Robust local testability of tensor products of LDPC codes[⋆]

Irit Dinur[1], Madhu Sudan[2], and Avi Wigderson[3]

[1] Hebrew University, Jerusalem, Israel. dinuri@cs.huji.ac.il
[2] Massachusetts Institute of Technology, Cambridge, MA. madhu@mit.edu
[3] Institute for Advanced Study, Princeton, NJ. avi@ias.edu

Abstract. Given two binary linear codes R and C, their tensor product $R \otimes C$ consists of all matrices with rows in R and columns in C. We analyze the "robustness" of the following test for this code (suggested by Ben-Sasson and Sudan [6]): Pick a random row (or column) and check if the received word is in R (or C). Robustness of the test implies that if a matrix M is far from $R \otimes C$, then a significant fraction of the rows (or columns) of M are far from codewords of R (or C).
We show that this test *is* robust, provided one of the codes is what we refer to as *smooth*. We show that expander codes and locally-testable codes are smooth. This complements recent examples of P. Valiant [13] and Coppersmith and Rudra [9] of codes whose tensor product is not robustly testable.

1 Introduction

A binary linear code is a linear subspace $C \subseteq \{0,1\}^n$. A code is *locally testable* if given a word $x \in \{0,1\}^n$ one can verify whether $x \in C$ by reading only few (randomly chosen) bits from x. More precisely such a code has a *tester*, which is a randomized algorithm with oracle access to the received word x. The tester reads at most q symbols from x and based on this "local view" decides if $x \in C$ or not. It should accept codewords with probability one, and reject words that are "far" (in Hamming distance) from the code with "noticeable" probability.

Locally testable codes (LTCs) are related to probabilistically checkable proofs (PCPs). LTCs were first explicitly studied by Goldreich and Sudan [12], who describe them as the "combinatorial core of PCPs". They constructed LTCs relying on some of the PCP machinery [11, 2, 1]. Since locally testable codes are simpler than PCPs, it seems natural to seek alternative constructions for them, possibly departing from the PCP framework.

One of the most interesting challenges in constructing LTCs, is to come up with an LTC that has constant relative distance and highest possible (maybe linear?) rate. Several steps in this direction were made in recent years, see [12, 8, 3, 4, 6, 7, 10].

All known efficient constructions of LTCs rely on some form of "composition" of two (or more) codes. In this paper we focus on composition by tensor product, which is an elementary way to compose two codes. Given two binary codes $R \subseteq \{0,1\}^m$ and $C \subseteq \{0,1\}^n$, their tensor product is the code $R \otimes C$ consisting of all binary $n \times m$ matrices whose rows belong to R and whose columns belong to C.

[⋆] Most of the research was done while the authors were visiting Microsoft Research Theory group. Additionally, Irit Dinur's work was supported in part by ISF grant 984/04, Madhu Sudan's work was supported in part by NSF Award CCR-0514915, and Avi Wigderson's work was supported in part by NSF Award CCR-0324906.

Ben-Sasson and Sudan [6] suggested using the tensor operation for constructing LTCs. They introduce the notion of *robust LTCs*: An LTC is called robust if whenever the received word is far from the code, then with noticeable probability the local view of the tester is *far* from an accepting local view. It is very easy to compose testers for robust LTCs: If it so happens that restriction of the code to the local view of the tester is itself an LTC, then instead of reading the entire local view, a tester for the smaller LTC can be invoked thereby saving on the query complexity of the tester.

Ben-Sasson and Sudan [6] showed that a code obtained by tensoring three or more codes (i.e. a code of the form $C_1 \otimes C_2 \otimes C_3$) is robustly testable, and used this result to construct LTCs. For the tensor product of two codes R and C, they considered the following natural test, and asked whether it is robust:

Test for $R \otimes C$: Pick a random row (or column), accept iff it belongs to R (or C).

Rather than providing a general definition of robustness (which can be found in Section 2.2), let us spell out the meaning of robustness for this particular test. Let x be an $n \times m$ matrix. Let $\delta^{\text{row}}(x)$ denote the expected distance of a random row of x from R, and let $\delta^{\text{col}}(x)$ denote the expected distance of a random column of x from C. Let $\delta_{R\otimes C}(x)$ denote the distance of x from the tensor product code $R \otimes C$. The robustness of the test is the largest value of α that satisfies

$$\frac{\delta^{\text{row}}(x) + \delta^{\text{col}}(x)}{2} \geq \alpha \cdot \delta_{R\otimes C}(x)$$

for every x. We say that the test is *robust* if its robustness is bounded away from 0.

Paul Valiant [13] showed a surprising example of two linear codes R and C for which the test above is not robust, by exhibiting a word x that is far from $R \otimes C$ but such that the rows of x are very close to being in R (i.e. $\delta^{\text{row}}(x)$ is small) and the columns of x are very close to being in C (i.e. $\delta^{\text{col}}(x)$ is small). An additional example of [9] gives a code whose tensor product with itself is not robust, and a similar result is shown for some non-linear code.

Results. Despite these examples, in this paper we show that the test above is robust for two important classes of Low Density Parity Check (LDPC) codes: Expander codes, and LTCs (see Proposition 1). We note that these are almost disjoint classes, as [5] prove that random expander LDPC codes are *not* locally testable.

We do this by introducing *smooth* codes which are a class of low density parity check codes. The smoothness property captures how badly the code is affected if some of the parity checks are removed from it.

We first show that if either R or C are smooth, then $R \otimes C$ has the following property. Any given word x that has small $\delta^{\text{row}}(x)$ and small $\delta^{\text{col}}(x)$, must have a large sub-matrix that completely agrees with some word in $R \otimes C$ (so x is close to $R \otimes C$). This implies that $R \otimes C$ is robust. We then argue that both LTCs and expander codes are smooth.

2 Notation, Definitions, and Results

All codes we consider will be binary linear codes. A binary linear code is a linear subspace $C \subseteq \{0,1\}^n$, whose dimension is denoted by $\dim(C)$. Every member of C is called a codeword.

We define the *distance* between two words $x, y \in \{0,1\}^n$ to be $\delta(x,y) = \Pr_i[x_i \neq y_i]$. We also define the weight of a string to be $\mathrm{wt}(x) = \delta(x, \mathbf{0})$. The distance of a code is denoted $\delta(C)$, and defined to be the minimal value of $\delta(x,y)$ for two distinct codewords $x, y \in C$. Clearly the distance of a linear code is equal to weight of the minimal-weight non-zero codeword.

Let $I_n = \{0,1\}^n$ denote the trivial code. For $x \in I_n$ and $C \subseteq I_n$, let $\delta_C(x) = \min_{\{y \in C\}}\{\delta(x,y)\}$ denote the distance of x from the code C.

2.1 Tensor Products of Codes

For $x \in I_m$ and $y \in I_n$ we let $x \otimes y$ denote the tensor product of x and y (i.e., the $n \times m$ matrix xy^T).

Let $R \subseteq I_m$ and $C \subseteq I_n$ be linear codes. We define the tensor product code $R \otimes C$ to be the linear subspace spanned by words $r \otimes c \in \{0,1\}^{n \times m}$ for $r \in R$ and $c \in C$. The following facts are immediate:

- The code $R \otimes C$ consists of all $n \times m$ matrices whose rows belong to R and whose columns belong to C.
- $\dim(R \otimes C) = \dim(R) \cdot \dim(C)$
- $\delta(R \otimes C) = \delta(R) \cdot \delta(C)$.

Fix $R \subseteq I_m$ and $C \subseteq I_n$ of distance δ_R and δ_C respectively for the rest of the manuscript.

Let $M \in I_m \otimes I_n$ and let $\delta(M) = \delta_{R \otimes C}(M)$. Let $\delta^{\mathrm{row}}(M) = \delta_{R \otimes I_n}(M)$ denote its distance from the space of matrices whose rows are codewords of R. This is the expected distance of a random row in x from R. Similarly let $\delta^{\mathrm{col}}(M) = \delta_{I_m \otimes C}(M)$.

2.2 Robust Locally Testable Codes

Locally testable codes, as described in the introduction, are codes for which one can test whether a given word x is in the code by reading only few (randomly chosen) symbols from x. We discuss here only *non adaptive* and *bi-regular* testers. Non adaptive means that which queries are read is determined before any query is made, and bi-regular means that every test queries the same number of bits, and every bit is queried by the same number of tests. It would be interesting to extend our result for locally testable codes without these restrictions.

Definition 1 ((Non adaptive, bi-regular) Locally Testable Code). *We say that a code $C \subseteq I_n$ is $(d, \delta, \epsilon, \rho)$-locally-testable if $\delta(C) \geq \delta$ and there is a randomized algorithm (called a* tester*) T, which selects d indices from $[n]$, and for any given word $x \in I_n$, T reads the bits of x in these locations, satisfying:*

- *If $x \in C$ then $\Pr[T^x \text{ accepts}] = 1$.*
- *If $\delta_C(x) \geq \rho$ then $\Pr[T^x \text{ rejects}] > \epsilon$.*

Moreover, the probability that a given index is chosen to be read by T is the same for all indices in $[n]$.

A somewhat stronger notion of LTCs is that of robust-LTCs. Such a code has a stronger soundness requirement: Whenever $x \notin C$ the local view of the tester is *far* (in expectation) from an accepting view. For a formal definition let us introduce a little notation. The tester algorithm T has two inputs: the random string r, and the word x that is being tested. The tester reads the string r and computes a predicate T_r and a d-tuple of indices $i_1, \ldots, i_d$ in which it queries the word x. It accepts iff $T_r(x[i_1], \ldots, x[i_d]) = 1$. Let $acc(T_r) = \left\{ w \in \{0,1\}^d \,\middle|\, T_r(w) = 1 \right\}$ be the set of local-views on which the tester accepts. Define the robustness of T on x to be

$$\rho^T(x) = \mathbb{E}_r[\delta((x[i_1], \ldots, x[i_d]) \, , \, acc(T_r))] \, ,$$

which is the expected distance of the local view from an accepting one. The robustness of T is the minimal ratio between the robustness of T on x, and the distance of x from the code:

$$\rho^T = \min_{x \notin C} \frac{\rho^T(x)}{\delta_C(x)}.$$

Definition 2 (Robust Code). *We say that a code $C \subseteq I_n$ is α-robust if there is a tester T that accepts every word in C with probability 1, such that $\rho^T \geq \alpha$.*

2.3 Low Density Parity Check (LDPC) Codes

A bipartite graph $([n], [m], E)$ is a parity check graph for a code $C \subseteq I_n$ if the following holds (let $\Gamma(j)$ denote the neighbors of j in the graph):

$$x \in C \qquad \Longleftrightarrow \qquad \forall j \in [m] \quad \sum_{i \in \Gamma(j)} x_i = 0 \; mod \; 2$$

In other words, every right-hand-side vertex $j \in [m]$ corresponds to a parity constraint, and a word is in the code if and only if it satisfies all of the constraints.

A code is referred to as an LDPC code if it has a "low-density" parity check graph, e.g. a graph with constant[4] average degree.

We first remark that LTCs are low density parity check codes, since a parity check graph can be constructed from the tester algorithm. Moreover, since our LTCs are bi-regular, so is their parity check graph.

Proposition 1. *Every $(d, \delta, \epsilon, \rho)$-LTC C with $\rho < \delta$ has a parity check graph with right degree d and such that for every word x, if $\delta_C(x) \geq \rho$ then it violates at least ϵ fraction of the parity checks.*

Proof. Let T be a tester for C. The predicates computed by T are parity checks (perhaps redundant) of C, since the code is linear. The construction of a parity graph (L, R, E) from T is immediate, with the nodes of R corresponding to the enumeration of the random strings of T.

[4] Implicit throughout this manuscript is the notion that we are working with infinite families of codes/graphs, where the parameters such as the degree or the distance do not change with the length of the code/graph etc.

Another important class of LDPC codes is that of expander codes.

Definition 3 (**(c,d)-regular (γ,δ)-expander**). *Let $c,d \in \mathbb{N}$ and let $\gamma,\delta \in (0,1)$. Define a* (c,d)-regular (γ,δ)-expander *to be a bipartite graph (L,R,E) with vertex sets L,R such that all vertices in L have degree c, and all vertices in R have degree d; and the additional property that every set of vertices $L' \subset L$, such that $|L'| \leq \delta\,|L|$, has at least $(1-\gamma)c\,|L'|$ neighbors.*

We say that a code C is an (c,d,γ,δ)-expander code if it has a parity check graph that is a (c,d)-regular (γ,δ)-expander.

The following is an important (and straightforward) property of expander codes,

Proposition 2. *If C is a (c,d,γ,δ)-expander code and $\gamma < \frac{1}{2}$, then $\delta(C) \geq \delta$.*

Proof. We prove that every non-zero word in C must have weight more than δn. Indeed let (L,R,E) be a parity check graph of C that is a (c,d)-regular (γ,δ)-expander. The proposition follows by examining the unique neighbor structure of the graph. Let $x \in C$ be a non-zero codeword, and let $L' \subseteq L$ be the set of indices in which x is 1. If $|L'| \leq \delta n$ then L' has at least $(1-\gamma)c\,|L'| > \frac{c}{2}\,|L'|$ neighbors in R. At least one of these sees *only one* element of L', so the parity of its neighbors is one, violating the corresponding constraint and contradicting $x \in C$.

2.4 Results

Let R,C be codes. We study the robustness of the following test (described also in the introduction) for a given word $M \in I_m \otimes I_n$.

Test T for $R \otimes C$:

1. Select $b \in \{0,1\}$ at random.
2. If $b = 0$ select $i \in [n]$ at random, and accept iff the i-th row of M is in R.
3. If $b = 1$ select $j \in [m]$ at random, and accept iff the j-th column of M is in C.

Obviously, T accepts every word of $R \otimes C$ with probability 1. We are interested in studying the robustness of T which we sometimes refer to as ρ instead of ρ^T.

Recall our notation $\delta(M) = \delta_{R\otimes C}(M)$ and our definition of $\delta^{\text{row}}(M) = \delta_{R\otimes I_n}(M)$ and $\delta^{\text{col}}(M) = \delta_{I_m\otimes C}(M)$. In other words $\delta^{\text{row}}(M)$ equals the average distance of a row of M from R, and similarly $\delta^{\text{col}}(M)$ equals the average distance of a column of M from C. The following proposition is immediate:

Proposition 3. *The robustness of T on input M is $\rho(M) = \frac{\delta^{\text{row}}(M)+\delta^{\text{col}}(M)}{2}$.* □

In order to establish robustness for T, say $\rho^T \geq \alpha > 0$, we must be able to prove for all M that $\frac{(\delta^{\text{row}}(M)+\delta^{\text{col}}(M))/2}{\delta(M)} \geq \alpha$.

As already mentioned in the introduction, for general codes R and C this is false. Paul Valiant [13] described a pair of codes R and C and a word M that is very far from $R \otimes C$, yet both $\delta^{\text{row}}(M)$ and $\delta^{\text{col}}(M)$ are very small.

Nevertheless, we observe that if C (or R) is somewhat "nice", then such a bound can be proven.

Theorem 1 (Tensoring Expander-codes). *Let $R \subset I_m$ be a code of distance at least $\delta_R > 0$. Let $C \subset I_n$ be a (c, d, γ, δ)-expander code for some $c, d \in \mathbb{N}, \delta > 0$, and $0 < \gamma < 1/6$. Then*

$$\rho^T \geq \frac{(\frac{1}{3} - 2\gamma)\delta\delta_R}{4d}\,.$$

Theorem 2 (Tensoring LTCs). *Let $R \subset I_m$ and $C \subset I_n$ be codes of relative distance at least δ_R, δ_C respectively. Furthermore, let C be a $(d, \delta_C, \epsilon, \rho)$-LTC, with $\rho \leq \frac{\delta_C}{16}$. Then,*

$$\rho^T \geq \min\left\{\frac{\epsilon\delta_R}{2d^2}, \frac{\delta_R\delta_C}{16}\right\}.$$

3 Smooth codes

We prove the two theorems by a common technique, where we show that the tensor product has nice testing properties if the underlying codes are nice in a certain sense that we refer to as "smooth". To motivate this notion, consider a code $C \subseteq I_n$ given by a (possibly redundant[5]) parity check graph $B = (L, R, E)$, where every vertex of R has degree d.

We consider how badly the code is affected if we remove some constraints $R_0 \subseteq R$. Let $C(R_0)$ denote the resulting code. $C(R_0)$ clearly contains C, but may now contain codewords of lesser weight. For instance we may remove all the neighbors of some vertex $u \in L$ (for the vertex u of minimum degree, this only requires us to remove a $d/|L|$ fraction of the right vertices), and now u is unconstrained, leading to a code of distance one. However if we delete the uth coordinate of $C(R_0)$ one may hope that the resulting code still has large distance. More generally, we may hope that the negative effect of deleting some subset R_0 of the constraints may be recovered by dropping some subset L_0 of the coordinate vertices. If a code exhibits such a property, we call it smooth, defined quantitatively below.

For a set $S \subset [n]$ we always denote $\overline{S} = [n] - S$. For a code $C \subseteq I_n$ and $L_0 \subseteq L = [n]$ let $C|_{L_0}$ be the projection of the codewords of C to the coordinates of $\overline{L_0}$. (Such a code is called a punctured code. For reasons that will be evident later, it is nicer to highlight the set of coordinates that are being deleted.)

For a code C defined by a bipartite graph $B = (L, R, E)$, let $C(R_0)$ denote the "supercode" given by the parity check graph $B' = (L = [n], R - R_0, E' = E \cap (L \times (R - R_0)))$.

[5] A parity check graph is redundant if removing a node from the right still results in a parity check graph for the same code.

Definition 4 (Smooth Code). *A code $C \subseteq I_n$ is $(d, \alpha, \beta, \delta)$-smooth if it has a parity check graph $B = (L, R, E)$ where all the right vertices R have degree d, the left vertices have degree $c = d|R|/|L|$, and for every set $R_0 \subseteq R$ such that $|R_0| \leq \alpha|R|$, there exists a set $L_0 \subseteq L$, $|L_0| \leq \beta|L|$ such that the code $C(R_0)|_{L_0}$ has distance at least δ.*

We next turn to prove that the test T described in the previous section is robust when one of the codes being tensored is smooth. More specifically we prove that for any word M, if $\rho(M) = (\delta^{\text{row}}(M) + \delta^{\text{col}}(M))/2$ is small then $\delta(M)$ is proportionally small.

Lemma 1 (Main Lemma). *Let $R \subseteq I_m$ and $C \subseteq I_n$ be codes of distance δ_R and δ_C. Let C be $(d, \alpha, \frac{\delta_C}{2}, \frac{\delta_C}{2})$-smooth, and let $M \in I_m \otimes I_n$. If $\rho(M) \leq \min\left\{\alpha\frac{\delta_R}{2d^2}, \frac{\delta_R\delta_C}{8}\right\}$ then $\delta(M) \leq 8\rho(M)$.*

Proof. For row $i \in [n]$, let $r_i \in R$ denote the codeword of R closest to the ith row of M. For column $j \in [m]$, let $c^{(j)} \in C$ denote the codeword of C closest to the jth column of M. Let M_R denote the $n \times m$ matrix whose ith row is r_i, and let M_C denote the matrix whose jth column is $c^{(j)}$. Let $E = M_R - M_C$.

In what follows the matrices M_R, M_C and (especially) E will be the central objects of attention. We refer to E as the error matrix. Note that $\delta(M, M_R) = \delta^{\text{row}}(M)$ and $\delta(M, M_C) = \delta^{\text{col}}(M)$ and so

$$\text{wt}(E) = \delta(M_R, M_C) \leq \delta(M, M_R) + \delta(M, M_C) = \delta^{\text{row}}(M) + \delta^{\text{col}}(M) = 2\rho(M)\,. \qquad (1)$$

Our proof strategy is to show that the error matrix E is actually very structured. We do this in two steps. First we show (Proposition 4) that its columns satisfy most constraints of the column code. Then we show (Proposition 5) that E contains a large submatrix which is all zeroes. Finally using this structure of E we show (Proposition 6) that M is close to some codeword of $R \otimes C$. Proposition 4 is the crux of our analysis (while Proposition 5 follows more or less in a straightforward way from the definition of smoothness, and Proposition 6 is a standard property of tensor product codes).

Proposition 4. *Let $\{i_1, \ldots, i_d\}$ be a constraint of C (i.e., every codeword of $y \in C$ satisfies $y_{i_1} + \ldots + y_{i_d} = 0$). Let e_i denote the ith row of E. Suppose $\text{wt}(e_{i_j}) < \delta_R/d$ for every $j \in [d]$. Then $e_{i_1} + \cdots + e_{i_d} = \mathbf{0}$.*

Proof. Let c_i denote the i-th row of the matrix M_C. (Recall that these rows are not necessarily codewords of any nice code - it is only the columns of M_C that are codewords of C). For every column j, we have $(c_{i_1})_j + \cdots + (c_{i_d})_j = 0$ (since the columns of M_C are codewords of C). Thus we conclude that $c_{i_1} + \cdots + c_{i_d} = \mathbf{0}$ as a vector.

Now consider $r_{i_1} + \cdots + r_{i_d}$ (recall that r_i is the i-th row of M_R). Since each one of the r_i's is a codeword of R, we have $r_{i_1} + \cdots + r_{i_d} \in R$. But this implies

$$e_{i_1}+\cdots+e_{i_d} = (r_{i_1}-c_{i_1})+\cdots+(r_{i_d}-c_{i_d}) = (r_{i_1}+\cdots+r_{i_d})-(c_{i_1}+\cdots+c_{i_d}) = (r_{i_1}+\cdots+r_{i_d})-\mathbf{0} \in R$$

Now we use the fact that the e_is have small weight. This implies that $\text{wt}(e_{i_1} + \cdots + e_{i_d}) \leq \sum_j \text{wt}(e_{i_j}) < \delta_R$. But R is an error-correcting code of minimum distance δ_R so the only word of weight less than δ_R in it is the zero codeword, yielding $e_{i_1} + \cdots + e_{i_d} = \mathbf{0}$.

Combined with the smoothness of C, the above proposition gives us sufficient structure to show that E has a large clean submatrix. We argue this below.

Proposition 5. *There exist subsets $U \subseteq [m]$ and $V \subseteq [n]$ with $|U|/m < \delta_R/2$ and $|V|/n < \delta_C/2$ such that $E(i,j) \neq 0$ implies $i \in V$ or $j \in U$.*

Proof. First, we consider the rows of E that have weight above δ_R/d. Let

$$V_1 = \{i \in [n] \mid \ \mathrm{wt}(e_i) \geq \delta_R/d\} \ .$$

We use $\delta^{\mathrm{row}}(M) \leq 2\rho(M) \leq \frac{\alpha\delta_R}{d^2}$ and Markov's inequality to deduce $|V_1|/n \leq \frac{2\rho(M)}{\delta_R/d} \leq \frac{\alpha}{d}$. Next, we consider every constraint of C that involves an index in V_1. Recall that the code C is $(d, \alpha, \frac{\delta_C}{2}, \frac{\delta_C}{2})$-smooth, and let $B = ([n], [\ell], F)$ be the corresponding parity check graph of C (with right degree d and left degree $c = \frac{d\ell}{n}$). Viewing V_1 as a subset of the left vertices of B, let $W \subseteq [\ell]$ be the set of neighbors of V_1 in B. First notice that $|W| \leq c\,|V_1| \leq c \cdot \alpha n/d = \alpha\ell$. Next, observe that constraints in $[\ell] - W$ touch only indices outside V_1, i.e., indices j with $w(e_j) < \delta_R/d$. By Proposition 4, such constraints are satisfied by the rows of E. It is clear that if an equality holds for row-vectors, it also holds for each column separately. Thus, *every column* of the error matrix E, denoted $e^{(j)}$, is contained in the code $C(W)$.

Now we use the smoothness of C to define the sets V and U. Since $|W| \leq \alpha\ell$, there must be a set $V \subseteq [n]$ of cardinality at most $\frac{\delta_C}{2}n$ such that the code $C(W)|_{\overline{V}}$ has distance at least $\frac{\delta_C}{2}n$. Let U be the set of indices corresponding to columns of E that have $\frac{\delta_C}{2}n$ or more non-zero elements in the rows outside V. This means that for every j, $e^{(j)}$ is either all zero on $\overline{V}$ or has at least $\frac{\delta_C}{2}n$ non-zero values on $\overline{V}$. If also $j \notin U$ then $e^{(j)}$ must be zero outside V. We conclude that if we throw away from the matrix E all the rows corresponding to V and all the columns corresponding to U, we are left with the zero matrix.

The fraction of rows thrown away is at most $\frac{|V|}{n} \leq \delta_C/2$. The fraction of columns thrown away is at most $\frac{\delta^{\mathrm{col}}(M)}{\delta_C/2} \leq \frac{4\rho(M)}{\delta_C} \leq \delta_R/2$, where we used Markov's inequality and $\delta^{\mathrm{col}}(M) \leq 2\rho(M) \leq \frac{\delta_C\delta_R}{4}$.

We now use a standard property of tensor products to claim M_R (and M_C and M) is close to a codeword of $R \times C$. Recall that $M \in \{0,1\}^{n \times m}$ and that $\delta(M_C, M_R) \leq 2\rho(M)$.

Proposition 6. *Assume there exist sets $U \subseteq [m]$ and $V \subseteq [n]$, $|U|/m \leq \delta_R/2$ and $|V|/n \leq \delta_C/2$ such that $M_R(i,j) \neq M_C(i,j)$ implies $j \in U$ or $i \in V$. Then $\delta(M) \leq 8\rho(M)$.*

Proof. This is a standard proposition. First we note that there exists a matrix $N \in R \otimes C$ that agrees with M_R and M_C on $\overline{V} \times \overline{U}$ (See [6, Proposition 3][6]). Recall also that $\delta(M, M_R) = \delta^{\mathrm{row}}(M) \leq 2\rho(M)$. So it suffices to show $\delta(M_R, N) \leq 6\rho(M)$. We do so in two steps. First we show that $\delta(M_R, N) \leq 2\rho(M_R)$. We then show that $\rho(M_R) \leq 3\rho(M)$ concluding the proof.

[6] Erase from the matrix M_R entries in rows V or columns U. Observe that decoding from erasures first each row and then each column, must result in the same matrix as decoding first each column and then each row (due to the distances of the codes).

For the first part we start by noting that M_R and N agree on every row in $\overline{V}$. This is the case since both rows are codewords of R which may disagree only on entries from the columns of U, but the number of such columns is less that $\delta_R m/2$. Next we claim that for every column $j \in [m]$ the closest codeword of C to the $M_R(\cdot, j)$, the jth column of M_R, is $N(\cdot, j)$, the jth column of N. This is true since $M_R(i,j) \neq N(i,j)$ implies $i \in V$ and so the number of such i is less than $\delta_C n/2$. Thus for every j, we have $N(\cdot, j)$ is the (unique) decoding of the jth column of M_R. Averaging over j, we get that $\delta^{\text{col}}(M_R) = \delta(M_R, N)$. In turn this yields $\rho(M_R) \geq \delta^{\text{col}}(M_R)/2 = \delta(M_R, N)/2$. This yields the first of the two desired inequalities.

Now to bound $\rho(M_R)$, note that for any pair of matrices M_1 and M_2 we have $\rho(M_1) \leq \rho(M_2) + \delta(M_1, M_2)$. Indeed it is the case that $\delta^{\text{row}}(M_1) \leq \delta^{\text{row}}(M_2) + \delta(M_1, M_2)$ and $\delta^{\text{col}}(M_1) \leq \delta^{\text{col}}(M_2) + \delta(M_1, M_2)$. To see the former, for instance, note that if the ith row of M_2 is within ρ_i of some codeword of R, then the ith row of M_1 is within $\rho_i + \delta(M_1(i,\cdot), M_2(i,\cdot))$ of the same codeword of R. Averaging over i yields $\delta^{\text{row}}(M_1) \leq \delta^{\text{row}}(M_2) + \delta(M_1, M_2)$. A similar argument yields $\delta^{\text{col}}(M_1) \leq \delta^{\text{col}}(M_2) + \delta(M_1, M_2)$, when combined the two yield $\rho(M_1) \leq \rho(M_2) + \delta(M_1, M_2)$. Applying this inequality to $M_1 = M_R$ and $M_2 = M$ we get $\rho(M_R) \leq \rho(M) + \delta(M_R, M) \leq 3\rho(M)$. This yields the second inequality and thus the proof of the proposition as well as Lemma 1.

In what follows we will show that expander codes, as well as LTCs are smooth.

4 Expander codes are smooth

Lemma 2. *Every (c, d, γ, δ)-expander code C is $(d, \alpha, \beta, \delta)$-smooth, provided $\gamma < \frac{1}{6}$, $\alpha < (\frac{1}{3} - 2\gamma)\delta d$ and $\beta = \frac{\alpha}{(\frac{1}{3} - 2\gamma)d}$.*

Proof. Let $B = (L, R, E)$ be the (c, d) regular (γ, δ)-expanding parity check graph of the code C. Let $R_0 \subseteq R$ of size $|R_0| \leq \alpha \cdot |R|$ be given. We will construct sets L', R' satisfying $L' \subseteq L$, $|L'| \leq \beta |L|$ and $R_0 \subseteq R' \subseteq R$ such that every subset of $L - L'$ of size at most δn expands sufficiently in the induced subgraph on $(L - L') \cup (R - R')$. This will suffice to prove that $C(R_0)|_{L'} \subseteq C(R')|_{L'}$ has distance at least δn.

We construct the sets L' and R' iteratively. Initially we set $L' = \emptyset$ and $R' = R_0$. We then iterate as follows: While there exists a vertex $u \in L - L'$ such that u has more than $\frac{1}{3}c$ neighbors in R', we add u' to L' and add all the neighbors of u' to R'. We prove below that this process stops in $t \leq \beta n$ steps, and that the induced graph on $(L - L') \cup (R - R')$ is a (good) expander.

We claim that this process must stop after at most βn steps. To see this, we count the number of unique neighbors of the set L' in the graph B. Initially this number is at most $|R_0|$. At each iteration this number goes up by at most $\frac{2}{3}c$. Assume we have completed some $t \leq \delta n$ iterations (and recall $\beta n < \delta n$). We have $|L'| = t$. Denote $\mathit{\Gamma}_{\text{unique}}(L')$ the set of vertices in R that have exactly one neighbor in L'. So $|\mathit{\Gamma}_{\text{unique}}(L')| \leq |R_0| + \frac{2}{3}ct$. Observe that $|\mathit{\Gamma}_{\text{unique}}(L')| \geq (1 - 2\gamma)c|L'|$, otherwise L' couldn't have $(1 - \gamma)c\,|L'|$ distinct neighbors (here we use $t \leq \delta n$). Putting these inequalities together we have

$$(1 - 2\gamma - \frac{2}{3})ct \leq |\mathit{\Gamma}_{\text{unique}}(L')| - \frac{2}{3}ct \leq |R_0|$$

and so $t \leq \frac{1}{(\frac{1}{3}-2\gamma)c}|R_0| \leq \frac{\alpha}{(\frac{1}{3}-2\gamma)c}|R| = \frac{\alpha}{(\frac{1}{3}-2\gamma)d}|L| = \beta n$.

Now we claim that the induced subgraph on $(L-L') \cup (R-R')$ is an expander. For this part consider any set $S \subseteq L-L'$ with $|S| \leq \delta n$. Let T be the neighborhood of S in the graph B. Then $|T| \geq (1-\gamma)c|S|$. Now each vertex of S may have upto $\frac{1}{3}c$ neighbors in R'. Even allowing for these neighborhoods to be disjoint, we get $|T \cap (R-R')| \geq (1-\gamma)c|S| - \frac{1}{3}c|S| = (\frac{2}{3}-\gamma)c|S|$. Since $\frac{2}{3}-\gamma > \frac{1}{2}$, we have that the induced subgraph on $(L-L') \cup (R-R')$ has the property that every set of size at most δn expands by more than a factor of $c/2$, thus implying that $C(R')|_{L'}$ is a code of minimum distance at least δn (see Proposition 2). This concludes the proof.

Proof (Theorem 1). Note that C is a code of distance at least δ (by Proposition 2). By Lemma 2 it follows that C is (d,α,β,δ)-smooth for any $\alpha \leq (\frac{1}{3}-2\gamma)d\delta$ and $\beta = \frac{\alpha}{(\frac{1}{3}-2\gamma)d}$. Set $\alpha = (\frac{1}{3}-2\gamma)d\delta/2$, and so $\beta = \frac{\alpha}{(\frac{1}{3}-2\gamma)d} = \delta/2$. The code is certainly $(d,\alpha,\frac{\delta}{2},\frac{\delta}{2})$-smooth.

Fix any $M \notin R \otimes C$, and let us lower bound $\frac{\rho(M)}{\delta(M)}$. Set $\rho_0 = \min\left\{\alpha\frac{\delta_R}{2d^2}, \frac{\delta_R\delta}{8}\right\}$. If $\rho(M) \geq \rho_0$ then surely $\frac{\rho(M)}{\delta(M)} \geq \rho_0$. Otherwise, we note that the conditions necessary for the application of Lemma 1 are satisfied, and we get $\delta(M) \leq 8\rho(M)$. All in all, we have proven that

$$\rho^T = \min_{M \notin R\otimes C} \frac{\rho(M)}{\delta(M)} \geq \min\left\{\rho_0, \frac{1}{8}\right\} = \rho_0 = \frac{(\frac{1}{3}-2\gamma)\delta\delta_R}{4d}$$

where the last equality follows by plugging the value for α into ρ_0 and assuming $d \geq 2$.

5 LTCs are smooth

Lemma 3. *Every (d,δ,ϵ,ρ)-LTC code C is $(d,\epsilon,\delta',\delta')$-smooth, provided $\rho \leq \delta'/4$ and $\delta' \leq \delta/4$.*

Proof. Let $B = (L,R,E)$ be a parity check graph for C whose right-hand-side corresponds to the tests of a tester for C (Proposition 1). Fix $R_0 \subseteq R$ of size $|R_0| \leq \epsilon \cdot |R|$ and consider the code $C(R_0)$. If all the non-zero words in $C(R_0)$ have weight at least δ' then setting $L_0 = \emptyset$ satisfies the definition of smoothness and so we have nothing to prove. So we assume $C(R_0)$ has some non-zero words of weight at most δ'. Let $\{c_1,\ldots,c_m\}$ be the set of all codewords of $C(R_0)$ whose weight is at most $2\delta'$. Let S_i be the set of coordinates where c_i is non-zero, and let $L_0 = \cup_i S_i$.

If $|L_0| \leq \delta' n$, we claim that $C(R_0)|_{L_0}$ has distance at least $\delta' n$ as needed. This is true since every codeword of $C(R_0)$ of weight less than $2\delta' n$ is non-zero only on some subset of L_0 and so projects to the zero codeword in $C(R_0)|_{L_0}$. On the other hand, codewords of weight greater than $2\delta' n$ in $C(R_0)$ project to words of weight at least $\delta' n$ when we delete the $\delta' n$ coordinates corresponding to L_0. Thus $C(R_0)|_{L_0}$ is a code of weight at least $\delta' n$. Thus it remains to show below that $|L_0| \leq \delta' n$.

Assume for contradiction that $|L_0| > \delta' n$. We show first that $C(R_0)$ must have a codeword of weight between $\frac{\delta'}{4}n$ and $2\delta' n$. We then show that this violates the local testability of C.

For the first part, note that if one of the c_i's has weight between $\frac{\delta'}{2}n$ and $2\delta' n$, then we are already done. So we may assume each c_i has weight less than $\frac{\delta'}{2}n$. Now pick a subset $\{c_1, \ldots, c_j\}$ of the low weight codewords so that $\frac{\delta'}{2}n \le |\cup_{i=1}^{j} S_i| \le \delta' n$. This is obviously possible since the cardinality of this union starts at 0, as j varies from 0 to m, ends at $|L_0| > \delta' n$ and goes up by at most $\frac{\delta'}{2}n$ in each step. For this setting of j, consider words of the form $\sum_{i=1}^{j} x_i c_i$ where $x_i \in \{0,1\}$. For every choice of x_i's we get a codeword of $C(R_0)$ of weight at most $|\cup_{i=1}^{j} S_i| \le \delta' n$. The expected weight of such a word, when $x_i \in \{0,1\}$ are chosen uniformly and independently is $\frac{1}{2}|\cup_{i=1}^{j} S_i| \ge \frac{\delta'}{4}n$. Thus the maximum weight codeword in this set has weight between $\frac{\delta'}{4}n$ and $\delta' n$, as desired.

Now let $c_1 \in C(R_0)$ be a codeword of weight between $\frac{\delta'}{4}n$ and $2\delta' n$. Since $2\delta' < \delta/2$ we have that c_1 is a word at distance more than $\delta' n \ge \rho n$ from C but is rejected only by the tests in R_0 which form at most ϵ fraction of all parity checks in B, contradicting the assumption that C is a $(d, \delta, \epsilon, \rho)$-LTC.

Theorem 2 follows from Lemma 3 analogous to the way Theorem 1 followed from Lemma 2.

Proof (Theorem 2). The code C is a $(d, \delta_C, \epsilon, \rho)$-LTC, with $\rho \le \delta_C/16$. By Lemma 3, it must be $(d, \epsilon, \frac{\delta_C}{4}, \frac{\delta_C}{4})$-smooth. Fix any $M \notin R \otimes C$, and let us lower bound $\frac{\rho(M)}{\delta(M)}$.

Set $\rho_0 = \min\{\frac{\epsilon\delta_R}{2d^2}, \frac{\delta_R\delta_C}{16}\}$. If $\rho(M) \ge \rho_0$ then surely $\frac{\rho(M)}{\delta(M)} \ge \rho_0$. Otherwise, we apply Lemma 1 and deduce that $\rho(M) < \rho_0$ implies that $\delta(M) \le \frac{6}{\max\{\delta_R, \delta_C/2\}}\rho(M)$.

All in all, we have proven that

$$\rho^T = \min_{M \notin R \otimes C} \frac{\rho(M)}{\delta(M)} \ge \min\left\{\rho_0, \frac{1}{8}\right\} = \min\left\{\frac{\epsilon\delta_R}{2d^2}, \frac{\delta_R\delta}{16}\right\} .$$

References

1. Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, May 1998.
2. Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM*, 45(1):70–122, January 1998.
3. Eli Ben-Sasson, Oded Goldreich, Prahladh Harsha, Madhu Sudan, and Salil Vadhan. Robust PCPs of proximity, shorter PCPs and applications to coding. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, page (to appear), 2004.
4. Eli Ben-Sasson, Oded Goldriech, Prahladh Harsha, Madhu Sudan, and Salil Vadhan. Short PCPs verifiable in polylogarithmic time. In *Proceedings of the Twelfth Annual IEEE Conference on Computational Complexity*, pages 120–134, June 12–15 2005.
5. E. Ben-Sasson and P. Harsha and S. Raskhodnikova, Some 3CNF properties are hard to test. In SIAM Journal on Computing, 35(1):1-21.
6. E. Ben-Sasson and M. Sudan. Robust locally testable codes and products of codes. In *Proc. RANDOM: International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 286–297, 2004.
7. Eli Ben-Sasson and Madhu Sudan. Short PCPs with poly-log rate and query complexity. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 266–275, 2005.
8. Eli Ben-Sasson, Madhu Sudan, Salil Vadhan, and Avi Wigderson. Randomness efficient low-degree tests and short PCPs via ϵ-biased sets. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, pages 612–621, 2003.

9. D. Coppersmith and A. Rudra. On the robust testability of product of codes. ECCC TR05-104, 2005.
10. Irit Dinur. The PCP theorem by gap amplification. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 241–250, 2006.
11. Uriel Feige, Shafi Goldwasser, Laszlo Lovasz, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *Journal of the ACM*, 43(2):268–292, 1996.
12. O. Goldreich and M. Sudan. Locally testable codes and PCPs of almost-linear length. In *Proc. 43rd IEEE Symp. on Foundations of Computer Science*, pages 13–22, 2002.
13. P. Valiant. The tensor product of two codes is not necessarily robustly testable. In *APPROX-RANDOM*, pages 472–481, 2005.