

# Quantum LDPC codes with positive rate and minimum distance proportional to $n^{1/2}$

Jean-Pierre Tillich\*      Gilles Zémor†

January 11, 2013

## Abstract

The current best asymptotic lower bound on the minimum distance of quantum LDPC codes with fixed non-zero rate is logarithmic in the blocklength. We propose a construction of quantum LDPC codes with fixed non-zero rate and prove that the minimum distance grows proportionally to the square root of the blocklength.

## 1 Introduction

LDPC codes [18] and their variants are one of the most satisfying answers to the problem of devising codes guaranteed by Shannon's theorem. They display outstanding performance for a large class of error models with a fast decoding algorithm. Generalizing these codes to the quantum setting seems a promising way to devise powerful quantum error correcting codes for protecting, for instance, the very fragile superpositions manipulated in a quantum computer. It should be emphasized that a fast decoding algorithm could be even more crucial in the quantum setting than in the classical one. In the classical case, when error correction codes are used for communication over a noisy channel, the decoding time translates directly into communication delays. This has been the driving motivation to devise decoding schemes of low complexity, and is likely to be important in the quantum setting as well. However, there is an important additional motivation for efficient decoding in the quantum setting. Quantum computation is likely to require active stabilization. The decoding time thus translates into computation delays, and most importantly in error suppression delays. If errors accumulate faster than they can be identified, quantum computation may well become infeasible: fast decoding is an essential ingredient to fault-tolerant computation.

Besides the search for efficiently decodable codes, there is additional theoretical motivation for studying quantum LDPC codes that is totally absent

---

\*INRIA, Projet Secret, BP 105, Domaine de Voluceau F-78153 Le Chesnay, France. Email: jean-pierre.tillich@inria.fr

†Institut de Mathématiques de Bordeaux, UMR 5251, Université Bordeaux 1, 351, cours de la Libération, F-33405 Talence Cedex, France Email: Gilles.Zemor@math.u-bordeaux1.fr

Material in this paper was presented in part at ISIT 2009, 799-803.

from the classical setting. The capacity of the depolarizing channel, the quantum equivalent of the binary symmetric channel, is unknown. What is clear however, is that capacity can only be achieved by making use of degeneracy: the same syndrome must be able to correct different error patterns. For this to be possible, the stabilizer group must contain elements of reasonably small weight. In other words, the quantum code must be at least partially low-density. This intuition is confirmed by the concatenated construction [14] (and the subsequent improvements [36, 16]) improving upon the hashing lower bound on the capacity of the depolarizing channel. All these constructions have many elements in the stabilizer group of very low Hamming weight. Quantum LDPC codes are not just about fast decoding schemes, but may be the way towards a better understanding of quantum channels from a purely information theory viewpoint.

Quantum generalizations of LDPC codes have been proposed in [31]. However, it has turned out that the design of high performance quantum LDPC codes is much more complicated than in the classical setting. This is due to several reasons, the most obvious of which being that the parity-check matrix of quantum LDPC codes must satisfy certain orthogonality constraints. This complicates significantly the construction of such codes. In particular, the plain random constructions that work so well in the classical setting are pointless here. There have been a number of attempts at overcoming this difficulty and a variety of methods for constructing quantum LDPC codes have been proposed [33, 27, 31, 10, 11, 30, 19, 21, 24, 15, 35, 1, 2, 23, 38, 25, 12, 3, 4]. However, all of these constructions suffer from disappointingly small minimum distances, namely whenever they have non-vanishing rate and parity-check matrices with bounded row-weight, their minimum distance is either proved to be *bounded*, or unknown and with little hope for unboundedness. The point has been made several times that minimum distance is not everything, because there are complex decoding issues involved, whose behavior depends only in part on the minimum distance, and also because a poor asymptotic behavior may be acceptable when one limits oneself to practical lengths. This is illustrated for instance in our case by the codes constructed in [25, 4] whose performance under iterative decoding are outstanding. Nevertheless, very poor minimum distances will imply significant error floors. We note also that it has recently been proved [28] that a sufficiently large growing minimum distance – for quantum LDPC codes – is enough to imply a non-zero decoding threshold, i.e. that the code corrects almost all error patterns of weight up to a value linear in the block length. Finally, the minimum distance has been the most studied parameter of error-correcting codes and given that asymptotically good (dimension and minimum distance both linear in the blocklength) quantum LDPC codes are expected to exist, it is of great theoretical interest, and possibly also practical, to devise quantum LDPC codes with large, growing, minimum distance. This is the problem that we address in the present paper, leaving aside decoding issues for discussion elsewhere.

Besides the above constructions, we must mention the design of quantum LDPC codes based on tessellations of surfaces [27, 6, 7, 13, 34], among which the most prominent example is the toric code of [27]. Toric codes have minimum

distances which grow like the square root of the blocklength and parity-check equations of weight 4 but unfortunately have fixed dimension which is 2, and hence zero rate asymptotically. It turns out that by taking appropriate surfaces of large genus, quantum LDPC codes of non vanishing rate can be constructed with minimum distance logarithmic in the blocklength, this has actually been achieved in [17, Th. 12.4], see also [40], [26]. To the best of our knowledge, this is until now the only known family of quantum LDPC codes of non-vanishing rate that yields a (slowly) growing minimum distance.

We improve here on these surface codes in several ways, by providing a flexible construction of quantum LDPC codes from any pair  $(\mathbf{H}_1, \mathbf{H}_2)$  of parity-check matrices of binary LDPC codes  $\mathcal{C}_1$  and  $\mathcal{C}_2$ . Although the constructed quantum code belongs to the CSS class [9, 37], there is no restriction on  $\mathcal{C}_1$  and  $\mathcal{C}_2$ . For instance, they do not need to be mutually orthogonal spaces as in the CSS construction. In particular we can choose  $\mathcal{C}_1 = \mathcal{C}_2$ , in which case our main result reads:

**Theorem 1** *Let  $\mathbf{H}$  be a full-rank  $(n - k) \times n$  parity-check matrix of a classical LDPC code  $\mathcal{C}$  of parameters  $[n, k, d]$ . There is a construction of a quantum LDPC code with  $\mathbf{H}$  as building block, of length  $N = n^2 + (n - k)^2$ , dimension  $k^2$ , and quantum minimum distance  $d$ . The quantum code has a stabilizer (parity-check) matrix with row weights of the form  $i + j$ , where  $i$  and  $j$  are respectively row and column weights of the original parity-check matrix  $\mathbf{H}$ .*

In particular, any family of classical asymptotically good LDPC codes of fixed rate yields a family of quantum LDPC codes of fixed rate and minimum distance proportional to a square root of the block length.

It should also be mentioned that the rate of this construction can be further improved while keeping the same minimum distance as has been observed in [29].

## 2 Basic facts about CSS codes and Tanner graphs

In this section, we recall a few basic facts about quantum codes and give some terminology about LDPC codes.

**CSS codes.** The codes constructed in this paper fall into the category of Calderbank-Shor-Steane (CSS) codes [9, 37] which belong to a more general class of quantum codes called stabilizer codes [20, 8]. The first class is described with the help of a pair of mutually orthogonal binary codes, whereas the second class is given by an additive self-orthogonal code over  $GF(4)$  with respect to the trace Hermitian inner product. Quantum codes on  $n$  qubits are linear subspaces of a Hilbert space of dimension  $2^n$  and do not necessarily have a compact representation in general. The nice feature of stabilizer codes is that they allow to define such a space with the help of a very short representation, which is given here by a set of generators of the aforementioned additive code. Each generator is viewed as an element of the Pauli group on  $n$  qubits and the quantum code is then nothing but the space stabilized by these Pauli group elements. Moreover, the set of errors that such a quantum code can correct

can also be deduced directly from this discrete representation. For the subclass of CSS codes, this representation in terms of additive self-orthogonal codes is equivalent to a representation in terms of a pair  $(\mathcal{C}_X, \mathcal{C}_Z)$  of binary linear codes satisfying the condition  $\mathcal{C}_Z^\perp \subset \mathcal{C}_X$ . The *quantum minimum distance* of such a CSS code is given by

$$\begin{aligned} d_Q &\stackrel{\text{def}}{=} \min\{d_X, d_Z\}, \text{ where} \\ d_X &\stackrel{\text{def}}{=} \min\{|x|, x \in \mathcal{C}_X \setminus \mathcal{C}_Z^\perp\}, \\ d_Z &\stackrel{\text{def}}{=} \min\{|x|, x \in \mathcal{C}_Z \setminus \mathcal{C}_X^\perp\}. \end{aligned} \quad (1)$$

Such a code allows one to protect a subspace of  $k_Q$  qubits against errors where

$$k_Q \stackrel{\text{def}}{=} \dim(\mathcal{C}_X / \mathcal{C}_Z^\perp). \quad (2)$$

$k_Q$  is called the *quantum dimension* of the CSS code. Notice that this quantity can be expressed in different ways in order to show its symmetric nature.

$$\dim(\mathcal{C}_X / \mathcal{C}_Z^\perp) = \dim \mathcal{C}_X - \dim(\mathcal{C}_Z^\perp) \quad (3)$$

$$= \dim(\mathcal{C}_X) + \dim(\mathcal{C}_Z) - n \quad (4)$$

$$= n - \dim(\mathcal{C}_X^\perp) + \dim(\mathcal{C}_Z) - n$$

$$= \dim \mathcal{C}_Z - \dim(\mathcal{C}_X^\perp)$$

$$= \dim(\mathcal{C}_Z / \mathcal{C}_X^\perp), \quad (5)$$

where  $n$  denotes the length of  $\mathcal{C}_X$  (or of  $\mathcal{C}_Z$ ).

If  $\mathbf{H}_X$  and  $\mathbf{H}_Z$  are parity-check matrices of the binary codes  $\mathcal{C}_X$  and  $\mathcal{C}_Z$  respectively, the pair  $(\mathbf{H}_X, \mathbf{H}_Z)$  is referred to either as the *stabilizer matrix* or as the parity-check matrix of the quantum code, by analogy with the classical case. Its rows are also referred to as *generators* (of the stabilizer group).

**LDPC codes.** LDPC (Low Density Parity Check) codes are linear codes which have a sparse parity-check matrix. They can be decoded by using the *Tanner graph* associated to such a parity-check matrix  $\mathbf{H}$ . This graph is defined as follows. Assume that  $\mathbf{H} = (h_{ij})_{\substack{1 \leq i \leq r \\ 1 \leq j \leq n}}$  is an  $r \times n$  matrix (where  $n$  is the length of the code). The Tanner graph, which is denoted by  $\mathcal{T}(V, C, E)$  is bipartite and has:

- (i) vertex set  $V \cup C$ , where the first set  $V$  is in bijection with the indices of the columns of  $\mathbf{H}$ , say  $V = \{1, \dots, n\}$  and is called the set of *variable nodes*, whereas the second set  $C$  is called the set of *check nodes* and is in bijection with the indices of the rows of  $\mathbf{H}$ :  $C = \{\oplus_1, \dots, \oplus_r\}$ .

- (ii) edge set  $E$ ; there is an edge between  $\oplus_i$  and  $j$  if and only if  $h_{ij} = 1$ .

A CSS code defined by a couple of binary code  $(\mathcal{C}_X, \mathcal{C}_Z)$  is said to be a *quantum LDPC code* if  $\mathcal{C}_X$  and  $\mathcal{C}_Z$  are LDPC codes, i.e. if  $\mathcal{C}_X$  and  $\mathcal{C}_Z$  have parity-check matrices  $\mathbf{H}_X$  and  $\mathbf{H}_Z$  that are both sparse.

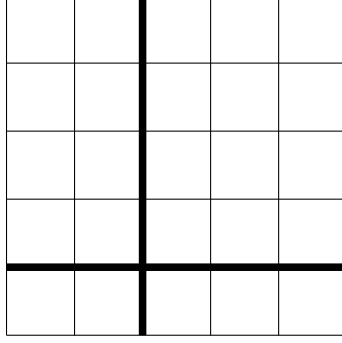


Figure 1: a two-dimensional torus: identify opposing sides of the outer square.

### 3 The toric code and its generalization

Our construction borrows both from classical LDPCs and Kitaev's toric quantum code [27]. It is not a coincidence that we obtain minimum distances that grow like the square root of the blocklength, similarly to the toric code, but we shall achieve much larger dimensions that can grow linearly in the blocklength. To get a clear picture of the construction it is desirable to take a close look at the toric code and explain how we shall generalize it.

**The toric code.** The toric code is based on the graph  $\mathcal{G}$  represented on Figure 1 which is a tiling of the 2-dimensional torus. The vertex set of the graph is  $\mathcal{V} = \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$  and there is an edge between every vertex  $(x, y)$  and the four vertices  $(x \pm 1, y)$ ,  $(x, y \pm 1)$ . In the whole subsection, addition and subtraction are performed modulo  $m$ . Now number the edges from 1 to  $n = 2m^2$  so as to identify the edge set with  $\{1, 2, \dots, n\}$ . The ambient space  $\mathbb{F}_2^n$  is therefore identified with subsets of edges. The matrix  $\mathbf{H}_X = (h_{ij})$  is the vertex-edge incident matrix, rows are indexed by vertices of the graph  $\mathcal{G}$ , and  $h_{ij} = 1$  iff vertex  $i$  is incident to edge  $j$ . The associated code  $\mathcal{C}_X$  is the *cycle code* of  $\mathcal{G}$ , a cycle being by definition a set of edges  $\mathcal{Z}$  such that every vertex is incident to an even number of edges of  $\mathcal{Z}$ . Elements of the row-space  $\mathcal{C}_X^\perp$  are called *cocycles*, rows of  $\mathbf{H}_X$  are called *elementary cocycles*, and the row-space itself  $\mathcal{C}_X^\perp$  is also known as the *cocycle code* of  $\mathcal{G}$ .

The second matrix  $\mathbf{H}_Z$  of the quantum code is defined as the face-edge incidence matrix. The faces are defined as the 4-cycles  $(x, y), (x + 1, y), (x + 1, y + 1), (x, y + 1)$ .

The rowspace  $\mathcal{C}_Z^\perp$  of  $\mathbf{H}_Z$  is therefore a subspace of the cycle code  $\mathcal{C}_X$ , and the quotient  $\mathcal{C}_X/\mathcal{C}_Z^\perp$  is readily seen to have dimension 2, coset leaders of the quotient being given by cycles of the form  $(a, 0), (a, 1), \dots, (a, m - 1)$  and  $(0, a), (1, a), \dots, (m - 1, a)$ , as represented by the thick lines on Figure 1. The dimension of the quantum code is therefore equal to 2 and the minimum weight of a vector of  $\mathcal{C}_X$  not in  $\mathcal{C}_Z^\perp$  is therefore equal to  $m$ .

To conclude that the minimum distance of the quantum code is actually  $m$ , it remains to determine the minimum weight of a vector of  $\mathcal{C}_Z$  that is not in  $\mathcal{C}_X^\perp$ , i.e. that is not a cocycle. This particular graph  $\mathcal{G}$  has the nice property of being

a tiling of a surface (the torus). This embedding into a surface allows one to define its *dual graph*. The (Poincaré) dual graph  $\mathcal{G}'$  has vertex set equal to the faces of  $\mathcal{G}$ , and there is an edge between two vertices of  $\mathcal{G}'$  if the corresponding faces of  $\mathcal{G}$  have a common edge in  $\mathcal{G}$ . Furthermore the dual graph  $\mathcal{G}'$  of  $\mathcal{G}$  is isomorphic to  $\mathcal{G}$  itself, and given that the edges of  $\mathcal{G}$  define the edges of  $\mathcal{G}'$ , the ambient space  $\mathbb{F}_2^n$  can be identified with the edge set of the dual graph  $\mathcal{G}'$ . With this identification, the elementary cocycles of  $\mathcal{G}$  become the faces of  $\mathcal{G}'$  and the faces of  $\mathcal{G}$  become the elementary cocycles of  $\mathcal{G}'$ . Hence the minimum weight of a vector of  $\mathcal{C}_Z$  that is not in  $\mathcal{C}_X^\perp$  is exactly the same as the minimum weight of a vector of  $\mathcal{C}_X$  not in  $\mathcal{C}_Z^\perp$  and the minimum distance of the quantum code is exactly  $m$ .

This duality argument is quite powerful because it ensures that whatever we prove on the weight of codewords of  $\mathcal{C}_X$  not in  $\mathcal{C}_Z^\perp$  is also valid for the weight of codewords of  $\mathcal{C}_Z$  not in  $\mathcal{C}_X^\perp$ : for this reason a number of quantum codes that arise by replacing the graph  $\mathcal{G}$  by different tilings of different surfaces have been investigated (surface codes). Here we shall consider a different generalization that does not destroy graph duality but generalizes it.

Our first remark is that the graph  $\mathcal{G}$  is a product graph: it is the product of two graphs each equal to an elementary cycle of length  $m$ . The (Cartesian) product of two graphs is namely defined as follows.

**Definition 2 (graph product)** *The product  $\mathcal{G}_1 \times \mathcal{G}_2$  of two graphs  $\mathcal{G}_1$  and  $\mathcal{G}_2$  has vertex set made up of couples  $(x, y)$ , where  $x$  is a vertex of  $\mathcal{G}_1$  and  $y$  of  $\mathcal{G}_2$ . The edges of the product graph connect two vertices  $(x, y)$  and  $(x', y')$  if either  $x = x'$  and  $\{y, y'\}$  is an edge of  $\mathcal{G}_2$  or  $y = y'$  and  $\{x, x'\}$  is an edge of  $\mathcal{G}_1$ .*

Note that any two edges  $\{a, b\}$  and  $\{x, y\}$  of  $\mathcal{G}_1$  and  $\mathcal{G}_2$  define the 4-cycle of  $\mathcal{G}_1 \times \mathcal{G}_2$  :

$$\begin{array}{ccc} (a, y) & \text{---} & (b, y) \\ | & & | \\ (a, x) & \text{---} & (b, x) \end{array} \quad (\star)$$

Now, we are tempted to define a quantum code by, as before, declaring  $\mathbf{H}_X$  to be the vertex-edge incident matrix of a product graph  $\mathcal{G} = \mathcal{G}_1 \times \mathcal{G}_2$  of two arbitrary graphs, and by declaring  $\mathbf{H}_Z$  to be the matrix whose rows are the characteristic vectors of all faces, i.e. the 4-cycles of the form  $(\star)$ . This is a quantum code which generalizes the toric code, since the latter corresponds to the case when  $\mathcal{G}_1$  and  $\mathcal{G}_2$  are two cycles of length  $m$ . This construction loses graph duality however, and our objective was to preserve it. In particular, it is not clear at all why the argument used for obtaining the value of  $d_X$  can also be used for deriving  $d_Z$ . But a closer look shows us that graph duality has not completely gone: the dual has simply become a *hypergraph*, whose vertex set is the set of faces of  $\mathcal{G}$  and where the hyperedges are the subsets of those faces of  $\mathcal{G}$  that meet in a common edge of  $\mathcal{G}$ . This observation shows us that we really should consider products of hypergraphs rather than graph products

to start with. This is the approach which was followed in a preliminary version of this work [39] where the construction is described in terms of products of hypergraphs. This way of viewing the toric code highlights the connections with algebraic topology and is also a natural generalization of the quantum code construction known under the name of surface codes.

However, we shall slightly change our point of view and detail our construction in a way that stays with the more familiar notion of graph product. This is the approach which we follow in the next paragraphs. The hypergraph connection will become apparent in Sections 4.3 and 4.4.

**Another way of viewing the toric code in terms of a graph product construction.** There arises another interpretation of the toric code as a product of two cycles, of length double the previous length, by considering simultaneously the Tanner graphs of  $\mathcal{C}_X$  and  $\mathcal{C}_Z$  and by putting the qubits on the vertices instead of the edges. Indeed, let us consider for instance the Tanner graph corresponding to  $\mathbf{H}_X$  for  $m = 3$  which is depicted in Figure 2.

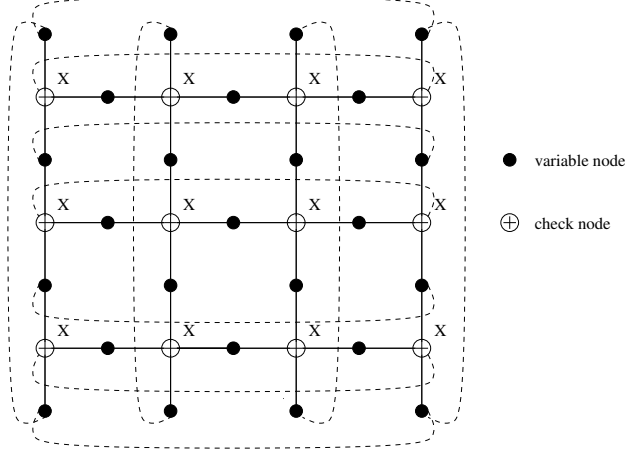


Figure 2: The Tanner graph of  $\mathcal{C}_X$  for  $m = 3$ : the dashed lines indicate the pair of vertices which are identified.

It is insightful to consider the union of both Tanner graphs of  $\mathcal{C}_X$  and  $\mathcal{C}_Z$  (see Figure 3).

The Tanner graph of  $\mathcal{C}_X$  does not have a product graph structure, whereas the union of both Tanner graphs has now the structure of the product of two cycles of length 6 in the example and  $2m$  in general. Notice that such a cycle can be viewed as the Tanner graph of the repetition code of length  $m$ . Moreover, it is clear from this picture why the rows of  $\mathbf{H}_X$  and  $\mathbf{H}_Z$  are orthogonal : two parity-check nodes of type  $X$  and  $Z$  respectively which are adjacent to a same variable node are also adjacent to a second variable node as shown in Figure 4. This comes from the very definition of a product graph as explained before.

This discussion strongly suggests to consider the following generalization of the toric code.

**Definition 3 (CSS code  $\mathcal{Q}(\mathcal{G}_1 \times \mathcal{G}_2)$  associated to a graph product)**

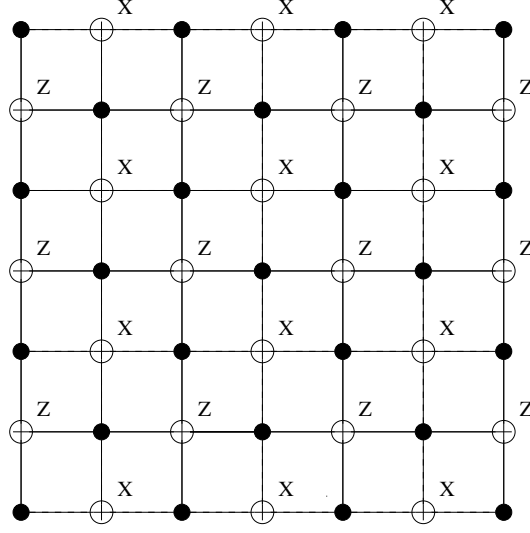


Figure 3: The union of the Tanner graph of  $\mathcal{C}_X$  and  $\mathcal{C}_Z$  for  $m = 3$ : identify the opposing vertices of the outer square, the dashed lines indicate here the edges of the Tanner graph of  $\mathcal{C}_Z$ , the solid lines the edges of the Tanner graph of  $\mathcal{C}_X$ . The check nodes marked with an “X”, respectively with a “Z” belong to the Tanner graph of  $\mathcal{C}_X$ , respectively  $\mathcal{C}_Z$ .

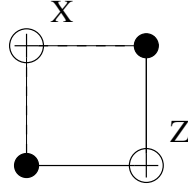


Figure 4: Two check nodes of different types sharing two different variable nodes.

Let  $\mathcal{G}_1 = \mathcal{T}(V_1, C_1, E_1)$  and  $\mathcal{G}_2 = \mathcal{T}(V_2, C_2, E_2)$  be two Tanner graphs. Let

$$V \stackrel{\text{def}}{=} V_1 \times V_2 \cup C_1 \times C_2$$

$$C \stackrel{\text{def}}{=} C_1 \times V_2 \cup V_1 \times C_2,$$

so that the product graph  $\mathcal{G}_1 \times \mathcal{G}_2$  is a bipartite graph with vertex set  $V \cup C$ . Let  $\mathcal{G}_1 \times_X \mathcal{G}_2$  be the Tanner graph defined as the subgraph of  $\mathcal{G}_1 \times \mathcal{G}_2$  with set of variable nodes  $V$  and set of check nodes  $C_1 \times V_2$ . Similarly, define the Tanner graph  $\mathcal{G}_1 \times_Z \mathcal{G}_2$  as the subgraph of  $\mathcal{G}_1 \times \mathcal{G}_2$  with set of variable nodes  $V$  and set of check nodes  $V_1 \times C_2$ : the union of the two edge-sets of  $\mathcal{G}_1 \times_X \mathcal{G}_2$  and  $\mathcal{G}_1 \times_Z \mathcal{G}_2$  make up therefore the edge set of  $\mathcal{G}_1 \times \mathcal{G}_2$ . Finally, define the two classical codes  $\mathcal{C}_X = \mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)$  and  $\mathcal{C}_Z = \mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)$  as the codes associated to the Tanner graphs  $\mathcal{G}_1 \times_X \mathcal{G}_2$  and  $\mathcal{G}_1 \times_Z \mathcal{G}_2$  respectively.

The couple  $(\mathcal{C}_X, \mathcal{C}_Z)$  defines a CSS code that we denote by  $\mathcal{Q}(\mathcal{G}_1 \times \mathcal{G}_2)$ .

The construction is summarized on Figure 5.

It is readily checked that if  $\mathcal{G}_1 = \mathcal{G}_2$  are two cycles of even length  $2m$ , so that  $\mathcal{G}_1$  and  $\mathcal{G}_2$  are two isomorphic bipartite graphs with  $|V_1| = |V_2| = |C_1| =$



$|C_2| = m$ , then the two Tanner graphs  $\mathcal{G}_1 \times_X \mathcal{G}_2$  and  $\mathcal{G}_1 \times_Z \mathcal{G}_2$  are the familiar Tanner graphs of  $\mathcal{C}_X$  and  $\mathcal{C}_Z$  depicted on Figure 2.

It will be explained shortly in the next section why the 4-cycle joining  $(v_1, v_2), (c_1, v_2), (c_1, c_2)$  and  $(v_1, c_2)$  ensures that the couple  $(\mathcal{C}_X, \mathcal{C}_Z)$  always defines indeed a CSS code. It will also turn out that the minimum distance of the quantum code is related to the minimum distance of the classical binary codes with Tanner graphs  $\mathcal{G}_1$  and  $\mathcal{G}_2$ .

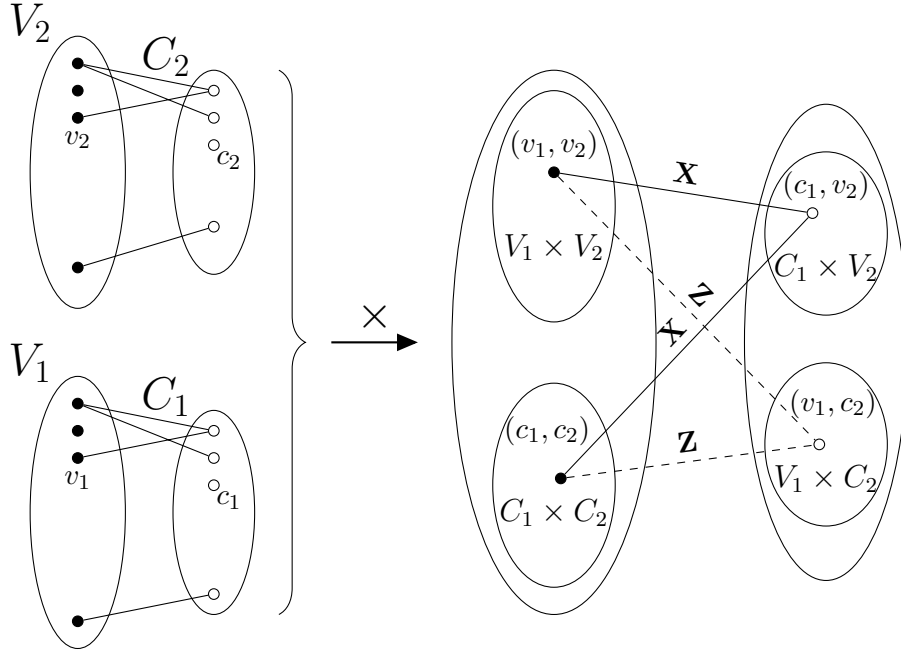


Figure 5: The Tanner graphs of  $\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)$  and  $\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)$  are given by the edges of  $\mathcal{G}_1 \times \mathcal{G}_2$  which join  $V$  to  $C_1 \times V_2$  for  $\mathcal{C}_X$  (the corresponding edges are denoted by solid lines) and which join  $V$  to  $V_1 \times C_2$  for  $\mathcal{C}_Z$  (the corresponding edges are indicated by dashed lines).

## 4 Dimension of the CSS code $\mathcal{Q}(\mathcal{G}_1 \times \mathcal{G}_2)$ and relationship with product codes

### 4.1 Validity of the construction of $\mathcal{Q}(\mathcal{G}_1 \times \mathcal{G}_2)$ .

We prove here that  $\mathcal{Q}(\mathcal{G}_1 \times \mathcal{G}_2)$  is indeed a CSS code. This follows at once from

**Proposition 4** *Let  $\mathcal{G}_1 = \mathcal{T}(V_1, C_1, E_1)$  and  $\mathcal{G}_2 = \mathcal{T}(V_2, C_2, E_2)$  be two Tanner graphs. We have:*

$$\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)^\perp \subset \mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)$$

*Proof :* Let  $\mathbf{H}_X$  and  $\mathbf{H}_Z$  be the parity-check matrices associated to the Tanner graphs  $\mathcal{G}_1 \times_X \mathcal{G}_2$  and  $\mathcal{G}_1 \times_Z \mathcal{G}_2$  respectively. For  $c_1 \in C_1$  and  $v_2 \in V_2$ , denote by

$\mathbf{h}_X(c_1, v_2)$  the row of  $\mathbf{H}_X$  corresponding to the check node  $(c_1, v_2)$  of  $\mathcal{G}_1 \times_X \mathcal{G}_2$ . It will be convenient to view vectors as sets of (variable) nodes of  $\mathcal{G}_1 \times \mathcal{G}_2$  by identifying them with their supports, so that  $\mathbf{h}_X(c_1, v_2)$  is the set of neighbors of  $(c_1, v_2)$  in  $\mathcal{G}_1 \times \mathcal{G}_2$ . Similarly, row  $\mathbf{h}_Z(v_1, c_2)$  of  $\mathbf{H}_Z$  should be thought of the set of neighbors of  $(v_1, c_2)$  in the graph  $\mathcal{G}_1 \times \mathcal{G}_2$ , for some  $v_1 \in V_1$  and  $c_2 \in C_2$ .

To prove that  $\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)^\perp \subset \mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)$  it is sufficient to prove that for any  $v_i \in V_i$ ,  $c_i \in C_i$ ,  $i \in \{1, 2\}$ , row  $\mathbf{h}_X(c_1, v_2)$  of  $\mathbf{H}_X$  is orthogonal to row  $\mathbf{h}_Z(v_1, c_2)$  of  $\mathbf{H}_Z$ . This is achieved as follows. We first notice that the inner product between  $\mathbf{h}_X(c_1, v_2)$  and  $\mathbf{h}_Z(v_1, c_2)$  can be expressed as

$$\langle \mathbf{h}_X(c_1, v_2), \mathbf{h}_Z(v_1, c_2) \rangle = \#S \pmod{2},$$

where  $S$  is the set of elements of  $V_1 \times V_2 \cup C_1 \times C_2$  which are adjacent to both  $(c_1, v_2)$  and  $(v_1, c_2)$  in  $\mathcal{G}_1 \times \mathcal{G}_2$ . Now the set  $S$  is clearly empty if either  $v_1$  is not adjacent to  $c_1$  in  $\mathcal{G}_1$  or if  $c_2$  is not adjacent to  $v_2$  in  $\mathcal{G}_2$ . When  $v_i$  is adjacent to  $c_i$  in  $\mathcal{G}_i$  for  $i = 1, 2$ , then there are exactly two vertices in  $V_1 \times V_2 \cup C_1 \times C_2$  which are both adjacent to  $(c_1, v_2)$  and  $(v_1, c_2)$ , namely  $(c_1, c_2)$  and  $(v_1, v_2)$ . This implies in both cases that  $\mathbf{h}_X(c_1, v_2)$  and  $\mathbf{h}_Z(v_1, c_2)$  are orthogonal. ■

#### 4.2 Degree structure of the Tanner graphs $\mathcal{G}_1 \times_X \mathcal{G}_2$ and $\mathcal{G}_1 \times_Z \mathcal{G}_2$ of $\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)$ and $\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)$ .

The Tanner graphs of all the constructions of quantum LDPC codes which have been proposed so far have in general a very specific degree structure. Unfortunately it is well known that in order to obtain classical LDPC codes which allow to operate successfully at rates very close to capacity, the degree structure has to be optimized very carefully. This kind of approach can not be carried out for any of the constructions which have been mentioned in the introduction but one (namely the the construction based on LDGM codes [30, 19]). However in this case, the minimum distance is constant and there is no threshold error probability for decoding. The construction which is proposed here allows in principle this kind of optimization up to some extent as we will see now.

**Proposition 5** *Let  $(\lambda_i(j))_j, (\rho_i(j))_j$  be respectively the right and left degree distributions of the Tanner graph  $\mathcal{G}_i = \mathcal{T}(V_i, C_i, E_i)$  for  $i \in \{1, 2\}$  (that is  $\lambda_i(j)$  is the fraction of vertices of  $V_i$  of degree  $j$ , whereas  $\rho_i(j)$  is the fraction of vertices of  $C_i$  of degree  $j$ ). We let  $\lambda_X(j)$ , respectively  $\rho_X(j)$ , be the fraction of vertices of  $V_1 \times V_2 \cup C_1 \times C_2$ , respectively of  $C_1 \times V_2$  of degree  $j$  in  $\mathcal{G}_1 \times_X \mathcal{G}_2$ .  $\lambda_Z(j)$ , respectively  $\rho_Z(j)$ , are defined similarly as the fraction of vertices of  $V_1 \times V_2 \cup C_1 \times C_2$ , respectively of  $V_1 \times C_2$  of degree  $j$  in  $\mathcal{G}_1 \times_Z \mathcal{G}_2$ . For  $i \in \{1, 2\}$  we define the following average left and right degrees*

$$\overline{\lambda_i} \stackrel{\text{def}}{=} \sum_j \lambda_i(j)j \tag{6}$$

$$\overline{\rho_i} \stackrel{\text{def}}{=} \sum_j \rho_i(j)j \tag{7}$$

We have

$$\lambda_X(j) = \frac{\lambda_1(j) + \frac{\bar{\lambda}_1 \bar{\lambda}_2}{\bar{\rho}_1 \bar{\rho}_2} \rho_2(j)}{\frac{\bar{\lambda}_1 \bar{\lambda}_2}{\bar{\rho}_1 \bar{\rho}_2} + 1} \quad (8)$$

$$\lambda_Z(j) = \frac{\lambda_2(j) + \frac{\bar{\lambda}_1 \bar{\lambda}_2}{\bar{\rho}_1 \bar{\rho}_2} \rho_1(j)}{\frac{\bar{\lambda}_1 \bar{\lambda}_2}{\bar{\rho}_1 \bar{\rho}_2} + 1} \quad (9)$$

$$\rho_X(k) = \sum_{i,j:i+j=k} \rho_1(i) \lambda_2(j) \quad (10)$$

$$\rho_Z(k) = \sum_{i,j:i+j=k} \lambda_1(i) \rho_2(j) \quad (11)$$

*Proof :* The number of vertices of degree  $j$  in  $V_1 \times V_2$  of the graph  $\mathcal{G}_1 \times_X \mathcal{G}_2$  is equal to  $|V_1||V_2|\lambda_1(j)$ , whereas the number of vertices of degree  $j$  in  $C_1 \times C_2$  of degree  $j$  is equal to  $|C_1||C_2|\rho_2(j)$ . This implies that

$$\begin{aligned} \lambda_X(j) &= \frac{|V_1||V_2|\lambda_1(j) + |C_1||C_2|\rho_2(j)}{|V_1||V_2| + |C_1||C_2|} \\ &= \frac{\lambda_1(j) + \frac{\bar{\lambda}_1 \bar{\lambda}_2}{\bar{\rho}_1 \bar{\rho}_2} \rho_2(j)}{\frac{\bar{\lambda}_1 \bar{\lambda}_2}{\bar{\rho}_1 \bar{\rho}_2} + 1} \end{aligned}$$

since  $\frac{|C_i|}{|V_i|} = \frac{\bar{\lambda}_i}{\bar{\rho}_i}$ . The number of vertices of degree  $k$  in  $C_1 \times V_2$  of the graph  $\mathcal{G}_1 \times_X \mathcal{G}_2$  is equal to the number of vertices  $(c_1, v_2) \in C_1 \times V_2$  such that the degree of  $c_1$  in  $\mathcal{G}_1$  plus the degree of  $v_2$  in  $\mathcal{G}_2$  is equal to  $k$ . This number is therefore equal to  $\sum_{i,j:i+j=k} |C_1||V_2|\rho_1(i)\lambda_2(j)$ . This yields immediately that

$$\rho_X(k) = \sum_{i,j:i+j=k} \rho_1(i) \lambda_2(j).$$

The formulas for  $\lambda_Z(j)$  and  $\rho_Z(k)$  are obtained in a similar fashion. ■

This result implies that there is a large degree of freedom for choosing the degree distributions of the Tanner graphs  $\mathcal{G}_1 \times_X \mathcal{G}_2$  and  $\mathcal{G}_1 \times_Z \mathcal{G}_2$ . This can potentially be used to devise quantum LDPC codes with very good iterative decoding performances.

### 4.3 The hypergraph connection, product codes

Our objective is to derive a formula for the dimension of the quantum code  $\mathcal{Q}(\mathcal{G}_1 \times \mathcal{G}_2)$ . For this purpose we shall relate the Tanner graphs  $\mathcal{G}_1 \times_X \mathcal{G}_2$  and  $\mathcal{G}_1 \times_Z \mathcal{G}_2$  to Tanner graphs of product codes.

We first need to define another product notion  $\mathcal{G}_1 \otimes \mathcal{G}_2$  for two Tanner graphs  $\mathcal{G}_1$  and  $\mathcal{G}_2$  that will make  $\mathcal{G}_1 \otimes \mathcal{G}_2$  the Tanner graph of the product code  $\mathcal{C}_1 \otimes \mathcal{C}_2$  when  $\mathcal{C}_1$  and  $\mathcal{C}_2$  are the codes associated to the Tanner graphs  $\mathcal{G}_1$  and  $\mathcal{G}_2$ . This product is the natural extension of Definition 2 when  $\mathcal{G}_1$  and  $\mathcal{G}_2$  are viewed as *hypergraphs*.

Recall that a *hypergraph*  $\mathcal{H} = (V, \mathcal{E})$  is simply a set  $V$  together with a collection  $\mathcal{E}$  of subsets called *hyperedges* or simply *edges*. A hypergraph is a graph when every edge has cardinality 2. Consider the following definition:

**Definition 6** Let  $\mathcal{H}_1 = (V_1, \mathcal{E}_1)$  and  $\mathcal{H}_2 = (V_2, \mathcal{E}_2)$  be two hypergraphs. The product hypergraph  $\mathcal{H}_1 \times \mathcal{H}_2$  is defined as the hypergraph  $\mathcal{H} = (V, \mathcal{E})$  such that  $V = V_1 \times V_2$  and  $\mathcal{E}$  is the collection of subsets of  $V$

- either of the form  $\{v_1\} \times e_2$  with  $v_1 \in V_1$  and  $e_2 \in \mathcal{E}_2$ ,
- or of the form  $e_1 \times \{v_2\}$  with  $e_1 \in \mathcal{E}_1$  and  $v_2 \in V_2$ .

It should be clear that when  $\mathcal{H}_1$  and  $\mathcal{H}_2$  are graphs then  $\mathcal{H}_1 \times \mathcal{H}_2$  reduces to the graph product of Definition 2.

Now hypergraphs are often described by their associated bipartite graph: for a hypergraph  $\mathcal{H} = (V, \mathcal{E})$  this bipartite graph has vertex set  $V \cup \mathcal{E}$  and we put an edge between  $v \in V$  and  $e \in \mathcal{E}$  whenever  $v \in e$  in  $\mathcal{H}$ . Conversely, any bipartite graph, in other words any Tanner graph  $\mathcal{T}(V, C, E)$ , generates a hypergraph on vertex set  $V$  by declaring that the hyperedges are the neighborhoods of the check vertices  $c$ , when  $c$  ranges over  $C$ .

Tanner graphs are of course simply bipartite graphs. The terminology “Tanner” serves the sole purpose of reminding us that we are interested in the associated error-correcting code. Translated into Tanner graph terminology, Definition 6 becomes:

**Definition 7** Let  $\mathcal{G}_1 = \mathcal{T}(V_1, C_1, E_1)$  and  $\mathcal{G}_2 = \mathcal{T}(V_2, C_2, E_2)$  be two Tanner graphs. The hypergraph product  $\mathcal{G}_1 \otimes \mathcal{G}_2$  of  $\mathcal{G}_1$  and  $\mathcal{G}_2$  is defined as the induced subgraph of  $\mathcal{G}_1 \times \mathcal{G}_2$  with variable node set  $V_1 \times V_2$  and check node set  $C_1 \times V_2 \cup V_1 \times C_2$ .

Now the hypergraph product of Tanner graphs is directly related to the standard product construction for codes. We recall the definition:

**Definition 8 (Product code)** Let  $\mathcal{C}_1$  and  $\mathcal{C}_2$  be two binary codes of length  $n_1$  and  $n_2$  respectively. The product code  $\mathcal{C}_1 \otimes \mathcal{C}_2$  of  $\mathcal{C}_1$  and  $\mathcal{C}_2$  is the binary code of length  $n_1 \times n_2$  whose codewords may be viewed as binary matrices of size  $n_1 \times n_2$  and such that a matrix belongs to  $\mathcal{C}_1 \otimes \mathcal{C}_2$  if and only if all its columns belong to  $\mathcal{C}_1$  and all its rows to  $\mathcal{C}_2$ .

It is well known (see [32, Ch. 18. §2] for instance) that the dimension of the product code is given by

**Proposition 9** Let  $\mathcal{C}_1$  and  $\mathcal{C}_2$  be two binary linear codes.

$$\dim(\mathcal{C}_1 \otimes \mathcal{C}_2) = \dim \mathcal{C}_1 \dim \mathcal{C}_2.$$

The following proposition is straightforward:

**Proposition 10** For  $i \in \{1, 2\}$ , let  $\mathcal{G}_i = \mathcal{T}(V_i, C_i, E_i)$  be a Tanner graph of the binary linear code  $\mathcal{C}_i$ . A Tanner graph for  $\mathcal{C}_1 \otimes \mathcal{C}_2$  is given by  $\mathcal{G}_1 \otimes \mathcal{G}_2$ .

Now to describe the Tanner graphs  $\mathcal{G}_1 \times_X \mathcal{G}_2$  and  $\mathcal{G}_1 \times_Z \mathcal{G}_2$  in terms of hypergraph products and to derive the dimensions of the associated codes  $\mathcal{C}_X$  and  $\mathcal{C}_Z$ , we need one extra tool:

#### 4.4 The Transpose Tanner Graph

**Definition 11 (Transpose of a Tanner graph)** *The transpose of a Tanner graph  $\mathcal{T}(V, C, E)$  is the Tanner graph  $\mathcal{T}(C, V, E)$ .*

In other words, transposing a Tanner graph amounts to exchanging the role of the variable node set with the check node set. For a binary linear code  $\mathcal{C}$ , we shall abuse notation somewhat and denote by  $\mathcal{C}^T$  (and call it a transpose code of  $\mathcal{C}$ ) the binary code specified by  $\mathcal{G}^T$  where  $\mathcal{G}$  is a Tanner graph for  $\mathcal{C}$ , even if this notion assumes implicitly some choice for the Tanner graph of  $\mathcal{C}$ . We do this whenever the choice of the underlying Tanner graph is obvious from the context. Note that the length of  $\mathcal{C}^T$  can be varied by adding or removing redundant parity-checks to the parity-check matrix/Tanner graph for  $\mathcal{C}$ . The transpose code can be viewed as the code associated to the linear combinations of the rows of a parity-check matrix of the code which are equal to 0. This interpretation directly leads to the following relationship between the dimension of a code and its transpose:

**Fact 12** *Let  $\mathcal{C}$  be a binary code specified by a Tanner graph  $\mathcal{G} = \mathcal{T}(V, C, E)$ . Then the dimension  $\dim(\mathcal{C})$  of  $\mathcal{C}$  and the dimension  $\dim(\mathcal{C}^T)$  of the code associated to the transpose  $\mathcal{G}^T$  are related by*

$$\dim(\mathcal{C}) = |V| - |C| + \dim(\mathcal{C}^T). \quad (12)$$

We are now ready now to derive the aforementioned relationship between the Tanner graphs  $\mathcal{G}_1 \times_X \mathcal{G}_2$  and  $\mathcal{G}_1 \times_Z \mathcal{G}_2$  and Tanner graphs of product codes.

**Proposition 13** *For  $i \in \{1, 2\}$ , let  $\mathcal{G}_i = \mathcal{T}(V_i, C_i, E_i)$  be the Tanner graph of a binary code  $\mathcal{C}_i$ , then*

$$\mathcal{G}_1 \times_X \mathcal{G}_2 = (\mathcal{G}_1^T \otimes \mathcal{G}_2)^T \quad (13)$$

$$\mathcal{G}_1 \times_Z \mathcal{G}_2 = (\mathcal{G}_1 \otimes \mathcal{G}_2^T)^T. \quad (14)$$

*In other words,*

$$\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)^T = \mathcal{C}_1^T \otimes \mathcal{C}_2 \quad (15)$$

$$\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)^T = \mathcal{C}_1 \otimes \mathcal{C}_2^T. \quad (16)$$

*Proof :* Remember that  $\mathcal{G}_1 \times_X \mathcal{G}_2$  is the subgraph of  $\mathcal{G}_1 \times \mathcal{G}_2$  induced by vertex node set  $V_1 \times V_2 \cup C_1 \times C_2$  and check node set  $C_1 \times V_2$ . This gives (13) by definition of the hypergraph product: (14) is obtained analogously. Equalities (15) and (16) follow directly. ■

From this last proposition, Proposition 9 and Fact 12, we obtain immediately that the dimension of the quantum code  $\mathcal{Q}(\mathcal{G}_1 \times \mathcal{G}_2)$  is given by

**Proposition 14** *For  $i \in \{1, 2\}$ , let  $\mathcal{G}_i$  be the Tanner graph of a binary code  $\mathcal{C}_i$  and let  $k_i = \dim(\mathcal{C}_i)$ ,  $k_i^T = \dim(\mathcal{C}_i^T)$ . The quantum dimension  $k_{\mathcal{Q}}$  of  $\mathcal{Q}(\mathcal{G}_1 \times \mathcal{G}_2)$  is given by*

$$k_{\mathcal{Q}} = k_1 k_2 + k_1^T k_2^T.$$

*Proof :* Let  $r_i$  be the number of check nodes of  $\mathcal{G}_i$  and let  $n_i$  be the length of  $\mathcal{C}_i$ . By using first Fact 12, then Proposition 13 and finally Proposition 9, we obtain

$$\begin{aligned}
\dim(\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)) &= n_1 n_2 + r_1 r_2 - r_1 n_2 + \dim(\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)^T) \\
&= n_1 n_2 + r_1 r_2 - r_1 n_2 + \dim(\mathcal{C}_1^T \otimes \mathcal{C}_2) \\
&= n_1 n_2 + r_1 r_2 - r_1 n_2 + \dim(\mathcal{C}_1^T) \dim(\mathcal{C}_2) \\
&= n_1 n_2 + r_1 r_2 - r_1 n_2 + k_1^T k_2
\end{aligned}$$

We derive in a similar way that

$$\dim(\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)) = n_1 n_2 + r_1 r_2 - r_2 n_1 + k_2^T k_1$$

By using Formula (4) for the quantum dimension we obtain

$$\begin{aligned}
k_Q &= \dim(\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)) - n_1 n_2 - r_1 r_2 + \dim(\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)) \\
&= n_1 n_2 + r_1 r_2 - r_1 n_2 + k_1^T k_2 - n_1 n_2 - r_1 r_2 \\
&\quad + n_1 n_2 + r_1 r_2 - r_2 n_1 + k_2^T k_1 \\
&= (n_1 - r_1)(n_2 - r_2) + k_1^T k_2 + k_2^T k_1 \\
&= (k_1 - k_1^T)(k_2 - k_2^T) + k_1^T k_2 + k_2^T k_1 \\
&= k_1 k_2 + k_1^T k_2^T.
\end{aligned}$$

■

**Comment.** If we go back to our initial definition of Kitaev's toric code, with  $\mathcal{C}_X$  being defined as the cycle code of the (graph) product of a cycle with itself, and if we wish to express the Tanner graph of  $\mathcal{C}_X$  as a function of the Tanner graph  $\mathcal{G}$  of the original cycle, then the more natural expression for  $\mathcal{C}_X$  will be that it has the Tanner graph

$$(\mathcal{G} \otimes \mathcal{G})^T$$

and then  $\mathcal{C}_Z$  will come out as:

$$(\mathcal{G}^T \otimes \mathcal{G}^T)^T.$$

However, since  $\mathcal{G}^T$  is isomorphic to  $\mathcal{G}$  (a cycle of even length), the more twisted Tanner graph expressions in (13) and (14) generalize the toric code just as well. Definition 3, equivalently Proposition 13, has the following two advantages:

- the quantum code  $\mathcal{Q}(\mathcal{G} \times \mathcal{G})$  always has positive dimension for  $\mathcal{G}$  the Tanner graph of any non-trivial code, by Proposition 14,
- $\mathcal{C}_X(\mathcal{G} \otimes \mathcal{G})$  and  $\mathcal{C}_Z(\mathcal{G} \otimes \mathcal{G})$  are clearly isomorphic.

## 5 Minimum distance

In this section, we show that the minimum distance of the quantum code  $\mathcal{Q}(\mathcal{G}_1 \times \mathcal{G}_2)$  has a very simple expression, when we adopt the convention that the minimum distance of a code reduced to the all-zero codeword is  $\infty$

**Theorem 15** *For  $i \in \{1, 2\}$ , let  $d_i$  be the minimum distance of a code with Tanner graph  $\mathcal{G}_i$  and let  $d_i^T$  denote the minimum distance of the code specified by the transpose Tanner graph  $\mathcal{G}_i^T$ . The minimum distance  $d_{\mathcal{Q}}$  of the quantum code  $\mathcal{Q}(\mathcal{G}_1 \times \mathcal{G}_2)$  satisfies*

$$d_{\mathcal{Q}} \geq \min(d_1, d_2, d_1^T, d_2^T).$$

and is given by

$$d_{\mathcal{Q}} = \min(d_1, d_2, d_1^T, d_2^T).$$

in the following cases

- $d_i = \min(d_1, d_2, d_1^T, d_2^T)$  for some  $i \in \{1, 2\}$  and  $d_{3-i} \neq \infty$ ,
- or  $d_i^T = \min(d_1, d_2, d_1^T, d_2^T)$  for some  $i \in \{1, 2\}$  and  $d_{3-i}^T \neq \infty$ .

In other words, the minimum distance of the quantum code is governed by the minimum distance of the underlying binary codes  $\mathcal{C}_i$  with Tanner graphs  $\mathcal{G}_1$  and  $\mathcal{G}_2$  when  $\mathcal{G}_1^T$  and  $\mathcal{G}_2^T$  are Tanner graphs of trivial codes. This also implies that by choosing the binary linear codes  $\mathcal{C}_i$  with linear minimum distance we will be able to construct families of quantum codes with minimum distance behaving like the square root of the blocklength (and with a dimension which is linear in the blocklength when the rates of  $\mathcal{C}_i$  are chosen appropriately). Moreover, by choosing  $\mathcal{G}_1$  and  $\mathcal{G}_2$  to be sparse it turns out by Proposition 5 that  $\mathcal{G}_1 \times_X \mathcal{G}_2$  and  $\mathcal{G}_1 \times_Z \mathcal{G}_2$  are also sparse. In other words we obtain in this way quantum LDPC codes with a minimum distance which can be of the form  $\Omega(\sqrt{N})$  where  $N$  is the blocklength of the quantum code. We will prove this theorem by first proving that the righthand-side is a lower bound on the minimum distance, then we will prove that this lower bound is attained by exhibiting suitable codewords of the quantum code.

### 5.1 A lower bound on the minimum distance

A common strategy for obtaining a lower bound on the minimum distance of a quantum CSS code is to simply look for lower bounds on the minimum distances of the classical codes  $\mathcal{C}_X$  and  $\mathcal{C}_Z$ . However this approach will necessarily fail here because of the LDPC nature of the CSS code. Recall that since  $\mathcal{C}_Z^\perp \subset \mathcal{C}_X$  and  $\mathcal{C}_X^\perp \subset \mathcal{C}_Z$ , any lower bound on the minimum weight of  $\mathcal{C}_X$  and  $\mathcal{C}_Z$  will not exceed the minimum weight of the parity-check matrices  $\mathbf{H}_X$  and  $\mathbf{H}_Z$  of  $\mathcal{C}_X$  and  $\mathcal{C}_Z$ , which is a constant since  $\mathbf{H}_X$  and  $\mathbf{H}_Z$  are precisely constructed to have small row weights.

To obtain any interesting lower bound on the quantum minimum distance, we must therefore use its full definition, namely that the minimum distance is the smallest weight of a codeword of  $\mathcal{C}_X$  not in  $\mathcal{C}_Z^\perp$  or of a codeword of  $\mathcal{C}_Z$

not in  $\mathcal{C}_X^\perp$ . Our strategy will therefore be to consider a non-zero element of  $\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)$  and prove that when its weight is too small, then it has to belong to  $\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)^\perp$ .

By using the same notation as for Theorem 15 we will prove in this way that

**Lemma 16**

$$d_Q \geq \min(d_1, d_2, d_1^T, d_2^T).$$

*Proof:* We first prove that any element  $\mathbf{x}$  of  $\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)$  which is of weight less than  $\min(d_1, d_2^T)$  belongs to  $\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)^\perp$ . We denote by  $\text{supp}(\mathbf{x})$  the support of  $\mathbf{x}$  which is a subset of  $V_1 \times V_2 \cup C_1 \times C_2$ . Let  $V_1' \stackrel{\text{def}}{=} \{v' \in V_1 : \exists v \in V_2, (v', v) \in \text{supp}(\mathbf{x})\}$  and  $C_2' \stackrel{\text{def}}{=} \{c' \in C_2 : \exists c \in C_1, (c, c') \in \text{supp}(\mathbf{x})\}$ . Let  $\mathcal{G}_1'$  be the subgraph of  $\mathcal{G}_1$  induced by  $V_1' \cup C_1$  and let  $\mathcal{G}_2'$  be the subgraph of  $\mathcal{G}_2$  induced by  $V_2 \cup C_2'$ .

Let  $\mathcal{C}_i$  (respectively  $\mathcal{C}_i'$ ) be the binary code defined by the Tanner graph  $\mathcal{G}_i$  (respectively  $\mathcal{G}_i'$ ), and let  $\mathcal{C}_i'^T$  be the binary code described by  $\mathcal{G}_i'^T$ . Since a codeword of  $\mathcal{C}_1'$  can be viewed as a codeword of  $\mathcal{C}_1$  by extending it with zeros on the positions of  $V_1 \setminus V_1'$  and since  $|V_1'| < d_1$  we necessarily have  $\dim(\mathcal{C}_1') = 0$ . A similar reasoning shows that  $\dim(\mathcal{C}_2'^T) = 0$ . By using Proposition 14 we see that the dimension of  $\mathcal{Q}(\mathcal{G}_1' \times \mathcal{G}_2')$  is equal to zero. This is equivalent to

$$\mathcal{C}_X(\mathcal{G}_1' \times \mathcal{G}_2') = \mathcal{C}_Z(\mathcal{G}_1' \times \mathcal{G}_2')^\perp. \quad (17)$$

Notice now that the restriction  $\mathbf{x}'$  of  $\mathbf{x}$  to the positions in  $V_1' \times V_2 \cup C_1 \times C_2'$  belongs to  $\mathcal{C}_X(\mathcal{G}_1' \times \mathcal{G}_2')$ . Therefore it also belongs to  $\mathcal{C}_Z(\mathcal{G}_1' \times \mathcal{G}_2')^\perp$  and therefore  $\mathbf{x}$  belongs to  $\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)^\perp$ .

We obtain similarly that any element  $\mathbf{x}$  of  $\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)$  which is of weight less than  $\min(d_2, d_1^T)$  will belong to  $\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)^\perp$ . Therefore any element in the union of  $\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2) \setminus \mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)^\perp$  and of  $\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2) \setminus \mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)^\perp$  should have weight at least  $\min(d_1, d_2, d_1^T, d_2^T)$ . ■

## 5.2 An upper bound on the minimum distance

**Lemma 17** *Let  $i$  belong to  $\{1, 2\}$ . Assume that  $d_{3-i} \neq \infty$ . Then  $d_Q \leq d_i$ . If  $d_{3-i}^T \neq \infty$  then  $d_Q \leq d_i^T$ .*

*Proof:* We will prove here that if  $d_2 \neq \infty$  then  $d_Q \leq d_1$ . The other inequalities are proved in a similar fashion.

As before,  $\mathcal{G}_1 = \mathcal{T}(V_1, C_1, E_1)$  and  $\mathcal{G}_2 = \mathcal{T}(V_2, C_2, E_2)$  denote the two Tanner graphs used to define the quantum code  $\mathcal{Q}(\mathcal{G}_1 \times \mathcal{G}_2)$  and  $\mathcal{C}_1$  and  $\mathcal{C}_2$  are the binary codes of Tanner graphs  $\mathcal{G}_1$  and  $\mathcal{G}_2$  respectively.

It will again be convenient to identify codewords with their supports, and we will allow ourselves the use of set operations  $\cap$ ,  $\subset$  and  $\times$  on vectors and likewise we will use vector addition on sets.

Consider a codeword  $\mathbf{x}_1$  of  $\mathcal{C}_1$  of weight  $d_1$ . Since  $d_2 \neq \infty$ ,  $\mathcal{C}_2$  is not reduced to the zero word, and therefore  $\mathcal{C}_2^\perp$  does not contain the whole of  $\{0, 1\}^{|V_2|}$ .



Therefore there exists an element  $y$  of  $V_2$  such that  $\{y\}$  does not belong to  $\mathcal{C}_2^\perp$ . Let  $\mathbf{x} = \mathbf{x}_1 \times \{y\}$ . Our objective is to show that  $\mathbf{x}$  is a weight  $d_1$  codeword of  $\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)$  that is not in  $\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)^\perp$ .

The vector (set)  $\mathbf{x}$  is clearly an element of  $\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)$  since the only check nodes of the Tanner graph  $\mathcal{G}_1 \times_X \mathcal{G}_2$  incident to  $\mathbf{x}$  are nodes of the form  $(c_1, y)$  with  $c_1$  incident to  $\mathbf{x}_1$  in  $\mathcal{G}_1$ .

We now assume that  $\mathbf{x}$  is also an element of  $\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)^\perp$  and work towards a contradiction.

$\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)^\perp$  is generated by the rows  $\mathbf{h}_Z(v_1, c_2)$  of the parity-check matrix  $\mathbf{H}_Z$ , with  $(v_1, c_2)$  ranging over  $V_1 \times C_2$ : viewed as a set,  $\mathbf{h}_Z(v_1, c_2)$  is the neighborhood in  $\mathcal{G}_1 \times \mathcal{G}_2$  of vertex  $(v_1, c_2)$ . Recall from the definition of  $\mathcal{G}_1 \times \mathcal{G}_2$  that  $\mathbf{h}_Z(v_1, c_2)$  is the union of all nodes  $(v_1, v_2)$  such that  $v_2$  is adjacent to  $c_2$  in  $\mathcal{G}_2$  and all nodes  $(c_1, c_2)$  such that  $c_1$  is adjacent to  $v_1$  in  $\mathcal{G}_1$ .

If  $\mathbf{x}$  is also a codeword of  $\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)^\perp$ , then there exists a subset  $U$  of  $V_1 \times C_2$  such that

$$\mathbf{x} = \bigoplus_{(v_1, c_2) \in U} \mathbf{h}_Z(v_1, c_2).$$

Notice now that since  $\mathbf{x} \subset \mathbf{x}_1 \times V_2$ , we can write

$$\mathbf{x} = \bigoplus_{(v_1, c_2) \in U} \mathbf{h}_Z(v_1, c_2) \cap (\mathbf{x}_1 \times V_2) \quad (18)$$

$$= \bigoplus_{\substack{(x, c_2) \\ (x, c_2) \in U, x \in \mathbf{x}_1}} \mathbf{h}_Z(x, c_2) \cap (\mathbf{x}_1 \times V_2), \quad (19)$$

since  $\mathbf{h}_Z(v_1, c_2) \cap (\mathbf{x}_1 \times V_2) = \emptyset$  for  $v_1 \notin \mathbf{x}_1$ . For an element  $x$  in  $\mathbf{x}_1$  we denote by

$$A(x) = \bigoplus_{\substack{c_2 \\ (x, c_2) \in U}} \mathbf{h}_Z(x, c_2) \cap (\mathbf{x}_1 \times V_2)$$

Notice that

$$A(x) \subset \{x\} \times V_2 \quad (20)$$

By combining this remark with  $\mathbf{x}_1 \times \{y\} = \mathbf{x} = \bigoplus_{x \in \mathbf{x}_1} A(x)$ , we obtain that for any  $x$  in  $\mathbf{x}_1$ ,

$$\{(x, y)\} = A(x).$$

On the other hand, we notice that

$$A(x) = \bigoplus_{\substack{c_2 \\ (x, c_2) \in U}} \bigoplus_{v_2 \sim c_2} \{(x, v_2)\}$$

where by  $v_2 \sim c_2$  we mean that  $v_2$  is adjacent to  $c_2$  in  $\mathcal{G}_2$ . This implies

$$\{(x, y)\} = \bigoplus_{\substack{c_2 \\ (x, c_2) \in U}} \bigoplus_{v_2 \sim c_2} \{(x, v_2)\}$$

This in turn implies that

$$\{y\} = \bigoplus_{\substack{c_2 \\ (x, c_2) \in U}} \bigoplus_{v_2 \sim c_2} \{v_2\}$$

which means that  $\{y\}$  is in  $C_2^\perp$ . This contradicts the assumption made on  $y$ . ■

Lemmas 16 and 17 together prove Theorem 15. Theorem 1 is obtained by applying Theorem 15 with  $\mathcal{G}_1 = \mathcal{G}_2$  the Tanner graph of an  $[n, k, d]$  code.

## 6 Comparison with other constructions of quantum codes based on binary linear codes

Our construction can be viewed as a way of producing a quantum code from two classical binary codes  $\mathcal{C}_1$  and  $\mathcal{C}_2$ . The CSS construction achieves the same purpose but requires that  $\mathcal{C}_2^\perp \subset \mathcal{C}_1$ . Using this construction directly to obtain quantum LDPC codes is delicate due to the aforementioned orthogonality constraint. Our construction does not require this constraint and gives a quantum LDPC code when  $\mathcal{C}_1$  and  $\mathcal{C}_2$  are classical LDPC codes. If we denote the length of  $\mathcal{C}_i$  by  $n_i$ , its dimension by  $k_i$ , its co-dimension  $n_i - k_i$  by  $r_i$ , its minimum distance by  $d_i$  and if we choose a full rank parity-check matrix  $\mathbf{H}_i$  for it which describes the Tanner graph  $\mathcal{G}_i$  used in the construction, then a straightforward application of the previous results leads to a quantum code  $\mathcal{Q}(\mathcal{G}_1 \times \mathcal{G}_2)$  with parameters

$$[[n_1 n_2 + r_1 r_2, k_1 k_2, \min(d_1, d_2)]].$$

Notice that  $\mathcal{C}_X(\mathcal{G}_1 \times \mathcal{G}_2)$  has a parity-check matrix with the block form

$$\mathbf{H}_X = (\mathbf{H}_1 \otimes \mathbf{I}_{n_2} \quad \mathbf{I}_{r_1} \otimes \mathbf{H}_2^T)$$

whereas  $\mathcal{C}_Z(\mathcal{G}_1 \times \mathcal{G}_2)$  has a parity-check matrix of the form

$$\mathbf{H}_Z = (\mathbf{I}_{n_1} \otimes \mathbf{H}_2 \quad \mathbf{H}_1^T \otimes \mathbf{I}_{r_2})$$

where  $\mathbf{I}_t$  stands for the  $t \times t$  identity matrix. Notice that under this form, the property ensuring what we indeed define in this way a valid CSS code which is  $\mathbf{H}_X \mathbf{H}_Z^T = 0$  can be verified directly

$$\begin{aligned} \mathbf{H}_X \cdot \mathbf{H}_Z^T &= \mathbf{H}_1 \otimes \mathbf{I}_{n_2} \cdot (\mathbf{I}_{n_1} \otimes \mathbf{H}_2)^T + \mathbf{I}_{r_1} \otimes \mathbf{H}_2^T \cdot (\mathbf{H}_1^T \otimes \mathbf{I}_{r_2})^T \\ &= \mathbf{H}_1 \otimes \mathbf{H}_2^T + \mathbf{H}_1 \otimes \mathbf{H}_2^T \\ &= 0. \end{aligned}$$

This construction displays some similarities with the generalized Shor code construction, see [5], which produces with the help of two binary linear codes  $\mathcal{C}_1$  and  $\mathcal{C}_2$  a quantum code with parameters  $[[n_1 n_2, k_1 k_2, \min(d_1, d_2)]]$ . It is a CSS code like our construction and is associated to a couple  $(\mathcal{C}_X, \mathcal{C}_Z)$  of binary codes satisfying  $\mathcal{C}_X^\perp \subset \mathcal{C}_Z$  defined by the following parity check matrices

$$\mathbf{H}_X = \mathbf{H}_1 \otimes \mathbf{I}_{n_2}, \quad \mathbf{H}_Z = \mathbf{G}_1 \otimes \mathbf{H}_2.$$

where  $\mathbf{G}_1$  is a generator matrix for  $\mathcal{C}_1$ . Notice that our construction has the same dimension and minimum distance as the generalized Shor code construction but has an additional term  $r_1 r_2$  in the length which compares favorably to Shor's construction. Moreover, whereas our construction when applied to sparse matrices  $\mathbf{H}_i$ 's yields sparse matrices  $\mathbf{H}_X$  and  $\mathbf{H}_Z$ , this is not the case in Shor's construction:  $\mathbf{H}_X$  stays sparse, however as soon as the minimum distance of  $\mathcal{C}_1$  is large, this is not the case anymore for  $\mathbf{H}_Z$ . Unlike our construction, the generalized Shor code construction is unable to yield quantum LDPC code families with non constant minimum distance.

## References

- [1] S. A. Aly. A class of quantum LDPC codes derived from Latin squares and combinatorial objects. Technical report, Department of Computer Science, Texas A&M University, April 2007.
- [2] S. A. Aly. A class of quantum LDPC codes constructed from finite geometries. In *Proceedings of IEEE GLOBECOM 2008*, pages 1–5, December 2008.
- [3] I. Andriyanova, D. Maurice, and J. P. Tillich. Quantum LDPC codes obtained by non-binary constructions. In *Proc. IEEE Int. Symp. Info. Theo.*, pages 343–347, 2012.
- [4] I. Andriyanova, D. Maurice, and J. P. Tillich. Spatially coupled quantum LDPC codes. In *Proc. of Inf. Theor. Workshop ITW2012*, pages 327–331, Lausanne, Switzerland, 2012.
- [5] D. Bacon and A. Casaccino. Quantum error correcting subsystem codes from two classical linear codes. In *Proceedings of the 44th Allerton conference on Communication, Control and Computing*. Curran Associates, Inc., 2006. <http://arxiv.org/abs/quant-ph/0610088>
- [6] H. Bombin and M. A. Martin-Delgado. Topological quantum distillation. *Phys. Rev. Lett.*, 97(180501), 2006.
- [7] H. Bombin and M. A. Martin-Delgado. Homological error correction: classical and quantum codes. *J. Math. Phys.*, 48:052105–1–052105–35, 2007.
- [8] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane. Quantum error correction via codes over  $\text{GF}(4)$ . *IEEE Trans. Info. Theor.*, 44(4), 1369–1387, 1998.
- [9] A. R. Calderbank and P. W. Shor. Good quantum error-correcting codes exist. *Phys. Rev. A*, 54:1098–1105, 1996.
- [10] T. Camara, H. Ollivier, and J.-P. Tillich. Constructions and performance of classes of quantum LDPC codes, 2005. <http://arxiv.org/abs/quant-ph/0502086>

- [11] T. Camara, H. Ollivier, and J.-P. Tillich. A class of quantum LDPC codes: construction and performances under iterative decoding. In *Proc. IEEE Int. Symp. Info. Theo.*, pages 811–815, Nice, June 2007.
- [12] A. Couvreur, N. Delfosse, and G. Zémor. A construction of quantum LDPC codes from Cayley graphs. In *Proc. IEEE Int. Symp. Info. Theo.*, pages 643–647, St. Petersburg, August 2011. Long version available at <http://arxiv.org/abs/1206.2656>.
- [13] C. D. de Albuquerque, R. Palazzo, and E. B. da Silva. Construction of topological quantum codes on compact surfaces. In *Proceedings of ITW 2008*, pages 391–395, Porto, May 2008.
- [14] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin. Quantum-channel capacity of very noisy channels. *Phys. Rev. A*, 57:830–839, 1998.
- [15] I. B. Djordjevic. Quantum LDPC codes from incomplete block designs. *IEEE Communication Letters*, 12(5):389–391, May 2008.
- [16] J. Fern and K. B. Whaley. Lower bounds on the nonzero capacity of Pauli channels. *Phys. Rev. A*, 78:062335, December 2008.
- [17] M. H. Freedman, D. A. Meyer, and F. Luo.  $\mathbb{Z}_2$ -systolic freedom and quantum codes. In *Mathematics of quantum computation*, Chapman & Hall/CRC, pages 287–320, Boca Raton, FL, 2002.
- [18] R. G. Gallager. *Low Density Parity Check Codes*. M.I.T. Press, Cambridge, Massachusetts, 1963.
- [19] J. Garcia-Frias and K. Liu. Design of near-optimum quantum error-correcting codes based on generator and parity-check matrices of LDGM codes. In *Proceedings of CISS*, pages 562–567, Princeton, March 2008.
- [20] D. Gottesman. *Stabilizer codes and quantum error correction*. PhD thesis, California Institute of Technology, Pasadena, CA, 1997.
- [21] M. Hagiwara and H. Imai. Quantum quasi-cyclic LDPC codes. In *Proc. IEEE Int. Symp. Info. Theo. (ISIT'07)*, pages 806–811, Nice, June 2007.
- [22] S. L. Hakimi and J. G. Bredeson. Graph theoretic error-correcting codes. *IEEE Trans. Info. Theor.*, 14:584–591, 1968.
- [23] M.-H. Hsieh, T. A. Brun, and I. Devetak. Quantum quasi-cyclic low-density parity check codes, March 2008. <http://arxiv.org/abs/0803.0100> [quant-ph].
- [24] L. Ioffe and M. Mézard. Asymmetric quantum error-correcting codes. *Phys. Rev. Lett. A*, 75, 032345, 2007.
- [25] K. Kasai, M. Hagiwara, H. Imai, and Sakaniwa K. Quantum error correction beyond the bounded distance decoding limit. *IEEE Trans. Info. Theor.*, vol. 58, 1223–1230, 2012.

- [26] I. H. Kim. Quantum codes on Hurwitz surfaces. Master's thesis, MIT, 2007. available at <http://dspace.mit.edu/handle/1721.1/40917>.
- [27] A. Y. Kitaev. Fault-tolerant quantum computation by anyons. *Ann. Phys.*, 303:2, 2003.
- [28] A. Kovalev and L. Pryadko. Fault-tolerance of "bad" quantum low-density parity check codes, August 2012. <http://arxiv.org/abs/1208.2317>
- [29] A. Kovalev and L. Pryadko. Improved quantum hypergraph-product LDPC codes. In *Proc. IEEE Int. Symp. Info. Theo.*, pages 348–352, Boston, July 2012.
- [30] H. Lou and J. Garcia-Frias. On the application of error-correcting codes with low-density generator matrix over different quantum channels. In *Proceedings of Turbo-coding 2006*, Munich, April 2006.
- [31] D. J. C. MacKay, G. Mitchison, and P. L. MacFadden. Sparse graph codes for quantum error-correction. *IEEE Trans. Info. Theor.*, 50(10):2315–2330, 2004.
- [32] F.J. MacWilliams and Sloane N.J.A. *The theory of error-correcting codes*. North-Holland, 1986.
- [33] M. S. Postol. A proposed quantum low density parity check code, 2001. available at <http://arxiv.org/abs/quant-ph/0108131>.
- [34] P. Sarvepalli. Topological color codes over higher alphabets. In *Proceedings of the IEEE Information Theory Workshop*, Dublin, Ireland, september 2010.
- [35] K. P. Sarvepalli, M. Rötteler, and A. Klappenecker. Asymmetric quantum LDPC codes. In IEEE, editor, *Proc. IEEE Int. Symp. Info. Theo.*, pages 305–309, Toronto, Canada, July 2008.
- [36] G. Smith and J. A. Smolin. Degenerate coding for Pauli channels. *Phys. Rev. Lett.*, 98, 2007.
- [37] A. M. Steane. Multiple particle interference and quantum error correction. *Proc. R. Soc. Lond. A*, 452:2551–2577, 1996.
- [38] P. Tan and J. Li. Efficient quantum stabilizer codes: LDPC and LDPC-convolutional constructions. *IEEE Trans. Info. Theor.*, 56(1):476–491, 2010.
- [39] J.-P. Tillich and G. Zémor. Quantum LDPC codes with positive rate and minimum distance proportional to  $n^{\frac{1}{2}}$ . In *Proc. IEEE Int. Symp. Info. Theo.*, pages 799–803, July 2009.
- [40] G. Zémor. On Cayley graphs, surface codes and the limits of homological coding for quantum error correction. In *Coding and Cryptology, 2nd international Workshop IWCC 2009*, volume 5557 of *LNCS*, pages 259–273, 2009.