Efficient Randomized Protocol for every Karchmer-Wigderson Relation with Three Rounds — An Alternative Proof

Or Meir*
February 12, 2018

Abstract

One of the important challenges in circuit complexity is proving strong lower bounds for constant-depth circuits. One possible approach to this problem is to use the framework of Karchmer-Wigderson relations: Karchmer and Wigderson [KW90] observed that for every Boolean function $f$ there is a corresponding communication problem $KW_f$, called the Karchmer-Wigderson relation of $f$, whose deterministic communication complexity is tightly related to the depth complexity of $f$. In particular, if we could prove that every deterministic constant-round protocol for $KW_f$ must transmit at least $c$ bits, then this would imply a lower bound of $2^{Ω(c)}$ on the size of constant-depth circuits computing $f$.

Jowhari, Ságlam, and Tardos [JST11] showed that there is a randomized three-round protocol that solves every Karchmer-Wigderson relation $KW_f$ by transmitting only $O(\log n)$ bits. This means that if we wish to use Karchmer-Wigderson relations in order to prove super-polynomial lower bounds for constant-depth circuits, then we cannot use techniques that work against randomized protocols.

The protocols of [JST11] are based on a result from the sketching literature. In this note, we replace this tool with a simple protocol that is based on the Hamming code. This results in protocols that require no background in sketching to understand. The aforementioned simple protocol is a deterministic two-round protocol that solves a special case of Karchmer-Wigderson relations, and may be of independent interest.

Clarification: Following an earlier revision of this work, it has been brought to our attention that a very similar protocol to ours has already appeared in the M.Sc. thesis of Mert Ságlam [S11].

1 Introduction

Proving circuit lower bounds is a central challenge of complexity theory. Unfortunately, proving even super-linear lower bounds for general circuits seems to be beyond our reach at this stage. In order to make progress and develop new proof techniques, much of the current research focuses on proving lower bounds for restricted models of circuits. One of the simplest restricted models that are not yet fully understood is circuits of constant depth (with unbounded fan-in).

By a standard counting argument, we know that there exists a non-explicit function that requires such circuits of size $Ω(2^n)$. On the other hand, the strongest lower bound we have for an explicit function $\text{Ajt83, FSS84, Hås86}$ says that circuits of depth $d$ computing the parity of $n$ bits must be of size $2^{Ω(n^{1/(d-1)})}$. Hence, while strong lower bounds are known in this model, there is still a

*Department of Computer Science, University of Haifa, Haifa 3498838, Israel. ormeir@cs.haifa.ac.il. Partially supported by the Israel Science Foundation (grant No. 1445/16).
significant gap in our understanding. In particular, it is an outstanding open problem to prove a lower bound of \(2^{\Omega(n)}\) even for depth-3 circuits computing an explicit function (or, indeed, any lower bound that is better than \(2^{\Omega(\sqrt{n})}\)).

One possible approach for attacking this problem is a framework due to Karchmer and Wigderson \[KW90\]. This framework was originally developed for proving lower bounds on the depth of circuits with bounded fan-in. Given a function \(f\), we define the depth complexity of \(f\) to be the smallest depth of a circuit with fan-in 2 that computes \(f\). Karchmer and Wigderson observed that for every Boolean function \(f\) there is a corresponding communication problem \(KW_f\), called the Karchmer-Wigderson relation of \(f\), such that the deterministic communication complexity of \(KW_f\) is exactly equal to the depth complexity of \(f\). Hence, one can prove lower bounds on the depth complexity of a function \(f\) by proving lower bounds on the communication complexity of \(KW_f\). This approach has proved very fruitful in the setting of monotone circuits \[KW90, GS91, RW92, KRW95\].

The framework of Karchmer and Wigderson could also be used to prove lower bounds on constant-depth circuits with unbounded fan: it is implicit in the work of \[KW90\] that lower bounds on the deterministic communication complexity of constant-round protocols for \(KW_f\) imply lower bounds for constant-depth circuits. More specifically, if every deterministic \(r\)-round protocol for \(KW_f\) must transmit at least \(c\) bits, then every depth-\(r\) circuit (with unbounded fan-in) that computes \(f\) must be of size at least \(2^{c/r}\). Hence, if we could find an explicit function \(f\) such that every constant-round protocol for \(KW_f\) must transmit \(\Omega(n)\) bits, we would obtain a lower bound of \(2^{\Omega(n)}\) on the size of circuits computing \(f\).

Soon after the introduction of Karchmer-Wigderson relations, Håstad observed a severe limitation of this framework (see \[RW89\]): there is a randomized protocol that solves every Karchmer-Wigderson relation by transmitting \(O(\log n)\) bits. This means that if one wishes to use Karchmer-Wigderson relations in order to prove super-logarithmic lower-bounds on depth complexity, then one has to use proof techniques that cannot prove lower bounds against randomized protocols. Since the most powerful techniques in the field of communication complexity are effective against randomized protocols, this limitation makes the use of Karchmer-Wigderson relations quite difficult.

Håstad’s protocol uses a logarithmic number of rounds, so it is not clear a priori that this limitation applies to proving constant-depth lower bounds. Jowhari, Ságlam, and Tardos \[JST11\] showed that a similar limitation applies in the setting of constant-depth lower-bounds on depth complexity, then one has to use proof techniques that cannot prove lower bounds against randomized protocols. Specifically, they showed that there is a randomized three-round protocol that solves every Karchmer-Wigderson relation \(KW_f\) by transmitting only \(O(\log n)\) bits. This means that proof techniques that are effective against randomized protocols cannot prove super-polynomial lower bounds for constant-depth circuits.

The protocol of \[JST11\] is based on a result from the on-line sketching literature. In particular, the crux of their protocols is a lemma that says that a random sparse real vector can be reconstructed from a random linear sketch. In this note, we show that it is possible to replace this lemma with a simple protocol, presented in Section \[JST11\] that solves a special case of Karchmer-Wigderson relations. We then present in Sections \[JST11\] and \[JST11\] reductions of the general case to the latter special case using the same underlying ideas of \[JST11\]. The resulting protocols have the advantage that they can be understood without background in sketching. We note that the protocol for the aforementioned special case is a deterministic two-round protocol and may be of independent interest.

**Clarification.** Following an earlier revision of this work, it has been brought to our attention that a very similar protocol to ours has already appeared in the M.Sc. thesis of Mert Ságlam \[S11\].

\[1\] It follows more explicitly from the discussions of Karchmer-Wigderson relations in \[Raz90, KKN95\], and a similar observation was also made in \[KPPY84\].
2 Preliminaries and the Result of [JST11]

For $n \in \mathbb{N}$, we denote $[n] \overset{\text{def}}{=} \{1, \ldots, n\}$. For a string $x \in \{0, 1\}^n$ and a set of coordinates $S \subseteq [n]$, we denote by $x|_S$ the projection of $x$ to the coordinates in $S$. We use the standard definitions of communication complexity — see the book of Kushilevitz and Nisan [KN97] for more details.

Our proof uses the Hamming code, which we present next. Given two strings $x, y \in \{0, 1\}^n$, the (Hamming) distance between $x$ and $y$ is the number of coordinates on which they differ. Given a string $x \in \{0, 1\}^n$ and $r \in \mathbb{N}$, the Hamming ball of radius $r$ around $x$ is the set of all strings whose Hamming distance from $x$ is at most $r$. The Hamming code is a partition of $\{0, 1\}^n$ to balls of radius 1, and it exists for every $n \in \mathbb{N}$ for which $n + 1$ is a power of 2 (see, e.g., Lecture 2 in [Sud01] for the construction of the Hamming code).

We also use the following two simple facts from probability theory.

Claim 1. Consider $n$ independent Bernoulli trials with success probability $p$ such that $\frac{1}{4n} \leq p \leq \frac{1}{2n}$. The probability that the number of successes is exactly 1 is at least $\frac{1}{8}$.

Proof. The probability that there is exactly one success is

$$n \cdot p \cdot (1 - p)^{n-1} \geq n \cdot \frac{1}{4 \cdot n} \cdot \left(1 - \frac{1}{2 \cdot n}\right)^n \geq \frac{1}{4} \cdot \left(1 - \frac{n}{2 \cdot n}\right) = \frac{1}{8},$$

as required. ■

Claim 2. Consider $n$ independent Bernoulli trials with success probability $p$ such that $p \leq \frac{1}{2n}$. Conditioned on having at least one success, the probability that the number of successes is exactly 1 is at least $\frac{3}{8}$.

Proof. Let $p, n$ be as in the claim. For every number $k > 0$, the probability that there are exactly $k$ successes is at most

$$\left(\frac{n}{k}\right) \cdot p^k \cdot (1 - p)^{n-k} \leq \frac{n}{k!} \cdot p \cdot \left(\frac{1}{2n}\right)^{k-1} \leq n \cdot p \cdot \frac{1}{4^{k-2}},$$

where the second inequality is due to the fact that $k! \geq 2^{k-1}$. Therefore, the probability of having at least one success is at most

$$n \cdot p \cdot \sum_{k=1}^{\infty} \frac{1}{4^{k-2}} \leq n \cdot p \cdot 4 \cdot \sum_{k=0}^{\infty} 1 \cdot 4^k \leq \frac{4}{3} \cdot n \cdot p.$$

On the other hand, the probability of exactly one success is

$$n \cdot p \cdot (1 - p)^{n-1} \geq n \cdot p \cdot (1 - \frac{1}{2 \cdot n})^{n-1} \geq n \cdot p \cdot (1 - \frac{n-1}{2 \cdot n}) \geq \frac{1}{2} \cdot n \cdot p.$$

Combining the two inequalities, it follows that the probability of exactly one success conditioned on a non-zero number of successes is at least

$$\frac{1}{2} \cdot n \cdot p \geq \frac{3}{8},$$

as required. ■
2.1 Karchmer-Wigderson relations and the universal relation

Let \( f : \{0,1\}^n \to \{0,1\} \) be a Boolean function. The Karchmer-Wigderson relation \( KW_f \) is defined as follows: Alice gets an input \( x \in f^{-1}(0) \), and Bob gets as input \( y \in f^{-1}(1) \). Clearly, it holds that \( x \neq y \). The goal of Alice and Bob is to find a coordinate \( i \) such that \( x_i \neq y_i \). Note that there may be more than one possible choice for \( i \), which means that \( KW_f \) is a relation rather than a function. As noted above, Karchmer and Wigderson observed that the communication complexity of \( KW_f \) is exactly equal to the depth complexity of \( f \).

In order to study Karchmer-Wigderson relations, Karchmer, Raz, and Wigderson defined the universal relation, which is the following computational problem: Alice and Bob take as inputs two distinct strings \( x, y \in \{0,1\}^n \) respectively, and their goal is to find a coordinate \( i \in [n] \) such that \( x_i \neq y_i \). It is easy to see that every Karchmer-Wigderson relation reduces to the universal relation. Thus, in order to prove the results of [JST11], it suffices to devise an efficient randomized protocol for the universal relation. The main result of [JST11] that we reprove in this note can now be stated as follows.

**Theorem 3.** There is a randomized three-round protocol that solves the universal relation over \( n \) bits with probability at least \( \frac{2}{3} \) by transmitting at most \( O(\log n) \) bits.

In addition to the foregoing result, [JST11] also proved that there is a slightly less efficient two-round protocol.

**Theorem 4.** There is a randomized two-round protocol that solves the universal relation over \( n \) bits with probability at least \( \frac{2}{3} \) by transmitting at most \( O(\log^2 n) \) bits.

We note that [JST11] also proved a lower bound of \( \Omega(\log^2 n) \) on two-round protocols, so the gap between two-round and three-round protocols is essential (their lower bound was subsequently improved by [KNP+17]). We also note that the deterministic communication complexity of the universal relation is between \( n + 1 \) and \( n + 2 \) [TZ97].

As explained in the introduction, the crux of both protocols is a deterministic protocol that solves a special case of the universal relation. Specifically, this protocol solves the universal relation in the special case where the inputs differ only on one coordinate (i.e., the Hamming distance between the inputs is 1), and is presented in Section 3. We then use this protocol to prove Theorems 3 and 4.

We start by proving Theorem 4 in Section 4 since its proof is a bit simpler. The idea of the proof of Theorem 4 is to reduce the general case to the foregoing special case along the lines of the Valiant-Vazirani reduction [VV86]. We then prove Theorem 3 in Section 5 by using a slightly more sophisticated variant of the Valiant-Vazirani reduction.

We note that the protocol of Section 3 is the main point in which we deviate from [JST11], while the proofs in Sections 4 and 5 follow more or less the same lines as [JST11].

3 The Case of Hamming Distance 1

In this section, we present a deterministic two-round protocol that solves the universal relation in the special case where Alice and Bob get inputs that disagree on exactly one coordinate. The protocol transmits \( O(\log n) \) bits. Let us denote by \( x, y \in \{0,1\}^n \) the inputs of Alice and Bob respectively. Without loss of generality, we may assume that \( n + 1 \) is a power of 2, so the Hamming code exists over \( \{0,1\}^n \) (otherwise, the players pad their inputs with 0s in order to satisfy this restriction, and this increases \( n \) by a factor of at most 2). The protocol is as follows:

---

2 We note that the aforementioned protocol of Håstad solves the universal relation as well.
1. In the beginning of the protocol, Alice finds the ball in the Hamming code to which \( x \) belongs, and denotes its center by \( c_x \in \{0,1\}^n \). Bob does similarly for \( y \), thus obtaining a center \( c_y \in \{0,1\}^n \).

2. Alice sends the first message in the protocol, which is an integer from 0 to \( n \) that she determines as follows:

   (a) If \( x = c_x \), then Alice sends 0.
   
   (b) Otherwise, Alice sends the unique coordinate \( j \) in \([n]\) on which \( x \) and \( c_x \) disagree.

3. If Alice sent 0:

   (a) Observe that in this case it holds that \( c_y = x = c_x \) (since \( x \) and \( y \) are within distance 1 and \( x \) is the center of a ball).
   
   (b) Thus, Bob sends to Alice the unique coordinate \( i \) on which \( y \) and \( c_y = x \) disagree, and this is the output of the protocol.

4. If Alice sent \( j \in [n] \) and \( y = c_y \):

   (a) Observe that in this case, it holds that \( c_x = y = c_y \) (since \( x \) and \( y \) are within distance 1 and \( y \) is the center of a ball).
   
   (b) Thus, Bob can deduce that \( j \) is the coordinate on which \( x \) and \( y \) differ.
   
   (c) Hence, Bob sends \( j \) back to Alice, and this is the output of the protocol.

5. If Alice sent \( j \in [n] \) and \( y \neq c_y \):

   (a) Let us denote by \( i \in [n] \) the unique coordinate on which \( x \) and \( y \) disagree.
   
   (b) Observe that \( i \neq j \), since otherwise it would follow that \( y = c_y \).
   
   (c) This implies that \( y \) and \( c_x \) disagree exactly on the coordinates \( i \) and \( j \).
   
   (d) Bob computes the string \( y' \) obtained by flipping the \( j \)-th coordinate of \( y \), so \( y' \) disagrees with \( c_x \) only on the coordinate \( i \).
   
   (e) Then, \( y' \) is within Hamming distance 1 of \( c_x \), and therefore must be in the ball around \( c_x \) in the Hamming code.
   
   (f) Bob now determines \( c_x \) by finding the ball of \( y' \) in the Hamming code, and deduces \( i \) by finding the unique coordinate on which \( y' \) and \( c_x \) disagree.
   
   (g) Bob sends \( i \) to Alice, and this is the output of the protocol.

The correctness of the protocol is explained within the foregoing description, and it is not hard to see that it indeed sends \( O(\log n) \) bits, as required.

**Remark 5.** By personal communication, we know that this protocol was discovered independently by Mauricio Karchmer and Avi Wigderson, and by Benjamin Rossman. However, to the best of our knowledge, this is the first time this protocol is published.

In addition, Tardos and Zwick [T97] used the Hamming code to give a deterministic protocol for the universal relation in the general case (rather than the case of Hamming distance 1), which has communication complexity \( n + 2 \).
4 A Two-Round Protocol

In this section we describe a randomized two-round protocol that solves the universal relation in the general case, thus proving Theorem 4. We describe a public-coin protocol, and it can be converted into a private-coin protocol using Newman’s lemma [New91].

The idea is to reduce the general case to the special case of Section 3 along the lines of the Valiant-Vazirani reduction: For start, suppose that the parties knew that their inputs disagree on exactly \( \ell \) coordinates. In this case, the parties could choose a random set of coordinates of size \( n/2^\ell \), and with constant probability this set would contain exactly one coordinate on which they disagree. Thus, the parties could project their inputs to this set and use the protocol of Section 3. The next step in the argument is to observe that this idea works even if the parties only have an estimate of \( \ell \) up to a factor of 2. Finally, since the parties do not have such an estimate of \( \ell \), they try different values of \( \ell = 1, 2, 4, 8, \ldots, n \) and apply the foregoing protocol in parallel for each of those values. Details follow.

Formally, the protocol is defined as follows. Suppose that Alice and Bob get as inputs the strings \( x, y \in \{0, 1\}^n \) respectively, so \( x \neq y \). They perform the following steps in parallel for every \( t \in \{1, \ldots, \lceil \log n \rceil + 1\} \):

1. Using the public coins, they choose a random set of coordinates \( S \subseteq [n] \) by putting each coordinate in \( S \) with probability \( 2^{-t} \) independently.
2. They execute the protocol of Section 3 on \( x|_S \) and \( y|_S \), thus obtaining a coordinate \( i \in [n] \).
3. Alice and Bob send \( x_i \) and \( y_i \) to each other.
4. If \( x_i \neq y_i \), they output the coordinate \( i \) and end the protocol.

It is easy to see that the protocol indeed uses two rounds and transmits \( O(\log^2 n) \) bits. It remains to show that it outputs a coordinate \( i \in [n] \) on which \( x_i \neq y_i \) with good probability.

Fix specific inputs to the players \( x, y \in \{0, 1\}^n \), and let \( I \subseteq [n] \) be the set of coordinates on which \( x \) and \( y \) differ. Observe that the protocol succeeds whenever, in the foregoing steps, the random set \( S \) contains exactly one coordinate on which \( x \) and \( y \) differ (i.e., \( |S \cap I| = 1 \)). When \( t = \lceil \log |I| \rceil + 1 \), the probability that the latter event happens is at least \( \frac{1}{2} \) by Claim 1 (since there are \( |I| \) Bernoulli trials with success probability \( 2^{-t} \) which is between \( \frac{1}{4|I|} \) and \( \frac{1}{2|I|} \)).

It follows that the protocol succeeds with probability at least \( \frac{1}{8} \). Note that the protocol is a zero-error protocol, i.e., the parties always know whether they succeeded or not. Hence, the success probability can be amplified to \( \frac{2}{3} \) by repeating the protocol a constant number of times in parallel, while maintaining a communication complexity of \( O(\log^2 n) \). This concludes the proof.

5 A Three-Round Protocol

In this section we present the protocol that has communication complexity \( O(\log n) \) and three rounds, thus proving Theorem 3. This variant was suggested to us by Oded Goldreich, and follows the underlying ideas of [JST11]. The idea of this protocol is that instead of invoking the protocol of Section 3 for every value of \( t \), the parties first find the correct value of \( t \), and then invoke the protocol of Section 3 only for this value of \( t \). Formally, the protocol is defined as follows:

1. Suppose that Alice and Bob get as inputs the strings \( x, y \in \{0, 1\}^n \) respectively, so \( x \neq y \).
2. The parties perform the following steps in parallel for every \( t \in \{1, \ldots, \lceil \log n \rceil + 1\} \):
(a) Using the public coins, choose a random set of coordinates $S \subseteq [n]$ by putting each coordinate in $S$ with probability $2^{-t}$ independently.

(b) Bob sends the parity of $y|_S$.

(c) Alice sends the parity of $x|_S$.

3. If the parities of $x|_S$ and $y|_S$ agree for all the values of $t$, the protocol fails.

4. Let $t^*$ be the maximal value of $t$ on which the parity of $x|_S$ and $y|_S$ disagree, and let $S^*$ be the corresponding set.

5. The parties execute the protocol of Section 3 on $x|_{S^*}$ and $y|_{S^*}$, thus obtaining a coordinate $i \in [n]$.

6. Alice and Bob send $x_i$ and $y_i$ to each other.

7. If $x_i \neq y_i$, the protocol outputs $i$, and otherwise the protocol fails.

Again, it is not hard to see that this protocol transmits $O(\log n)$ bits. To see that the protocol uses only three rounds, observe that in Step 2 the protocol uses two rounds where Alice speaks last, and in Step 5 the protocol uses two rounds where Alice speaks first. Hence, we can merge the second round of Step 2 and the first round of Step 5 into a single round, thus obtaining a three-round protocol. It remains to show that the protocol succeeds with constant probability.

We first observe that with probability at least $\frac{1}{8}$, the parities of $x|_S$ and $y|_S$ disagree for $t = \lceil \log |I| \rceil + 1$. The reason is that for this value of $t$, the set $S$ contains exactly one coordinate on which $x$ and $y$ disagree with probability at least $\frac{1}{8}$, as we have shown in the previous section. Now, let us condition on this event, and consider the value $t^*$ and set $S^*$ from the protocol, where $t^*$ may be equal to or greater than $\lceil \log |I| \rceil + 1$. Then, the set $S^*$ contains at least one coordinate on which $x$ and $y$ differ. Hence, by Claim 2 with probability at least $\frac{3}{8}$ the set $S^*$ contains exactly one coordinate on which $x$ and $y$ differ, in which case the protocol succeeds (since there are $|I|$ Bernoulli experiments and success probability is $2^{-t^*} \leq \frac{1}{2|I|}$). It follows that the protocol succeeds with probability at least $\frac{1}{8} \cdot \frac{3}{8} = \frac{3}{64}$, as required.

Acknowledgement. We are grateful to Jelani Nelson for referring us to the papers of [JST11, KNP+17] and for helpful explanations on sketching algorithms, and to Oded Goldreich for suggesting the protocol of Section 3 and for comments that improved the presentation of this note. We are also grateful to Hossein Jowhari and Gábor Tardos for helpful explanations of their work, and to Mert Səğələm for bringing [S11] to our attention. We would also like to thank Avi Wigderson, Benjamin Rossman and Daniel Kane for valuable discussions and ideas. Finally, we would like to thank anonymous referees for suggestions that improved the presentation of this work.

References


