On the Hardness of Decoding the Gale–Berlekamp Code

Ron M. Roth, Krishnamurthy Viswanathan
Advanced Studies
HP Laboratories Palo Alto
HPL-2007-2
January 8, 2007*

Gale–Berlekamp switching game, Hadamard matrices, intractable problems, maximum-likelihood decoding, NP-complete problems

The Gale–Berlekamp (in short, GB) code is the dual code of the binary product code in which the horizontal and vertical constituent codes are both the parity code. It is shown that the problem of deciding whether there is a codeword of the GB code within a prescribed distance from a given received word, is NP-complete. The problem remains hard (in a well-defined sense) even if the decoder is allowed unlimited preprocessing that depends only on the code length. While the intractability of maximum-likelihood decoding for specific codes has already been shown by Bruck and Naor and Lobstein, the result herein seems to be the first that shows hardness for familiar (or "natural") codes. In contrast, it is also shown that, with respect to any memoryless binary symmetric channel with crossover probability less than 1/2, maximum-likelihood decoding can be implemented in linear time for all error events except for a portion that occurs with vanishing probability.

* Internal Accession Date Only
Approved for External Publication

On the Hardness of Decoding the Gale–Berlekamp Code

Ron M. Roth*
Computer Science Department
Technion
Haifa 32000, Israel
ronny@cs.technion.ac.il

Krishnamurthy Viswanathan
Hewlett–Packard Laboratories
1501 Page Mill Road
Palo Alto, CA 94304, USA
krishnamurthy.viswanathan@hp.com

Abstract

The Gale–Berlekamp (in short, GB) code is the dual code of the binary product code in which the horizontal and vertical constituent codes are both the parity code. It is shown that the problem of deciding whether there is a codeword of the GB code within a prescribed distance from a given received word, is NP-complete. The problem remains hard (in a well-defined sense) even if the decoder is allowed unlimited preprocessing that depends only on the code length. While the intractability of maximum-likelihood decoding for specific codes has already been shown by Bruck and Naor and Lobstein, the result herein seems to be the first that shows hardness for familiar (or "natural") codes. In contrast, it is also shown that, with respect to any memoryless binary symmetric channel with crossover probability less than 1/2, maximum-likelihood decoding can be implemented in linear time for all error events except for a portion that occurs with vanishing probability.

Keywords: Gale–Berlekamp switching game; Hadamard matrices; Intractable problems; Maximum-likelihood decoding; NP-complete problems.

1 Introduction

Denote by Φ the subset $\{1, -1\}$ of the real field $\mathbb{R}$ and by $\Phi^{n\times n}$ the set of all $n \times n$ real matrices with entries in Φ. We consider the following optimization problem:

*Work done while visiting Hewlett–Packard Laboratories, Palo Alto, CA.

Optimization Problem 1.1. *Given a matrix $A \in \Phi^{n\times n}$, flip the sign of entire rows and columns in A so that the resulting matrix has the largest possible number of 1's.*

This problem is known as the *Gale–Berlekamp* (in short, GB) *switching game*; see for example Brown and Spencer [4], Fishburn and Sloane [6], or Spencer [22, Lecture 6]. The subject of this work is showing that this game is hard to solve. In the more general formulation of the game, the number of rows and columns in A does not have to be the same, yet for the purpose of demonstrating the hardness, we will restrict ourselves to the special case where the matrix A is square.

1.1 Equivalent formulations of the GB switching game

There are several equivalent formulations of the GB switching game, as shown next. For a vector $\boldsymbol{x} \in \mathbb{R}^n$, denote by $D(\boldsymbol{x})$ the $n \times n$ real diagonal matrix whose main diagonal consists of the entries of $\boldsymbol{x}$. Flipping entire rows and columns of A can be represented as the product $D(\boldsymbol{x})AD(\boldsymbol{y})$ for some row vectors $\boldsymbol{x}, \boldsymbol{y} \in \Phi^n$. Denote by J_n the matrix in $\Phi^{n\times n}$ whose entries are all 1. The number of (-1)'s in $D(\boldsymbol{x})AD(\boldsymbol{y})$ equals the number of 1's in the 0–1 matrix

$$B = \tfrac{1}{2}(J_n - D(\boldsymbol{x})AD(\boldsymbol{y}))\ . \tag{1}$$

But this number is also the sum of the entries of B and this sum, in turn, is equal to $(n^2 - \boldsymbol{x}A\boldsymbol{y}^T)/2$ (hereafter $(\cdot)^T$ denotes transposition). Hence, we have the following reformulation of the GB switching game:

Optimization Problem 1.2. *Given a matrix $A \in \Phi^{n\times n}$, find row vectors $\boldsymbol{x}, \boldsymbol{y} \in \Phi^n$ that maximize $\boldsymbol{x}A\boldsymbol{y}^T$.*

The number of nonzero entries in the matrix B in (1) equals the number of nonzero entries in $2D(\boldsymbol{x})BD(\boldsymbol{y})$, for every $\boldsymbol{x}, \boldsymbol{y} \in \Phi^n$. But

$$\begin{aligned} 2D(\boldsymbol{x})BD(\boldsymbol{y}) &= D(\boldsymbol{x})(J_n - D(\boldsymbol{x})AD(\boldsymbol{y}))D(\boldsymbol{y}) \\ &= D(\boldsymbol{x})J_nD(\boldsymbol{y}) - A\ . \end{aligned}$$

We now observe that $D(\boldsymbol{x})J_nD(\boldsymbol{y})$ has rank 1; in fact, the set

$$\mathcal{R}(n) = \Big\{ M \in \Phi^{n\times n}\ : \\ M = D(\boldsymbol{x})J_nD(\boldsymbol{y})\ \text{ for some }\ \boldsymbol{x}, \boldsymbol{y} \in \Phi^n \Big\}$$

(which is of size 2^{2n-1}) consists of all the matrices of rank 1 in $\Phi^{n\times n}$. Denoting by $\mathsf{d}(\cdot,\cdot)$ the Hamming distance between two matrices—or two vectors—of the same order (namely, the number of entries in which the two matrices differ), we get the next reformulation of the GB switching game:

Optimization Problem 1.3. *Given a matrix $A \in \Phi^{n\times n}$, find a rank-1 matrix $M \in \Phi^{n\times n}$ that minimizes* $\mathsf{d}(M, A)$.

We mention that Optimization Problem 1.3 is a constrained form of the problem of computing the rigidity of a matrix: see Lokam [14] and the references therein.

Optimization Problem 1.3 can be translated into a problem in which the objects are matrices over the binary field $\mathbb{F}_2$, simply by applying to each entry the bijection $\varphi : \Phi \to \mathbb{F}_2$ which sends 1 to 0 and -1 to 1. Next, we characterize the set of matrices to which $\mathcal{R}(n)$ is mapped under this bijection.

Given a positive integer n, let $\mathbf{1}_n$ be the all-1 row vector of length n and, for $i = 1, 2, \ldots, n$, denote by $\boldsymbol{e}_i$ the row unit-vector in $\mathbb{F}_2^n$ whose ith coordinate equals 1. Also, let $\mathbb{F}_2^{n\times n}$ stand for the set of all $n \times n$ matrices over $\mathbb{F}_2$. The $n \times n$ *Gale–Berlekamp code*, denoted hereafter by $\mathcal{C}_{\mathrm{GB}}(n)$, is the set of all matrices in $\mathbb{F}_2^{n\times n}$ that belong to the linear span of the following set of $2n$ matrices over $\mathbb{F}_2$:

$$\mathcal{L}(n) = \left\{\boldsymbol{e}_i^T \cdot \mathbf{1}_n\right\}_{i=1}^{n} \cup \left\{\mathbf{1}_n^T \cdot \boldsymbol{e}_j\right\}_{j=1}^{n} \ .$$

It is easy to verify that any $2n-1$ matrices in $\mathcal{L}(n)$ are linearly independent over $\mathbb{F}_2$ (but the sum of all $2n$ matrices in $\mathcal{L}(n)$ is zero). In addition, it is not hard to see that, with respect to Hamming weight, the elements in $\mathcal{L}(n)$ are the minimum-weight nonzero elements in the linear span of $\mathcal{L}(n)$. Hence, $\mathcal{C}_{\mathrm{GB}}(n)$ is a linear $[n^2, 2n-1, n]$ code over $\mathbb{F}_2$. The elements of $\mathcal{C}_{\mathrm{GB}}(n)$, each being a matrix in $\mathbb{F}_2^{n\times n}$, will be referred to as the codewords of $\mathcal{C}_{\mathrm{GB}}(n)$. From the definition of the code $\mathcal{C}_{\mathrm{GB}}(n)$ we get that it is the dual code the $[n^2, (n-1)^2, 4]$ product code over $\mathbb{F}_2$ with the horizontal and vertical constituent codes both being the $[n, n-1, 2]$ parity code over $\mathbb{F}_2$.

Under the bijection $\varphi : \Phi \to \mathbb{F}_2$, the set $\mathcal{R}(n)$ maps to $\mathcal{C}_{\mathrm{GB}}(n)$. Thus, Optimization Problems 1.1–1.3 can be expressed as a maximum-likelihood decoding (MLD) problem of the GB code, with respect to any memoryless binary symmetric channel (BSC) with crossover probability less than 1/2. Specifically:

Optimization Problem 1.4. *Given a matrix $Z \in \mathbb{F}_2^{n\times n}$, find a codeword $\Gamma \in \mathcal{C}_{\mathrm{GB}}(n)$ that minimizes* $\mathsf{d}(\Gamma, Z)$.

There has been a substantial body of work published on the problem of computing and bounding the covering radius, $\rho_{\mathrm{GB}}(n)$, of $\mathcal{C}_{\mathrm{GB}}(n)$: see [4], [6], [9, pp. 396–397], and [22, Lecture 6]. In particular, it is known that

$$\frac{n^2}{2} - \frac{n^{3/2}}{2} + o(n^{3/2}) \le \rho_{\mathrm{GB}}(n) \le \frac{n^2}{2} - \frac{n^{3/2}}{\sqrt{2\pi}} + o(n^{3/2}) \ . \tag{2}$$

As such, GB codes have a small covering radius given their (fairly low) rate: from the sphere-covering bound one gets that the covering radius of any linear $[N, k]$ code over $\mathbb{F}_2$ is greater

than

$$\frac{N}{2} - \sqrt{\frac{Nk}{2\log_2 \mathrm{e}}} ,$$

where $\mathrm{e} = 2.71828\cdots$ (see [19]), and plugging $N = n^2$ and $k = 2n-1$ into this expression yields the lower bound

$$\frac{n^2}{2} - \frac{n^{3/2}}{\sqrt{\log_2 \mathrm{e}}} .$$

Furthermore, the upper bound in (2), which was proved using probabilistic arguments by Brown and Spencer in [4, pp. 47–49], was established also algorithmically by Spencer in [22] through derandomization. He presented a deterministic polynomial-time algorithm which finds for every matrix $Z \in \mathbb{F}_2^{n\times n}$, a codeword $\Gamma \in \mathcal{C}_{\mathrm{GB}}(n)$ such that $\mathtt{d}(\Gamma, Z)$ is at most the right-hand side of (2) (see also Berger [2] and Pach and Spencer [19]). Note, however, that this algorithm does not necessarily find a *closest* codeword in $\mathcal{C}_{\mathrm{GB}}(n)$ to Z, i.e., this algorithm is not a maximum-likelihood decoder.

Relating $\rho_{\mathrm{GB}}(n)$ to the notation of Optimization Problem 1.2, we also have

$$\min_{A\in\Phi^{n\times n}} \max_{\boldsymbol{x},\boldsymbol{y}\in\Phi^n} \boldsymbol{x}A\boldsymbol{y}^T = n^2 - 2\rho_{\mathrm{GB}}(n) .$$

1.2 Complexity of MLD

The complexity of MLD of general linear codes was first studied by Berlekamp *et al.* in [3]. To show that MLD is intractable, it was stated as a decision problem:

Decision Problem 1.5. MLD OF LINEAR CODES.

Instance: Linear $[N, k]$ code $\mathcal{C}$ over $\mathbb{F}_2$ (represented, say, by its parity-check matrix), a word $\boldsymbol{z} \in \mathbb{F}_2^N$, and an integer τ.

Question: Is there a codeword in $\boldsymbol{c} \in \mathcal{C}$ such that $\mathtt{d}(\boldsymbol{c}, \boldsymbol{z}) \le \tau$?

Berlekamp *et al.* showed in [3] that this problem is NP-complete, using a reduction from THREE-DIMENSIONAL MATCHING (see the book of Garey and Johnson [8] about the theory of NP-completeness; the latter problem is described on pp. 50–53]). In Problem 1.5, the code $\mathcal{C}$ is part of the input, even though, in practice, the code is usually known in advance. Studying this more realistic version of the problem was the subject of the papers by Bruck and Naor [5] and Lobstein [13], who considered MLD of *specific* linear codes $\mathcal{C}_{\mathrm{BN}}(n)$ and $\mathcal{C}_{\mathrm{Lob}}(n)$, of parameters

$$[N{=}3n(n{-}1)/2, k{=}(n(n{+}1)/2){-}1, d{=}2]$$

and

$$[N{=}3n^3, k{=}n^3{-}3n{+}2, d{=}12] ,$$

respectively. The codes $\mathcal{C}_{\mathrm{BN}}(n)$ and $\mathcal{C}_{\mathrm{Lob}}(n)$ are explicitly described in the respective papers and, in particular, generator (or parity-check) matrices of these codes can be constructed in polynomial time. The results in [5] and [13] imply that the following decision problem is NP-complete (the problem is stated here for the parameters of [5]):

Decision Problem 1.6. MLD OF $\mathcal{C}_{\mathrm{BN}}(n)$.

Instance: Word $\boldsymbol{z} \in \mathbb{F}_2^{3n(n-1)/2}$ and an integer τ.

Question: Is there a codeword $\boldsymbol{c} \in \mathcal{C}_{\mathrm{BN}}(n)$ such that $\mathtt{d}(\boldsymbol{c}, \boldsymbol{z}) \leq \tau$?

In fact, since the codes are now specific, an even stronger statement was made in those two papers: Problem 1.6 is unlikely to become easy even if the decoder is allowed unlimited preprocessing that depends only on n (but not on $\boldsymbol{z}$): here "unlikely" means that if Problem 1.6 could be solved in polynomial time and unlimited preprocessing, then the polynomial hierarchy would collapse (albeit not necessarily collapsing NP with P).

One drawback of the results in [5] and [13] is that the codes $\mathcal{C}_{\mathrm{BN}}(n)$ and $\mathcal{C}_{\mathrm{Lob}}(n)$ are not "natural": they are artificially crafted for the proofs to work. Thus, attempts have been made to show the intractability of MLD of more widely-known codes. Indeed, Barg showed in [1] that MLD of product codes is NP-complete, and a similar result was obtained by Guruswami and Vardy in [10] for the class of generalized Reed–Solomon codes. However, in both these results the ensembles of codes at any given parameter range are super-polynomially large, and the results do not specify which of the codes within the ensemble is hard to decode.

1.3 Results of this work

In this work, we prove that the following decision-problem version of Optimization Problem 1.2, is NP-complete:

Decision Problem 1.7. BILINEAR FORM OVER Φ.

Instance: Matrix $A \in \Phi^{n \times n}$ and an integer τ.

Question: Are there vectors $\boldsymbol{x}, \boldsymbol{y} \in \Phi^n$ such that $\boldsymbol{x} A \boldsymbol{y}^T \geq \tau$?

A direct consequence of this result is that the following decision-problem version of Optimization Problem 1.4, is NP-complete:

Decision Problem 1.8. MLD OF GB CODES.

Instance: Matrix $Z \in \mathbb{F}_2^{n \times n}$ and an integer τ.

Question: Is there a codeword $\Gamma \in \mathcal{C}_{\mathrm{GB}}(n)$ such that $\mathtt{d}(\Gamma, Z) \leq \tau$?

Furthermore, we invoke the arguments in [5] to claim that these problems remain hard (in the sense that otherwise the polynomial hierarchy collapses) even with unlimited preprocessing. Our result seems to be the first to exhibit the intractability of MLD of a specific and familiar code, which we can comfortably refer to as "natural": the GB code has been studied before in several contexts—some of which motivated by its favorable covering radius.

Our NP-completeness proof consists of two reductions, which will be presented in Sections 3 and 4.

In contrast to our hardness result, we present in Section 5 a linear-time algorithm which, almost always, implements MLD of the GB code, with respect to any BSC with (fixed) crossover probability less than 1/2. The MLD implementation and, indeed, also the decoding fail only for a portion of error events which occurs with probability that goes to zero as the code length increases.

The rate of $\mathcal{C}_{\mathrm{GB}}(n)$, which equals $(2n-1)/n^2$ and is therefore inversely proportional to the square root of the length of $\mathcal{C}_{\mathrm{GB}}(n)$, vanishes as the code length increases (this is why it is at all possible to decode this code reliably even over a BSC with crossover probability arbitrarily close to 1/2). However, using $\mathcal{C}_{\mathrm{GB}}(n)$ as a building block, it is easy to come up with higher-rate codes for which MLD is intractable. We demonstrate how this can be done in Section 4.4.

We mention that our result on the NP-completeness of Problem 1.7 (Bilinear Form over Φ) improves on an earlier result by Poljak and Rohn [20], where a similar result was obtained for a less constrained problem: the entries of the matrix A therein can take integer values rather than values only from Φ (the particular reduction in [20] constructs matrices in which some of the entries may grow with the matrix order n).

As another application of the NP-completeness of Problem 1.7, we show in Section 6 that the following problem is NP-complete:

Decision Problem 1.9. Quadratic Form over Φ.

Instance: Symmetric matrix $Q \in \Phi^{n \times n}$ and an integer σ.

Question: Is there a row vector $\boldsymbol{v} \in \Phi^n$ such that $\boldsymbol{v}Q\boldsymbol{v}^T \geq \sigma$?

2 Preliminaries

We present here several definitions and quote several known results.

Let $\mathcal{G} = (V, E)$ be a (finite undirected) graph with a vertex set V and an edge set E. We will further assume that a graph has neither self-loops nor parallel edges. For a subset $S \subseteq V$, denote by $\partial(S)$ the set of edges each having one endpoint in S and one endpoint in $V \setminus S$. A *cut-set* is a subset of E that equals $\partial(S)$ for some $S \subseteq V$. The following decision problem is well-known to be NP-complete [8, p. 210]:

Decision Problem 2.1. MAX-CUT.

Instance: Graph: $\mathcal{G} = (V, E)$ and an integer τ.

Question: Is there a cut-set $\partial(S) \subseteq E$ of size at least τ?

The *incidence matrix* of a graph $\mathcal{G}$ is a $|V| \times |E|$ matrix $U_{\mathcal{G}} = (u_{i,e})$ over $\mathbb{F}_2$ whose rows (respectively, columns) are indexed by V (respectively, E), and $u_{i,e} = 1$ if and only if i is one of the endpoints of e in $\mathcal{G}$. Representing each cut-set in $\mathcal{G}$ by its characteristic vector in $\mathbb{F}_2^{|E|}$, we obtain a set of vectors over $\mathbb{F}_2$ which is called the *cut-set code* of $\mathcal{G}$ (see Hakimi and Frank [11]). This set is a linear space over $\mathbb{F}_2$ and is spanned by the rows of $U_{\mathcal{G}}$. When $\mathcal{G}$ is connected then $\mathrm{rank}(U_{\mathcal{G}}) = |V| - 1$.

The code $\mathcal{C}_{\mathrm{BN}}(n)$ in Problem 1.6 was constructed by Bruck and Naor in [5] as a cut-set code of a certain graph over $n(n+1)/2$ vertices. They then showed that Problem 1.6 is NP-complete by a reduction from Problem 2.1 (MAX-CUT).

A graph $\mathcal{G} = (V, E)$ is called bipartite if the set of vertices V can be partitioned into disjoint subsets V' and V'' such that each edge in E has one endpoint in V' and one endpoint in V''; we then write $\mathcal{G} = (V' : V'', E)$. A bipartite graph is balanced if $|V'| = |V''|$ and it is complete if there is an edge connecting each vertex in V' with each vertex in V''.

It is rather easy to see that the code $\mathcal{C}_{\mathrm{GB}}(n)$ is the cut-set code of a complete balanced bipartite graph where $|V'| = |V''| = n$ (see Solé and Zaslavsky [21]).

Problem 2.1 is trivial to solve for the case of bipartite graphs. However, if edges may be assigned negative weights then the problem (as formulated next) becomes NP-complete:

Decision Problem 2.2. BIPARTITE MAX-CUT OVER Φ.

Instance: Balanced bipartite graph: $\mathcal{G} = (V' : V'', E)$, a weight function $\omega : E \to \Phi$, and an integer τ.

Question: Is there a cut-set $\partial(S) \subseteq E$ such that $\sum_{e \in \partial(S)} \omega(e) \geq \tau$?

The NP-completeness of Problem 2.2 was proved in McCormick *et al.* [17] using a reduction from Problem 2.1. (The problem as stated in [17] does not assume that the graph is balanced, yet this restriction can be easily incorporated into the reduction by adding dummy vertices. In addition, while the edge weights in [17] can be arbitrary, one can verify that the NP-completeness proof therein still holds even when the weights are restricted to Φ.)

3 Relaxed problem

In this section, we prove the NP-completeness of a relaxation of Problem 1.7 (BILINEAR FORM OVER Φ) where we allow the matrix to have also zero entries. Our main hardness

result, which will be proved in Section 4, will be based on a reduction from the problem considered here.

Hereafter, Φ_0 denotes the set $\Phi \cup \{0\}$, and the formal statement of the relaxed problem is as follows:

Decision Problem 3.1. BILINEAR FORM OVER Φ_0.

Instance: Matrix $B \in \Phi_0^{n \times n}$ and an integer σ.

Question: Are there row vectors $\boldsymbol{x}, \boldsymbol{y} \in \Phi^n$ such that $\boldsymbol{x} B \boldsymbol{y}^T \geq \sigma$?

Proposition 3.1. *Problem 3.1 is NP-complete.*

Proof. First, it is easy to see that Problem 3.1 is in NP. Our proof of completeness borrows ideas from Poljak and Rohn [20] and will be by a reduction from Problem 2.2 (BIPARTITE MAX-CUT OVER Φ). Let

$$(\mathcal{G}=(V' : V'', E)\,,\ \omega : E \to \Phi\,,\ \tau)$$

be an instance of the latter problem and denote by n the size of V' (which is also the size of V''). We map this instance to an instance (B, σ) of Problem 3.1, where

$$\sigma = 2\tau - \sum_{e \in E} \omega(e)$$

and $B = (b_{i,j})$ is a matrix in $\Phi_0^{n \times n}$ whose rows (respectively, columns) are indexed by the elements of V' (respectively, V''), and

$$b_{i,j} = \begin{cases} -\omega(e) & \text{if } i \text{ and } j \text{ are connected by an edge } e \\ 0 & \text{otherwise} \end{cases} .$$

With any two subsets $S \subseteq V'$ and $T \subseteq V''$, we associate the following two vectors $\boldsymbol{x} = \boldsymbol{x}(S) = (x_i)_{i \in V'}$ and $\boldsymbol{y} = \boldsymbol{y}(S) = (y_j)_{j \in V''}$ in Φ^n:

$$x_i = \begin{cases} 1 & \text{if } i \in S \\ -1 & \text{otherwise} \end{cases}$$

and

$$y_j = \begin{cases} 1 & \text{if } j \in T \\ -1 & \text{otherwise} \end{cases} .$$

Clearly, the mapping $S \mapsto \boldsymbol{x}(S)$ (respectively, $T \mapsto \boldsymbol{y}(T)$) is a bijection from the set of subsets of V' (respectively, V'') onto Φ^n. Denoting by $\overline{S}$ and $\overline{T}$ the sets $V' \setminus S$ and $V'' \setminus T$,

respectively, we next compute the total weight of the edges in the cut-set $\partial(S \cup T)$:

$$\begin{aligned}
\sum_{e \in \partial(S \cup T)} \omega(e) &= \Big(-\sum_{(i,j) \in S \times \overline{T}} b_{i,j}\Big) + \Big(-\sum_{(i,j) \in \overline{S} \times T} b_{i,j}\Big) \\
&= -\frac{1}{4} \sum_{(i,j) \in V' \times V''} b_{i,j}(x_i - y_j)^2 \\
&= \frac{1}{2}\Big(\boldsymbol{x} B \boldsymbol{y}^T - \sum_{(i,j) \in V' \times V''} b_{i,j}\Big) \\
&= \frac{1}{2}\Big(\boldsymbol{x} B \boldsymbol{y}^T + \sum_{e \in E} \omega(e)\Big) \ .
\end{aligned}$$

Hence,

$$\sum_{e \in \partial(S \cup T)} \omega(e) \geq \tau \qquad \text{if and only if} \qquad \boldsymbol{x} B \boldsymbol{y}^T \geq \sigma \ .$$

The result follows. □

4 Main hardness result

In this section, we prove our main hardness result:

Theorem 4.1. *Problem 1.7* (Bilinear Form over Φ) *is NP-complete.*

4.1 Kronecker product and Hadamard matrices

The proof will make use of two lemmas, which we state next.

For the the first lemma we need the following definition. Let $X = (x_{i,j})$ and Y be real matrices of orders $k \times \ell$ and $p \times q$, respectively. The *Kronecker product* $X \otimes Y$ is defined as the $(kp) \times (\ell q)$ matrix that has the following block form:

$$X \otimes Y = \begin{pmatrix} x_{1,1}Y & x_{1,2}Y & \cdots & x_{1,\ell}Y \\ x_{2,1}Y & x_{2,2}Y & \cdots & x_{2,\ell}Y \\ \vdots & \vdots & \cdots & \vdots \\ x_{k,1}Y & x_{k,2}Y & \cdots & x_{k,\ell}Y \end{pmatrix} .$$

Among the properties of Kronecker product, it is known that for every four matrices X, Y, Z, and W,

$$(X \otimes Z)(Y \otimes W) = (XY) \otimes (ZW) \ , \tag{3}$$

provided that the (ordinary) matrix multiplications are allowed, namely, the number of columns of X (respectively, Z) equals the number of rows of Y (respectively, W); see [16, Theorem 43.4].

Lemma 4.2. *Let B be a matrix in $\mathbb{R}^{n\times n}$ and let m be a positive integer. Then*

$$\max_{\boldsymbol{x},\boldsymbol{y}\in\Phi^{mn}} \boldsymbol{x}(B\otimes J_m)\boldsymbol{y}^T = m^2\cdot \max_{\boldsymbol{r},\boldsymbol{s}\in\Phi^n} \boldsymbol{r}B\boldsymbol{s}^T\ . \tag{4}$$

Furthermore, the left-hand side of (4) is maximized for vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in Φ^{mn} of the form

$$\boldsymbol{x} = \boldsymbol{r}\otimes \mathbf{1}_m \qquad \textit{and} \qquad \boldsymbol{y} = \boldsymbol{s}\otimes \mathbf{1}_m \tag{5}$$

where $\boldsymbol{r}$ and $\boldsymbol{s}$ are vectors in Φ^n that maximize the right-hand side of (4).

Proof. First, for every two vectors $\boldsymbol{x},\boldsymbol{y}\in\Phi^{mn}$ of the form (5) we have

$$\begin{aligned}
\boldsymbol{x}(B\otimes J_m)\boldsymbol{y}^T &= (\boldsymbol{r}\otimes \mathbf{1}_m)(B\otimes J_m)(\boldsymbol{s}\otimes \mathbf{1}_m)^T\\
&= (\boldsymbol{r}\otimes \mathbf{1}_m)(B\otimes J_m)(\boldsymbol{s}^T\otimes \mathbf{1}_m^T)\\
&= (\boldsymbol{r}B\boldsymbol{s}^T)\otimes(\mathbf{1}_m J_m \mathbf{1}_m^T)\\
&= m^2\cdot(\boldsymbol{r}B\boldsymbol{s}^T)\ ,
\end{aligned}$$

where the third equality follows from two applications of the rule (3). Hence, it remains to show that the maximum in the left-hand side of (4) is indeed attained by vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ of the form (5).

Let $\boldsymbol{x} = (x_1\ x_2\ \ldots\ x_{mn})$ and $\boldsymbol{y}$ attain that maximum and consider the real vector

$$\boldsymbol{v}^T = (v_1\ v_2\ \ldots\ v_{mn})^T = (B\otimes J_m)\boldsymbol{y}^T\ .$$

Fixing $\boldsymbol{v}$, the vector $\boldsymbol{x}$ must be one of those in Φ^{mn} that maximize the scalar product $\boldsymbol{x}\cdot\boldsymbol{v}^T$. Such vectors, in turn, are characterized by

$$x_i = \operatorname{sgn}(v_i)\ , \quad 1\le i\le mn$$

(for indexes i where $v_i = 0$, the value x_i can be arbitrarily set to either 1 or -1). On the other hand, by the particular form of the matrix $B\otimes J_m$ we get that for every $j = 0, 1, \ldots, n-1$,

$$v_{jm+1} = v_{jm+2} = \ldots = v_{(j+1)m}\ ;$$

consequently, a maximizing vector $\boldsymbol{x}$ satisfies

$$x_{jm+1} = x_{jm+2} = \ldots = x_{j(m+1)-1} = \operatorname{sgn}(v_{jm+1})$$

whenever $v_{jm+1}\neq 0$, and can always be assumed to satisfy these equalities when $v_{jm+1} = 0$. It follows that for every vector $\boldsymbol{y}$ that belongs to a pair $(\boldsymbol{x},\boldsymbol{y})$ that maximizes the left-hand

side of (4), we can always assume that the respective vector $\boldsymbol{x}$ takes the form $\boldsymbol{r} \otimes \mathbf{1}_m$ for some vector $\boldsymbol{r} \in \Phi^n$. Reversing the roles of $\boldsymbol{x}$ and $\boldsymbol{y}$, we conclude that $\boldsymbol{y}$ can be assumed to take the form $\boldsymbol{s} \otimes \mathbf{1}_m$ for some $\boldsymbol{s} \in \Phi^n$. □

The second lemma, taken from [4], presents a property of Hadamard matrices. Recall that an $m \times m$ matrix H over $\mathbb{R}$ is called a *Hadamard matrix* if the entries of H are in Φ and $HH^T = m \cdot I$.

Lemma 4.3. *Let H be an $m \times m$ Hadamard matrix. For every two row vectors $\boldsymbol{r}, \boldsymbol{s} \in \Phi^m$,*

$$\left|\boldsymbol{r} H \boldsymbol{s}^T\right| \leq m^{3/2} \ .$$

Proof. For a row vector $\boldsymbol{r} \in \mathbb{R}^m$, denote by $\|\boldsymbol{r}\|$ its L_2 norm $\sqrt{\boldsymbol{r} \cdot \boldsymbol{r}^T}$. We have,

$$\|\boldsymbol{r}H\|^2 = \boldsymbol{r}HH^T\boldsymbol{r}^T = m \cdot \|\boldsymbol{r}\|^2 \ .$$

Applying the Cauchy–Schwartz inequality yields for any two row vectors $\boldsymbol{r}, \boldsymbol{s} \in \Phi^m$,

$$\left|\boldsymbol{r} H \boldsymbol{s}^T\right| \leq \|\boldsymbol{r}H\| \cdot \|\boldsymbol{s}\| = \sqrt{m} \cdot \|\boldsymbol{r}\| \cdot \|\boldsymbol{s}\| = m^{3/2} \ .$$

(For an alternate proof of the lemma, see [4].) □

Polynomially-constructible symmetric Hadamard matrices are known for infinitely many orders m; in particular, Sylvester-type Hadamard matrices exist whenever $m = 2^h$ and take the form

$$\underbrace{\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}}_{h \text{ times}}$$

(see, for example [15, Section 2.3]).

4.2 Proof of main hardness result

We are now ready to prove Theorem 4.1.

Proof of Theorem 4.1. Problem 1.7 (Bilinear Form over Φ) is easily verified to be in NP. The completeness will be established by a reduction from Problem 3.1 (Bilinear Form over Φ_0).

Let (B, σ) be an instance of the latter problem, where $B = (b_{i,j})$ is a matrix in $\Phi_0^{n \times n}$. Take m to be the smallest power of 2 that is greater than $2n^4$, and let H be an $m \times m$ Hadamard matrix. We now map (B, σ) to an instance (A, τ) of Problem 1.7, where

$$\tau = \left(\sigma - \frac{1}{2}\right) m^2$$

and A is a matrix in $\Phi^{mn\times mn}$ of the block form $(A_{i,j})_{i,j=1}^{n}$ in which each $m \times m$ block $A_{i,j}$ is given by

$$A_{i,j} = \begin{cases} b_{i,j}J_m & \text{if } b_{i,j} \neq 0 \\ H & \text{if } b_{i,j} = 0 \end{cases} , \quad 1 \leq i,j \leq n \ .$$

Notice that $A_{i,j}$ equals the respective block in $B \otimes J_m$ whenever $b_{i,j} \neq 0$; otherwise, $A_{i,j} = H$ whereas the respective block in $B \otimes J_m$ is all-zero.

For convenience, we introduce the notation α and β for the following maximal values:

$$\alpha = \max_{\boldsymbol{x},\boldsymbol{y}\in\Phi^{mn}} \boldsymbol{x}A\boldsymbol{y}^T \qquad \text{and} \qquad \beta = \max_{\boldsymbol{r},\boldsymbol{s}\in\Phi^{n}} \boldsymbol{r}B\boldsymbol{s}^T \ .$$

By Lemma 4.3 we get that for every two vectors $\boldsymbol{x}, \boldsymbol{y} \in \Phi^{mn}$,

$$\begin{aligned} \left|(\boldsymbol{x}A\boldsymbol{y}^T) - (\boldsymbol{x}(B \otimes J_m)\boldsymbol{y}^T)\right| &= \left|\boldsymbol{x}(A - (B \otimes J_m))\boldsymbol{y}^T\right| \\ &\leq n^2 \cdot \max_{\boldsymbol{r},\boldsymbol{s}\in\Phi^{m}} \left|\boldsymbol{r}H\boldsymbol{s}^T\right| \\ &\leq n^2 \cdot m^{3/2} \ ; \end{aligned}$$

so, by Lemma 4.2,

$$\left|\alpha - \beta m^2\right| \leq n^2 \cdot m^{3/2} \ .$$

Hence, if $\beta \geq \sigma$ then

$$\alpha \geq \beta m^2 - n^2 \cdot m^{3/2} \geq \left(\sigma - \frac{n^2}{\sqrt{m}}\right) m^2 > \tau \ ,$$

where the last inequality follows from the requirement that $m > 2n^4$. Conversely, if $\alpha \geq \tau$ then

$$\beta \geq \frac{\alpha}{m^2} - \frac{n^2}{\sqrt{m}} > \frac{\tau}{m^2} - \frac{1}{2} \geq \sigma - 1 \ ,$$

namely, $\beta \geq \sigma$. We conclude that

$$\alpha \geq \tau \qquad \text{if and only if} \qquad \beta \geq \sigma \ ,$$

thereby completing the proof. □

As Problem 1.7 (Bilinear Form over Φ) and Problem 1.8 (MLD of GB Codes) are equivalent, the following corollary immediately follows.

Corollary 4.4. *Problem 1.8* (MLD of GB Codes) *is NP-complete.*

4.3 Hardness with preprocessing

In practice, the design of any decoder—say, a decoder of $\mathcal{C}_{\mathrm{GB}}(n)$—is carried out only once, as opposed to the number of applications of the decoder (to received words) which, in turn, can be very large. Therefore, when measuring the decoding complexity, one can attribute computations that depend only on the parameter n (and not on the received word) to the decoding design stage, rather than to the time the decoder is actually applied. Computations that depend only on n are referred to as *preprocessing*, and the question is whether MLD of $\mathcal{C}_{\mathrm{GB}}(n)$ remains hard even if we ignore the complexity of preprocessing. Another way of posing this question is whether, for any given n, a Boolean circuit that implements MLD of $\mathcal{C}_{\mathrm{GB}}(n)$ is still unlikely to be polynomially large in n, even if we ignore the time it takes to design that circuit. By invoking a result of Karp and Lipton [12], Bruck and Naor provided a positive answer to this question, for the case of the code $\mathcal{C}_{\mathrm{BN}}(n)$; in their result, "unlikely" means that if there were polynomially large circuits that implement MLD of that code, then the polynomial hierarchy would collapse (see [5, Section 3] for more details). Using the arguments made by Bruck and Naor, the very same conclusion can be drawn from Corollary 4.4 with respect to circuits that implement MLD of $\mathcal{C}_{\mathrm{GB}}(n)$.

4.4 Codes at higher rates

While the rate of $\mathcal{C}_{\mathrm{GB}}(n)$ vanishes as n increases, we next use $\mathcal{C}_{\mathrm{GB}}(n)$ as a building block to obtain higher-rate codes for which MLD is intractable: simply take the code

$$\begin{aligned} \mathcal{C} &= \mathcal{C}_0 \times \mathcal{C}_{\mathrm{GB}}(n) \\ &= \left\{ (\boldsymbol{c}_0 \,|\, \Gamma) \;:\; \boldsymbol{c}_0 \in \mathcal{C}_0 \text{ and } \Gamma \in \mathcal{C}_{\mathrm{GB}}(n) \right\}, \end{aligned}$$

where $\mathcal{C}_0$ is a linear $[N, k, d]$ code over $\mathbb{F}_2$ with length N that is polynomially large in n but also greater than n^3, the rate k/N is at least a prescribed value R (< 1), and $d > n^2$. Explicit constructions of such codes $\mathcal{C}_0$ are known (e.g., Justesen codes [15, Section 10.11]) and, for these parameters, the rate of $\mathcal{C}$ is greater than $(1 - (1/n))R$. Given a received word of the form $\boldsymbol{z} = (\boldsymbol{0} \,|\, Z)$ (where $Z \in \mathbb{F}_2^{n \times n}$) and an integer $\tau \leq n^2$, any codeword in $\mathcal{C}$ at Hamming distance at most τ from $\boldsymbol{z}$ must take the form $\boldsymbol{c} = (\boldsymbol{0} \,|\, \Gamma)$ for some $\Gamma \in \mathcal{C}_{\mathrm{GB}}(n)$. Thus, any polynomial-time implementation of MLD of $\mathcal{C}$ would imply such an implementation for $\mathcal{C}_{\mathrm{GB}}(n)$.

5 Decoding algorithm over the BSC

In this section, we present a linear-time decoding algorithm for $\mathcal{C}_{\mathrm{GB}}(n)$. We show that, with respect to any BSC with crossover probability less than $1/2$, the algorithm errs with

probability that decays exponentially with n. Since a similar behavior of the error probability is achieved also by a maximum-likelihood decoder for $\mathcal{C}_{\mathrm{GB}}(n)$, our analysis will lead to the conclusion that, with respect to the probability measure which is induced by the channel, the decoding algorithm that we present here implements MLD with probability approaching 1 as the code length goes to infinity.

We will use the notation BSC(p) for a BSC with crossover probability $p \in [0,1]$. We assume that a codeword $\Gamma = (\Gamma_{i,j})$ of $\mathcal{C}_{\mathrm{GB}}(n)$ is transmitted through BSC(p) and an $n \times n$ matrix $Z = (Z_{i,j})$ over $\mathbb{F}_2$ is received at the channel output, such that $\mathsf{Prob}\{Z_{i,j} \neq \Gamma_{i,j}\} = p$, independently for distinct pairs (i,j).

Given a decoder $\mathcal{D} : \mathbb{F}_2^{n \times n} \to \mathcal{C}_{\mathrm{GB}}(n)$, we denote by $P_{\mathrm{err}}(\mathcal{D}|\Gamma)$ the probability that the decoder returns the incorrect codeword, given that $\Gamma \in \mathcal{C}_{\mathrm{GB}}(n)$ is transmitted; namely,

$$P_{\mathrm{err}}(\mathcal{D}|\Gamma) = \mathsf{Prob}\left\{\mathcal{D}(Z) \neq \Gamma \,|\, \Gamma \text{ was transmitted}\right\} \ ,$$

where the conditional probability is the one induced by BSC(p) (the dependence of $P_{\mathrm{err}}(\mathcal{D}|\Gamma)$ on p is kept implicit for the sake of simplicity of the notation). Also, we let $P_{\mathrm{err}}(\mathcal{D})$ denote the decoding error probability for the worst-case codeword:

$$P_{\mathrm{err}}(\mathcal{D}) = \max_{\Gamma \in \mathcal{C}_{\mathrm{GB}}(n)} P_{\mathrm{err}}(\mathcal{D}|\Gamma) \ .$$

5.1 Linear-time decoding algorithm

The decoder that we present for $\mathcal{C}_{\mathrm{GB}}(n)$ is the function

$$\mathcal{D}_{\mathrm{GB}}^{(n)} : \mathbb{F}_2^{n \times n} \to \mathcal{C}_{\mathrm{GB}}(n)$$

whose value for every given $Z \in \mathbb{F}_2^{n \times n}$ is given by the return value of the algorithm GB_DECODER shown in Figure 1 (in the figure, we use our earlier notation $\mathsf{d}(\cdot,\cdot)$ for Hamming distance). The algorithm GB_DECODER can be viewed as a variant of an algorithm that has been recently suggested for jointly compressing similar files [23].

We first provide the intuition behind the algorithm. Recall from Section 1.1 that $\mathcal{C}_{\mathrm{GB}}(n)$ is the linear span of

$$\mathcal{L}(n) = \left\{\boldsymbol{e}_i^T \cdot \mathbf{1}_n\right\}_{i=1}^n \cup \left\{\mathbf{1}_n^T \cdot \boldsymbol{e}_j\right\}_{j=1}^n \ ,$$

and that every $2n-1$ matrices in $\mathcal{L}(n)$ form a basis of this span over $\mathbb{F}_2$. It follows that for every $\Gamma \in \mathcal{C}_{\mathrm{GB}}(n)$ there exist unique row vectors $\boldsymbol{a} = (a_i)_{i=1}^n$ and $\boldsymbol{b} = (b_j)_{j=1}^n$ over $\mathbb{F}_2$ such that $a_1 = 0$ and

$$\Gamma = \boldsymbol{a}^T \cdot \mathbf{1}_n \ + \ \mathbf{1}_n^T \cdot \boldsymbol{b} \ . \tag{6}$$

Thus, the ith row of the transmitted codeword Γ equals either $\boldsymbol{b}$ (if $a_i = 0$) or $\boldsymbol{b} + \mathbf{1}_n$ (if $a_i = 1$); and, since $a_1 = 0$, the first row always equals $\boldsymbol{b}$. Based on this simple observation,

Algorithm GB_DECODER (*Input:* $Z \in \mathbb{F}_2^{n\times n}$):
/* $\boldsymbol{r}_i$ denotes the ith row of Z and $\boldsymbol{c}_j^T$ denotes its jth column. */

1. Compute a vector $\widehat{\boldsymbol{a}} = (0\ \hat{a}_2\ \hat{a}_3\ \ldots\ \hat{a}_n)$ over $\mathbb{F}_2$ by:

$$\hat{a}_i \leftarrow \begin{cases} 1 & \text{if } \mathsf{d}(\boldsymbol{r}_1, \boldsymbol{r}_i) > n/2 \\ 0 & \text{otherwise} \end{cases}, \quad 2 \le i \le n\ .$$

2. Compute a vector $\widehat{\boldsymbol{b}} = (\hat{b}_1\ \hat{b}_2\ \ldots\ \hat{b}_n)$ over $\mathbb{F}_2$ by:

$$\hat{b}_j \leftarrow \begin{cases} 1 & \text{if } \mathsf{d}(\widehat{\boldsymbol{a}}, \boldsymbol{c}_j) > n/2 \\ 0 & \text{otherwise} \end{cases}, \quad 1 \le j \le n\ .$$

3. Recompute $\widehat{\boldsymbol{a}}$ by

$$\hat{a}_i \leftarrow \begin{cases} 1 & \text{if } \mathsf{d}(\widehat{\boldsymbol{b}}, \boldsymbol{r}_i) > n/2 \\ 0 & \text{otherwise} \end{cases}, \quad 2 \le i \le n\ .$$

4. Return $\widehat{\Gamma} = \widehat{\boldsymbol{a}}^T \cdot \mathbf{1}_n + \mathbf{1}_n^T \cdot \widehat{\boldsymbol{b}}$.

Figure 1: Algorithm for computing $\mathcal{D}^{(n)}_{\mathrm{GB}}(Z)$.

Step 1 of the algorithm GB_DECODER in Figure 1 finds an initial estimate for $\boldsymbol{a}$ by comparing each row of the *noisy* matrix Z to its first row: if the ith row of Z is close (with respect to Hamming distance) to the first row of Z then the estimate for a_i is $\hat{a}_i = 0$; otherwise, the estimate is $\hat{a}_i = 1$. (Clearly, the choice of the first row in Z as a reference in these comparisons is arbitrary; any other row in Z could play that role just as well.)

Once we have (an estimate for) $\boldsymbol{a}$, we can proceed with the recovery of the entries of $\boldsymbol{b}$: we see from (6) that the jth column of Γ is equal to $\boldsymbol{a}^T$ when $b_j = 0$, and to $(\boldsymbol{a} + \mathbf{1}_n)^T$ when $b_j = 1$. Therefore, we can estimate b_j according to the Hamming distance between the jth column of Z and our estimate for $\boldsymbol{a}^T$; this is done in Step 2 of GB_DECODER.

While the estimates for $\boldsymbol{a}$ and $\boldsymbol{b}$ that are computed in Steps 1 and 2 already yield a decoding error probability that decays exponentially with n, Step 3 was added to GB_DECODER to accelerate that decay so that it matches that of a maximum-likelihood decoder (see Propositions 5.1 and 5.5 below).

The operations used during the execution of GB_DECODER are additions in $\mathbb{F}_2$ and increments of counters and indexes of length $O(\log_2 n)$. The number of applications of these operations is quadratic in n, i.e., it is linear in the code length of $\mathcal{C}_{\mathrm{GB}}(n)$.

We next turn to analyzing the decoding error probability of the decoder $\mathcal{D}_{\mathrm{GB}}^{(n)}$ implemented by GB_Decoder. For two reals $p, \theta \in (0,1)$, denote by $\delta_p(\theta)$ the value

$$\delta_p(\theta) = \left(\frac{p}{\theta}\right)^{\theta} \left(\frac{1-p}{1-\theta}\right)^{1-\theta} .$$

It can be readily verified that for every fixed $p \in (0,1)$, the function $\theta \mapsto \delta_p(\theta)$ is continuous over $(0,1)$ and it attains a unique maximum in that interval at $\theta = p$ (in which case $\delta_p(p) = 1$).

Let $\boldsymbol{\epsilon} = (\epsilon_\ell)_{\ell=1}^n$ be a random vector over $\mathbb{F}_2$ whose entries are independent Bernoulli random variables taking on $\mathbb{F}_2$ with $\mathsf{Prob}\{\epsilon_\ell = 1\} = p$ (e.g., $\boldsymbol{\epsilon}$ can be taken as the error vector that is added by BSC(p) to each row or column of the transmitted codeword). It is known that, with respect to this probability measure, the Hamming weight of $\boldsymbol{\epsilon}$, which we denote by $\mathsf{w}(\boldsymbol{\epsilon})$, satisfies

$$\mathsf{Prob}\left\{\mathsf{w}(\boldsymbol{\epsilon}) \geq \theta n\right\} \leq \left(\delta_p(\theta)\right)^n , \tag{7}$$

whenever $\theta > p$ (see [7, p. 531]). We will use the notation γ_p for the value $\delta_p(1/2)$: it is easy to see that

$$\gamma_p = \delta_p(1/2) = 2\sqrt{p(1-p)} . \tag{8}$$

Proposition 5.1. *With respect to* BSC(p) *with any fixed* $p < 1/2$,

$$P_{\mathrm{err}}\left(\mathcal{D}_{\mathrm{GB}}^{(n)}\right) \leq \gamma_p^{n(1-o_p(1))} ,$$

where $o_p(1)$ *stands for a positive expression that converges to* 0 *as* n *goes to infinity (at a rate that may depend on* p*). In particular,* $P_{\mathrm{err}}\left(\mathcal{D}_{\mathrm{GB}}^{(n)}\right)$ *decays exponentially with* n.

We break the proof into three lemmas.

Lemma 5.2. *Let the codeword* $\Gamma = \boldsymbol{a}^T \cdot \mathbf{1}_n + \mathbf{1}_n^T \cdot \boldsymbol{b}$ *of* $\mathcal{C}_{\mathrm{GB}}(n)$ *be transmitted through* BSC(p) *with* $p < 1/2$, *and let* $Z \in \mathbb{F}_2^{n \times n}$ *be the (random) matrix received at the channel output. Fix* $m : \mathbb{Z}^+ \to \mathbb{Z}^+$ *to be any integer function such that*

$$\lim_{n\to\infty} m(n) = \infty \qquad \textit{and} \qquad \lim_{n\to\infty} \frac{m(n)}{n} = 0 . \tag{9}$$

Then the random vector $\widehat{\boldsymbol{a}} = \widehat{\boldsymbol{a}}(Z)$ *that is computed in Step 1 of* GB_Decoder *satisfies*

$$\mathsf{Prob}\left\{\mathsf{d}(\boldsymbol{a}, \widehat{\boldsymbol{a}}) \geq m(n)\right\} \leq \gamma_p^{n(1-o_p(1))} ,$$

where $\mathsf{Prob}\{\cdot\}$ *is the probability measure induced by the channel on its output* Z, *conditioned on* Γ *being transmitted.*

Proof. Let $\mathcal{I}$ be a subset of $\{2, 3, \ldots, n\}$ of size $m = m(n)$. We compute an upper bound on the probability that the values $\hat{a}_i$ that are computed in Step 1 are erroneous for all $i \in \mathcal{I}$.

For $i = 1, 2, \ldots, n$, let

$$\boldsymbol{\epsilon}_i = (\epsilon_{i,1}\ \epsilon_{i,2}\ \ldots\ \epsilon_{i,n})$$

be the error vector (over $\mathbb{F}_2$) that is added by the channel to the ith row of Γ (to form the ith row of Z). For $j = 1, 2, \ldots, n$, denote by X_j the number of rows, among the rows of Z that are indexed by $\mathcal{I}$, in which the jth entry of the error vector differs from the respective entry in $\boldsymbol{\epsilon}_1$, namely,

$$X_j = \left|\{i \in \mathcal{I}\ :\ \epsilon_{i,j} \neq \epsilon_{1,j}\}\right| \ .$$

By the definition of BSC(p) we have, for every $1 \leq j \leq n$ and $1 \leq k \leq m$,

$$\mathsf{Prob}\left\{X_j = k\right\} = \binom{m}{k}\left(p^k(1-p)^{m+1-k} + p^{m+1-k}(1-p)^k\right). \tag{10}$$

Moreover, the random variables $X_1, X_2, \ldots, X_n$ are statistically independent.

Now,

$$\begin{aligned}
\mathsf{Prob}\left\{\bigcap_{i\in\mathcal{I}}(\hat{a}_i \neq a_i)\right\} &\leq \mathsf{Prob}\left\{\bigcap_{i\in\mathcal{I}}\left(\mathsf{w}(\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_i) \geq \frac{n}{2}\right)\right\} \\
&\leq \mathsf{Prob}\left\{\left(\sum_{i\in\mathcal{I}}\mathsf{w}(\boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_i)\right) \geq \frac{mn}{2}\right\} \\
&= \mathsf{Prob}\left\{\sum_{j=1}^{n} X_j \geq \frac{mn}{2}\right\}.
\end{aligned}$$

Let z be a real in $(0, 1)$ and denote by $\mathsf{E}\{\cdot\}$ an expectation value taken with respect to the probability measure $\mathsf{Prob}\{\cdot\}$. By the Chernoff bound (see [7, p. 127]) we get

$$\begin{aligned}
\mathsf{Prob}\left\{\sum_{j=1}^{n} X_j \geq \frac{mn}{2}\right\} &\leq \mathsf{E}\left\{z^{mn-2\sum_{j=1}^{n} X_j}\right\} \\
&\leq z^{mn} \cdot \mathsf{E}\left\{\prod_{j=1}^{n} z^{-2X_j}\right\} \\
&= \left(z^m \cdot \mathsf{E}\left\{z^{-2X_1}\right\}\right)^n \ ,
\end{aligned}$$

where, from (10),

$$\begin{aligned}
z^m \cdot \mathsf{E}\left\{z^{-2X_1}\right\} &= z^m \cdot \sum_{k=0}^{m}\binom{m}{k}\left(p^k(1-p)^{m+1-k}\right. \\
&\qquad \left. + p^{m+1-k}(1-p)^k\right) z^{-2k} \\
&= (1-p)\left(pz^{-1} + (1-p)z\right)^m \\
&\qquad + p\left(pz + (1-p)z^{-1}\right)^m \ .
\end{aligned}$$

The last three chains of (in)equalities can be summarized by

$$\begin{aligned}\mathsf{Prob}\left\{\bigcap_{i\in\mathcal{I}}(\hat{a}_i \neq a_i)\right\} &\leq \mathsf{Prob}\left\{\sum_{j=1}^{n} X_j \geq \frac{mn}{2}\right\} \\ &\leq (\gamma_p(z,m))^n \ , \end{aligned} \tag{11}$$

where

$$\begin{aligned}\gamma_p(z,m) &= (1-p)\left(pz^{-1}+(1-p)z\right)^m \\ &\quad + p\left(pz+(1-p)z^{-1}\right)^m \ . \end{aligned}$$

Let $c = c(p)$ be defined by

$$c = \frac{\ln\left((1/p)-1\right)}{2(1-2p)} \ , \tag{12}$$

and select $z = z_m = 1/(1+(c/m))$. In this case,

$$\begin{aligned}&\left(pz_m^{-1}+(1-p)z_m\right)^m \\ &\quad = \left(p\Big(1+\frac{c}{m}\Big)+(1-p)\Big(1+\frac{c}{m}\Big)^{-1}\right)^m \\ &\quad \leq \left(p\Big(1+\frac{c}{m}\Big)+(1-p)\Big(1-\frac{c}{m}+\frac{c^2}{m^2}\Big)\right)^m \\ &\quad = \left(1-\frac{(1-2p)c}{m}+\frac{pc^2}{m^2}\right)^m \\ &\quad = \mathrm{e}^{-(1-2p)c}+\tilde{o}_p(1) \ , \end{aligned}$$

where $\tilde{o}_p(1)$ stands for an expression that goes to 0 as m goes to infinity (at a rate that may depend on p). Similarly,

$$\begin{aligned}\left(pz_m+(1-p)z_m^{-1}\right)^m &\leq \left(1+\frac{(1-2p)c}{m}+\frac{(1-p)c^2}{m^2}\right)^m \\ &= \mathrm{e}^{(1-2p)c}+\tilde{o}_p(1) \ . \end{aligned}$$

It follows from the last two chains of inequalities that

$$\gamma_p(z_m,m) = (1-p)\cdot\mathrm{e}^{-(1-2p)c}+p\cdot\mathrm{e}^{(1-2p)c}+\tilde{o}_p(1) \ .$$

Now, from (12) we get that

$$\mathrm{e}^{(1-2p)c} = \sqrt{\frac{1-p}{p}}$$

and, so,

$$\gamma_p(z_m,m) = 2\sqrt{p(1-p)}+\tilde{o}_p(1) = \gamma_p+\tilde{o}_p(1) \ .$$

Combining the last equation with (11), we conclude that

$$\begin{aligned}\mathsf{Prob}\left\{\bigcap_{i\in\mathcal{I}}(\hat{a}_i \neq a_i)\right\} &\leq (\gamma_p(z_m, m))^n \\ &= (\gamma_p + \tilde{o}_p(1))^n \ ,\end{aligned}$$

and, taking the union bound over all subsets $\mathcal{I}$ of $\{2, 3, \ldots, n\}$ of size m, we obtain

$$\begin{aligned}\mathsf{Prob}\left\{\mathsf{d}(\boldsymbol{a}, \widehat{\boldsymbol{a}}) \geq m\right\} &\leq \binom{n-1}{m} (\gamma_p(z_m, m))^n \\ &\leq 2^{n\mathsf{H}(m/n)} (\gamma_p + \tilde{o}_p(1))^n \ ,\end{aligned}$$

where the last inequality follows from known upper bounds on the binomial coefficients, in terms of the binary entropy function $\mathsf{H}(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ (see, for example [15, p. 309]). Finally, since we assume that m grows with n yet $m/n = o_p(1)$, we get

$$\begin{aligned}\mathsf{Prob}\left\{\mathsf{d}(\boldsymbol{a}, \widehat{\boldsymbol{a}}) \geq m\right\} &\leq 2^{n\mathsf{H}(o_p(1))} (\gamma_p + o_p(1))^n \\ &= \gamma_p^{n(1-o_p(1))} \ ,\end{aligned}$$

as claimed. □

Lemma 5.3. *Under the conditions of Lemma 5.2, the random vector* $\widehat{\boldsymbol{b}} = \widehat{\boldsymbol{b}}(Z)$ *that is recomputed in Step 2 of* GB_DECODER *satisfies*

$$\mathsf{Prob}\left\{\widehat{\boldsymbol{b}} \neq \boldsymbol{b}\right\} \leq \gamma_p^{n(1-o_p(1))} \ .$$

Proof. For $j = 1, 2, \ldots, n$, denote by $\boldsymbol{\epsilon}'_j$ the error vector (over $\mathbb{F}_2$) that is added by the channel to the jth column of Γ. Also, let $\widehat{\boldsymbol{a}}$ be the vector computed in Step 1 of GB_DECODER. Then, for every positive integer m, the event

$$\left\{\widehat{\boldsymbol{b}} \neq \boldsymbol{b}\right\}$$

is contained in the following union of $n+1$ events:

$$\left\{\mathsf{d}(\boldsymbol{a}, \widehat{\boldsymbol{a}}) \geq m\right\} \bigcup \left\{\bigcup_{j=1}^{n}\left(\mathsf{w}(\boldsymbol{\epsilon}'_j) > \frac{n}{2} - m\right)\right\} .$$

Therefore, by the union bound,

$$\begin{aligned}\mathsf{Prob}\left\{\widehat{\boldsymbol{b}} \neq \boldsymbol{b}\right\} &\leq \mathsf{Prob}\left\{\mathsf{d}(\boldsymbol{a}, \widehat{\boldsymbol{a}}) \geq m\right\} \\ &\quad + \sum_{j=1}^{n} \mathsf{Prob}\left\{\mathsf{w}(\boldsymbol{\epsilon}'_j) > \frac{n}{2} - m\right\} . \qquad (13)\end{aligned}$$

For the remaining part of the proof, we assume that $m = m(n)$ satisfies the conditions in (9). Lemma 5.2 then provides an upper bound on the first term in the right-hand side of (13). As for each of the remaining terms in (13), from (7) we get

$$\begin{aligned}\mathsf{Prob}\Big\{\mathsf{w}(\boldsymbol{\epsilon}'_j) > \frac{n}{2} - m\Big\} &\leq \left(\delta_p\Big(\frac{1}{2} - \frac{m}{n}\Big)\right)^n \\ &= \left(\delta_p\Big(\frac{1}{2} - o_p(1)\Big)\right)^n \\ &= (\gamma_p + o_p(1))^n \ ,\end{aligned}$$

where the last equality follows from the continuity of $\theta \mapsto \delta_p(\theta)$. Hence,

$$\mathsf{Prob}\Big\{\mathsf{w}(\boldsymbol{\epsilon}'_j) > \frac{n}{2} - m\Big\} \leq \gamma_p^{n(1-o_p(1))} \ ,$$

and the result follows from (13) and Lemma 5.2. □

Lemma 5.4. *Under the conditions of Lemma 5.2, the vector $\widehat{\boldsymbol{a}} = \widehat{\boldsymbol{a}}(Z)$ that is computed in Step 3 of* GB_Decoder *satisfies*

$$\mathsf{Prob}\,\{\widehat{\boldsymbol{a}} \neq \boldsymbol{a}\} \leq \gamma_p^{n(1-o_p(1))} \ .$$

Proof. As we did in the proof of Lemma 5.2, we denote by $\boldsymbol{\epsilon}_i$ the error vector that is added by the channel to the ith row of Γ. The event

$$\{\widehat{\boldsymbol{a}} \neq \boldsymbol{a}\}$$

is contained in the union

$$\Big\{\boldsymbol{b} \neq \widehat{\boldsymbol{b}}\Big\} \bigcup \left\{\bigcup_{i=1}^{n}\Big(\mathsf{w}(\boldsymbol{\epsilon}_i) > \frac{n}{2}\Big)\right\} \ ,$$

where $\widehat{\boldsymbol{b}}$ is the vector computed in Step 2 of GB_Decoder. Therefore,

$$\mathsf{Prob}\,\{\widehat{\boldsymbol{a}} \neq \boldsymbol{a}\} \leq \mathsf{Prob}\left\{\widehat{\boldsymbol{b}} \neq \boldsymbol{b}\right\} + \sum_{i=1}^{n} \mathsf{Prob}\Big\{\mathsf{w}(\boldsymbol{\epsilon}_i) > \frac{n}{2}\Big\} \ ,$$

and the claim follows from Lemma 5.3 and (7). □

Proof of Proposition 5.1. Conditioning on the transmitted codeword being $\Gamma = \boldsymbol{a}^T \cdot \mathbf{1}_n + \mathbf{1}_n^T \cdot \boldsymbol{b}$, we have,

$$\begin{aligned}\mathsf{Prob}\left\{\mathcal{D}_{\mathrm{GB}}^{(n)}(Z) \neq \Gamma\right\} &= \mathsf{Prob}\left\{(\widehat{\boldsymbol{a}} \neq \boldsymbol{a}) \cup (\widehat{\boldsymbol{b}} \neq \boldsymbol{b})\right\} \\ &\leq \mathsf{Prob}\,\{\widehat{\boldsymbol{a}} \neq \boldsymbol{a}\} + \mathsf{Prob}\left\{\widehat{\boldsymbol{b}} \neq \boldsymbol{b}\right\} \ ,\end{aligned}$$

where $\widehat{\boldsymbol{a}}$ and $\widehat{\boldsymbol{b}}$ are the vectors computed in Steps 3 and 2, respectively, of GB_Decoder. The result now follows from Lemmas 5.3 and 5.4. □

As our analysis in Section 5.2 will reveal, the error exponent (i.e., the rate of the exponential decay in n) in Proposition 5.1 matches that of a maximum-likelihood decoder.

5.2 Error probability of MLD

In this section, we compare the decoding error probability of $\mathcal{D}_{\mathrm{GB}}^{(n)}$ to that of a maximum-likelihood decoder, with respect to $\mathrm{BSC}(p)$ with any fixed crossover probability $p < 1/2$.

For an indeterminate ξ, let

$$W_{\mathrm{GB}}^{(n)}(\xi) = \sum_{t=0}^{n^2} W_t \xi^t$$

denote the weight distribution of $\mathcal{C}_{\mathrm{GB}}(n)$; that is, W_t is the number of codewords in $\mathcal{C}_{\mathrm{GB}}(n)$ of Hamming weight t. Using the characterization (6) of the codewords of $\mathcal{C}_{\mathrm{GB}}(n)$, it is easy to verify that the Hamming weight of a codeword Γ can be written as

$$\mathsf{w}(\Gamma) = n \cdot (\mathsf{w}(\boldsymbol{a}) + \mathsf{w}(\boldsymbol{b})) - 2\mathsf{w}(\boldsymbol{a}) \cdot \mathsf{w}(\boldsymbol{b}) \ .$$

Ranging over all $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{F}_2^n$ such that $a_1 = 0$, we obtain

$$\begin{aligned} W_{\mathrm{GB}}^{(n)}(\xi) &= \sum_{\Gamma \in \mathcal{C}_{\mathrm{GB}}(n)} \xi^{\mathsf{w}(\Gamma)} \\ &= \sum_{\boldsymbol{a} \,:\, a_1 = 0} \sum_{\boldsymbol{b}} \xi^{n \cdot (\mathsf{w}(\boldsymbol{a}) + \mathsf{w}(\boldsymbol{b})) - 2\mathsf{w}(\boldsymbol{a}) \cdot \mathsf{w}(\boldsymbol{b})} \\ &= \sum_{k=0}^{n-1} \sum_{\ell=0}^{n} \binom{n-1}{k} \binom{n}{\ell} \xi^{n(k+\ell) - 2k\ell} \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} (\xi^k + \xi^{n-k})^n \ . \end{aligned}$$

Let

$$\mathcal{D}_{\mathrm{ML}}^{(n)} : \mathbb{F}_2^{n \times n} \to \mathcal{C}_{\mathrm{GB}}(n)$$

be a maximum-likelihood decoder for $\mathcal{C}_{\mathrm{GB}}(n)$. It is known that, with respect to $\mathrm{BSC}(p)$ with $p < 1/2$, the decoding error probability of $\mathcal{D}_{\mathrm{ML}}^{(n)}$ satisfies

$$P_{\mathrm{err}}\left(\mathcal{D}_{\mathrm{ML}}^{(n)}\right) \le W_{\mathrm{GB}}^{(n)}(\gamma_p) - 1 \ , \tag{14}$$

where γ_p is given by (8) (see [18, p. 153]). An inspection of the expression for $W_{\mathrm{GB}}^{(n)}(\gamma_p)$ yields

$$W_{\mathrm{GB}}^{(n)}(\gamma_p) = 1 + 2n\gamma_p^n + O(n^2 \gamma_p^{2n-2}) \ , \tag{15}$$

which means that the main term in the upper bound (14) is $W_n \gamma_p^n$, where $W_n = |\mathcal{L}(n)| = 2n$.

Proposition 5.5. *With respect to* $\mathrm{BSC}(p)$ *with any fixed* $p < 1/2$,

$$P_{\mathrm{err}}\left(\mathcal{D}_{\mathrm{ML}}^{(n)}\right) = \gamma_p^{n(1 - o_p(1))} \ .$$

Proof. Equations (14)–(15) imply that

$$P_{\text{err}}\left(\mathcal{D}_{\text{ML}}^{(n)}\right) \leq \gamma_p^{n(1-o_p(1))} \ .$$

To show the inequality in the other direction, let us assume that the decoder is told the value of the vector $\boldsymbol{a}$ that is associated (by (6)) to the transmitted codeword Γ, and all the decoder needs to find is an estimate $\widehat{\boldsymbol{b}}$ for $\boldsymbol{b}$. As in the proof of Lemma 5.3, we denote by $\boldsymbol{\epsilon}'_j$ the error vector (over $\mathbb{F}_2$) that is added by the channel to the jth column of Γ. We have,

$$\begin{aligned}
\mathsf{Prob}\left\{\widehat{\boldsymbol{b}} \neq \boldsymbol{b}\right\} &\geq 1 - \mathsf{Prob}\left\{\bigcap_{j=1}^{n}\left(\mathsf{w}(\boldsymbol{\epsilon}'_j) \leq \frac{n}{2}\right)\right\} \\
&= 1 - \prod_{j=1}^{n}\mathsf{Prob}\left\{\mathsf{w}(\boldsymbol{\epsilon}'_j) \leq \frac{n}{2}\right\} \\
&= 1 - \left(1 - \mathsf{Prob}\left\{\mathsf{w}(\boldsymbol{\epsilon}'_1) > \frac{n}{2}\right\}\right)^n \ .
\end{aligned}$$

Now,

$$\begin{aligned}
\mathsf{Prob}\left\{\mathsf{w}(\boldsymbol{\epsilon}'_1) > \frac{n}{2}\right\} &\geq \binom{n}{\lfloor n/2 \rfloor + 1} p^{(n/2)+1}(1-p)^{(n/2)-1} \\
&\geq \frac{p}{1-p} \cdot \Omega\left(\frac{\gamma_p^n}{\sqrt{n}}\right) \ ,
\end{aligned}$$

where the last inequality follows from the Stirling approximation of the binomial coefficients (see [15, p. 309]; here $\Omega(\cdot)$ stands for an expression that grows at least linearly with its argument). From the last two chains of inequalities we deduce that

$$\begin{aligned}
\mathsf{Prob}\left\{\widehat{\boldsymbol{b}} \neq \boldsymbol{b}\right\} &\geq 1 - \left(1 - \frac{p}{1-p} \cdot \Omega\left(\frac{\gamma_p^n}{\sqrt{n}}\right)\right)^n \\
&= \frac{p}{1-p} \cdot \Omega\left(\sqrt{n} \cdot \gamma_p^n\right) \\
&= \gamma_p^{n(1-o_p(1))} \ ,
\end{aligned}$$

thereby completing the proof. □

From Propositions 5.1 and 5.5 we conclude that the error exponents of $\mathcal{D}_{\text{GB}}^{(n)}$ and $\mathcal{D}_{\text{ML}}^{(n)}$ are the same. We also have the following corollary, which states that $\mathcal{D}_{\text{GB}}^{(n)}$ approximates $\mathcal{D}_{\text{ML}}^{(n)}$ well, over BSC(p) with any $p < 1/2$.

Corollary 5.6. *With respect to* BSC(p) *with any* $p < 1/2$*, and for every transmitted codeword* $\Gamma \in \mathcal{C}_{\text{GB}}(n)$*,*

$$\mathsf{Prob}\left\{\mathcal{D}_{\text{ML}}^{(n)}(Z) = \mathcal{D}_{\text{GB}}^{(n)}(Z) = \Gamma\right\} = 1 - \gamma_p^{n(1-o_p(1))} \ ,$$

where $\mathsf{Prob}\{\cdot\}$ *is the probability measure induced by the channel on its output* $Z \in \mathbb{F}_2^{n \times n}$*, conditioned on* Γ *being transmitted.*

6 Quadratic forms

As another application of the NP-completeness of Problem 1.7 (Bilinear Form over Φ), we prove here the following result:

Proposition 6.1. *Problem 1.9* (Quadratic Form over Φ) *is NP-complete.*

Proof. Problem 1.9 is clearly in NP. The proof of completeness will be carried out by a reduction from Problem 1.7.

Let (A, τ) be an instance of the latter problem, where A is a matrix in $\Phi^{n\times n}$. Take ℓ to be the smallest integer such that $\ell \geq 16n^3$ and $\ell = \lceil 2^h/n \rceil$ for some integer h. Denote by H a symmetric matrix in $\Phi^{\ell n\times \ell n}$ which is obtained from a $2^h \times 2^h$ symmetric Hadamard matrix by padding $\ell n - 2^h$ ($< n$) all-1 rows and columns. We now map the instance (A, τ) to an instance (Q, σ) of Problem 1.9 where

$$\sigma = (2\tau - 1)\ell^2$$

and Q is the following symmetric matrix in $\Phi^{2\ell n\times 2\ell n}$:

$$Q = \begin{pmatrix} H & J_\ell \otimes A \\ (J_\ell \otimes A)^T & H \end{pmatrix} .$$

Let

$$\boldsymbol{x} = (\boldsymbol{x}_1 | \, \boldsymbol{x}_2 | \ldots | \, \boldsymbol{x}_\ell) \qquad \text{and} \qquad \boldsymbol{y} = (\boldsymbol{y}_1 | \, \boldsymbol{y}_2 | \ldots | \, \boldsymbol{y}_\ell)$$

be two vectors in $\Phi^{\ell n}$, where each block, $\boldsymbol{x}_i$ or $\boldsymbol{y}_j$, is a vector in Φ^n. Denoting by $\boldsymbol{v}$ the vector $(\boldsymbol{x} \,|\, \boldsymbol{y})$ in $\Phi^{2\ell n}$, we have,

$$\begin{aligned} \boldsymbol{v}Q\boldsymbol{v}^T &= \boldsymbol{x}H\boldsymbol{x}^T + \boldsymbol{y}H\boldsymbol{y}^T + 2\boldsymbol{x}(J_\ell \otimes A)\boldsymbol{y}^T \\ &= \boldsymbol{x}H\boldsymbol{x}^T + \boldsymbol{y}H\boldsymbol{y}^T + 2\Big(\sum_{i,j=1}^{\ell} \boldsymbol{x}_i A \boldsymbol{y}_j^T\Big) . \end{aligned} \tag{16}$$

By Lemma 4.3 we get that

$$\begin{aligned} \left|\boldsymbol{x}H\boldsymbol{x}^T\right|, \left|\boldsymbol{y}H\boldsymbol{y}^T\right| &< 2^{3h/2} + 2\ell n(\ell n - 2^h) \\ &< (\ell n)^{3/2} + 2\ell n^2 \\ &\leq \left(\sqrt{\frac{n^3}{\ell}} + \frac{2n^3}{\ell}\right)\ell^2 \\ &\leq \left(\frac{1}{4} + \frac{2}{16}\right)\ell^2 \;<\; \frac{\ell^2}{2} , \end{aligned} \tag{17}$$

where the second and fourth inequalities follow, respectively, from the choice of ℓ so that $\ell = \lceil 2^h/n \rceil$ and $\ell \geq 16n^3$. From (16) and (17) we obtain

$$\left| \Big(\max_{\boldsymbol{v} \in \Phi^{2\ell n}} \boldsymbol{v} Q \boldsymbol{v}^T \Big) - 2 \Big(\max_{\boldsymbol{v} \in \Phi^{2\ell n}} \sum_{i,j=1}^{\ell} \boldsymbol{x}_i A \boldsymbol{y}_j^T \Big) \right| < \ell^2 \ . \tag{18}$$

Denote

$$\alpha = \max_{\boldsymbol{v} \in \Phi^{2\ell n}} \boldsymbol{v} Q \boldsymbol{v}^T \qquad \text{and} \qquad \beta = \max_{\boldsymbol{r}, \boldsymbol{s} \in \Phi^n} \boldsymbol{r} A \boldsymbol{s}^T \ .$$

Observing that

$$\max_{\boldsymbol{v} \in \Phi^{2\ell n}} \sum_{i,j=1}^{\ell} \boldsymbol{x}_i A \boldsymbol{y}_j^T = \beta \ell^2 \ ,$$

we get from (18) that

$$(2\beta - 1)\ell^2 < \alpha < (2\beta + 1)\ell^2 \ .$$

Hence, if $\beta \geq \tau$ then

$$\alpha > (2\tau - 1)\ell^2 = \sigma \ .$$

Conversely, if $\alpha \geq \sigma$ then

$$\beta > \frac{1}{2}\Big(\frac{\alpha}{\ell^2} - 1 \Big) \geq \frac{1}{2}\Big(\frac{\sigma}{\ell^2} - 1 \Big) = \tau - 1 \ ,$$

namely, $\beta \geq \tau$. We conclude that

$$\alpha \geq \sigma \qquad \text{if and only if} \qquad \beta \geq \tau \ .$$

This completes the proof. □

Recall that Problem 1.7 (Bilinear Form over Φ) is equivalent to MLD of the GB code which, in turn, is the cut-set code of a complete balanced bipartite graph. It can be shown that, in analogy, Problem 1.9 (Quadratic Form over Φ) is equivalent to MLD of the $[n(n-1)/2, n-1, n-1]$ cut-set code of a complete graph over n vertices (in which every two distinct vertices are connected by an edge). The latter code thus serves as yet another example of a case where MLD is NP-complete.

Acknowledgments

We would like to thank Ram Swaminathan for discussions that led us to this problem.

References

[1] A. Barg, *Some new NP-complete coding problems, Probl. Inform. Transm.,* 30 (1994), 298–214.

[2] B. Berger, *The fourth moment method, SIAM J. Comput.,* 26 (1997), 1188–1207.

[3] E.R. Berlekamp, R.J. McEliece, H.C.A. van Tilborg, *On the inherent intractability of certain coding problems, IEEE Trans. Inform. Theory,* 24 (1978), 384–386.

[4] T. Brown, J. Spencer, *Minimization of ± 1 matrices under line shifts, Colloq. Math.,* 23 (1971), 165–171.

[5] J. Bruck, M. Naor, *The hardness of decoding linear codes with preprocessing, IEEE Trans. Inform. Theory,* 36 (1990), 381–385.

[6] P.C. Fishburn, N.J.A. Sloane, *The solution to Berlekamp's switching game, Discrete Math.,* 74 (1989), 263–290.

[7] R.G. Gallager, *Information Theory and Reliable Communication,* Wiley, New York, 1968.

[8] M.R. Garey, D.S. Johnson, *Computers and Intractability: a Guide to the Theory of NP-Completeness,* Freeman, New York, 1979.

[9] R.L. Graham, N.J.A. Sloane, *On the covering radius of codes, IEEE Trans. Inform. Theory,* 31 (1985), 385–401.

[10] V. Guruswami, A. Vardy, *Maximum-likelihood decoding of Reed–Solomon codes is NP-hard, IEEE Trans. Inform. Theory,* 51 (2005), 2249–2256.

[11] S.L. Hakimi, H. Frank, *Cut-set matrices and linear codes, IEEE Trans. Inform. Theory,* 11 (1965), 457–458.

[12] R.M. Karp, R.J. Lipton, *Some connections between nonuniform and uniform complexity classes, Proc. 12th Annual Symp. Theory of Computing (STOC'1980),* Los Angeles, California, 1980, pp. 302–309.

[13] A. Lobstein, *The hardness of solving subset sum with preprocessing, IEEE Trans. Inform. Theory,* 36 (1990), 943–946.

[14] S.V. Lokam, *Spectral methods for matrix rigidity with applications to size-depth tradeoffs and communication complexity, Proc. 36th Annual Symp. Foundations of Computer Science (FOCS'1995),* Milwaukee, Wisconsin, 1995, pp. 6–15.

[15] F.J. MacWilliams, N.J.A. Sloane, *The Theory of Error-Correcting Codes,* North-Holland, Amsterdam, 1977.

[16] C.C. MacDuffee, *The Theory of Matrices,* Chelsea, New York, 1946.

[17] S.T. McCormick, M.R. Rao, G. Rinaldi, *Easy and difficult objective functions for max cut, Math. Program., Ser. B,* 94 (2003), 459–466.

[18] R.J. McEliece, *The Theory of Information and Coding,* Second Edition, Cambridge University Press, Cambridge, 2002.

[19] J. Pach, J. Spencer, *Explicit codes with low covering radius, IEEE Trans. Inform. Theory,* 34 (1988), 1281–1285.

[20] S. Poljak, J. Rohn, *Checking robust nonsingularity is NP-hard, Math. Control Signals Systems,* 6 (1993), 1–9.

[21] P. Solé, T. Zaslavsky, *A coding approach to signed graphs, SIAM J. Disc. Math.,* 7 (1994), 544–553.

[22] J. Spencer, *Ten Lectures on the Probabilistic Method,* Second Edition, SIAM, Philadelphia, Pennsylvania, 1994.

[23] K. Viswanathan, R. Swaminathan, M. Uysal, *Collaborative compression, In preparation,* 2006.