Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 62 (2007)

The Tensor Product of Two Good Codes Is Not Necessarily Robustly Testable[*]

Oded Goldreich[†] Or Meir[‡]

June 16, 2011

Abstract

Given two codes R and C, their tensor product $R \otimes C$ consists of all matrices whose rows are codewords of R and whose columns are codewords of C. The product $R \otimes C$ is said to be robust if for every matrix M that is far from $R \otimes C$ it holds that the rows and columns of M are far on average from R and C respectively. Ben-Sasson and Sudan (ECCC TR04-046) have asked under which conditions the product $R \otimes C$ is robust.

Addressing this question, Paul Valiant (APPROX-RANDOM 2005) constructed two linear codes with constant relative distance whose tensor product is not robust. However, one of those codes has a sub-constant rate. We show that this construction can be modified such that both codes have both constant rate and constant relative distance. We also provide an alternative proof for the non-robustness of the tensor product of those codes, based on the inverse direction of the "rectangle method" that was presented by the second author (ECCC TR07-061). We believe that this proof gives an additional intuition for why this construction works.

1 Introduction

An error correcting code is said to be *locally testable* if there is a test that can check whether a given string is a codeword of the code or far from the code, by reading only a constant number of

[*] A preliminary version of this paper was published as ECCC TR07-062. This research was partially supported by the Israel Science Foundation (grant No. 460/05).

[†] Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100 Israel. Email: oded.goldreich@weizmann.ac.il

[‡] Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100 Israel. Email: or.meir@weizmann.ac.il

ISSN 1433-8092

symbols of the string. Locally Testable Codes (LTCs) were first systematically studied by Goldreich and Sudan [GS06] and since then several constructions of LTCs were suggested (See [Gol05] for an extensive survey of some constructions, as well as [Din07, BSS08, Mei09] for a few later constructions).

Ben-Sassson and Sudan [BSS06] suggested using the tensor product operation for the construction of LTCs. Given two linear error correcting codes R, C, their tensor product $R \otimes C$ is the code that consists of all matrices whose rows are codewords of R and whose columns are codewords of C. If R and C are locally testable, we would like $R \otimes C$ to be locally testable. [BSS06] suggested using the following test for testing the tensor product $R \otimes C$.

The Row/Column Test. Choose a random row (or column) of the matrix and accept if and only if it is a codeword of R (C, respectively).

In order to study the conditions under which $R \otimes C$ is locally testable, [BSS06] introduced the notion of "robust" tensor product (which is a special case of the notion of robustness of [BSGH$^+$06, DR06]). The tensor product $R \otimes C$ is said to be robust if, for every matrix M that is far from $R \otimes C$, it holds that the rows and columns of M are far on average from R and C respectively. It is not hard to see that if both R and C are locally testable, and $R \otimes C$ is robust, then $R \otimes C$ is locally testable.

This gives rise to the question under which conditions the tensor product is robust. On the positive side, it was proven that for three important families of codes that their tensor product is robust: The Reed-Solomon codes (this is the bivariate low degree test of [PS94]), "smooth" and "weakly smooth" LDPC codes [DSW06, BSV09b], and "semi-LTCs" [BSV09a][1]. In addition, [BSS06] showed that the tensor product of three codes or more ($C_1 \otimes C_2 \otimes C_3 \otimes \ldots$) is robust under a more general notion of robustness, and their work was imroved in [Vid11].

The focus of this note is on the negative side. In this context, Paul Valiant [Val05] constructed two linear codes of constant relative distance whose tensor product is not robust, and this construction was extended in [CR05]. However, one of the codes constructed by [Val05] has sub-constant rate. In this note, we show the construction of [Val05] can be modified such that both codes have constant rate. Using Theorem 2 of [CR05], it follows that there exists a linear code of constant rate and relative distance, whose tensor product *with itself* is not robust.

We provide two proofs that the tensor product of our codes is not robust: the first proof is an adaptation of the proof of [Val05]. The second proof is an alternative proof that is based on the inverse direction of the "rectangle method" that was presented in [Mei07], and which can also be adapted to the original codes of [Val05]. We believe that our alternative proof gives an additional intuition for why this construction works.

[1] The work of [BSV09a] uses a little different notion of robustness, but it is relevant for our work as well.

Organization. The rest of this work is organized as follows. In Section 2, we provide the required preliminaries. In Section 3, we describe our construction of the aforementioned codes, and show that they have the required rate and relative distance. In Section 4, we adapt the proof of [Val05] to show that the tensor product of our codes is not robust. Finally, In Section 5, we provide the alternative proof that the tensor product of our codes is not robust.

2 Preliminaries

For any two binary strings x, y of the same length, we denote by $\delta(x, y)$ the *relative* Hamming distance between x and y. We say that x and y are τ-close if $\delta(x, y) \leq \tau$, and otherwise we say that they are τ-far.

2.1 Error Correcting Codes

We review the basics of error correcting codes [MS88]. A linear code C is linear subspace of $\{0,1\}^n$, where n is called the block length of C. The vectors of C are called codewords. The code C has relative distance δ_C if for any two distinct codewords $c^1 \neq c^2 \in C$ it holds that $\delta(c^1, c^2) \geq \delta_C$. It is well-known that C has relative distance δ_C if and only if every non-zero codeword $c \in C$ has at least δ_C fraction of non-zero coordinates.

A generator matrix for C is a matrix G whose rows form a basis of C. The dual code $C^\perp$ is the code defined by

$$C^\perp \stackrel{\text{def}}{=} \{v : \langle v, c\rangle = 0 \text{ for every } c \in C\}$$

where the inner product is over $\mathbf{GF}(2)$.

We are usually interested in infinite families of codes. An infinite family of codes $C = \{C_k\}_k$ is an infinite sequence of codes such that the code C_k has dimension k. We say that the family C has block length $n(k)$ (where $n : \mathbb{N} \to \mathbb{N}$) if the code C_k has block length $n(k)$ for every k, and say that the family has rate r_C (for $r_C \in (0,1)$) if for every k it holds that $\frac{k}{n(k)} \geq r_C$. Finally, the family C has relative distance δ_C (for $\delta_C \in (0,1)$) if for every k it holds that C_k has relative distance at least δ_C.

2.2 Tensor Product Codes

Let R, C denote linear codes with block lengths m, n and relative distances δ_R, δ_C respectively. The tensor product $R \otimes C \subseteq \{0,1\}^{n \cdot m}$ is the linear code that consists of all the binary $n \times m$ matrices whose rows are codewords of R and whose columns are codewords of C. It is well known that the relative distance of $R \otimes C$ is $\delta_R \cdot \delta_C$.

For any binary $n \times m$ matrix M, we denote by $\delta_{R \otimes C}(M)$ the relative distance of M to $R \otimes C$. We also denote by $\delta_{\text{row}}(M)$ the average relative distance of a row of M to R, and define δ_{col} similarly. Finally, we denote by $\rho(M)$ the average of $\delta_{\text{row}}(M)$ and $\delta_{\text{col}}(M)$, that is,

$$\rho(M) \stackrel{\text{def}}{=} \frac{\delta_{\text{row}}(M) + \delta_{\text{col}}(M)}{2}$$

We can now state the definition of robustness.

Definition 2.1. We say that $R \otimes C$ is α-robust if for every M it holds that $\rho(M) \geq \alpha \cdot \delta_{R \otimes C}(M)$.

Now, let $R = \{R_k\}_k C = \{C_k\}_k$ be two infinite families of codes. We say that the infinite family $\{R_k \otimes C_k\}_k$ is α-robust (for $\alpha \in (0,1)$) if for every $k \in \mathbb{N}$ it holds that $R_k \otimes C_k$ is α-robust.

3 The codes

Our main result in this work is the following.

Theorem 3.1. *There exists two infinite families of codes $R = \{R_k\}_k C = \{C_k\}_k$ such that both R and C have rate and relative distance at least $\frac{1}{1000}$, but the infinite family $\{R_k \otimes C_k\}_k$ is not α-robust for any constant $\alpha \in (0,1)$.*

Fix $k \in \mathbb{N}$. The rest of this section is devoted to describing the construction of the codes R_k and C_k, and establishing their rate and relative distance. We will show that the infinite family $\{R_k \otimes C_k\}_k$ is not α-robust for any constant $\alpha \in (0,1)$ in Sections 4 and 5.

With a slight abuse of notation, we abbreviate $R \stackrel{\text{def}}{=} R_k$ and $C \stackrel{\text{def}}{=} C_k$. Let $n \stackrel{\text{def}}{=} 100 \cdot k$. We choose the code C to be an arbitrary code with dimension k, block length n, and relative distance at least $1/100$, such that the dual code $C^{\perp}$ has relative distance at least $1/100$ as well. The existence of such code existence can be established rather easily using the probabilistic method. In order to construct the code Rm we define some additional notation.

Notation 3.2. Let $S \subseteq \{0,1\}^n$ be the set of vectors that consist of $\frac{n}{10}$ "homogenous" blocks of 10 bits (recall that n is divisible by 10). By "homogenous" we mean that in each block all the bits are equal.

In order to construct the code R, we use the following code D, whose existence can be established rather easily using the probabilistic method.

Fact 3.3. *There exists a code D with dimension k and block length n, such that every non-zero codeword of D is $\frac{1}{100}$-far from every vector in S.*

Now, let G_C and G_D be $k \times n$ generator matrices of C and D respectively. We define $J \overset{\text{def}}{=} G_C^T \cdot G_D$. Let J^{10} be the $n \times 10n$ matrix that consists of 10 consecutive copies of J and let I_n^{10} be the n-rank identity matrix with each column duplicated to appear 10 times consecutively. That is, I_n^{10} is a matrix of the form

$$n\left\{\begin{pmatrix} \overbrace{1\ \ 1\ \ \dots\ \ 1}^{10} & \overbrace{0\ \ 0\ \ \dots\ \ 0}^{10} & & \overbrace{0\ \ 0\ \ \dots\ \ 0}^{10} \\ 0\ \ 0\ \ \dots\ \ 0 & 1\ \ 1\ \ \dots\ \ 1 & & 0\ \ 0\ \ \dots\ \ 0 \\ 0\ \ 0\ \ \dots\ \ 0 & 0\ \ 0\ \ \dots\ \ 0 & \dots & \vdots\ \ \vdots\ \ \vdots\ \ \vdots \\ \vdots\ \ \vdots\ \ \vdots\ \ \vdots & \vdots\ \ \vdots\ \ \vdots\ \ \vdots & & 0\ \ 0\ \ \dots\ \ 0 \\ 0\ \ 0\ \ 0\ \ 0 & 0\ \ 0\ \ 0\ \ 0 & & 1\ \ 1\ \ \dots\ \ 1 \end{pmatrix}\right.$$

We define the code R to be the space spanned by the rows of the matrix $M = J^{10} + I_n^{10}$. It remains to show that R has dimension k and relative distance at least 1/1000. We will actually show a better bound, namely, that R has relative distance at least 1/100. The following two claim and corollaries will be useful for both ends.

Claim 3.4. *The matrix J has rank k.*

Proof. On one hand, the columns of J are linear combinations of rows of G_C, so its rank can be at most k. On the other hand, both G_C and G_D are matrices of rank k and have k rows, so each of them contains a full rank $k \times k$ submatrix, denote those submatrices K_C, K_D respectively. Observe that the matrix $K_C^T \cdot K_D$ has full rank, i.e., has rank k. Now, note that $K_C^T \cdot K_D$ is a submatrix of $G_C^T \cdot G_D$, so the rank of J is at least k. It follows that the rank of J is exactly k. ■

Corollary 3.5. *The columns of J span the code C.*

Proof. The corollary follows immediately from the facts that the columns of J are linear combinations of rows of G_C and that J has rank k. ■

Corollary 3.6. *The matrix J^{10} has rank k.*

3.1 The dimension of R

We first show that R has dimension k. Since J^{10} has rank k, it has k independent rows $u_1, \dots, u_k \in \{0,1\}^{10n}$. Let $w_1, \dots, w_k$ be the corresponding rows of M, and let $v_1, \dots, v_k$ be the corresponding rows of I_n^{10}. We prove that the rows $w_1, \dots, w_k$ of M are linearly independent, and this will imply that R has dimension at least k (since R is spanned by the rows of M).

Let $w_{i_1}, \ldots, w_{i_m}$ for $m \leq k$ be a subset of $w_1, \ldots, w_k$. We show that the sum $w_{i_1}, \ldots, w_{i_m}$ is non-zero. Since this will hold for any choice of $w_{i_1}, \ldots, w_{i_m}$, this will show that $w_1, \ldots, w_k$ are independent. We first observe that

$$\sum_{j=1}^{m} w_{i_j} = \sum_{j=1}^{m} u_{i_j} + \sum_{j=1}^{m} v_{i_j}$$

We now make two observations:

Claim 3.7. *The sum $\sum_{j=1}^{m} u_{i_j}$ consists of 10 concatenated copies of a non-zero codeword of the code D.*

Claim 3.8. *The sum $\sum_{j=1}^{m} v_{i_j}$ consists of 10 concatenated vectors from the set S.*

Since by Fact 3.3, every non-zero codeword of D is 1/100-far from every vector in S, it holds that the sum

$$\sum_{j=1}^{m} w_{i_j} = \sum_{j=1}^{m} u_{i_j} + \sum_{j=1}^{m} v_{i_j}$$

has 1/100 fraction of non-zero coordinates, and in particular it is a non-zero vector, as required. It remains to prove Claims 3.7 and 3.8. The proof of Claim 3.8 is trivial. To see that Claim 3.7 holds, observe that:

1. Every row of $J \stackrel{\text{def}}{=} G_C^T \cdot G_D$ is a linear combination of rows of G_D.
2. Therefore, every row of J is a codeword of D.
3. Hence, every row of J^{10} consists of 10 concatenated copies of a codeword of D.

This concludes the proof that R has dimension k.

3.2 The distance of C_2

We show that every non-zero codeword of R has at least 1/100 fraction of non-zero coordinates, and this will imply that R has relative distance at least 1/100. Let c be any nonzero codeword of R. Since R is spanned by the rows of M, there exists a non-zero vector $v \in \{0,1\}^n$ such that $c = v \cdot M$. Now, recall that $M = J^{10} + I_n^{10}$. We consider two cases: the case where $v \cdot J^{10} = 0$, and the case where $v \cdot J^{10} \neq 0$.

The case where $v \cdot J^{10} = 0$. If $v \cdot J^{10} = 0$, then in particular it holds that $v \cdot J = 0$. Recall that by Corollary 3.5, the columns of J span C. Thus, the assumtion that $v \cdot J = 0$ implies that v is orthogonal to all the codewords in C, or in other words, $v \in C^{\perp}$.

Since, by assumption, $C^{\perp}$ has relative distance at least 1/100, it follows that v has at least 1/100 fraction of non-zero coordinates. This implies, in turn, that $v \cdot I_n^{10}$ has at least 1/100 fraction of non-zero coordinates. Finally, it holds that

$$c = v \cdot M = v \cdot J^{10} + v \cdot I_n^{10} = v \cdot I_n^{10},$$

and therefore c has 1/100 fraction of non-zero coordinates, as required.

The case where $v \cdot J^{10} \neq 0$. Denote $d \overset{\text{def}}{=} v \cdot J$. We proceed as in the proof that C_2 has dimension k. Observe that

1. d is a non-zero codeword of D.
2. $v \cdot J^{10}$ is a concatenation of 10 copies of d
3. $v \cdot I_n^{10}$ is the concatenation of 10 elements of S.

It follows that

$$c = v \cdot M = v \cdot J^{10} + v \cdot I_n^{10}$$

consists of 10 blocks, each of them is the sum of d with an element of S. Since d is a non-zero codeword of D, so by Fact 3.3 it holds that d is 1/100 far from every vector of S. It follows that c has at least 1/100 fraction of non-zero coordinates..

4 The non-robustness of $R \otimes C$

In this section, we prove that $R \otimes C$ is not α-robust for any constant $\alpha \in (0, 1)$ using a variant of the proof of [Val05]. To this end, we show that the matrix M that was constructed as part of the definition of R in Section 3 is a counter-example to the robustness of $R \otimes C$. In other words, M is a $n \times 10n$ matrix that it is far from $R \otimes C$, while the rows and columns of M are close to R and C.

Every row of M is a codeword of R, so $\delta_{\text{row}}(M) = 0$. Furthermore, every column of $J^{10} = M - I_n^{10}$ is a codeword of C, so $\delta_{\text{col}}(M) = \frac{1}{n}$. We thus have that $\rho(G_2) \leq \frac{1}{2n}$. We now observe the following.

Claim 4.1. *It holds that*

$$\delta_{R \otimes C}(M) \geq \frac{99}{100}$$

Proof. Consider an arbitrary $N \in R \otimes C$. Every row of $M - N$ is a codeword of R. Furthermore, each column of $M - I_n^{10} - N$ is a codeword of C, so the rank of $M - I_n^{10} - N$ is at most k. This implies that the rank of $M - N$ must be at least $n - k$: Otherwise, the rank of $-I_n^{10} = (M - I_n^{10} - N) + (N - M)$ would have been less than n (since rank is sub-additive and the ranks of $M - N$ and $N - M$ are equal).

Thus, there are at least $n - k$ non-zero rows in $M - N$, each of which is a codeword of R. Each of those non-zero rows of $M - N$ has at least $1/100$ fraction of non-zero coordinates, i.e., at least $n/10$ non-zero coordinates. It follows that M and N differ on at least $(n - k) \cdot n/10$ coordinates, so

$$\delta_{R\otimes C}(M) = \min_{N \in R\otimes C} \{\delta(M, N)\} \geq \frac{(n-k) \cdot n/10}{10n^2} \geq \frac{99}{100},$$

as required. ■

It follows that

$$\rho(M) \leq \frac{100}{198 \cdot n} \cdot \delta_{R\otimes C}(M),$$

and therefore $R \otimes C$ is not α-robust for any constant $\alpha \in (0, 1)$.

5 An alternative proof for the non-robustness of $R \otimes C$

In this section, we prove that $R \otimes C$ is not α-robust for any constant $\alpha \in (0, 1)$ using the rectangle method of [Mei07]. For any $n \times 10n$ matrix N, let N_R denote the matrix obtained from decoding every row of N to nearest codeword of R, and let N_C be defined similarly for the columns and C. We say that N_R and N_C agree on a large rectangle if there exist sets $U \subseteq [n]$, $V \subseteq [10n]$ such that $|U| \geq 0.99 \cdot n$, $|V| \geq 9.9 \cdot n$, and such that N_R and N_C agree on all the entries in $U \times V$. The following fact is a corollary of [Mei07].

Fact 5.1. *$R \otimes C$ is α-robust only if for every $n \times 10n$ matrix N that satisfies $\rho(M) < \frac{1}{60000} \cdot \alpha$ it holds that N_R and N_C agree on a large rectangle.*

Thus, in order to prove that $R \otimes C$ is not α-robust for every constant $\alpha \in (0, 1)$, it suffices to prove that for every constant $\alpha_0 \in (0, 1)$ there exists a matrix N with $\rho(N) \leq \alpha_0$ such that N_R and N_C do not agree on a large rectangle. As in Section 4, the matrix N we will use will be the matrix M that was constructed as part of the definition of R in Section 3. Recall that $\rho(M) \leq \frac{1}{2n}$. We now prove the following claim, which will conclude the proof that $R \otimes C$ is not α-robust.

Claim 5.2. *M_R and M_C do not agree on a large rectangle.*

Proof. Let U and V be sets such that $|U| \geq 0.99 \cdot n$, $|V| \geq 9.9 \cdot n$. Observe that $M_R = M$ and $M_C = J^{10} = M - I_n^{10}$, so $M_R - M_C = I_n^{10}$. We show that I_n^{10} has an entry with value 1 in $U \times V$, and therefore M_R and M_C do not agree on $U \times V$.

We know that every column of I_n^{10} contains exactly one entry with value 1, so the total number of entries with value 1 contained in the columns of V is $|V|$. We also know that every row of I_n^{10} contains exactly ten entries with value 1, so there are at least $\frac{1}{10}|V|$ rows that contain 1's *in their intersection with* V. Now, note that $\frac{1}{10}|V| > n - |U|$, and thus the rows containing entries with value 1 *in their intersection with* V can not all be in $[n] \setminus U$. It follows that I_n^{10} has 1 on at least one of the coordinates in $U \times V$, as required. ■

References

[BSGH$^+$06] Eli Ben-Sasson, Oded Goldreich, Prahladh Harsha, Madhu Sudan, and Salil Vadhan. Robust PCPs of proximity, shorter PCPs and applications to coding. *SIAM Journal of Computing*, 36(4):120–134, 2006.

[BSS06] Eli Ben-Sasson and Madhu Sudan. Robust locally testable codes and products of codes. *Random Struct. Algorithms*, 28(4):387–402, 2006. Preliminary version in APPROX-RANDOM 2004.

[BSS08] Eli Ben-Sasson and Madhu Sudan. Short PCPs with polylog query complexity. *SIAM J. Comput.*, 38(2):551–607, 2008. Preliminary version in STOC 2005.

[BSV09a] Eli Ben-Sasson and Michael Viderman. Composition of semi-ltcs by two-wise tensor products. In *APPROX-RANDOM*, pages 378–391, 2009.

[BSV09b] Eli Ben-Sasson and Michael Viderman. Tensor products of weakly smooth codes are robust. *Theory of Computing*, 5(1):239–255, 2009.

[CR05] Don Coppersmith and Atri Rudra. On the robust testability of tensor products of codes. *Electronic Colloquium on Computational Complexity (ECCC)*, (104), 2005.

[Din07] Irit Dinur. The PCP Theorem by gap amplification. *Journal of ACM*, 54(3):241–250, 2007. Preliminary version in STOC 2006.

[DR06] Irit Dinur and Omer Reingold. Assignment testers: Towards combinatorial proof of the PCP theorem. *SIAM Journal of Computing*, 36(4):155–164, 2006.

[DSW06] Irit Dinur, Madhu Sudan, and Avi Wigderson. Robust local testability of tensor products of ldpc codes. In *APPROX-RANDOM*, pages 304–315, 2006.

[Gol05] Oded Goldreich. Short locally testable codes and proofs (survey). *Electronic Colloquium on Computational Complexity (ECCC)*, (014), 2005.

[GS06] Oded Goldreich and Madhu Sudan. Locally testable codes and PCPs of almost linear length. *Journal of ACM*, 53(4):558–655, 2006. Preliminary version in FOCS 2002, pages 13-22.

[Mei07] Or Meir. On the rectangle method in proofs of robustness of tensor products. *Electronic Colloquium on Computational Complexity (ECCC)*, 14(061), 2007.

[Mei09] Or Meir. Combinatorial construction of locally testable codes. *SIAM J. Comput.*, 39(2):491–544, 2009. Preliminary version appeared in STOC 2008, full version can be retrieved as ECCC TR07-115.

[MS88] Florence Jessie MacWilliams and Neil James Alexander Sloane. *The theory of error correcting codes*. Elsevier/North-Holland, Amsterdam, 1988.

[PS94] Alexander Polishchuk and Daniel A. Spielman. Nearly-linear size holographic proofs. In *STOC*, pages 194–203, 1994.

[Val05] Paul Valiant. The tensor product of two codes is not necessarily robustly testable. In *APPROX-RANDOM*, pages 472–481, 2005.

[Vid11] Michael Viderman. Linear time decoding of regular expander codes. *Electronic Colloquium on Computational Complexity (ECCC)*, 18:58, 2011.