Trellises for stabilizer codes: definition and uses

Harold Ollivier\(^1\) and Jean-Pierre Tillich\(^2\)

\(^1\)Perimeter Institute, 31 Caroline St. N, Waterloo, ON N2L 2Y5, Canada.
\(^2\)INRIA, Projet Codes, Domaine de Voluceau BP 105, F-78153 Le Chesnay cedex, France.

(Dated:)

Trellises play an important theoretical and practical role for classical codes. Their main utility is to devise complexity-efficient error estimation algorithms. Here, we describe trellis representations for quantum stabilizer codes. We show that they share the same properties as their classical analogs. In particular, for any stabilizer code it is possible to find a minimal trellis representation. Our construction is illustrated by two fundamental error estimation algorithms.

Introduction. Since the discovery of efficient quantum algorithms for solving hard classical problems, many efforts have been devoted to building quantum processing devices. While small scale prototypes are readily available, scalability remains a practical issue because of the extreme sensibility of quantum devices to external noise. Fortunately, theoretical advances, such as the discovery of error correction schemes and fault-tolerant implementations, have notably cleared the future of quantum computation. One way of building quantum codes is through quantum algorithms for solving hard classical problems, many efforts have been devoted to building quantum processing devices. While small scale prototypes are readily available, scalability remains a practical issue because of the extreme sensibility of quantum devices to external noise. Fortunately, theoretical advances, such as the discovery of error correction schemes and fault-tolerant implementations, have notably cleared the future of quantum computation.

In classical coding theory, one often relies on a graph-theoretical representation of the code, called a trellis, to perform error estimation. For instance, trellises yield many complexity-efficient error estimation schemes for memoryless channels as well as means of estimating the noise parameters. In particular, it can be used to calculate with linear complexity the most likely error for a convolutional code of bounded memory over any memoryless channel. In this article, we apply general results concerning group codes for classical communication to show that a similar representation is available for quantum stabilizer codes. Two error estimation schemes that exploit trellises are introduced for memoryless channels. We show that the complexity of these algorithms is related to the number of trellis vertices, and provide a construction of a trellis which minimizes this quantity. One of these algorithms achieves the performance of an algorithm for convolutional codes that was proposed in [4], but with a significantly lower complexity.

Stabilizer codes: elementary facts. In the rest of this paper, some familiarity with quantum computation is assumed. This section provides a brief introduction to stabilizer codes, for a more detailed introduction the reader is redirected to [1, 2, 5] and references therein.

In what follows, without loss of generality, the quantum register of interest has \(n\) physical qubits.

Preliminaries. Stabilizer codes rely heavily on properties of \(G_n\), the \(n\)-qubit Pauli group. This group is defined in terms of the Pauli matrices for a single qubit: \(I = (0 0 \ 0 0), X = (0 1 \ 1 0), Y = (0 i 0 i)\), and \(Z = (0 1 0 0)\). The group \(G_n\) is the multiplicative group generated by the \(n\)-fold tensor products of single qubit Pauli matrices.

For our purpose here, phases are irrelevant and it will be more convenient to work with the effective Pauli group \(G_n \equiv G_n/\{ \pm I^\otimes n, \pm i Z^\otimes n \}\) (see [2]). The elements of \(G_1\) will be denoted by \(I \equiv [I], X \equiv [X], Y \equiv [Y],\) and \(Z \equiv [Z]\). Here, \([\mathcal{P}]\) denotes the equivalence class of \(\mathcal{P} \in G_n\), that is \(\{ \pm \mathcal{P}, \pm i \mathcal{P}\}\). Note that \(G_n\) is Abelian, so that we will use the additive notation for its group operation. Since \(G_n \cong G_n\), we often view \(P \in G_n\) as an \(n\)-tuple \((P_i)_{i=1}^n\) with entries in \(G_1\).

The crucial fact about \(G_n\), is that any pair of elements \(\mathcal{P}, \mathcal{Q}\) either commutes or anti-commutes. This leads to the definition of an inner product “\(\ast\)” for elements of \(G_n\) such that \((P_i) \ast (Q_i) = \sum_i P_i \ast Q_i \mod 2\). Here, \(P_i \ast Q_i = 1\) if \(P_i \neq Q_i\), \(P_i \neq I\) and \(Q_i \neq I;\) and \(P_i \ast Q_i = 0\) otherwise. One can then check easily that \(\mathcal{P}, \mathcal{Q} \in G_n\) commute if and only if \([\mathcal{P}] \ast [\mathcal{Q}] = 0\).

Error model. Stabilizer codes can accommodate for a broad class of channels. For simplicity, only memoryless Pauli channels will be considered, although the tools presented in this paper extend to other memoryless channels. Memoryless Pauli channels act on the whole \(n\)-qubit register as \(\sigma \rightarrow \Psi(\sigma) = (\Psi^1 \otimes \Psi^2 \otimes \ldots \otimes \Psi^n)(\sigma)\). Above \(\Psi^i\) is a 1-qubit channel whose action on \(\rho\), a single qubit density operator, is given by \(\Psi^i(\rho) = \sum_{E \in \{Z, X, Y, Z\}} \text{Pr}_i([E]) E \rho E\), with \(\text{Pr}_i(\cdot)\) a probability distribution over \(G_1\).
Definition of the code subspace. The code subspace \( C \) of an \((n, k)\) stabilizer code is the largest subspace stabilized by the action of \( \mathcal{S} \), an Abelian subgroup of \( \mathbb{G}_n \). For the code to protect \( k \) qubits using \( n \), i.e., to be of rate \( k/n \), \( \mathcal{S} \) must be generated by \( n-k \) independent operators \( S_j \), and be such that \(-I^\otimes n \notin \mathcal{S} \). The code subspace is equivalently defined by \( n-k \) eigenvalue equations: \( \ket{\psi} \in C \) if and only if \( \forall j \in \{1, \ldots, n-k\}, S_j \ket{\psi} = \ket{\psi} \).

To study the main properties of these codes, phases are again irrelevant. More precisely, it is sufficient to represent the set of generators of the stabilizer group \( \{S_j\}_j \) by the set of equivalence classes \( \{S_j\} \), where \( S_j = [S_j] \) which generate a subgroup \( \mathcal{S} \) of \( \mathcal{G}_n \). Using a slight abuse in terminology, we also call \( \mathcal{S} \) the stabilizer group of the code and \( \{S_j\}_j \) the stabilizer set of the code.

Quantum convolutional codes. Following the definition of \( \mathcal{G}_n \), an \((n, k)\) stabilizer code is convolutional with parameters \((\eta, \kappa)\) if there exists a set of generators \( \{S_j\}_j \) of its stabilizer group \( \mathcal{S} \) with an \( \eta \)-qubit shift invariance property. More precisely, the values \( S_j^i \) must be equal to the entries \( H^i_j \in G_1 \) of an infinite matrix \( H \) which satisfies \( H^{i+n-k} = H^i \) for every \( i, j \).

For a convolutional code given by the stabilizer set \( \{S_j\}_j \), the memory is defined as \( m = \max_i \sum_j \sharp(S_j^i) \).

Here, \( \sharp(S_j^i) = 0 \) if either \( S_j^i = I \) for all \( k > i \) or \( S_j^k = I \) for all \( k \leq i \); and \( \sharp(S_j^i) = 1 \) otherwise. It is important to note that the memory depends on the set \( \{S_j\}_j \) and is thus not an intrinsic property of the code.

Error estimation. The goal of error estimation is to infer channel errors from their action on the state of the quantum register. In the context of stabilizer codes, the necessary information is provided by the measurement of the Hermitian operators \( S_j \).

Let \( \mathcal{E} \in \mathbb{G}_n \) be the actual, yet unknown, quantum error that affected the state \( \ket{\psi} \in C \). The measurement of the operators \( S_j \) on \( \ket{\psi'} = \mathcal{E} \ket{\psi} \) defines a binary vector of length \( n-k \) called the syndrome of \( \mathcal{E} \): \( s(\mathcal{E}) \triangleq \left(s^j(\mathcal{E})\right)_{j=1}^{n-k} \triangleq \frac{1}{2} (1 - \bra{\psi|S_j\ket{\psi'}}) = (S_j \mathcal{E})_{j=1}^{n-k} \).

Among all possible error operators \( \mathcal{F} \in \mathbb{G}_n \), only some of them are compatible with \( s(\mathcal{E}) \) (i.e., they satisfy \( s(\mathcal{F}) = s(\mathcal{E}) \)). Therefore, they belong to a coset of \( N(\mathcal{S}) \), the normalizer of \( \mathcal{S} \) in \( \mathbb{G}_n \). Equivalently, these compatible errors \( \mathcal{F} \) satisfy that \( [\mathcal{F}] \) belongs to a coset of \( S^\perp \) containing \( \{P \in \mathcal{G}_n : P \otimes Q = 0, \forall Q \in \mathcal{S} \} \).

The knowledge of the syndrome restricts the class of errors that could have happened during the transmission, the error model further discriminates between these elements by assigning them probabilities. Error recovery then uses these probabilities to find a best guess \( \mathcal{E} \) for the actual error \( \mathcal{E} \). For instance, maximum likelihood error estimation consists in finding a most likely \( \mathcal{E} \) compatible with the measured syndrome \( s(\mathcal{E}) \). Trellises are both aimed at computing these probabilities and at choosing a best guess efficiently.

Trellises for stabilizer codes. Considering the obvious remarks about error estimation, it is natural to seek a representation of cosets of \( S^\perp \) in which it is easy to search for their most likely element. This is the main motivation for the definition of trellises for quantum as well as for classical codes. In a broader context, this motivation extends to the calculation of quantities on which error estimation is based, e.g. the likelihood function, the a posteriori qubit error probability, etc.

Definition. In the rest of this section, \( \mathcal{S} \) denotes the stabilizer group of a quantum code with parameters \((n, k)\), and \( \{S_j\}_{j=1}^{n-k} \) its stabilizer set.

An \( n \)-section trellis relative to the stabilizer set \( \{S_j\}_{j=1}^{n-k} \) and syndrome \( s \in \mathbb{F}_2^{n-k} \) is a directed graph with the following properties:

1. Its vertices can be grouped into \( n+1 \) sets \( V_i \), with \( |V_0| = |V_n| = 1 \). The set \( V_i \) is called the \( i \)-th state space of the trellis.

2. Its edges are directed and can be grouped into \( n \) sets \( E_i \). An edge \( e \in E_i \) is said to be issued at vertex \( v \) and ending at vertex \( w \). Edges in \( E_i \) are issued from a vertex of \( V_{i-1} \) and end at a vertex of \( V_i \). The set \( E_i \) is called the \( i \)-th section of the trellis.

3. An edge \( e \in E_i \) bears a label \( l(e) \) such that \( l(e) \in G_1 \). We say that \( l(e) \) is the Pauli-label of \( e \).

4. Each element \( P \in \mathcal{G}_n \) with syndrome \( s \) is associated to a unique directed path \( (e^1, \ldots, e^n) \) such that \( l(e^i) = P^i \).

An example of trellis is considered in Figure 4 for the 4-qubit code with stabilizer group generated by \( \{XXXX, ZZZZ\} \).

Construction. Given a code stabilized by \( \mathcal{S} \), there are many possible trellises representing the coset of \( S^\perp \) of syndrome \( s \). We provide here a simple construction for which the number of vertices in each \( V_i \) is bounded by \( 2^{n-k} \). In analogy with classical codes, this trellis will be called the Wolff trellis of the code defined by the stabilizer set \( \{S_j\}_j \) relative to the syndrome \( s \).

For every \( i \in \{0, \ldots, n\} \), let \( \pi_i \) be a mapping from \( \mathcal{G}_n \) to itself defined by \( \pi_i(P^1, P^2, \ldots, P^n) = (P^1, \ldots, P^i, I, \ldots, I) \) (with the convention that \( \pi_0(P^1, P^2, \ldots, P^n) = (I, \ldots, I) \)). Let \( P_s \) be an arbitrary, but fixed, element of \( \mathcal{G}_n \) with syndrome \( s \). For every \( i \in \{0, \ldots, n\} \), \( V_i \) is a subset of \( \mathbb{F}_2^{n-k} \) defined by \( \{(S_j \ast \pi_i(P))_{j=1}^{n-k} : P \in P_s + S^\perp \} \). A vertex \( v \in V_i \) is con-
nected to vertex \( w \in V_{i+1} \) with an edge labelled by \( E \) iff there exists \( P \in P_r + S^\perp \) such that \( v = (S_j \ast \pi_t(P))_{j=1}^{n-k}, w = (S_j \ast \pi_{i+1}(P))_{j=1}^{n-k} \) and \( P^{i+1} = E \).

An example of trellis obtained in this way is given in Figure 2 for the 5-qubit code associated to the stabilizer set \( \{ZXII, XZZXI, IXZXI, 1IXXX \} \).

As hinted by this example, the Wolff trellis of convolutional codes has state-spaces that are bounded in size by \( 2^m \), where \( m \) is the memory associated to the chosen stabilizer set \( \{S_j \} \). Note that this bound is achieved for all indices \( i \) such that \( m = \sum_j z(S_j^u) \).

Minimality. As it will be seen below, the complexity of many useful algorithms using trellises is linear in their number of vertices. This raises the issue of finding a trellis which minimizes this quantity. If one is willing to change the stabilizer set for the code in order to put it in a trellis oriented form, then the Wolff trellis is minimal. Indeed, the trellis obtained in this way does not only minimize the number of vertices but also the state space profile. As for classical trellises, we define the state profile of the trellis of an \((n,k)\)-tuple \((\xi_0, \xi_1, \ldots, \xi_n)\) where \( \xi_i = \log_2 |V_i| \) where \( V_i \) is the \( i \)-th state space of the trellis. In other words, we are going to prove that the Wolff trellis applied to a stabilizer set in trellises oriented form minimizes each \( \xi_i \) individually. Without loss of generality, we now assume that the trellis is associated to the syndrome \( s = (0, \ldots, 0) \).

First, we define the trellis oriented form of a stabilizer set. For \( 1 \leq j \leq n-k \), let \( c(j) \) and \( d(j) \) be respectively the position of the first (resp. last) component of \( S_j \) which is different from \( I \). We say that the stabilizer set \( \{S_j \} \) is in trellis oriented form if and only if for all \( j \) (1) \( S_j^{c(j)} I \) for all \( j > 1 \) and \( S_j^{c(j)} \neq S_j^{c(j)} \), and (2) there is at most one \( j' \neq j \) such that \( b(j) = b(j') \) and in such case \( S_j^{c(j)} = S_j^{c(j')} \). Note that any stabilizer set \( \{S_j \} \) can be put in trellis oriented form by indices permutations and group additions.

Second, we show a lower bound on \( \xi_i \). For an \((n,k)\) stabilizer code given by the stabilizer set \( \{S_j \} \), and for \( i \in \{0, \ldots, n\} \), let \( C_i^f \) (the future subgroup) be the subgroup of \( S^\perp \) whose elements have their first \( i \) components equal to \( I \); and let \( C_i^p \) (the past subgroup) be the subgroup of \( S^\perp \) whose elements have their last \( n-i \) components equal to \( I \). We then have the following lemma.

**Lemma 1.** \( \xi_i \geq n + k - \log_2 |C_i^p| - \log_2 |C_i^f| \).

**Proof.** Let \( v \in V_i \) and \( C_v \) be the set of elements of \( S^\perp \) that correspond to a path in the trellis associated to the all-zero syndrome and passing through \( v \). Let \( P \) and \( F \) be the set of all paths that go from \( v_0 \) to \( v \) (resp. from \( v \) to \( v_n \)). Let \( Q \) be a fixed element of \( C_v \) and \( p_{qe} \) its corresponding path in the trellis, where \( q_e \in P \) and \( q_F \in F \). By construction of the trellis, we have \( |C_v| = |P||F| \). Note that any path \( p \in P \) can be extended into a path \( p_{qe} \) from \( v_0 \) to \( v_n \). Since any path \( p_{qe} \) can be associated to an element of \( Q + C_i^p \), we have \( |P| \leq |Q + C_i^p| = |C_i^p| \). Similarly \( |P| \leq |C_i| \), which yields \( 2^n + k = |S^\perp| = \sum_{v 
 \in V} |C_v| \leq |V_i||C_i^p||C_i^f| \). This gives the desired bound on \( \xi_i \) by taking the logarithm.

Finally we conclude with the following theorem.

**Theorem 1.** The Wolff trellis achieves the previous bound on \( \xi_i \) for each \( i \) when the stabilizer set is in trellis oriented form.

**Proof.** Let \( S_{\text{start} \leq i} \) be the subset of operators of \( \{S_j \} \) which have at least one of their \( i \)-first components different from \( I \). Let \( S_{\text{end} \geq i} \) be the subset of operators of \( \{S_j \} \) which have all their last \( n-i \) components equal to \( I \). It is straightforward to check that \( \xi_i = |S_{\text{start} \leq i}| - |S_{\text{end} \geq i}| \). Note that, \( C_i^p \) is the subgroup of \( G_n \) with \( I \) on their \( n-i \) last components and orthogonal to all elements of \( S_{\text{start} \leq i} \). This implies \( \log_2 |C_i^p| = 2r - |S_{\text{start} \leq i}| \). Using a similar argument, we get that \( \log_2 |C_i^f| = 2(n-i) - (n-k) - |S_{\text{end} \geq i}| \). Adding these equalities, we conclude that \( \xi_i + \log_2 |C_i^p| + \log_2 |C_i^f| = n + k \).

**Using trellises of stabilizer codes.**

**Min-Sum (Viterbi) algorithm.** The Min-Sum algorithm is certainly one of the most widely employed algorithm that benefits from the trellis representation of classical codes. Here, we present a Min-Sum algorithm for stabilizer codes that computes the most likely error for memoryless Pauli channels given a measured syndrome \( s \) by using a trellis associated to the code.

Consider an \((n,k)\) stabilizer code with stabilizer set \( \{S_j \} \). Define the likelihood of \( P \in G_n \) as \( \sum_{i=1}^{n} \log Pr_t(P^i) \). Consider the \( n \)-section trellis for this quantum code associated to the syndrome \( s \). The naming conventions for the vertices and edges are set as in previous sections. For each edge \( e \) of \( E_v \), define its weight \( wt(e^s) = -\log Pr_t((e^s)) \). By construction, the sum of weights along the path in the trellis that represents \( P \) is equal to the opposite of the likelihood. The task which consists in finding a most likely error \( \hat{E} \in G_n \) with syndrome \( s \) is thus equivalent to finding a lowest weight path \( (e^1, \ldots, e^n) \) in the trellis associated to \( s \).

This can be done by constructing recursively some sets \( C_i \) of lowest weight error candidates. More precisely, \( C_i \) contains couples \((c, w)\) where \( c \) is a path issued from \( v_0 \) that ends on a vertex of \( V_i \) and where \( w = wt(c) \).

| Min-Sum Algorithm |
|-------------------|
| **Initialization:** |
| \( C_0 := \{(v_0, 0)\} \) and \( C_i := \emptyset \) for \( i \leq 1 \) |
| **Main step:** |
| for \( i \) from 1 to \( n \) do |
| for all \( v \in V_i \) do |
| Put in \( C_i \) the pair \((c', wt(c'))\), where \( c' \) is a path of minimum weight (ties are broken at random) among all paths that: (1) end in \( v \); (2) have their \( n-i \) first vertices given by a path \( c \) of \( C_{i-1} \). |
Note that the time complexity of this algorithm is linear in the number of vertices in the trellis.

Sum-Product algorithm. While the Min-Sum finds a most likely error compatible with the observed syndrome $s$, the Sum-Product aims at calculating marginal error probabilities for physical qubits. That is, $p_i(P) \triangleq \Pr(\text{error at qubit } i = P|s)$ where $P \in G_1$. By definition of the trellis, this probability is equal to the probability that a path $(v^i)_{i=1}^n$ from $v_0$ to $v_n$ is such that $l(v^i) = P$.

The Sum-Product algorithm computes for each vertex $v$ of the trellis associated with $s$ a “forward” probability $f(v)$ and a “backward” probability $b(v)$. Both are then used to calculate the marginal probabilities $p_i(P)$.

### Computation of $A(x, y, z)$

**Initialization:** $A_{v_0}(x, y, z) = 1$

**Main step:**

for $i$ from 1 to $n$ do

for all $v \in V_i$ do

$A_v(x, y, z) := \sum_{w \in V_{i-1}^+(v)} A_w(x, y, z)Q_l(wv)(x, y, z)$

$A(x, y, z) := A_{v_n}(x, y, z)$

Above, $V_i^+(v)$ is the set of vertices $w$ in $V_{i-1}$ which are adjacent to $v$; $(2)$ $V_i^+(v)$ is the set of vertices $w$ in $V_{i+1}$ which are adjacent to $v$; and $(3)$ $E_i(P)$ is the set of edges between $V_{i-1}$ and $V_i$ that bear the Pauli-label $P$.

Once again, the practical relevance of this algorithm is due to the fact that its complexity is linear in the number of vertices in the trellis.

**Conclusion.** From a practical point of view, we have proposed a definition for the trellis of a stabilizer code together with two constructions and three algorithms that take advantage of this representation. Following the same path, several algorithms for classical codes running on trellises can be generalized to the quantum case. Among them, estimation of noise parameters (or equalization) seems a promising avenue for enhancing the performance of quantum optics fiber communications using near-future quantum technology.
[1] D. Gottesman, Ph.D. thesis, California Institute of Technology, Pasadena, CA (1997), quant-ph/9705052.
[2] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, IEEE Trans. Inf. Theory 44, 1369 (1998), quant-ph/9608006.
[3] G. D. Forney Jr. and M. D. Trott, IEEE Trans. Info. Theor. 39, 1491 (1993).
[4] H. Ollivier and J.-P. Tillich, Phys. Rev. Lett. 91, 177902 (2003), quant-ph/0304189.
[5] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, Cambridge, UK, 2000).
[6] H. Ollivier and J.-P. Tillich, Quantum convolutional codes: fundamentals (2004), quant-ph/0401134.
[7] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, Phys. Rev. A 57, 830 (1998), quant-ph/9706061.
[8] F.-R. Kschischang, B. J. Frey, and H.-A. Loeliger, IEEE Transactions on Information Theory 47, 498 (2001).