Hadamard Tensors and Lower Bounds on Multiparty Communication Complexity *

Jeff Ford and Anna Gál

Department of Computer Sciences
University of Texas at Austin
Austin, TX 78712-1188, USA
{jefford, panni}@cs.utexas.edu

Abstract

We develop a new method for estimating the discrepancy of tensors associated with multiparty communication problems in the "Number on the Forehead" model of Chandra, Furst and Lipton. We define an analogue of the Hadamard property of matrices for tensors in multiple dimensions and show that any k-party communication problem represented by a Hadamard tensor must have $\Omega(n/2^k)$ multiparty communication complexity. We also exhibit constructions of Hadamard tensors, giving $\Omega(n/2^k)$ lower bounds on multiparty communication complexity for a new class of explicitly defined Boolean functions.

1 Introduction

Communication complexity was introduced by Yao [26] in 1979. Two players wish to compute $f(x, y)$. One player knows x, and the other knows y. Both have unlimited computational power. The communication complexity of f is the number of bits they must exchange on an arbitrary input in order to determine the value of f. This model and many of its variants have been widely studied [15]. Communication complexity arguments have been used to derive results in circuit complexity and in other computational models.

[*] A preliminary version of this paper appeared in the Proceedings of ICALP (2005) [12]. The authors were supported in part by NSF CAREER Award CCR-9874862, NSF Grant CCF-0430695 and an Alfred P. Sloan Research Fellowship.

Dagstuhl Seminar Proceedings 06111
Complexity of Boolean Functions
http://drops.dagstuhl.de/opus/volltexte/2006/607

We consider the multiparty model of Chandra, Furst and Lipton [7] usually called the "Number on the Forehead" model. With k players, the input is partitioned into k parts: $x_1, \ldots, x_k$. The i-th player has access to every x_j except x_i. The Number on the Forehead model is stronger than the 2-party model, and sometimes the overlap between the players' inputs can be used to obtain surprising upper bounds (e.g. [21, 20]). This model is harder to analyze than the 2-party model, and very few lower bounds are known. On the other hand, lower bounds in this model have many applications in complexity theory, including constructions of pseudorandom generators for space bounded computation, universal traversal sequences, and time-space tradeoffs [2], as well as circuit complexity lower bounds [14, 19, 21].

Currently the largest known lower bounds for explicit functions are of the form $\Omega(n/2^k)$ where k is the number of players, and n is the number of bits each player misses. The first bounds of this form were given by Babai, Nisan and Szegedy [2] for the "quadratic character of the sum of coordinates" (QCS) function. $\Omega(n/c^k)$ lower bounds were also proved for the "generalized inner product" (GIP) function in [2] and by Chung and Tetali [10]. Chung [9] and Raz [22] generalized the method of [2] to give a sufficient condition for a function to have $\Omega(n/c^k)$ multiparty communication complexity. Raz [22] also obtained $\Omega(\sqrt{n}/2^k)$ lower bounds for a new function based upon matrix multiplication over GF(2). Babai, Hayes and Kimmel [3] obtained further examples of functions with $\Omega(n/2^k)$ multiparty communication complexity. All of these lower bounds were obtained by estimating discrepancy, and so they also hold in the distributional and randomized communication complexity models.

The known bounds all decrease exponentially as the number of players grows, becoming trivial for $k > \log n$. It is a major open problem, with important implications in circuit complexity, to prove nontrivial lower bounds on multiparty communication problems for a large number of players. The class ACC^0, defined by Barrington [4], consists of languages recognized by constant depth, unbounded fan-in polynomial size circuit families with AND, OR, NOT and MOD$_m$ gates for a fixed m. By the results of [27, 5, 14], families of functions that belong to ACC^0 can be computed by multiparty protocols with polylogarithmic (in n) communication by polylogarithmic (in n) number of players (where n is the number of bits each player misses). Separating ACC^0 from other complexity classes (e.g. NP) is a major open problem, and a sufficiently large multiparty communication complexity lower bound would resolve it.

Chor and Goldreich showed in [8] that any Boolean function defined by a Hadamard matrix has $\Omega(n)$ 2-party communication complexity. Their proof uses a lemma by Lindsey (see [11] p. 88) that estimates the largest possible sum of entries in a submatrix of a Hadamard matrix. Lindsey's lemma implies upper bounds on the discrepancy of functions defined by Hadamard matrices and "nearly" Hadamard matrices. Babai, Nisan and Szegedy [2] generalized the proof of Lindsey's lemma to obtain upper bounds on the discrepancy of tensors associated with certain multiparty communication problems. The lower bounds that followed (e.g. [9, 10, 22, 3]) all used this approach. These papers did not consider generalizing the Hadamard property to tensors. In fact, [10] mentions that it is not clear how to generalize Hadamard matrices to tensors.

In this paper we propose a generalization of the Hadamard property of matrices to tensors of arbitrary dimension. We show that any k-party communication problem represented by a Hadamard tensor must have $\Omega(n/2^k)$ multiparty communication complexity. We construct families of Hadamard tensors, giving $\Omega(n/2^k)$ lower bounds for a new class of explicitly defined Boolean functions. Our Hadamard property is stronger than the sufficient condition of Chung [9] and Raz [22] for $\Omega(n/2^k)$ bounds, and could yield larger than $\Omega(n/2^k)$ lower bounds. There are no matching upper bounds known for functions represented by Hadamard tensors. We show how the Chung-Raz condition and some previous lower bounds fit into a "nearly" Hadamard framework. We believe that Hadamard tensors may also be of independent interest.

Our approach is based upon a new general upper bound on the discrepancy of tensors in terms of the largest possible value achievable by multiplying a collection of lines of the tensor by -1 and taking the sum of the entries of the resulting tensor. We refer to this value as the *weight*. This measure has been analyzed for matrices (see e.g. [1, 23]), and the corresponding matrix problem is sometimes called the "switching lights game". Generalizing the switching lights game to tensors was previously suggested in [10]. As far as we know, the general upper bound we give for the discrepancy of a tensor in terms of its weight is new. We also show that this upper bound is not too much larger than the actual discrepancy. Thus, the weight will give good bounds and may be easier to use than directly computing discrepancy. Since our lower bounds are based on discrepancy, they also hold in the distributional and randomized models.

2 Preliminaries

In the k-party model of Chandra, Furst and Lipton [7], k players with unlimited computational power wish to compute the value of a function $f : X_1 \times \cdots \times X_k \to \{-1, 1\}$ on input $\vec{x} = (x_1, \ldots, x_k)$. Usually we assume that $X_1 = \cdots = X_k = \{0,1\}^n$. The function f is known to each player, and player P_i gets all of the input except $x_i \in X_i$. Players communicate by broadcasting messages, so all players receive all messages. If each player misses n bits of input, then $n + 1$ bits of communication is sufficient: Player P_2 broadcasts x_1, and then player P_1, who now has the entire input, broadcasts the answer.

Definition 2.1 *The deterministic k-party* communication complexity *of f (denoted $C(f)$) is the number of bits communicated by the players on the worst input $\vec{x}$ using the best protocol for computing f.*

Definition 2.2 *Let μ be a probability distribution over the input of f. The* bias *achieved by a protocol P is defined as $|Pr[P(\vec{x}) = f(\vec{x})] - Pr[P(\vec{x}) \neq f(\vec{x})]|$, where $\vec{x}$ is chosen according to the distribution μ.*

The ϵ-distributional communication complexity *of f (denoted $C_{\epsilon,\mu}(f)$) is the number of bits communicated by the players on the worst input $\vec{x}$ using the best protocol for computing f that achieves bias at least ϵ under the distribution μ. When μ is the uniform distribution we abbreviate to $C_\epsilon(f)$.*

Definition 2.3 *[2] A subset $Z_i \subseteq X_1 \times \cdots \times X_k$ is called a* cylinder *in the i-th dimension, if membership in Z_i does not depend on the i-th coordinate; that is, for every $(x_1, \ldots, x_i, \ldots, x_k) \in Z_i$ and every $x'_i \in X_i$ we have $(x_1, \ldots, x'_i, \ldots, x_k) \in Z_i$ as well. A subset $Z \subseteq X_1 \times \cdots \times X_k$ is called a* cylinder intersection *if it can be represented as $Z = \cap_{i=1}^k Z_i$, where each Z_i is a cylinder in the i-th dimension.*

A protocol can be thought of as reducing the space of possible inputs at each step until all the remaining possibilities give the same output. A message from player P_i winnows the input space, but not along the i-th dimension. Thus it causes the space of possible inputs to be intersected with a cylinder in the i-th dimension. After each message the consistent inputs form a cylinder intersection.

Definition 2.4 *The* discrepancy *of f on the cylinder intersection Z (denoted $\mathrm{Disc}_Z(f)$) is defined by*

$$\mathrm{Disc}_Z(f) = \left|\Pr[(\vec{x} \in Z) \wedge (f(\vec{x}) = 1)] - \Pr[(\vec{x} \in Z) \wedge (f(\vec{x}) \neq 1)]\right|,$$

where $\vec{x}$ is chosen according to the uniform distribution. The discrepancy *of f (denoted $\mathrm{Disc}(f)$) is the maximum value of $\mathrm{Disc}_Z(f)$ over all cylinder intersections Z.*

Since we define $\mathrm{Disc}(f)$ with respect to the uniform distribution, and the output of f is from $\{-1, 1\}$, we have the following:

$$\mathrm{Disc}_Z(f) = |\sum_{\vec{x} \in Z} f(\vec{x})| \;/\; |X_1 \times \cdots \times X_k|.$$

Lemma 2.1 *[2] For any function $f : X_1 \times X_2 \times \cdots \times X_k \to \{-1, 1\}$, $C(f) \geq \log_2(1/\mathrm{Disc}(f))$ and $C_\epsilon(f) \geq \log_2(\epsilon/\mathrm{Disc}(f))$.*

3 A General Upper Bound on Discrepancy

Problems in 2-party communication complexity can be represented as matrices with rows labeled by the possible inputs for player P_1 and columns labeled by the possible inputs for player P_2. An entry in the matrix at location (x, y) is given by $f(x, y)$.

A multiparty communication complexity problem can be represented by a tensor, the multidimensional analogue of a matrix. Each dimension of the tensor is labeled by the piece of input missed by a player. That is, given a function $f : X_1 \times \cdots \times X_k \to \{-1, 1\}$, the i-th dimension of the tensor representing f is indexed by the elements of X_i. We denote by $A(x_1, \ldots, x_k)$ the entry of the k-dimensional tensor A at location $(x_1, \ldots, x_k)$. For tensor A_f representing function f we have $A_f(x_1, \ldots, x_k) = f(x_1, \ldots, x_k)$. If $|X_1| = \cdots = |X_k| = N$, we say that the tensor has *order N*.

Definition 3.1 *Given a tensor A in k dimensions, a* line *of A is any vector formed by fixing all but one coordinate of A. A* face *of A is any $(k-1)$-dimensional tensor formed by fixing one coordinate of A.*

A tensor of order N has N entries in each line and N^{k-1} entries in each face. It has N^{k-1} lines and N faces along each of the k dimensions.

Definition 3.2 *Let A be a tensor with ± 1 entries. We say that a line of the tensor A is* flipped *if each entry in that line is multiplied by -1.*

Definition 3.3 *We say that a tensor is* cylindrical *in the i-th dimension, if it does not depend on the i-th coordinate x_i.*

If a tensor is cylindrical in the i-th dimension, the entries of any given line along the i-th dimension are identical, and the corresponding N faces are identical. Thus, a k-dimensional cylindrical tensor can be specified by a $(k-1)$-dimensional tensor (specifying the face that is repeated N times).

Definition 3.4 *We define the* excess *of a tensor A (denoted $S(A)$) to be the sum of its entries; that is, $S(A) = \sum_{\vec{x} \in X_1 \times \cdots \times X_k} A(\vec{x})$.*

Lemma 3.1 *(implicit in [10])* $\text{Disc}(f) = \max S(A_f \circ C_1 \circ \cdots \circ C_k)/N^k$*, where A_f is the ± 1 tensor representing f, and each C_i is a $0/1$ tensor that is cylindrical in the i-th dimension. ($A_f \circ C_1 \circ \cdots \circ C_k$ denotes the entry-wise product of the tensors $A, C_1, \ldots, C_k$.)*

Proof: Let $Z_i \subseteq X_1 \times \cdots \times X_k$ be a cylinder in the i-th dimension, and let C_i be the $0/1$ tensor representing the characteristic function of the cylinder Z_i. Then C_i is cylindrical in the i-th dimension. Conversely, every $0/1$ tensor which is cylindrical in the i-th dimension represents the characteristic function of some cylinder in the i-th dimension. The lemma immediately follows from the definitions and our notation. ∎

Definition 3.5 *We define the* weight *of a tensor A (denoted $W(A)$) to be the largest possible excess of a tensor A' where A' can be obtained from A by flipping an arbitrary collection of lines (in any direction). Note that the order in which the flips are performed does not matter.*

Alternatively, $W(A)$ can be described as $W(A) = \max S(A \circ T_1 \circ \cdots \circ T_k)$, where each T_i is a ± 1 tensor that is cylindrical in the i-th dimension. ($A \circ T_1 \circ \cdots \circ T_k$ denotes the entry-wise product of the tensors $A, T_1, \ldots, T_k$.)

Theorem 3.2 *Consider a function $f : X_1 \times \cdots \times X_k \to \{1, -1\}$ where for $i \in \{1, \ldots, k\}$, $|X_i| = N$. Let A_f be the tensor of order N representing f. Then* $\text{Disc}(f) \leq W(A_f)/N^k$.

Proof: For $i = 1, \ldots, k$, let C_i be an arbitrary 0/1 tensor which is cylindrical in the i-th dimension. We inductively define related ± 1 tensors $\hat{C}_i$ and T_i. For each $i = 1, \ldots, k$, we define a $(k-1)$-dimensional ± 1 tensor $\hat{C}_i$, where the i-th coordinate is left out. For example, $\hat{C}_1$ is a $(k-1)$-dimensional tensor that depends on the $k-1$ coordinates $x_2, \ldots, x_k$. To simplify notation, we will denote the entries of these tensors by $\hat{C}_i(\vec{x})$, with the understanding that for $\hat{C}_i$, x_i is not used for indexing. For example, $\hat{C}_1(\vec{x})$ represents $\hat{C}_1(x_2, \ldots, x_k)$.

We define $\hat{C}_1$ as follows: $\hat{C}_1(\vec{x}) = \text{sign}(\sum_{x_1} A_f(\vec{x}) \cdot C_2(\vec{x}) \cdots C_k(\vec{x}))$. In other words, to obtain $\hat{C}_1$, we collapse the k dimensional tensor $A_f \circ C_2 \circ \cdots \circ C_k$ to a $k-1$ dimensional tensor by summing the entries of each line along the first dimension and taking the sign of each line sum as an entry of $\hat{C}_1$. (If a given line sums to a negative number, the corresponding entry in $\hat{C}_1$ is -1, otherwise it is 1.) We use $\hat{C}_1$ to define the ± 1 tensor T_1, which is k-dimensional, and cylindrical in the first dimension. T_1 is obtained by taking N copies of $\hat{C}_1$ and using them as the faces of T_1 (along the first dimension).

Assume that $T_1, \ldots, T_{i-1}$ are already defined. We define $\hat{C}_i$ as follows:

$$\hat{C}_i(\vec{x}) = \text{sign}(\sum_{x_i} A_f(\vec{x}) \cdot T_1(\vec{x}) \cdots T_{i-1}(\vec{x}) \cdot C_{i+1}(\vec{x}) \cdots C_k(\vec{x}))$$

Once $\hat{C}_i$ is defined we use it to obtain T_i which is k-dimensional, and cylindrical in the i-th dimension. T_i is obtained by taking N copies of $\hat{C}_i$ and using them as the faces of T_i (along the i-th dimension).

First we show $S(A_f \circ C_1 \circ C_2 \circ \cdots \circ C_k) \leq S(A_f \circ T_1 \circ C_2 \circ \cdots \circ C_k)$. When we replace C_1 by T_1, the contribution of each line of the tensor $A_f \circ C_1 \circ C_2 \circ \cdots \circ C_k$ (along the first dimension) is replaced by a nonnegative value at least as large as the absolute value of the sum of the entries of the original line. To see this, notice that by definition, $\hat{C}_1$ and T_1 contain the signs of the sum of the entries of the corresponding lines of $A_f \circ C_2 \circ \cdots \circ C_k$. (If the sum is 0, we use 1 for the sign.) Obtaining $A_f \circ T_1 \circ C_2 \circ \cdots \circ C_k$ corresponds to multiplying each entry of a given line of $A_f \circ C_2 \circ \cdots \circ C_k$ by the sign of the sum of the entries of that line. Recall that each C_i is cylindrical, thus the lines of C_1 along the x_1 coordinate are constants (all 0 or all 1). If all entries of a given line of C_1 are 0 then the corresponding line of $A_f \circ C_1 \circ C_2 \circ \cdots \circ C_k$ did not contribute anything to the sum, while after the replacement it contributes a nonnegative value. For the lines of C_1 that are constant 1, the contribution of the corresponding line of $A_f \circ C_1 \circ \cdots \circ C_k$ is replaced by its absolute value. Thus, we never decrease the total sum. Similarly, at each inductive step above, we maintain that $S(A_f \circ T_1 \circ \cdots \circ T_{i-1} \circ C_i \circ \cdots \circ C_k) \leq S(A_f \circ T_1 \circ \cdots \circ T_i \circ C_{i+1} \circ \cdots \circ C_k)$.

It follows that $S(A_f \circ C_1 \circ \cdots \circ C_k) \leq S(A_f \circ T_1 \circ \cdots \circ T_k)$. By Lemma 3.1 and the definition of $W(A_f)$ the theorem follows. ∎

The following simple example shows that the discrepancy $\mathrm{Disc}(f)$ can be strictly smaller than $W(A_f)/N^k$. Let $k = 2$, and f be the parity function, i.e., f is 1 if the number of 1's among the input bits is even, and -1 otherwise. Then the discrepancy $\mathrm{Disc}(f) = 1/4$, while $W(A_f)/N^2 = 1$. To see this, note that in the matrix corresponding to the parity function the sum of entries in any rectangle is at most $N^2/4$. On the other hand, it is possible to flip the lines of the matrix so that we obtain the all 1 matrix. (Theorem 8 in [10] appears to claim that $\mathrm{Disc}(f) = W(A_f)/N^k$. However, this seems to be a mistake in notation, and they in fact prove Lemma 3.1.)

The following theorem shows that the discrepancy can not be too much smaller than the bound given by the weight. Thus, using the weight for bounding discrepancy will give good bounds.

Theorem 3.3 *Let $f : X_1 \times \cdots \times X_k \to \{1, -1\}$ be a function where for all i, $|X_i| = N$. Let A_f be the tensor of order N representing f. Then*

$$\mathrm{Disc}(f) \geq W(A_f)/(2^k N^k).$$

Proof: Consider the lines used to generate $W(A_f)$. Partition the entries of A_f into 2^k groups according to whether they were flipped by the lines along each of the k dimensions. Along each dimension the entries flipped by the lines form a cylinder, as do the unflipped entries. Thus the partition splits the entries of A_f into 2^k cylinder intersections. At least one of these cylinder intersections has entry sum with absolute value at least $W(A_f)/2^k$. Using that cylinder intersection in the discrepancy definition gives discrepancy at least $W(A_f)/(2^k N^k)$. ∎

It is known that $W(A) \geq N^{3/2}/\sqrt{2}$ for any N by N matrix A with ±1 entries (see Theorem 5.1 in [1]; see also [6] (c.f. [17])). We show the following extension of that result:

Theorem 3.4 *Let A be any k-dimensional ± 1 tensor of order N. Then*

$$W(A) \geq N^{k-\frac{1}{2}}/\sqrt{2}.$$

Proof: Consider the set of matrices formed by fixing all but the first two dimensions of A. Each of the matrices has weight at least $N^{3/2}/\sqrt{2}$. They do not intersect, so their lines can be flipped independently giving a tensor weight at least $N^{k-2}(N^{3/2}/\sqrt{2})$. ▮

The following standard probabilistic argument shows that there are tensors with weight $O(\sqrt{k}N^{k-\frac{1}{2}})$.

Theorem 3.5 *For any N and k, a $(1-2^{-kN^{k-1}})$-fraction of the k-dimensional tensors with order N have weight $O(\sqrt{k}N^{k-(1/2)})$.*

Proof: Consider the effect of flipping a particular set of lines on a random tensor A. The entries in the flipped tensor A' are all independent random variables with equal probability of being 1 or -1. Thus, using Chernoff's bound,

$$\Pr[\text{sum of entries of } A' \geq cN^{k-(1/2)}] \leq e^{-c^2N^{k-1}/2}.$$

The k players each have N^{k-1} lines available to flip, so there are $2^{k(N^{k-1})}$ sets of lines. Summing this probability over all of these sets gives the probability that one of them generates a tensor with entry sum above $cN^{k-(1/2)}$ is no more than

$$2^{k(N^{k-1})}e^{-c^2N^{k-1}/2} = 2^{N^{k-1}(k-c^2\log e/2)} \; .$$

Consider $c = 2\sqrt{(k/\log e)}$. Then the probability is at most $2^{-kN^{k-1}}$. Thus at least a $(1-2^{-kN^{k-1}})$-fraction of the tensors have weight at most $cN^{k-(1/2)}$. ▮

Proving a similar upper bound on the weight of explicitly defined tensors would yield lower bounds of the form $\Omega(n)$ on multiparty communication complexity, for any number of players. Thus, estimating the weight of tensors can potentially give close to optimal bounds on the discrepancy, and on the multiparty communication complexity of the corresponding functions.

4 Hadamard Tensors

An N by N matrix with ± 1 entries is called a *Hadamard matrix* if the inner product of any two distinct rows is 0. It is equivalent to state the condition for columns: The product of any two distinct rows is 0 if and only if the product of any two distinct columns is 0.

The Hadamard property is invariant under the arbitrary flipping of lines. Thus, Lindsey's lemma (see [11] p. 88) gives the following well known statement:

Lemma 4.1 *For any Hadamard matrix A of order N, $W(A) \leq N^{3/2}$.*

We define the *product of t lines* (along the same dimension) of a tensor as the sum of entries in their entry-wise product. For example, if $l_1, \ldots, l_t$ are lines along the first dimension, then their product is $\sum_{x_1} l_1(x_1) \cdots l_t(x_1)$.

Let A be a k-dimensional tensor of order N with ± 1 entries. For each of the first $k-1$ dimensions $i = 1, \ldots, k-1$, choose two distinct indices $y_i, z_i \in X_i$. Picking exactly one of y_i or z_i for each $i = 1, \ldots, k-1$ gives a point in $X_1 \times \cdots \times X_{k-1}$, and each such point specifies a line of A along the last coordinate x_k. There are 2^{k-1} possible choices for the selection described above, and since for each $i = 1, \ldots, k-1$, $y_i \neq z_i$, we get 2^{k-1} distinct lines this way. We say that the tensor A is Hadamard, if the product of any 2^{k-1} lines chosen in this way is 0. More formally, we define Hadamard tensors as follows:

Definition 4.1 *Let A be a k-dimensional tensor of order N with ± 1 entries. We say that A is a* Hadamard *tensor if for any $y_1, z_1 \in X_1, \ldots, y_{k-1}, z_{k-1} \in X_{k-1}$ such that $y_i \neq z_i$ for $i = 1, \ldots, k-1$, the following holds:*

$$\sum_{x_k \in X_k} \prod_{x_1 \in \{y_1, z_1\}, \ldots, x_{k-1} \in \{y_{k-1}, z_{k-1}\}} A(x_1, x_2, \ldots, x_k) = 0.$$

When $k = 2$ this definition is identical to the definition of Hadamard matrices.

Since the k-th coordinate plays a special role in the above definition of a Hadamard tensor, we can say that the definition is given with respect to the k-th dimension. Just as the Hadamard condition on the rows of a matrix implies the condition on the columns, we show the following for tensors:

Lemma 4.2 *Let A be a k-dimensional Hadamard tensor with respect to the i-th dimension for some $1 \leq i \leq k$. Then for any $j \neq i, 1 \leq j \leq k$, A is a k-dimensional Hadamard tensor with respect to the j-th dimension.*

Proof: Choose for all other dimensions m different from i and j two distinct indices y_m and z_m. Consider the matrix formed with columns indexed by dimension i and rows by dimension j where each entry is the product of the 2^{k-2} entries of A corresponding to that coordinate and the choices of y_m's and z_m's. Since the property holds for dimension i, we know that the inner product of any two of the rows of this matrix is 0. Thus it is a Hadamard matrix, and so the inner product of any two columns is 0. Since this is true for all choices of y_m's and z_m's, A is a Hadamard tensor with respect to the j-th dimension. ▮

Here is a different way to view the Hadamard property.

Definition 4.2 *A three dimensional tensor is Hadamard if the entry-wise product of any two distinct faces along the same dimension is a Hadamard matrix. A k-dimensional tensor for $k > 3$ is Hadamard if the entry-wise product of any two distinct faces along the same dimension is a Hadamard tensor in $k-1$ dimensions.*

Lemma 4.3 *Definition 4.1 and Definition 4.2 are equivalent.*

Proof: Let A^{x_i} denote the face of A obtained by fixing the i-th coordinate to the value x_i.

The statement follows by observing that for $k \geq 3$,

$$\sum_{x_k \in X_k} \prod_{x_1 \in \{y_1, z_1\}, \ldots, x_{k-2} \in \{y_{k-2}, z_{k-2}\}} A^{y_{k-1}} \circ A^{z_{k-1}}(x_1, \ldots, x_{k-2}, x_k) =$$
$$\sum_{x_k \in X_k} \prod_{x_1 \in \{y_1, z_1\}, \ldots, x_{k-1} \in \{y_{k-1}, z_{k-1}\}} A(x_1, \ldots, x_{k-2}, x_{k-1}, x_k),$$

and using induction on k. ▮

Next we prove that the Hadamard property of tensors is invariant under arbitrary flipping of lines.

Lemma 4.4 *Let A' be a tensor obtained from a Hadamard tensor A by flipping a collection of lines. Then A' is a Hadamard tensor.*

Proof: This follows by induction from the characterization of Hadamard tensors given by Definition 4.2. The result holds for matrices since after flipping a row or column any row or column product that was 0 remains 0. Suppose the result holds for tensors of dimension $k-1$. Consider any face product $A^{y_i} \circ A^{z_i}$ of a k-dimensional Hadamard tensor A. Flipping a line of A may miss A^{y_i} and A^{z_i} entirely, intersect both in one entry, or flip an entire line of A^{y_i} or A^{z_i}. In the first case the face product is unaffected. In the second case the face product is unchanged since the corresponding entry is negated twice. In the third case the face product has a line flipped. By the induction hypothesis this is still a Hadamard tensor. ▮

4.1 The Discrepancy of Hadamard Tensors

In light of Theorem 3.2, we can prove upper bounds on the discrepancy of any tensor A by proving upper bounds on $W(A)$. Let $W_k(N)$ denote the largest possible value of $W(A)$ if A is a k-dimensional Hadamard tensor of order N.

Lemma 4.5 *Let A be a k-dimensional Hadamard tensor of order N. Then*

$$(W(A))^2 \leq N^{2k-1} + N^{k+1}(W_{k-1}(N)).$$

Proof: Let A' be the k-dimensional tensor obtained from A by flipping a collection of lines that achieves maximal excess, that is $W(A) = S(A')$. By Lemma 4.4, A' is a Hadamard tensor, and by Definition 4.2 the entry-wise product of any two distinct faces of A' is a Hadamard tensor in $k-1$ dimensions. Thus, we have the following estimates (using the Cauchy-Schwartz inequality).

$$\begin{aligned}
(S(A'))^2 &= \left(\sum_{\vec{x}\in X_1\times\cdots\times X_k} A'(\vec{x})\right)^2 \\
&\leq N^{k-1}\sum_{x_1,\ldots,x_{k-1}}\left(\sum_{x_k} A'(\vec{x})\right)^2 \\
&= N^{k-1}\left(N^k + \sum_{i\neq j}\sum_{x_1,\ldots,x_{k-1}} A'(x_1,\ldots,x_{k-1},i)A'(x_1,\ldots,x_{k-1},j)\right) \\
&\leq N^{k-1}(N^k + (N^2-N)(W_{k-1}(N))) \\
&\leq N^{2k-1} + N^{k+1}(W_{k-1}(N))). \quad \blacksquare
\end{aligned}$$

Theorem 4.6 *Let A be a k-dimensional Hadamard tensor of order N. Then $W(A) \leq \phi N^{k-(1/2^{k-1})}$ where $\phi = (1+\sqrt{5})/2$.*

Proof: Let A be a three dimensional Hadamard tensor with order N. Using Lemma 4.5 and Lemma 4.1,

$$\begin{aligned}
W(A) &\leq (N^2(N^3 + (N^2-N)(N^{3/2})))^{1/2} \\
&\leq N(2N^{7/2})^{1/2} \\
&\leq \sqrt{2}N^{3-(1/4)} \\
&\leq \phi N^{3-(1/4)}.
\end{aligned}$$

Suppose the theorem holds for all Hadamard tensors of dimension $k-1$. Let A be a k-dimensional Hadamard tensor.

$$\begin{aligned}
W(A) &\leq (N^{k-1}(N^k + (N^2 - N)(\phi N^{(k-1)-(1/2^{k-2})})))^{1/2} \\
&\leq (N^{k-1}((1+\phi)(N^{(k+1)-(1/2^{k-2})})))^{1/2} \\
&= ((1+\phi)N^{2k-(1/2^{k-2})})^{1/2} \\
&= \sqrt{1+\phi}N^{k-(1/2^{k-1})} \\
&= \phi N^{k-(1/2^{k-1})}. \quad \blacksquare
\end{aligned}$$

Theorem 4.7 *Let* $f : (\{0,1\}^n)^k \to \{1,-1\}$ *be a function represented by a Hadamard tensor. Then* $\mathrm{Disc}(f) \leq \phi N^{-1/2^{k-1}}$ *where* $\phi = (1+\sqrt{5})/2$*.*

Proof: This follows directly from Theorem 4.6 and Theorem 3.2. $\blacksquare$

By the results of [2] (see Lemma 2.1) this yields the following:

Theorem 4.8 *Let* $f : (\{0,1\}^n)^k \to \{1,-1\}$ *be a function represented by a Hadamard tensor. Then* $C(f) = \Omega(n/2^k)$*, and* $C_\epsilon(f) = \Omega((n/2^k) + \log_2 \epsilon)$*.*

4.2 Constructions of Hadamard Tensors

Let $x_1, ..., x_k$ be n-bit strings. Consider each of these strings as an element of the finite field $\mathrm{GF}(2^n)$, representing the field elements as univariate polynomials over $\mathrm{GF}(2)$ modulo a fixed irreducible polynomial of degree n. (In this representation the i-th bit ($0 \leq i \leq n-1$) of a given n-bit string indicates whether the corresponding polynomial $p(a)$ contains the term a^i.)

Let χ_S stand for the function obtained by raising -1 to the parity of the bits with coordinates in S, such that χ_S is 1 when the parity is even, and -1 when the parity is odd. It is not hard to see that for any $x, y \in \{0,1\}^n$,

$$\chi_S(x)\chi_S(y) = \chi_S(x+y), \tag{1}$$

where $+$ represents addition in $\mathrm{GF}(2^n)$. (In fact the χ_S are the additive characters of $\mathrm{GF}(2^n)$.) By the definition of χ_S, $\chi_S(x)\chi_S(y) = \chi_S(x \oplus y)$, viewing x and y as strings and taking bitwise XOR, which is the same as $\chi_S(x+y)$ using addition in the field.

Definition 4.3 *Given a function $f : \{0,1\}^n \to \{1,-1\}$, we define the function $\mathrm{FFM}_f^{n,k} : (\{0,1\}^n)^k \to \{1,-1\}$ by*

$$\mathrm{FFM}_f^{n,k}(x_1, \ldots, x_k) = f(x_1 \cdot x_2 \cdot \ldots \cdot x_k),$$

where $x_1 \cdot x_2 \cdot \ldots \cdot x_k$ denotes the product of the field elements $x_1, \ldots, x_k$, and f is applied to the n-bit string representing the resulting field element.

For $S \subseteq \{0,1,\ldots,n-1\}$, we denote by $\mathrm{FFM}_S^{n,k}$ the function $\mathrm{FFM}_{\chi_S}^{n,k}$.

"FFM" is an abbreviation for "Finite Field Multiplication".

Theorem 4.9 *For every $\emptyset \neq S \subseteq \{0,1,\ldots,n-1\}$, the k-dimensional tensor associated with $\mathrm{FFM}_S^{n,k}$ is Hadamard.*

Proof: We need the following observation: For any t and for any $y_1, z_1, \ldots, y_t, z_t \in \mathrm{GF}(2^n)$ with $y_1 \neq z_1, \ldots, y_t \neq z_t$,

$$\sum_{x_1 \in \{y_1,z_1\},\ldots,x_t \in \{y_t,z_t\}} x_1 x_2 \cdots x_t \neq 0.$$

To see this, note that

$$\sum_{x_1 \in \{y_1,z_1\},\ldots,x_t \in \{y_t,z_t\}} x_1 x_2 \cdots x_t = \prod_{i=1}^{t} (y_i + z_i).$$

For distinct y_i and z_i, $y_i + z_i$ is nonzero since in $\mathrm{GF}(2^n)$ each element is its own additive inverse.

Consider the following sum from Definition 4.1:

$$\sum_{x_k} \prod_{x_1 \in \{y_1,z_1\},\ldots,x_{k-1} \in \{y_{k-1},z_{k-1}\}} \chi_S(x_1 x_2 \cdots x_k)$$

By (1) this is the same as

$$\sum_{x_k} \chi_S \left(\sum_{x_1 \in \{y_1,z_1\},\ldots,x_{k-1} \in \{y_{k-1},z_{k-1}\}} x_1 x_2 \cdots x_k \right)$$
$$= \sum_{x_k} \chi_S \left(x_k \sum_{x_1 \in \{y_1,z_1\},\ldots,x_{k-1} \in \{y_{k-1},z_{k-1}\}} x_1 x_2 \cdots x_{k-1} \right)$$

As shown above, the inner sum evaluates to a non-zero field element, so for some fixed non-zero w, we obtain $\sum_{x_k} \chi_S(x_k w) = \sum_{x_k} \chi_S(x_k) = 0$ ∎

Using Theorem 4.8 we immediately obtain the following:

Theorem 4.10 *For every $\emptyset \neq S \subseteq \{0, 1, \ldots, n-1\}$, $C(\mathrm{FFM}_S^{n,k}) = \Omega(n/2^k)$, and $C_\epsilon(\mathrm{FFM}_S^{n,k}) = \Omega(n/2^k + \log_2 \epsilon)$.*

Although all finite fields of order 2^n are isomorphic, it is necessary to specify exactly which one is being used for obtaining explicit constructions of Boolean functions this way. The deterministic algorithm developed by Shoup [24] can be used to construct an irreducible polynomial of degree n for any given n. Thus the family of Boolean functions associated with the tensors FFM_S belongs to the complexity class P. Note also that the polynomial $x^n + x^{n/2} + 1$ is irreducible over GF(2) when n is of the form $n = 2 \cdot 3^m$ (Theorem 1.1.28 in [18]).

Theorem 4.11 *The Boolean version of the function given by $\mathrm{FFM}_S^{n,k}$ for $S \subseteq \{0, 1, \ldots, n-1\}$ can be computed by polynomial size unbounded fan-in circuits with* AND, OR, NOT *and* MOD_2 *gates, in depth* $\log k$.

Proof: We use that multiplication in a finite field is bilinear. That is, for every n there is a set of n binary matrices $A^{(0)}, \ldots, A^{(n-1)}$, each of size $n \times n$, such that for any two field elements x and y viewed as binary vectors, the t-th bit of the product xy is given by $xA^{(t)}y$. (All operations are in GF(2).)

Multiplication of the finite field elements using these matrices can be done in constant depth, and k multiplications are sufficient, giving the $\log k$ depth. ∎

4.3 Relaxations of the Hadamard Property

Chung [9] and Raz [22] state a sufficient condition for a function to have $\Omega(n/c^k)$ multiparty communication complexity, generalizing the method of [2]. We can show that satisfying the condition of [9] and [22] is related to being nearly Hadamard in the following, relaxed sense: Instead of requiring that all the products of the 2^{k-1}-tuples of lines selected according to the Hadamard definition are 0, require that the products are small on average; e.g. that the sum of the squares of the line products is small. The tensor corresponding to the "generalized inner product" (GIP) function of [2] is nearly Hadamard in this relaxed sense, but it is not Hadamard.

For the tensor corresponding to the "quadratic character of the sum of coordinates" (QCS) function of [2] we can show that each (nontrivial) product of the selected 2^{k-1} tuples of lines is small (at most $2^k\sqrt{N}$).

Definition 4.4 *A matrix is* d-Hadamard *if the inner product of any two distinct rows is at most d. A tensor is* d-Hadamard *if the sum of the entries of any vector formed by taking the entry-wise product of lines as in Definition 4.1 is at most d.*

Similarly to the equivalence between the two definitions of Hadamard tensors, the following holds:

Lemma 4.12 *For $k \leq 3$, a k-dimensional tensor is d-Hadamard if and only if the entry-wise product of any two of its faces along the same dimension is a d-Hadamard tensor in $k-1$ dimensions.*

The proof is identical to the proof of Lemma 4.3.

Theorem 4.13 *Let A be a d-Hadamard matrix of order N. Then*

$$W(A) \leq N^{3/2}(1+d)^{1/2}.$$

Proof: Let A' be a matrix formed by flipping lines of A. We will apply the Cauchy-Schwartz inequality to an expression for the excess of A', eventually simplifying to an expression in terms of the products of the rows of A.

$$\sum_{i=1}^{N}\sum_{j=1}^{N} A'(i,j) \leq [N\sum_{i=1}^{N}(\sum_{j=1}^{N} A'(i,j))^2]^{1/2} \leq [N\sum_{j_1=1}^{N}\sum_{j_2=1}^{N}\sum_{i=1}^{N} A'(i,j_1)A'(i,j_2)]^{1/2}.$$

The sum over i, j_1, and j_2 is the sum over all choices of two rows of the matrix of the inner product of those two rows. Note that flipping rows or columns of A' does not affect any of these inner products except possibly in sign, so this is no more than

$$[N\sum_{j_1=1}^{N}\sum_{j_2=1}^{N}|\sum_{i=1}^{N} A(i,j_1)A(i,j_2)|]^{1/2}.$$

So the excess of A' is at most

$$[N(N^2+(N^2-N)d)]^{(1/2)} \leq N^{3/2}(1+d)^{1/2}. \quad \blacksquare$$

Theorem 4.14 *Let A be an d-Hadamard tensor in k dimensions. Then*

$$W(A) \leq \phi N^{k-1/2^{k-1}}(1+d)^{1/2^{k-1}}$$

The corresponding induction used in the proof of Theorem 4.6 gives the proof.

Theorem 4.15 *Let $f : (\{0,1\}^n)^k \rightarrow \{1,-1\}$ be a function represented by a d-Hadamard tensor A_f. Then $D(f) \geq (\log(N/(\phi(1+d))))/2^{k-1}$.*

Proof: This follows from Theorems 4.14 and 3.2 and Lemma 2.1. ∎

To show QCS is d-Hadamard we need a technical lemma.

Lemma 4.16 *Let p be any odd prime. Let $y_1, z_1, \ldots, y_k, z_k$ be elements of Z_p where for each i, $y_i \neq z_i$. Let $S_1 = \{y_1, z_1\}$. For $i > 1$, let $SY_i = y_i + S_{i-1}$, $SZ_i = z_i + S_{i-1}$, and $S_i = (SY_i \cup SZ_i) \setminus (SY_i \cap SZ_i)$. In other words, S_i contains elements formed by adding either y_i or z_i to an element of S_{i-1} but not those formed both ways. Then S_k is nonempty.*

Proof: Suppose S_{i-1} is nonempty and that $SY_i = SZ_i$. Let $s \in SY_{i-1}$. We know $s + y_i \in SY_i$ and $s + y_i \in SZ_i$. This means $s + y_i - z_i \in S_{i-1}$. Thus $s + (y_i - z_i) + y_i \in SY_i$. So $s + (y_i - z_i) + y_i \in SZ_i$. So $s + 2(y_i - z_i) \in S_{i-1}$. And for any n, $s + n(y_i - z_i) \in S_{i-1}$. But since $y_i - z_i$ is not 0, it generates Z_p, and so all elements of Z_p are in S_{i-1}. This is a contradiction since there are an odd number of elements of Z_p, but our construction builds sets with an even number of elements in each S_i. ∎

Definition 4.5 *[2] Let $\mathrm{QCS}_{p,k} : \{0,1,\ldots,p-1\}^k \rightarrow \{1,-1\}$ for prime p be the function which maps $(x_1,\ldots,x_k)$ to 1 if $x_1 + \cdots + x_k$ is a quadratic residue mod p and -1 otherwise.*

Theorem 4.17 $\mathrm{QCS}_{p,k}$ *is $(2^k\sqrt{p})$-Hadamard.*

Proof: Choose $y_1, z_1, \ldots y_{k-1}, z_{k-1} \in \{0,1,\ldots,p-1\}$ where for each i, $y_i \neq z_i$. Consider the following:

$$\sum_{x_k} \prod_{x_1 \in \{y_1,z_1\}} \prod_{x_2 \in \{y_2,z_2\}} \cdots \prod_{x_{k-1} \in \{y_{k-1},z_{k-1}\}} \mathrm{QCS}_{p,k}(x_1, x_2, \ldots, x_k)$$

Let $\chi : \{0,1,\ldots,p-1\} \rightarrow \{1,0,-1\}$ be the quadratic character function mapping 0 to 0, other squares to 1, and everything else to -1. It is a multiplicative character on GF(p), meaning that $\chi(x)\chi(y) = \chi(xy)$ for any elements of GF(p). The QCS function on input $x_1,\ldots,x_k$ is the same as

$\chi(x_1 + \cdots + x_k)$ for all inputs except those where $x_1 + \cdots + x_k$ is 0. However for any line, this occurs exactly once. So the sum above is no more than

$$2^{k-1} + \sum_{x_k} \prod_{x_1 \in \{y_1, z_1\}} \prod_{x_2 \in \{y_2, z_2\}} \cdots \prod_{x_{k-1} \in \{y_{k-1}, z_{k-1}\}} \chi(x_1 + \cdots + x_k)$$

where the first term accounts for the entries that are 0 after switching from QCS to χ. Since χ is a multiplicative character, the summation is a summation over all values of x_k of χ applied to a polynomial in x_k of degree 2^{k-1}. According to Weil's character sum estimate ([25], see also in [2]), if we can show that this polynomial is not a square of some smaller polynomial, then the sum is bounded by $(2^{k-1} - 1)\sqrt{p}$. The polynomial we are considering is already factored into terms of the form $(x_k + a)$. Since factoring is unique, this means that the only way this polynomial is a square is if after removing terms of the form $(x_k + a)^2$, no terms are left. Using Lemma 4.16 this cannot happen. So the original sum is bounded by $2^{k-1} + (2^{k-1} - 1)\sqrt{p}$ which is at most $2^k\sqrt{p}$. ∎

Theorem 4.18 *Let* $k \le \log(p^{1/4} - p^{-1/2})$. *Then* $C(\mathrm{QCS}_{p,k}) = \Omega((\log p)/2^k)$.

Proof: The statement follows by Theorem 4.17 and Theorem 4.15. ∎

Note that the property we prove for QCS is stronger than the condition required in [9, 22], but weaker than the Hadamard property.

Raz [22] considered the function defined as follows: each part of the input $x_i \in \{0,1\}^n$ is interpreted as a $\sqrt{n}$ by $\sqrt{n}$ matrix with $0, 1$ entries. The function is defined by the bit in the upper left corner of the matrix obtained by taking the product (over GF(2)) of the k matrices. Raz [22] proved that this function has (probabilistic) k-party communication complexity $\Omega(\sqrt{n}/2^k)$. The tensor associated with this function is not Hadamard, but we can show that it contains a subtensor of order $2^{\sqrt{n}}$ which is Hadamard. Thus, our methods give $\Omega(\sqrt{n}/2^k)$ lower bounds on the k-party communication complexity of the function.

Grolmusz [13] showed an $O(kn/2^k)$ upper bound on the multiparty communication complexity of GIP. There are no similar upper bounds known for any of the functions that we presented as examples of Hadamard tensors. The examples of Hadamard tensors we give and the QCS function from [2] are candidates for having $\Omega(n/poly(k))$ multiparty communication complexity.

References

[1] N. Alon, J. H. Spencer, “The Probabilistic Method”, Wiley-Interscience, 2000.

[2] L. Babai, N. Nisan, M. Szegedy, “Multiparty Protocols, Pseudorandom Generators for Logspace, and Time-Space Trade-Offs”, *JCSS*, 45(2):204-232, 1992.

[3] L. Babai, T. P. Hayes, P. G. Kimmel, “The Cost of the Missing Bit: Communication Complexity with Help”, *Proc. 30th ACM STOC*, 673-682, 1998.

[4] D. Barrington, “Bounded-width polynomial size branching programs recognize exactly those languages in NC_1”, *JCSS*, 38(1):150-164, 1989.

[5] R. Beigel, J. Tarui, “On ACC”, *Proc. 32nd IEEE FOCS*, 783-792, 1991.

[6] M. R. Best, “The Excess of a Hadamard Matrix”, *Indag. Math.*, 39(5):357-361, 1977.

[7] A. Chandra, M. Furst, R. Lipton: “Multiparty protocols”, *Proc. 15th ACM STOC*, 94-99, 1983.

[8] B. Chor, O. Goldreich, “Unbiased Bits from Sources of Weak Randomness and Probabilistic Communication Complexity”, *SIAM J. Comp.* 17:230-261, 1988.

[9] F. Chung, “Quasi-Random Classes of Hypergraphs”, *Random Structures and Algorithms*, 1(4):363-382, 1990.

[10] F. Chung, P. Tetali, “Communication complexity and quasi randomness”, *SIAM J. Discrete Math.*, 6(1):110-123, 1993.

[11] P. Erdős, J. H. Spencer, “Probabilistic methods in combinatorics”, Academic Press, 1974.

[12] J. Ford, A. Gál, “Hadamard tensors and lower bounds on multiparty communication complexity”, *Proc. ICALP’05*, 1163-1175, 2005.

[13] V. Grolmusz, “The BNS Lower Bound for Multi-Party Protocols is Nearly Optimal”, *Information and Computation*, 112:51-54, 1994.

[14] J. Håstad, M. Goldmann, "On the power of small depth threshold circuits", *Computational Complexity*, 113-129, 1991.

[15] E. Kushilevitz, N. Nisan, "Communication complexity", Cambridge, 1997.

[16] R. Lidl, H. Niederreiter, "Introduction to finite fields and their applications", Cambridge, 1994.

[17] J. H. van Lint, R. M. Wilson, "A Course in Combinatorics", Cambridge, 1992.

[18] J.H. van Lint, "Introduction to Coding Theory", Springer-Verlag, 1998.

[19] N. Nisan, A. Wigderson, "Rounds in communication complexity revisited" *SIAM J. Comp.*, 22:211-219, 1993.

[20] P. Pudlák, "Unexpected upper bounds on the complexity of some communication games", *Proc. ICALP'94*, 1-11, 1994.

[21] P. Pudlák, V. Rödl, J. Sgall, "Boolean circuits, tensor ranks and communication complexity", *SIAM J. Comp.*, 26:605-633, 1997.

[22] R. Raz, "The BNS-Chung criterion for multiparty communication complexity", *Computational Complexity*, 9(2):113-122, 2000.

[23] J. Spencer, "Ten lectures on the probabilistic method", Soc. for Industrial and Applied Math., 1987.

[24] V. Shoup, "New Algorithms for Finding Irreducible Polynomials over Finite Fields", *Mathematics of Computation*, 54:435-447, 1990.

[25] A. Weil, "On some exponential sums", *Proc. of the National Academy of Sciences of the USA*, 34 (5), pp. 204-207, 1948.

[26] A. Yao, "Some complexity questions related to distributed computing", *Proc. 11th ACM STOC*, 209-213, 1979.

[27] A. Yao, "On ACC and threshold circuits", *Proc. 31st IEEE FOCS*, 619-627, 1990.