On the Complexity of finding stopping set size in Tanner Graphs.

K. Murali Krishnan , Priti Shankar
Dept. of Computer Science and Automation
Indian Institute of Science, Bangalore - 560012, India
Email: {kmurali,priti}@csa.iisc.ernet.in

***Abstract*— The problem of determining whether a Tanner graph for a linear block code has a stopping set of a given size is shown to be NP-complete.**

I. Introduction

Stopping sets were introduced in [1] for the analysis of erasure decoding of LDPC codes, where, it was shown that the iterative decoder fails to decode to a codeword if and only if the set of error positions is a superset of some stopping set in the Tanner graph [6] for the code used in decoding. Since small stopping sets are directly responsible for poor performance of iterative algorithms, constructions of codes whose Tanner graphs do not contain small stopping sets have been studied as in [2], [3]. In particular the size of the smallest stopping set in a Tanner graph, called the *stopping distance* of the graph is of interest as it determines the fraction of erasures correctable by iterative decoding in the worst case on the graph. The relationship between stopping distance and other graph parameters like girth has been explored in [4] where it is shown that large girth implies high stopping distance. Schwartz and Vardy [5] conducts a deeper investigation into the structure of stopping sets on Tanner graphs and shows that addition of a sufficient number of parity checks can increase the stopping distance of a given Tanner graph to reach the limit possible, viz., the minimum distance of the code.

In this note, we show that the computational problem of determining whether a given Tanner graph has a stopping set of a given size is NP-complete. This is shown by reducing the well known NP-complete problem of determining whether a given graph contains a vertex cover of a given size to the above problem.

II. Background

Given parity check matrix $H = [h_{ij}] \in F_2^{(n-k)\times n}, 1 \leq k \leq n$ for an (n,k) binary linear code, the Tanner graph is defined by the bipartite graph $G = (L, R, E)$ where $L = \{x_i, 1 \leq i \leq n\}$, $R = \{c_j, 1 \leq j \leq n-k\}$ and $E = \{(x_i, c_j) : h_{ji} = 1\} \subseteq L \times R$. We call the set L and R as the set of left and right vertices and E the set of edges of G. For $S \subseteq L \cup R$, we define $\aleph(S) = \{y : (x,y) \in E, x \in S\}$. $S \subseteq L$ is a *stopping set* if for all $c_j \in \aleph(S)$, $|\aleph(\{c_j\}) \cap S| \geq 2$ ie., every vertex connected to a stopping set must have at least two edges going into the stopping set. The *stopping distance* of a Tanner graph is the size of the smallest stopping set in the graph. We define the decision problem STOPPING SET as the membership problem for the set $\{\ (G,t) : G = (L,R,E)$ is a Tanner graph with a stopping set of size $t\ \}$. It is easy to see that if there is an efficient algorithm for STOPPING SET, then by running the algorithm $|L|$ iterations at most, (or still faster using binary search) we can find the stopping distance of the graph.

The notion of NP-Completeness was introduced in [9] and is well established in Computer Science literature for the analysis of the computational complexity of problems (see [7], [8] for definition and a detailed account). A decision problem A belongs to the class NP if there exists a polynomial time algorithm Π such that for all $x \in A$, there exists, a string y (called a *certificate* for membership of x in A), with $|y|$ polynomially bounded in $|x|$ such that Π accepts (x,y), whereas, for all $x \notin A$, Π rejects (x,y) for any string y presented to the algorithm. In other words, problems in NP are precisely those for which membership verification is polynomially solvable. We say a decision problem A is *polynomial time many-one reducible* to a decision problem B if there exists a polynomial time algorithm Π' such that given an instance x of A, Π' produces an instance y of B satisfying $y \in B$ if and only if $x \in A$. In such case, we write $A \preceq_p B$. A problem $A \in$NP is NP-Complete if for every $X \in$NP, $X \preceq_p A$. It is generally believed that NP complete problems have no polynomial time algorithms.

Given a graph (not necessarily bipartite) $G = (V,E)$, $S \subseteq V$ is a *vertex cover* of G if for all $(u,v) \in E$ either $u \in S$ or $v \in S$ or both. The decision problem VERTEX COVER is defined as the membership problem for the set $\{(G,t) : G$ is a graph with a vertex cover of size $t\}$. It is well known that VERTEX COVER is NP-Complete [7]. In the following, we show that VERTEX COVER $\preceq_p$ STOPPING SET.

III. Hardness of STOPPING SET

The reduction from VERTEX COVER to STOPPING set is achieved by the graph construction that follows. All numerical quantities arising in this sequel are assumed to be non-negative integers unless stated otherwise.

Let $(G = (V,E), t)$ be an instance of the VERTEX COVER problem. Let $|V| = n$, $|E| = m$. Excluding trivial cases, we may assume $1 \leq t \leq n-1$. We construct a bipartite graph $G' = (L, R, E')$ as follows: $L = \bigcup_{i=0}^{n} L_i$, $R = \bigcup_{i=0}^{n-1} R_i$, where, $L_0 = V$, L_i=$\{e_{i,\ell} : e \in E\}$, $1 \leq i \leq n$ and R_i=$\{e_{i,r} : e \in E\}$, $0 \leq i \leq n-1$. E' is defined by the following rules:

1-4244-0350-2/06/$20.00 ©2006 IEEE

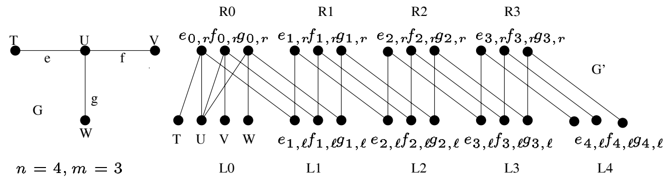

Fig. 1. Construction of G'

- For each $e = (u, v) \in E$ add the edges $(u, e_{0,r})$ and $(v, e_{0,r})$ between L_0 and R_0.
- For each $e \in E$, $0 \leq i \leq n-1$ add an edge between $e_{i,r} \in R_i$ and $e_{i+1,\ell} \in L_{i+1}$.
- For each $e \in E$, $1 \leq i \leq n-1$, connect $e_{i,r} \in R_i$ to $e_{i,\ell} \in L_i$.

The example in figure 1 illustrates the construction.

Lemma 1: If G has a vertex cover of size t, G' has a stopping set of size $t + nm$.

Proof: Let $X \subseteq V$ be a vertex cover in G with $|X| = t$. Consider the set $Y = \bigcup_{i=1}^{n} L_i \cup (L_0 \cap X)$. Clearly $Y \subseteq L$, $|Y| = t+nm$. Since $R_{i-1} \subseteq \aleph(L_i)$ for $1 \leq i \leq n$, $\aleph(Y) = R$. Hence it is sufficient to show that for all $e_{i,r} \in R$, $0 \leq i \leq n-1$ $|\aleph(\{e_{i,r}\}) \cap Y| \geq 2$. If $0 < i < n-1$, by construction of G', both $e_{i,\ell} \in L_i$ and $e_{i+1,\ell} \in L_{i+1}$ belong to $\aleph(\{e_{i,r}\})$ satisfying the required condition. When $i = 0$, consider an arbitrary $e_{0,r} \in R_0$. Since $e_{1,\ell} \in \aleph(\{e_{0,r}\})$, showing existence of one more neighbour in Y suffices. By construction of G', $e \in E$ in G. Let $e = (u, v)$. Since X is a vertex cover, either $u \in X$ or $v \in X$. or both. Hence, either $(u, e_{0,r})$ or $(v, e_{0,r})$ or both must be present in E'. ■

We now show that G has a vertex cover of size t only if G' has a stopping set of size $t+nm$. We begin with the following technical lemma.

Lemma 2: If $Y \subseteq L$ is a non-trivial stopping set in G', then $|Y| = p + qn$ for some $1 \leq p \leq n$, $0 \leq q \leq m$.

Proof: Suppose $e_{i,\ell} \in Y$ for some $1 \leq i < n$, then since $e_{i,r} \in \aleph(\{e_{i,\ell}\})$ and $\aleph(\{e_{i,r}\}) = \{e_{i,\ell}, e_{i+1,\ell}\}$, for Y to be a stopping set, both the vertices in $\aleph(\{e_{i,r}\})$ must be in Y and therefore, $e_{i+1,\ell} \in Y$. Similarly, if $e_{i,\ell} \in Y$ for some $1 < i \leq n$, it must be true that $e_{i-1,\ell} \in Y$. Extending the argument, it follows that $e_{i,\ell} \in Y$ for some $1 \leq i \leq n$ if and only if $e_{i,\ell} \in Y$ for every $1 \leq i \leq n$. Thus, $|(Y \cap L_i)| = |(Y \cap L_j)|$ for all $1 \leq i, j \leq n$. Set $p = |Y \cap L_0|$ and $q = |Y \cap L_i|$ for any $1 \leq i \leq n$, Clearly $0 \leq p \leq n$, $0 \leq q \leq m$. Since $|Y| = \sum_{i=0}^{n} |Y \cap L_i|$, we have $|Y| = p + qn$. Note that if $|Y| > 0$, $q = 0$ then $p > 0$. Finally if $q > 0$, then there exists some $e_{1,\ell} \in (Y \cap L_1)$. Therefore, $e_{0,r} \in \aleph(\{Y\})$. By construction of G', $\aleph(\{e_{0,r}\}) \cap (\bigcup_{i=1}^{n} L_i) = \{e_{1,\ell}\}$. since Y is a stopping set, we need $|\aleph(\{e_{0,r}\}) \cap Y| \geq 2$. Hence there must exist some $u \in L_0 \cap Y$ such that $u \in \aleph(\{e_{0,r}\})$ and consequently $p > 0$. ■

Lemma 3: If Y is a non-trivial stopping set in G' of size $t + nm$ for some $1 \leq t \leq n-1$, then $|Y \cap L_0| = t$ and $\bigcup_{i=1}^{n} L_i \subseteq Y$.

Proof: Suppose $|Y \cap L_0| = p$. By lemma 2, $n \geq p > 0$ and there exists some $0 \leq q \leq m$ such that $|Y| = t + nm = p + qn$ where $q = |Y \cap L_i|$ for all $1 \leq i \leq n$. Hence $t - p = n(q - m)$. ie., $n|(t-p)$. But since $1 \leq t, p \leq n$, $t - p = 0$ and hence $t = p$ and $q = m$. The lemma follows. ■

We are now ready to prove:

Lemma 4: If G' has a stopping set Y of size $t + nm$, $1 \leq t \leq n-1$, then G has a vertex cover of size t.

Proof: Let $X = Y \cap L_0$. By lemma 3, $L_1 \subseteq Y$. Hence, by construction of G', $R_0 \subseteq \aleph(Y)$. Further, in G', each $e_{0,r} \in R_0$ has exactly one neighbour in $\bigcup_{i=1}^{n} L_i$. But since Y is a stopping set, we have for every $e_{0,r} \in R_0$, $|N(\{e_{0,r}\}) \cap Y| \geq 2$. Hence, for each $e_{0,r} \in R_0$, there exists some $u \in X$ such that $(e_{0,r}, u) \in E'$. But then, by construction of G', $e = (u, v)$ for some $u \in V$, $v \in V$ in G. Thus, for every $e \in E$ there exists some $u \in X$ such that $e = (u, v)$ for some $v \in V$. Hence X is a vertex cover in G. By lemma 3, $|X| = t$, as required by the lemma. ■

Theorem 1: STOPPING SET is NP-Complete.

Proof: We have proved that $(G, t) \in$ VERTEX COVER if and only if $(G', t + nm) \in$ STOPPING SET. G' can be constructed from G in polynomial time ($O(nm)$ time suffices). Hence VERTEX COVER $\preceq_p$ STOPPING SET. It easy to verify whether a given set of left vertices of a bipartite graph forms a stopping set in time linear in the size of the graph. Hence STOPPING SET $\in$ NP. ■

IV. Acknowledgment

The authors would like to thank Dr. L. Sunil Chandran for helpful discussions.

References

[1] C. Di, D. Proietti, I. E. Telatar, T. J. Richardson and R. Urbanke, "Finite length analysis of low-density parity-check codes on the binary erasure channel", *IEEE Trans. Inf. Theory.*, Vol. 48, no. 6, pp. 1570-1579, June 2002.

[2] T. Tian, C. Jones, J. D. Villasenor, R. D. Wesel, "Construction of irregular LDPC codes with low error floors", *IEEE Intl. Conf. Comm.*, 2003, pp. 3125-3129.

[3] A. Ramamoorthy, R. Wesel, "Construction of short block length irregular LDPC codes", *ICC 2004*, Paris, June 2004.

[4] A. Orlitsky, R. Urbanke, K. Viswanathan, J. Shang, "Stopping sets and girth of Tanner graphs", *ISIT 2002*, June 2002.

[5] M. Schwartz, A. Vardy, "On the stopping distance and the stopping redudancy of codes", *submitted for publication to IEEE Trans. Info. Theory*, March 2005.

[6] M. Tanner, "A recursive approach to low-complexity codes", *IEEE Trans. Info. Theory*, Vol. 27, pp. 533-547, Sept 1981.

[7] T. H. Cormen, C. E. Leiserson, R. L. Rivest, *Introduction to Agorithms*, MIT Press, 1990.

[8] M. R. Garey, D. S. Johnson, *Computers and Intractability: A guide to the Theory of NP-Completeness*, W. H. Freeman, 1979.

[9] S. Cook, "The complexity of theorem proving procedures, *Proc. Third Ann. ACM Symp. on Theory of Computing*, pp. 151-158, 1971.