Iterative Decoding on Graphs with a Single Cycle[1]

Srinivas M. Aji, Gavin B. Horn[2] and Robert J. McEliece
Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA
E-mail: (mas,gavinh,rjm)@systems.caltech.edu

Abstract — **It is now understood [2, 3] that the turbo decoding algorithm is an instance of a probability propagation algorithm (PPA) on a graph with many cycles. In this paper we investigate the behavior of an PPA in graphs with a single cycle such as the graph of a tail-biting code. First, we show that for strictly positive local kernels, the iterations of the PPA converge to a unique fixed point, (which was also observed by Anderson and Hladik [1] and Weiss [5]). Secondly, we shall generalize a result of McEliece and Rodemich [4], by showing that if the hidden variables in the cycle are binary-valued, the PPA will always make an optimal decision. (This was also observed independently by Weiss [5]). When the hidden variables can assume 3 or more values, the behavior of the PPA is much harder to characterize.**

I. Message Passing Convergence

Consider a linear block code described by a tail-biting graph $G = (V, E)$ which consists of a single cycle. We can decode using this graph by passing messages $\mu_{i,j}$, between adjacent vertices v_i and v_j in the cycle. The PPA computes these messages as

$$\mu_{j,k} = \Phi_j \; \mu_{i,j} \tag{1}$$

where the matrix Φ_j is a function of the structure of the trellis and the received noisy codeword, and $(v_i, v_j), (v_j, v_k) \in E$. We can construct the matrix Φ_j for all $v_j \in V$.

A message passed in one direction will propagate through the vertices in the cycle due to the message passing schedule. Since the message is multiplied by a matrix at each vertex on the cycle, we can rewrite the updated message in terms of the old message as follows

$$\mu_{j,k}(new) = M_j \; \mu_{j,k}(old) \tag{2}$$

where $M_j = \Phi_j \Phi_i \cdots \Phi_k$ is the ordered product of the matrices associated with each vertex visited in the cycle. If we travel in the reverse direction we get the matrix M_j^T.

If M_j is strictly positive, or has only one eigenvalue of largest modulus, then by the Perron Frobenius Theorem, the PPA will converge to the unique, non-negative, principal right eigenvector of M_j in the forward direction and the unique, non-negative, principal left eigenvector of M_j in the reverse direction. So the iterations of the PPA converge but what is the significance of what they converge to?

II. PPA vs Optimal Decoding

Consider the following system. Define S to be the *signal matrix* consisting of a non-negative diagonal matrix with trace equal to 1. Define N to be the *noise matrix* consisting of a non-negative matrix with zero on its diagonal and all the off-diagonal elements less than 1. Define real non-negative constants a and b such that $a + b = 1$ so that we get the matrix

$$M = aS + bN. \tag{3}$$

Now let M be the matrix associated with vertex v and hidden variable x for a tail-biting graph. The PPA on the tail-biting graph will estimate the APP values of x as the component-wise product of the principal left and right eigenvalues of M, while the actual APP values of x are simply the values on the diagonal of M.

Let x be a binary valued hidden variable where m_{11} and m_{22} are the probabilities $Pr(x = 0)$ and $Pr(x = 1)$ respectively. For the component-wise product of the two eigenvectors to make an optimal APP decision we must satisfy $(\lambda_+ - m_{11})^2 \leq m_{12}m_{21}$, where λ_+ is the principal eigenvalue of M. It is easy to show that for a binary valued hidden variable the PPA will always make the correct decision, however the certainty of the decision is dependent on the product of the off-diagonal elements.

For the non-binary case, it is much harder to characterize the performance of the PPA in terms of the matrix M. When $b = 0$, there is no noise and the PPA will make a correct APP hard decision. When $a = 0$, the PPA decision is based entirely on the noise matrix N and can be correct or incorrect depending on the nature of the noise. For intermediate values of a and b, the PPA may or may not make a correct decision depending on the mean diagonal - off-diagonal ratio (DOR) of the elements of M. Generally the larger the DOR, the higher the likelihood of the PPA making a correct decision.

In the case where the off-diagonal elements are all the same, one can show that the PPA will always make the correct decision.

References

[1] J. B. Anderson, S. M. Hladik, "Tailbiting MAP decoders," *IEEE J. Selected Areas in Communication*, vol. 16, pp. 297–302, February 1998.

[2] F. R. Kschischang and B. Frey, "Iterative decoding of compound codes by probability propagation in graphical models," *IEEE J. Selected Areas in Communication*, vol. 16, pp. 219–230, February 1998.

[3] R. J. McEliece, D. J. C. MacKay, J-F. Cheng, "Turbo decoding as an instance of Pearl's 'belief propagation' algorithm," *IEEE J. Selected Areas in Communication*, vol. 16, pp. 140–153, February 1998.

[4] R. J. McEliece, E. R. Rodemich, and J-F. Cheng, "The turbo decision algorithm," Proc. *33rd Allerton Conference on Communication, Control and Computing* (October 1995), pp. 366–379.

[5] Y. Weiss, "Belief Propagation and Revision in Networks with Loops," M.I.T. A.I. Memo no. 1616, November 1997.

[1] This work was partially supported by NSF grant no. NCR-9505975, AFOSR grant no. 5F49620-97-1-0313, and a grant from Qualcomm, Inc.

[2] This work was partially supported by an NSERC Scholarship.