, l - 1$ or $z = y_i = 0$ for $0 = 1, \ldots, l - 1$.

Assume now that $0 \le k \le \frac{n}{2}$. Observation 1 and the analysis of the previous case imply that $A_{n-k}(n - X, y, z)^2$ is a quadratization of f_k. □

Corollary 1. *If $f : 0, 1^n \mapsto \mathbb{R}$ is a symmetric function the value of which is strictly above its minimum value for at most μ different values of X, then it can be quadratized with at most $\mu(1 + \lceil \log n \rceil)$ auxiliary variables.*

Consider now a slightly modified version of quadratization G_k, defined as follows. Let $l = \lceil \log k \rceil$, consider variables $z \in \{0, 1\}$ and $y = (y_1, ..., y_{l-1}) \in \{0, 1\}^{l-1}$ and define

$$A'_k(X, y, z) = X - (k - 2^l)z - (k + 1)(1 - z) - \sum_{i=1}^{l-1} 2^i y_i.$$

Let

$$G'_k(x, y, z) = \begin{cases} \binom{A'_k(X,y,z)}{2} & \text{if } \frac{n}{2} \le k \le n, \\ \binom{A'_{n-k}(n-X,y,z)}{2} & \text{if } 0 \le k \le \frac{n}{2}. \end{cases}$$

Theorem 2. *$G'_k(x, y, z)$ is a quadratization of f_k with $K + 1 \le \lceil \log n \rceil$ auxiliary variables, where $K = \max(\lceil \log(k) \rceil, \lceil \log(n - k) \rceil) - 1$.*

Proof. Note first that $G'_k(x, y, z) \ge 0$ for all for all (x, y, z) and k, since it is the half-product of two consecutive integers.

Assume that $\frac{n}{2} \le k \le n$. We show first that, for $X \ne k$, there exists (y, z) such that $G_k(x, y, z) = f_k(x) = 0$.

1. If $0 \le X < k$, set $z = 1$ so that $A'_k(X, y, 1) = X - k + 2^l - \sum_{i=1}^{l-1} 2^i y_i$. As before, $0 \le X - k + 2^l \le 2^l - 1$. Hence, depending on the parity of $X - k + 2^l$, one can choose y_i for $i = 1, \ldots, l-1$, in such a way that either $A'_k(X, y, z) = 0$ or $A'_k(X, y, z) - 1 = 0$, and hence $G'_k(x, y, z) = 0$.
2. If $k < X \le n$, set $z = 0$ so that $A'_k(X, y, 0) = X - k - 1 - \sum_{i=1}^{l-1} 2^i y_i$. As before, $0 \le X - k - 1 \le 2^l - 1$. Hence, depending on the parity of $X - k - 1$, one can choose y_i, for $i = 1, \ldots, l-1$, in such a way that either $A'_k(X, y, z) = 0$ or $A'_k(X, y, z) - 1 = 0$, and hence $G'_k(x, y, z) = 0$.

Consider finally the case where $X = k$. When $z = 1$, we obtain $A'_k(k, y, 1) = 2^l - \sum_{i=1}^{l-1} 2^i y_i \ge 2$, and hence $G'_k(x, y, z) \ge 1$. When $z = 0$, $A'_k(k, y, 0) = -1 - \sum_{i=1}^{l-1} \le -1$, and hence again $G'_k(x, y, z) \ge 1$. The minimum value $G'_k(x, y, z) = 1$ is obtained by setting either $z = y_i = 1$ for $i = 1, \ldots, l - 1$ or $z = y_i = 0$, for $i = 1, \ldots, l - 1$.

Assume now that $0 \le k \le \frac{n}{2}$. As before, by Observation 1 and by the analysis of the previous case, we have that $\frac{1}{2} A'_{n-k}(n - X, y, z)(A'_{n-k}(n - X, y, z) - 1)$ is a quadratization of f_k. □

Theorem 2 defines a quadratization using one less variable than the one of Theorem 1, but Theorem 1 has the advantage of unveiling the nature of the construction in a more transparent way.

The remainder of this section is devoted to proving a theorem giving a logarithmic lower bound on the minimal number of variables that are required to quadratize exact k-out-of-n functions, when requiring certain symmetry conditions on the quadratization.

Assume that $g(x, y)$ is a pseudo-Boolean function in $n + m$ variables $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$.

Definition 3. *A pseudo-Boolean function $g(x, y)$ is x-*symmetric *if for all $y^* \in \{0, 1\}^m$ the function $g(x, y^*)$ is a symmetric function of its variables $x_1, \ldots, x_n$, or equivalently, if there exists a function G such that $g(x, y) = G(X, y)$.*

Lemma 1. *If $g(x, y)$ is x-symmetric, then there exists a function G such that $g(x, y) = G(X, y)$, where G is a multilinear polynomial in the y variables and whose degree is the same as the degree of g.*

Proof. Let us introduce $[n] = \{1, 2, \ldots, n\}$ and write

$$g(x, y) = \sum_{S \subseteq [n]} \psi_S(y) \cdot \prod_{i \in S} x_i.$$

Since for any $y \in \{0, 1\}^m$ we have that $g(x, y)$ is symmetric in the x_i variables, we must have that $\psi_S(y) = \psi_T(y)$ for any two subsets $S, T \subseteq [n]$ with $|S| = |T|$. It follows that there exist pseudo-Boolean functions $\phi_k(y)$ for $k = 0, \ldots, n$ such that

$$g(x, y) = \sum_{k=0}^{n} \phi_k(y) \cdot \sum_{\substack{S \subseteq [n] \\ |S| = k}} \prod_{i \in S} x_i.$$

Each subexpression $T_k(x) = \sum_{\substack{S \subseteq [n] \\ |S| = k}} \prod_{i \in S} x_i$ defines an x-symmetric function of degree k, and since the expression $T_k(x)$ counts the number of subsets of size k of the set of x_i variables taking value one, $T_k(x)$ can be written as

$$T_k(x) = \binom{X}{k} = \frac{1}{k!} \prod_{j=0}^{k-1} (X - j). \tag{6}$$

Expression (6) is a polynomial of degree k in X, which can be substituted for T_k in the expression of $g(x, y)$. This completes the proof.

□

Definition 4. *A pseudo-Boolean function $g(x, y)$ is x-*linear *if every term in its unique multilinear expression contains at most one of the x_i variables.*

Observe that every pseudo-Boolean function $f(x)$ has an x-linear quadratization. Namely, let $q(x, y)$ be an arbitrary quadratization of $f(x)$. Then

$$M \sum_{i=1}^{n} (x_i \bar{x}'_i + \bar{x}_i x'_i) + q(x', y)$$

is an x-linear quadratization of $f(x)$ for a large enough M.

Now, let $Z = \{0, 1, \ldots, n\}$.

Definition 5. *A function* $r : Z \to \mathbb{Z}$ *is* concave *if for all* $0 < j < n$ *we have that* $r(j) \geq \frac{1}{2}r(j-1) + \frac{1}{2}r(j+1)$.

In the next lemma we show that if $f(x) = r(X)$ is a nonconcave symmetric function, then it does not have an x-linear x-symmetric quadratization.

Lemma 2. *If* $f(x) = r(X)$ *is a nonconcave symmetric function on* $n \geq 3$ *variables and* $g(x, y)$ *is an* x*-symmetric quadratization of* f*, then* g *is not* x*-linear.*

Proof. By Lemma 1 there exists a quadratic polynomial $G(X, y)$ that is multilinear in y such that $g(x, y) = G(X, y)$. G can be written as

$$G(X, y) = \alpha X^2 + X\left(\beta + \sum_{j=1}^{m} \gamma_j y_j\right) + \tag{7}$$
$$+ \left[\sum_{1 \leq i < j \leq m} \delta_{ij} y_i y_j + \sum_{j=1}^{m} \epsilon_j y_j + \phi\right]$$

The statement is equivalent to saying that $\alpha \neq 0$. Let us assume for a contradiction that $\alpha = 0$.

For each $X^* = 0, \ldots, n$ let us denote by $F_j(X^*)$ the minimizing binary values of the y_j, for $j = 1, \ldots, m$ in problem $min_{y \in \{0,1\}^m} G(X^*, y)$. Let us further introduce $a(X^*) = \beta + \sum_{j=1}^{m} \gamma_j F_j(X^*)$, and $b(X^*) = \sum_{1 \leq i < j \leq m} \delta_{ij} F_i(X^*) F_j(X^*) + \sum_{j=1}^{m} \epsilon_j F_j(X^*) + \phi$.

With these notations we have that there is a vector $y(X^*) = F(X^*)$, and two corresponding values $a(X^*)$, $b(X^*)$ such that

$$r(X^*) = \min_{X^{*'} = 0, \ldots, n} \{X^* \cdot a(X^{*'}) + b(X^{*'})\}.$$

Since $r(X^*)$ is the minimum of a finite number (2^m) of affine functions, it must be concave, which is a contradiction. □

Theorem 3. *If* $g(x, y) = G(X, y)$ *is an* x*-symmetric quadratization of* $f_k(x)$ *using* m *auxiliary variables, then* $m + 1 \geq \log n$.

Proof. Define a set $Z = \{0, 1, \ldots, n\}$, and a function $r : Z \to \{0, 1\}$ by $r(X) = f_k(x)$ for all $x \in \{0, 1\}^n$.

Since G is a quadratic polynomial and y_j, $j = 1, \ldots, m$ are binary variables, we can write

$$G(X, y) = \alpha X^2 + X\left(\beta + \sum_{j=1}^{m} \gamma_j y_j\right) + \tag{8}$$
$$+ \left[\sum_{1 \leq i < j \leq m} \delta_{ij} y_i y_j + \sum_{j=1}^{m} \epsilon_j y_j + \phi\right],$$

for some reals $\alpha, \beta, \gamma_j, \delta_{ij}, \epsilon_j$ and ϕ.

Since g is a quadratization of $f_k(x)$, we have that $f_k(x) = \min_{y \in \{0,1\}^m} g(x, y)$, for all $x \in \{0, 1\}^n$. For every x and every index $1 \leq j \leq m$, let us denote $F_j(x)$ a minimizing value of the variable y_j in the previous equality. Clearly, F_j is a Boolean function over $\{0, 1\}^n$ for every $j = 1, \ldots, m$. Furthermore, since $f_k(x) = \min_{y \in \{0,1\}^m} g(x, y) = \min_{y \in \{0,1\}^m} G(X, y)$, these Boolean functions can be chosen to be symmetric themselves. Consequently, it will be not misleading to use $F_j(X)$ to denote such a minimizing binary value.

Thus, we have the following equalities

$$r(X) = f_k(x) = \min_{y \in \{0,1\}^m} g(x, y) =$$
$$= \min_{y \in \{0,1\}^m} G(X, y) = G(X, F_1(X), \ldots, F_m(X)).$$

Therefore, by (8) we obtain the following system of equations

$$r(X) = \alpha X^2 + X\left(\beta + \sum_{j=1}^{m} \gamma_j F_j(X)\right) + \tag{9}$$
$$+ \left[\sum_{1 \leq i < j \leq m} \delta_{ij} F_i(X) F_j(X) + \sum_{j=1}^{m} \epsilon_j F_j(X) + \phi\right].$$

This equality implies that $r(X)$ can be viewed as a quadratic expression of X. Let us focus now on the three coefficients of this expression.

For a binary vector $\omega \in \{0, 1\}^m$ let us define $Z(\omega) = \{z \in Z \mid \omega = (F_1(x), \ldots, F_m(z))\}$, and observe that the three coefficients in (9) are the same for all $X \in Z(\omega)$. Namely, we have for all $X \in Z(\omega)$ that

$$r(X) = \alpha X^2 + X\left(\beta + \sum_{j=1}^{m} \gamma_j \omega_j\right) +$$
$$\left[\sum_{1 \leq i < j \leq m} \delta_{ij} \omega_i \omega_j + \sum_{j=1}^{m} \epsilon_j \omega_j + \phi\right].$$

Since for $n \geq 3$ function f_k is a nonconcave symmetric function, Lemma 2 implies that its x-symmetric quadratization is not x-linear, i.e., $\alpha \neq 0$. Since for a particular nonlinear quadratic function over $Z(\omega)$, we cannot have three pairwise different values $z, z', z'' \in Z(\omega)$ with $r(z) = r(z') = r(z'')$, and since this is true for all $\omega \in \{0, 1\}^m$, and r takes the same value zero for n different elements of Z, we can conclude that

$$2 \cdot 2^m \geq n,$$

which completes the proof. □

Theorems 1, 2 and 3 provide tight upper and lower bounds for the minimum number of auxiliary variables when requiring a quadratization of an exact k-out-of-n function to be x-symmetric. An important open question is to provide a lower bound for general non-symmetric quadratizations.

Quadratizations of the positive monomial

In this section we first write Theorems 1 and 2 for the positive monomial, which is an exact k-out-of-n function, with $k = n$.

Theorem 4. *Assume that* $n \leq 2^k$ *and et* $K = 2^k - n$*. Then,*

$$g(x, y) = \left(K + X - \sum_{i=0}^{k-1} 2^i y_i\right)^2$$

is a quadratization of the positive monomial P_n *using* $\lceil \log n \rceil$ *auxiliary variables.*

Theorem 5. *Assume that $n \leq 2^{k+1}$. Let $K = 2^{k+1} - n$. Then,*

$$g(x,y) = \frac{1}{2}(K + X - \sum_{i=1}^{k} 2^i y_i)(K + X - \sum_{i=1}^{k} 2^i y_i - 1) \quad (10)$$

is a quadratization of the positive monomial $f(x) = P_n(x) = \prod_{i=1}^n x_i$ using $\lceil \log n \rceil - 1$ auxiliary variables.

Theorems 4 and 5 are direct consequences of 1 and 2, respectively, when taking $k = n$ and fixing $z = 1$.

The quadratizations presented in Theorems 4 and 5 provide a significant improvement, of orders of magnitude, with respect to the linear bounds on the minimum number of auxiliary variables given by (5). Moreover, to the best of our knowledge, (10) is the quadratization of the positive monomial introducing the smallest number of auxiliary variables.

In the remainder of this section, we present a quadratization of the positive monomial using $\lceil \frac{n}{4} \rceil$ auxiliary variables. This quadratization uses a linear number of auxiliary variables, but we find it worth to present this result here due to computational reasons. Indeed, the quadratizations in Theorems 4 and 5 use many large coefficients, which might have a negative impact on computational performance even though they introduce a smaller number of auxiliary variables than quadratization (5) or the quadratization that we present below. Note also that $\lceil \log n \rceil - 1$ is equal to $\lceil \frac{n}{4} \rceil$ when $3 \leq n \leq 12$, so that the difference in the number of auxiliary variables only becomes relevant for very high degrees. Another factor that might affect computational performance is having a different number of positive quadratic terms. These are open computational questions that should be tested.

Theorem 6. *The positive monomial $P_n = \prod_{i=1}^n x_i$ has a quadratization using $m = \lceil \frac{n}{4} \rceil$ auxiliary variables.*

We will actually prove more. Consider X as in Definition 2 and let us define the following notation: for a vector of auxiliary variables $(y_1, \ldots, y_m)$, let $Y = \sum_{j=2}^m y_j$. Note that the sum defining Y starts with $j = 2$, so that $Y \leq m - 1$.

Theorem 7. *For all integers n, m, if $n \geq 2$, $\frac{n}{4} \leq m \leq \frac{n}{2}$, and $N = n - 2m$ then*

$$g(x,y) = \frac{1}{2}(X - Ny_1 - 2Y)(X - Ny_1 - 2Y - 1) \quad (11)$$

is a quadratization of the positive monomial $P_n = \prod_{i=1}^n x_i$ using m auxiliary variables.

Proof. Clearly, $g(x,y)$ is a quadratic function. Moreover, the right-hand side of (11) is one half of the product of two consecutive integers. Thus, $g(x,y)$ is nonnegative and is integral for all (x,y).

Consider now any vector $x \in \{0,1\}^n$ and let X be its Hamming weight. We must show that $P_n(x) = \min_y g(x,y)$.

1. Assume first that $X \leq 2m-1$. In particular, $X \leq n-1$, and hence $P_n(x) = 0 \leq g(x,y)$ for all y. To obtain a minimizer of $g(x,y)$, let $y_1 = 0$ and set the first $\lfloor \frac{X}{2} \rfloor$ components of $(y_2, \ldots, y_m)$ to 1. Note that this is always possible, since $\frac{X}{2} \leq m - \frac{1}{2}$, and hence $\lfloor \frac{X}{2} \rfloor \leq m - 1$. The Hamming weight of $(y_2, \ldots, y_m)$ is $Y = \lfloor \frac{X}{2} \rfloor$, that is, either $2Y = X$ or $2Y = X - 1$, depending on the parity of X. It immediately follows from definition (11) that $g(x,y) = 0$, as required.
2. Assume next that $2m \leq X \leq n-1$. Here again, $P_n(x) = 0 \leq g(x,y)$ for all y. In order to describe a minimizer of $g(x,y)$, let us first observe that $0 \leq 4m - n \leq X - N \leq 2m - 1$, and hence $0 \leq \lfloor \frac{X-N}{2} \rfloor \leq m - 1$. Thus, we can define y by setting $y_1^* = 1$, and the first $\lfloor \frac{X-N}{2} \rfloor$ components of $(y_2, \ldots, y_m)$ to 1. It follows that the Hamming weight of $(y_2, \ldots, y_m)$ satisfies either $2Y = X - N$ or $2Y = X - N - 1$, and in view of (11), $g(x,y) = 0$, as required.
3. Finally, assume that $X = n$, meaning that x is the all-one vector and $P_n(x) = 1$. For all $y \in \{0,1\}^m$, (11) becomes

$$g(x,y) = \frac{1}{2}(n - Ny_1 - 2Y)(n - Ny_1 - 2Y - 1). \quad (12)$$

Moreover, the following inequalities hold:

$$2Y \leq 2m - 2 = n - N - 2 \leq n - Ny_1 - 2,$$

which immediately implies that $g(x,y) \geq 1$ for all $y \in \{0,1\}^m$. Let now $y_j = 1$ for all $j = 1, \ldots, m$. From (12), we get $g(x,y) = 1$, and this completes the proof.

□

Remark 1. *As a side-remark, note that when $n = 2m$, $g(x,y)$ does not depend on y_1. So, $g(x,y)$ defines a quadratization of P_n using $m - 1 = \frac{n}{2} - 1$ auxiliary variables, as in* (5).

Remark 2. *Let us show that when, $m < \lceil \frac{n}{4} \rceil$, then g is not a quadratization of P_n. Choose a vector x^* with Hamming weight $X^* = 2m$. So, $P_n(x^*) = 0$ and*

$$g(x^*,y) = \frac{1}{2}(2m - Ny_1 - 2Y)(2m - Ny_1 - 2Y - 1).$$

If $y_1 = 0$, then $g(x^*,y) = \frac{1}{2}(2m - 2Y)(2m - 2Y - 1) > 0$, since $Y \leq m - 1$. If $y_1 = 1$, then

$$g(x^*,y) = \frac{1}{2}(2m - N - 2Y)(2m - N - 2Y - 1)$$
$$= \frac{1}{2}(4m - n - 2Y)(4m - n - 2Y - 1),$$

and again $g(x^*,y) > 0$ since $4m < n$. So, $\min_y g(x^*,y) > 0$, and g is not a quadratization.

A tight upper bound for quadratizations of symmetric functions

In this section we provide an upper bound of $O(\sqrt{n})$ auxiliary variables to quadratize symmetric functions that nicely matches the lower bound that was recently introduced in (Anthony et al. 2016).

Theorem 8. *There exist symmetric functions of n variables for which any quadratization must involve at least $\Omega(\sqrt{n})$ auxiliary variables. (Anthony et al. 2016)*

As in the previous section let $Z = \{0, 1, \ldots, n\}$ and $X = \sum_{i=1}^n x_i$. Now, let $r : Z \to \mathbb{Z}_+$ be a nonnegative function, and consider $f(x_1, \ldots, x_n) = r(X)$. Let $l = \lceil \sqrt{n+1} \rceil$, and choose a large integer M such that $M > r(k)$ for all $k \in Z$.

Let us consider auxiliary variables y_i, $i = 0, \ldots, l-1$ and z_j, $j = 0, \ldots, l-1$, and define

$$\begin{aligned} g(x,y,z) &= \sum_{i=0}^{l-1}\sum_{j=0}^{l-1} r(i \cdot l + j) \cdot y_i \cdot z_j \\ &+ M\left(1 - \sum_{i=0}^{l-1} y_i\right)^2 + M\left(1 - \sum_{j=0}^{l-1} z_j\right)^2 \\ &+ M\left(X - \left(l\sum_{i=0}^{l-1} i \cdot y_i + \sum_{j=0}^{l-1} j \cdot z_j\right)\right)^2 \end{aligned}$$

Theorem 9. *Function g is a quadratization of f with $2\lceil \sqrt{n+1}\rceil = O(\sqrt{n})$ auxiliary variables.*

Proof. Observe first that every integer $k \in Z$ has a unique representation $k = i \cdot l + j$ with $0 \leq i, j \leq l-1$. Since for every $x \in \{0,1\}^n$ we have $X \in Z$, let us define integers $i(x)$ and $j(x)$ such that $X = i(x) \cdot l + j(x)$, $0 \leq i(x) \leq l-1$ and $0 \leq j(x) \leq l-1$ hold. Note that every binary vector $x \in \{0,1\}^n$ defines uniquely such integers.

Let us then define auxiliary vectors $y^*, z^* \in \{0,1\}^l$ (we start indexing by 0), such that

$$y_i^* = \begin{cases} 1 \text{ if } i = i(x), \\ 0 \text{ otherwise,} \end{cases}$$

$$z_j^* = \begin{cases} 1 \text{ if } j = j(x), \\ 0 \text{ otherwise.} \end{cases}$$

Let us observe next that due to the three terms involving M in the definition of g, we have $g(x,y,z) < M$ if and only if $y = y^*$ and $z = z^*$. Due to the definition of the first part of g, in this case we have $g(x, y^*, z^*) = r(X) = f(x)$. □

Consider now a generalization of symmetric functions, where the value of the function f for a given x does not depend on X but on a weighted sum of the values of the components. More precisely, assume that l, n and $R < l^2$ are nonnegative integers, and $Z = \{0, 1, ..., R\}$. Assume also that $L : \{0,1\}^n \mapsto Z$ is a linear function, and that the pseudo-Boolean function $f : \{0,1\}^n \mapsto \mathbb{Z}$ is defined by $f(x) = r(L(x)) \geq 0$, where $r : Z \mapsto \mathbb{R}_+$ satisfies $r(0) = 0$.

We are concerned with the description of a quadratization of f with $O(l) = O(\sqrt{R})$ auxiliary variables.

Let us as consider binary variables, y_j, z_j, $j = 1, ...\ell - 1$, and define

$$Y = l\left(\sum_{j=1}^{l-1} y_j\right) + \sum_{j=1}^{l-1} z_j,$$

and set

$$X = L(x).$$

Furthermore, define $M = \max_{x \in \{0,1\}^n} f(x)$, and define reals $a_{i,j}$, $i = 1, ..., l-1$ and $j = 1, ..., l-1$ such that

$$r(\alpha \cdot l + \beta) = \sum_{i=1}^{\alpha}\sum_{j=1}^{\beta} a_{i,j} \tag{13}$$

holds for all $0 \leq \alpha < l$ and $0 \leq \beta < l$.

Theorem 10. *The function*

$$\begin{aligned} g(x,y,z) &= \sum_{i=1}^{l-1}\sum_{j=1}^{l-1} a_{i,j} \cdot y_i \cdot z_j \\ &+ M + M \cdot (X - Y - 1) \cdot (X - Y + 1) \\ &+ M \cdot \sum_{i=1}^{l-2} (1 - y_i) \cdot y_{i+1} \\ &+ M \cdot \sum_{j=1}^{l-2} (1 - z_j) \cdot z_{j+1} \end{aligned}$$

is a quadratization of f with $m = 2 \cdot l$ auxiliary variables.

Proof. The last two terms make sure that at the (y,z)-minimizing solution we have $y = (1..., 1, 0, ...0)$ and $z = (1, ..., 1, 0, ..., 0)$. The second term makes sure that $L(x) = l \cdot \sum_{i=1}^{l-1} y_i + \sum_{j=1}^{l-1} z_j$. Finally the first term guarantees that $f(x) = \min_{y,z} g(x,y,z)$, by (13). □

Conclusion

In this paper we have introduced new quadratizations for several types of symmetric pseudo-Boolean functions. For exact k-out-of-n functions, the quadratizations use a logarithmic number of auxiliary variables. Positive monomials are special cases of exact k-out-of-n functions. The quadratizations that we provide significantly improve the best known bound (roughly $\frac{n}{2}$) for the smallest number of auxiliary variables required to quadratize a positive monomial. A tight logarithmic lower bound is also provided for the case where we require certain symmetry properties to the quadratization. However, an important open question that remains is whether the lower bound of a general not necessarily symmetric quadratization is also logarithmic.

However, the quadratizations using a logarithmic number of variables presented here introduce many large coefficients, which might have a negative impact on their computational performance. A quadratization for the positive monomial using roughly $\frac{n}{4}$ variables is also defined, which does not introduce such large coefficients and still reduces the best known lower bound. All these quadratizations should be tested computationally, since it is not clear which one will be more beneficial, especially for polynomials of low degree.

Finally, we have also defined a quadratization that uses $O(\sqrt{n})$ auxiliary variables for symmetric functions and for a generalization of these, which nicely matches a recently published lower bound.

References

Anthony, M.; Boros, E.; Crama, Y.; and Gruber, A. 2016. Quadratization of symmetric pseudo-boolean functions. *Discrete Applied Mathematics* 203:1 – 12.

Anthony, M.; Boros, E.; Crama, Y.; and Gruber, A. 2017. Quadratic reformulations of nonlinear binary optimization problems. *Mathematical Programming* 162(1-2):115–144.

Boros, E., and Hammer, P. L. 2002. Pseudo-boolean optimization. *Discrete applied mathematics* 123(1):155–225.

Crama, Y., and Hammer, P. L. 2011. *Boolean functions: Theory, algorithms, and applications*. New York, N. Y.: Cambridge University Press.

Freedman, D., and Drineas, P. 2005. Energy minimization via graph cuts: settling what is possible. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, 939–946.

Hammer, P. L., and Rudeanu, S. 1968. *Boolean methods in operations research and related areas*. Ökonometrie und Unternehmensforschung. Berlin: Springer-Verlag.

Hammer, P. L.; Rosenberg, I.; and Rudeanu, S. 1963. On the determination of the minima of pseudo-boolean functions. *Studii si Cercetari Matematice* 14:359–364. in Romanian.

Ishikawa, H. 2009. Higher-order gradient descent by fusion-move graph cut. In *Computer Vision, 2009 IEEE 12th International Conference on*, 568–574.

Ishikawa, H. 2011. Transformation of general binary mrf minimization to the first-order case. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 33(6):1234–1249.

Kolmogorov, V., and Zabih, R. 2004. What energy functions can be minimized via graph cuts? *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 26(2):147–159.

Rosenberg, I. G. 1975. Reduction of bivalent maximization to the quadratic case. *Cahiers du Centre d'Études de Recherche Opérationnelle* 17:71–74.