Quantum Advantage in Classical Communications via Belief-Propagation with Quantum Messages

Narayanan Rengaswamy*, Kaushik Seshadreesan†, Saikat Guha†, and Henry D. Pfister*

*Rhodes Information Initiative at Duke (iiD), Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA

†College of Optical Sciences, University of Arizona, Tucson, AZ, USA

Motivation and Contributions

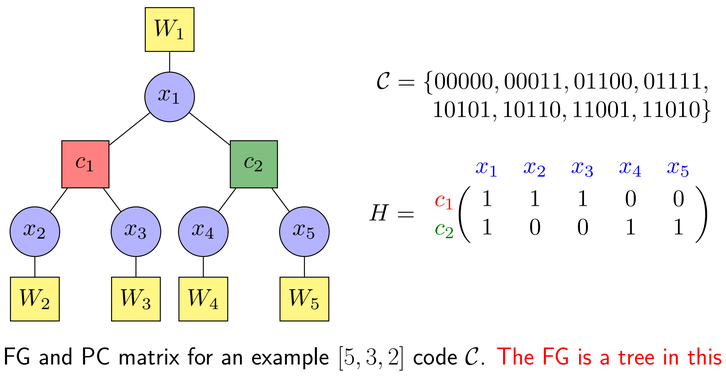
- Communication: Need to distinguish messages reliably at the receiver.
- We consider classical communication over classical-quantum (CQ) channels, e.g., deep-space optical communications (BPSK modulated pure-loss optical channel).
- In general, need a collective measurement on all received qubits to perform minimum probability of error measurement. But this is hard to realize in practice.
- Belief-propagation (BP) is a classical algorithm that can be used to efficiently decode classical codes over classical channels. It passes probabilities as messages over the code's factor graph to compute posterior marginal distributions.
- Renes [1] proposed a BP algorithm that passes quantum messages to decode classical codes over the pure-state CQ channel, e.g., deep-space optical comm.
- We analyze this algorithm for an example 5-bit code and show that it is actually quantum optimal, i.e., it achieves the minimum probability of error for distinguishing the codewords after transmission over the pure-state channel.
  - Better than current receivers in deep-space optical communications that measure each qubit and post-process the result classically. This provides a significant quantum advantage!
- We provide a full circuit decomposition for the algorithm into standard single-qubit, two-qubit and Toffoli gates. The BPQM circuit is very structured.

Near-Term Quantum Advantage!

Our results indicate a potentially near-term *communication* application for quantum technologies that does not require a universal fault-tolerant quantum computer!

Factor Graphs (FGs) for Linear Codes

- FG: A graphical representation for a joint distribution $f(x_1, x_2, \ldots, x_n)$ over n variables; bipartite with (circle) nodes for variables, (square) nodes for factors of f.
- If f can be factored, then the factor graph exhibits some (computational) structure.
- $[n, k, d]$ Linear Code: $\mathcal{C} \coloneqq \{\underline{x} \in \mathbb{F}_2^n \mid H \cdot \underline{x}^T \equiv \underline{0}^T (\mathrm{mod}\ 2)\}$, where $H \in \mathbb{F}_2^{(n-k)\times n}$ is a parity-check (PC) matrix for $\mathcal{C}$. Minimum Hamming weight of any $\underline{x} \in \mathcal{C}$ is d.



$$\mathcal{C} = \{00000, 00011, 01100, 01111, 10101, 10110, 11001, 11010\}$$

$$H = \begin{array}{c} \\ c_1 \\ c_2 \end{array} \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

Figure 1: FG and PC matrix for an example $[5, 3, 2]$ code $\mathcal{C}$. The FG is a tree in this case.

- Classical channel: $W(y|x)$, where $x \in \{0, 1\}$ and $y \in \mathcal{Y}$, some output alphabet. $W(y|x)$ is the transition probability that y is received upon transmitting x.

Belief-Propagation (BP) Algorithm

- Passes messages through the (bipartite) factor graph to efficiently compute the bitwise posterior marginal distributions. For an introduction to BP, see [2].
- Problem: Transmit $\underline{x} \in \mathcal{C}$, receive $\underline{y} \in \mathcal{Y}^n$, produce an estimate $\underline{\hat{x}}$ of the transmitted codeword that minimizes the average probability of error.
- Block Maximum-a-posteriori (MAP): Calculate the posterior distribution

$$\begin{aligned} p(\underline{x}|\underline{y}) &= \frac{p(\underline{y}|\underline{x}) \cdot p(\underline{x})}{\sum_{\underline{x}\in\{0,1\}^5} p(\underline{y}|\underline{x}) \cdot p(\underline{x})} \\ &\propto W(y_1|x_1) \cdot [\mathbb{I}(x_1 \oplus x_2 \oplus x_3 = 0) W(y_2|x_2) W(y_3|x_3)] \\ &\quad \cdot [\mathbb{I}(x_1 \oplus x_4 \oplus x_5 = 0) W(y_4|x_4) W(y_5|x_5)]\,. \end{aligned}$$

Then choose the vector that maximizes the posterior as the Block-MAP estimate:

$$\underline{\hat{x}}^{\mathsf{MAP}} \coloneqq \underset{\underline{x}\in\{0,1\}^5}{\mathsf{argmax}}\ p(\underline{x}|\underline{y}).$$

- This is optimal but computationally expensive; need to calculate $p(\underline{x}|\underline{y})$ for 2^k $\underline{x}$'s.
- Bit-MAP: Marginalize the posterior and estimate the vector $\underline{\hat{x}}$ bitwise, e.g., for $\hat{x}_1$,

$$\hat{x}_1^{\mathsf{MAP}} = \underset{x_1\in\{0,1\}}{\mathsf{argmax}} \left\{ W(y_1|x_1) \cdot \left[\sum_{x_2,x_3\in\{0,1\}^2} \mathbb{I}(x_1 \oplus x_2 \oplus x_3 = 0) W(y_2|x_2) W(y_3|x_3) \right] \cdot \left[\sum_{x_4,x_5\in\{0,1\}^2} \mathbb{I}(x_1 \oplus x_4 \oplus x_5 = 0) W(y_4|x_4) W(y_5|x_5) \right] \right\}.$$

- BP: Implements Bit-MAP efficiently by using distributivity of addition over multiplication. First, at nodes c_1 and c_2, compute the "local beliefs" (for x_1) in the two square brackets simultaneously. Then multiply these with the direct "belief" from the channel, $W(y_1|x_1)$, at node x_1. Finally, find the value of x_1 that maximizes the result. When the FG is not a tree, this approximates Bit-MAP (efficiently).

BP performs local inference matched to induced channels

- Variable Node (VN) Channel Convolution:

$$[W \circledast W'](y, z|x) = W(y|x) \cdot W'(z|x, y) = W(y|x) \cdot W'(z|x).$$

- Factor Node (FN) Channel Convolution:

$$[W \boxast W'](y, z|x) = \frac{1}{2} W(y|x) \cdot W'(z|0) + \frac{1}{2} W(y|x \oplus 1) \cdot W'(z|1).$$

Communication over the Pure-State CQ Channel

- Pure-state channel: Defined for classical inputs $x \equiv |x\rangle\langle x|, x \in \{0, 1\}$, as

$$W(x) \coloneqq \langle x|0\rangle \cdot |\theta\rangle\langle\theta| + \langle x|1\rangle \cdot |-\theta\rangle\langle-\theta| = |(-1)^x\theta\rangle\langle(-1)^x\theta|\,,$$
$$|\pm\theta\rangle \coloneqq \cos\frac{\theta}{2}|0\rangle \pm \sin\frac{\theta}{2}|1\rangle\,. \quad (\cos\theta = e^{-2N}, N = \text{mean photon number per mode})$$

- Distinguishing Quantum States: Two quantum states $|\psi_1\rangle, |\psi_2\rangle$ can be perfectly distinguished iff they are orthogonal: $\langle\psi_1|\psi_2\rangle = 0$. Otherwise, the minimum probability of error in distinguishing them is given by

$$P_e^* = \frac{1}{2}\left[1 - \sqrt{1 - |\langle\psi_1|\psi_2\rangle|^2}\right] = \frac{1}{2}\left[1 - \sqrt{1 - \cos^2\theta}\right] = \frac{1 - \sin\theta}{2} \quad (|\psi_1\rangle \coloneqq |\theta\rangle\,,\ |\psi_2\rangle \coloneqq |-\theta\rangle),$$

which is achieved by the Helstrom measurement [3], i.e., measuring in the X basis in this case.

- Dolinar Receiver: Performs qubit-wise Helstrom measurements, followed by classical post-processing (e.g., MAP). Quantum optimal receiver performs better via collective measurements on all qubits. This measurement is given by Yuen-Kennedy-Lax (YKL) conditions, but hard to find a structured circuit.

Belief-Propagation with Quantum Messages (BPQM)

- Renes [1] generalized BP channel convolutions to CQ channels $W(x) \equiv W(|x\rangle\langle x|), x \in \{0, 1\}$:

$$\text{VN Convolution}: [W \circledast W'](x) \coloneqq W(x) \otimes W'(x),$$
$$\text{FN Convolution}: [W \boxast W'](x) \coloneqq \frac{1}{2} W(x) \otimes W'(0) + \frac{1}{2} W(x \oplus 1) \otimes W'(1).$$

- Local inference strategy: Perform a unitary $U_{\circledast}(\theta, \theta')$ at a VN, and $\mathsf{CX}_{W\to W'}$ at a FN; these satisfy

$$U_{\circledast}(\theta, \theta')\left(|\pm\theta\rangle \otimes |\pm\theta'\rangle\right) = \left|\pm\theta^{\circledast}\right\rangle \otimes |0\rangle\,, \quad \mathsf{CX}\left([W \boxast W'](x)\right)\mathsf{CX}^\dagger = \sum_{j\in\{0,1\}} p_j \left|\pm\theta_j^{\boxast}\right\rangle\left\langle\pm\theta_j^{\boxast}\right| \otimes |j\rangle\langle j|\,.$$

- Messages: At VN, pass qubit from 1st system; at FN, measure 2nd system and pass resulting qubit from 1st system. But we need to perform tasks coherently and not discard any qubits along the way.

Quantum Optimality of BPQM for the 5-Bit Code

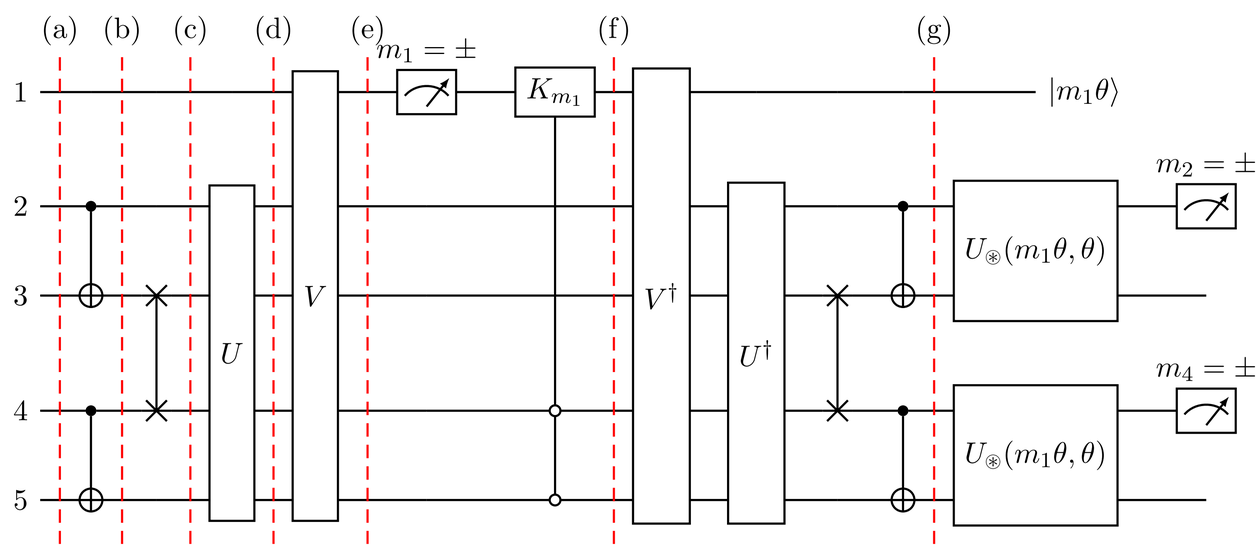


Figure 2: The full BPQM circuit for decoding all bits of the example 5-bit code $\mathcal{C}$ (on the left); Measurement $m_i = (-1)^{\hat{x}_i}$.

(a) $\rho_{\pm,a} = |\pm\theta\rangle\langle\pm\theta|_1 \otimes [W \boxast W](x_1)_{23} \otimes [W \boxast W](x_1)_{45} = |\pm\theta\rangle\langle\pm\theta| \otimes \frac{1}{4}\sum_{\underline{c}\in\mathcal{C}:\, c_1=x_1} \bigotimes_{i=2}^{5} W(c_i)$.

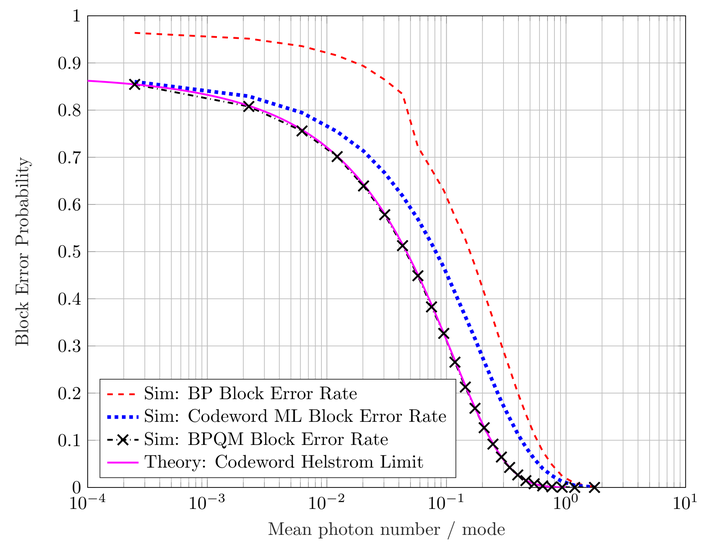
(b) $\rho_{\pm,b} = |\pm\theta\rangle\langle\pm\theta|_1 \otimes \left[\sum_{j\in\{0,1\}} p_j \left|\pm\theta_j^{\boxast}\right\rangle\left\langle\pm\theta_j^{\boxast}\right|_2 \otimes |j\rangle\langle j|_3\right] \otimes \left[\sum_{k\in\{0,1\}} p_k \left|\pm\theta_k^{\boxast}\right\rangle\langle\pm\theta_k^{\boxast}|_4 \otimes |k\rangle\langle k|_5\right]$.

(c) $\rho_{\pm,c} = |\pm\theta\rangle\langle\pm\theta|_1 \otimes \sum_{j,k\in\{0,1\}^2} p_j p_k \left|\pm\theta_j^{\boxast}\right\rangle\left\langle\pm\theta_j^{\boxast}\right|_2 \otimes \left|\pm\theta_k^{\boxast}\right\rangle\langle\pm\theta_k^{\boxast}|_3 \otimes |j\rangle\langle j|_4 \otimes |k\rangle\langle k|_5$.

(d) $\sigma_{\pm} = \sum_{j,k\in\{0,1\}^2} p_j p_k |\pm\theta\rangle\langle\pm\theta|_1 \otimes \left|\pm\theta_{jk}^{\circledast}\right\rangle\left\langle\pm\theta_{jk}^{\circledast}\right|_2 \otimes |0\rangle\langle 0|_3 \otimes |jk\rangle\langle jk|_{45}$, where the applied unitary operation is $U \coloneqq \sum_{j,k\in\{0,1\}^2} U_{\circledast}(\theta_j^{\boxast}, \theta_k^{\boxast})_{23} \otimes |jk\rangle\langle jk|_{45}$ and $\cos\theta_{jk}^{\circledast} \coloneqq \cos\theta_j^{\boxast}\cos\theta_k^{\boxast}$.

(e) $\Psi_{\pm} = \sum_{j,k\in\{0,1\}^2} p_j p_k \left|\pm\varphi_{jk}^{\circledast}\right\rangle\left\langle\pm\varphi_{jk}^{\circledast}\right|_1 \otimes |0\rangle\langle 0|_2 \otimes |0\rangle\langle 0|_3 \otimes |jk\rangle\langle jk|_{45}$, where the applied unitary operation is $V \coloneqq \sum_{j,k\in\{0,1\}^2} U_{\circledast}(\theta, \theta_{jk}^{\circledast})_{12} \otimes |jk\rangle\langle jk|_{45}$ and $\cos\varphi_{jk}^{\circledast} \coloneqq \cos\theta\cos\theta_{jk}^{\circledast}$.

- Optimality for x_1: Follows from unitary invariance of Helstrom success probability $\frac{1}{2} + \frac{1}{4}\|\rho_{+,a} - \rho_{-,a}\|_1$.
- Optimality for other bits and for full codeword: See [4] for detailed analysis and circuit decomposition.
- Questions: Goal of BPQM still unclear. Does optimality hold for a general code family (e.g., tree FGs)?



References

[1] J. M. Renes, "Belief propagation decoding of quantum channels by passing quantum messages," *New Journal of Physics*, vol. 19, no. 7, p. 072 001, 2017. [Online]. Available: http://arxiv.org/abs/1607.04833.

[2] T. J. Richardson and R. L. Urbanke, *Modern Coding Theory*. New York, NY: Cambridge Univ. Press, 2008.

[3] C. W. Helstrom, J. W. Liu, and J. P. Gordon, "Quantum-mechanical communication theory," *Proc. of the IEEE*, vol. 58, no. 10, pp. 1578–1598, 1970.

[4] N. Rengaswamy, K. Seshadreesan, S. Guha, and H. D. Pfister, "Quantum Advantage in Classical Communications via Belief-Propagation with Quantum Messages," *In preparation, to be posted on arXiv*, 2020.