MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Channel Decoding with Quantum Approximate Optimization Algorithm

Matsumine, T.; Koike-Akino, T.; Wang, Y.

TR2019-071 July 11, 2019

Abstract

Motivated by the recent advancement of quantum processors, we investigate quantum approximate optimizationalgorithm (QAOA) to employ quasi-maximum-likelihood (ML) decoding of classical channel codes. QAOA is a hybrid quantumclassical variational algorithm, which is advantageous for the near-term noisy intermediate-scale quantum (NISQ) devices, where the fidelity of quantum gates is limited by noise and decoherence. We first describe how to construct Ising Hamiltonian model to realize quasi-ML decoding with QAOA. For level-1 QAOA, we derive the systematic way to generate theoretical expressions of cost expectation for arbitrary binary linear codes. Focusing on [7, 4] Hamming code as an example, we analyze the impact of the degree distribution in associated generator matrix on the quantum decoding performance. The excellent performance of higher-level QAOA decoding is verified when Pauli rotation angles are optimized through meta-heuristic variational quantum eigensolver (VQE). Furthermore, we demonstrate the QAOA decoding performance in a real quantum device.

IEEE International Symposium on Information Theory (ISIT)

Channel Decoding with Quantum Approximate Optimization Algorithm

Toshiki Matsumine[‡†], Toshiaki Koike-Akino[‡], and Ye Wang[‡]
[‡] Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139, USA.
[†] Department of Electrical and Computer Engineering, Yokohama National University,
79-5 Tokiwadai, Hodogaya, Yokohama, Kanagawa, Japan
Email: matsumine-toshiki-tk@ynu.jp, {koike, yewang}@merl.com

***Abstract*—Motivated by the recent advancement of quantum processors, we investigate quantum approximate optimization algorithm (QAOA) to employ quasi-maximum-likelihood (ML) decoding of classical channel codes. QAOA is a hybrid quantum-classical variational algorithm, which is advantageous for the near-term noisy intermediate-scale quantum (NISQ) devices, where the fidelity of quantum gates is limited by noise and decoherence. We first describe how to construct Ising Hamiltonian model to realize quasi-ML decoding with QAOA. For level-1 QAOA, we derive the systematic way to generate theoretical expressions of cost expectation for arbitrary binary linear codes. Focusing on [7, 4] Hamming code as an example, we analyze the impact of the degree distribution in associated generator matrix on the quantum decoding performance. The excellent performance of higher-level QAOA decoding is verified when Pauli rotation angles are optimized through meta-heuristic variational quantum eigensolver (VQE). Furthermore, we demonstrate the QAOA decoding performance in a real quantum device.**

I. Introduction

Quantum computers can offer a significant potential to accomplish more efficient computations compared to traditional digital computers for various problems by exploiting quantum-mechanism, e.g., superposition, entanglement, and quantum tunneling, in terms of not only execution time but also energy consumption. It is highly expected that quantum computers could provide breakthroughs in wide range of research domain, such as chemical engineering, complex system optimizations, and artificial intelligence. Not only theoretical concept, several commercial quantum computers have been already built by several industries including IBM, Google, Honeywell, and so on. For instance, IBM has made 5 and 16 qubits quantum computers available to the public via cloud service.

In this paper, we consider decoding *classical* binary linear codes assisted by quantum computing. Since the maximum-likelihood (ML) decoding of channel codes, which has non-deterministic polynomial-time hardness [1], is feasible only when the block length is very short, the most popular approach in the communication standards is based on sub-optimal belief-propagation (BP) algorithm combined with probabilistic codes, such as turbo codes and low-density parity-check (LDPC) codes. Although these approaches already give the practical solutions, the decoding complexity may be further increasing when we consider improving the performance of channel codes with finite block lengths and finite iterations. For example, nonbinary codes have been shown to achieve significant performance gain over binary counterparts in the short-to-medium block length regime at the cost of increasing decoding complexity. Channel decoding with hybrid quantum-classical algorithm can be a potential new framework to significantly reduce the decoding complexity for those scenarios. Note that our focus is a classical error correction, e.g., as used in Wi-Fi channels, assisted by quantum processors and hence we do not discuss quantum error correction [2]–[6] for protecting quantum states from decoherence and control errors.

Applications of quantum algorithm to communication systems have been investigated, e.g., in [7]–[9]. In [8], the quantum-assisted iterative detection for multi-user systems was proposed. Joint channel estimation and data detection was studied in the uplink of multi-input multi-output (MIMO) systems [10]. Optimization of vector perturbation precoding for the multi-user transmission was proposed in [11]. However, they typically assume that long qubits (e.g., more than 1000) are available and that quantum-circuit gates have no errors, which may be beyond the capabilities of near-term noisy intermediate-scale quantum (NISQ) computers.

A hybrid quantum-classical algorithm called quantum approximate optimization algorithm (QAOA) was recently proposed to solve various NP-hard problems [12]. Because of high robustness against quantum errors, QAOA is expected to be a suitable candidate for NISQ devices [13] and a breakthrough driver towards *quantum supremacy* [14]. For level-p QAOA, classical discrete optimization can be probabilistically solved by mapping to an Ising Hamiltonian with $2p$ annealing parameters. The approximation quality of QAOA improves towards the ground state when increasing p. Even for $p = 1$, QAOA has guaranteed better probable performance than classical algorithms for certain problems such as MaxCut. In order to optimize QAOA ansatz, we typically use variational quantum eigensolver (VQE) where a quantum processor performs an expectation calculation of the cost function and the $2p$ parameters are optimized by a classical computer in a closed loop.

The main contributions of this paper are summarized below:

- We develop a new framework with QAOA quantum processor to decode classical channel codes;
- We introduce an Ising Hamiltonian for QAOA to realize quasi-ML decoding of arbitrary linear binary codes;

T. Matsumine conducted this research when he was an intern at MERL.

- We propose a theoretical method to derive quantum eigenvalue expressions for level-1 QAOA decoding;
- We analyze the impact of the degree distribution of $[7, 4]$ Hamming codes on quantum decoding performance;
- We implement QAOA decoder with both quantum simulators and real quantum computers.

II. Preliminaries

A. Quantum Bit (Qubit)

In quantum systems, a *qubit* is expressed as the following state superposing bases of $|0\rangle$ and $|1\rangle$: $|\phi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$, where α_1 and α_2 are complex numbers subject to $|\alpha_0|^2 + |\alpha_1|^2 = 1$. When qubits are measured, the classical bit 0 or 1 is observed with a probability of $|\alpha_0|^2$ or $|\alpha_1|^2$, respectively. The above *ket-notation* corresponds to column-vector operations of the two basis states $|0\rangle = [1, 0]^{\mathrm{T}}$ and $|1\rangle = [0, 1]^{\mathrm{T}}$, whereas the *bra-notation* is used for row-vector operations corresponds to its Hermitian transpose; i.e., $\langle\phi| = |\phi\rangle^\dagger = [\alpha_0^*, \alpha_1^*]$. Here, $[\cdot]^\dagger$, $[\cdot]^*$ and $[\cdot]^{\mathrm{T}}$ denote Hermitian transpose, complex conjugate and transpose, respectively. Note that multi-qubit state is represented by sum of Kronecker products of basis vectors such as $|000\rangle = |0\rangle^{\otimes 3}$.

B. Quantum Gates

The basic operations on a qubit is defined as a unitary matrix, which is called *gate*. Some of the most common gates are associated with Pauli matrices: $\mathbf{I} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, $\mathbf{X} = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, $\mathbf{Y} = \left[\begin{smallmatrix} 0 & -\jmath \\ \jmath & 0 \end{smallmatrix}\right]$, and $\mathbf{Z} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$, where $\jmath$ is the imaginary unit satisfying $\jmath^2 = -1$. The X gate is bit-flip (i.e., NOT operation), Z gate is phase-flip, and Y gate flips both bit and phase. The Hadamard (H) gate is used to generate a superposition state $|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$: $\mathbf{H} = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$. A controlled-NOT (CNOT or CX) gate is a multi-qubit gate that flips the target qubit if and only if the control qubit is $|1\rangle$.

C. QAOA Algorithm

QAOA algorithm [12] was proposed for discrete optimization problems, such as the MaxSat, MaxCut, and MaxClique, which are expressed as an unconstrained discrete optimization:

$$C(\mathbf{z}) = \sum_{\nu=1}^{n} C_\nu(\mathbf{z}), \tag{1}$$

where $\mathbf{z} = [z_1, z_2, \ldots, z_k] \in \mathbb{F}_2^k$ denotes binary label, and $C_\nu(\mathbf{z})$ is the νth binary function to satisfy, which are called *clause*. The QAOA tries to find a binary vector $\mathbf{z}$ that maximizes the number of satisfied clause $C_\nu(\mathbf{z})$, by associating $\mathbf{z}$ with quantum states in Z-direction of Bloch sphere.

For typical QAOA, quantum state is initialized to admixing superposition state $|+\rangle^{\oplus k}$, which produces equi-probable random bits $\mathbf{z}$ once measured when no other operations. Let $U(C, \gamma)$ denote a unitary operator for the cost Hamiltonian C with an angle $0 \le \gamma \le 2\pi$, which is defined as

$$U(C, \gamma) = \exp(-\jmath\gamma C) = \prod_{\nu=1}^{n} \mathrm{e}^{-\jmath\gamma C_\nu}. \tag{2}$$

Consider a driver Hamiltonian defined by $B = \sum_{\kappa=1}^{k} \mathbf{X}_\kappa$, which flips k qubits independently. We use the following unitary operator with an angle $0 \le \beta \le \pi$:

$$U(B, \beta) = \exp(-\jmath\beta B) = \prod_{\kappa=1}^{k} \mathrm{e}^{-\jmath\beta \mathbf{X}_\kappa}. \tag{3}$$

Note that this driver Hamiltonian has the ground-state eigenvector of $|\phi\rangle = |+\rangle^{\otimes k}$.

QAOA uses alternating quantum operator ansatz circuits of depth $p \in \mathbb{Z}_+$ with Hamiltonians B and C to maximize the expected cost function, with $2p$ angle parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$:

$$|\psi_{\boldsymbol{\gamma},\boldsymbol{\beta}}\rangle = U(B, \beta_p)U(C, \gamma_p)\cdots U(B, \beta_1)U(C, \gamma_1)|\phi\rangle. \tag{4}$$

Let F_p denote the expectation of the cost function C as

$$F_p(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \langle C\rangle(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \langle\psi_{\boldsymbol{\gamma},\boldsymbol{\beta}}|C|\psi_{\boldsymbol{\gamma},\boldsymbol{\beta}}\rangle, \tag{5}$$

and let $F_p^\star$ be the maximum of $F_p(\boldsymbol{\gamma}, \boldsymbol{\beta})$ over the angles: $F_p^\star = \max_{\boldsymbol{\gamma},\boldsymbol{\beta}} F_p(\boldsymbol{\gamma}, \boldsymbol{\beta})$. The objective of QAOA algorithm is to maximize $F_p^\star$ by properly choosing parameters $\boldsymbol{\gamma}, \boldsymbol{\beta}$. The quality of the approximation improves as p increases and the global optimum maximizing cost function $C(\mathbf{z})$ can be asymptotically achieved when infinite depth p, i.e., $\lim_{p\to\infty} F_p^\star = \max_{\mathbf{z}} C(\mathbf{z})$.

In QAOA, the calculation of the expectation $F_p(\boldsymbol{\gamma}, \boldsymbol{\beta})$ is performed by repeated measurements with quantum computers based on variational principle in the computational basis, which is infeasible for classical computers as p increases. The search for optimal variational parameters $\boldsymbol{\gamma}, \boldsymbol{\beta}$ that maximize $F_p(\boldsymbol{\gamma}, \boldsymbol{\beta})$ are efficiently performed by classical computers, e.g., employing Nelder–Mead (NM) method [14].

III. Channel Decoding with QAOA Algorithm

Here, we propose to use QAOA for performing quasi-ML decoding of linear error-correcting codes for digital communications over noisy classical channels.

A. Digital Communications Model

We consider an $[n, k]$ binary linear code, specified by a generator matrix of $\mathbf{G} \in \mathbb{F}_2^{k\times n}$, where n and k are the lengths of codewords and information bits, respectively. The codeword $\mathbf{x} \in \mathbb{F}_2^n$ is generated as $\mathbf{x} = \mathbf{u}\mathbf{G}$, where arithmetic operation based on modulo-2 is taken place and $\mathbf{u} \in \mathbb{F}_2^k$ is the binary information vector. Over digital transmission channels, the received signal is modeled as $\mathbf{y} = \mathbf{x} + \mathbf{w}$, where $\mathbf{y} \in \mathbb{F}_2^n$ and $\mathbf{w} \in \mathbb{R}^n$ are the received and noise vectors, respectively. In this paper, we focus on binary-symmetric channel (BSC) for simplicity since extension to other channels such as additive white Gaussian noise channel is straightforward.

For BSCs, the problem of channel decoding is to find the information bits $\mathbf{u}$ which generate codeword $\mathbf{x}$ whose Hamming distance from the received signal is minimized, i.e.,

$$\operatorname*{arg\,min}_{\mathbf{u}\,:\,\mathbf{x}=\mathbf{u}\mathbf{G}} d_{\mathrm{H}}(\mathbf{y}|\mathbf{x}) = \operatorname*{arg\,max}_{\mathbf{u}\,:\,\mathbf{x}=\mathbf{u}\mathbf{G}} \sum_{\nu=1}^{n} (1 - 2y_\nu)(1 - 2x_\nu), \tag{6}$$

where $d_{\mathrm{H}}(\mathbf{y}|\mathbf{x})$ is the number of elements of two vectors $\mathbf{x}$ and $\mathbf{y}$ which differ. This is equivalent to maximizing the correlation between the transmitted codeword and the received vector.

B. Construction of Cost Hamiltonian

The proposed quantum decoder operates on k-qubit space corresponding to information bits $\mathbf{u}$. The objective of quantum decoding is to find most-likely k-qubit states that achieves (6). To do so, we consider the following cost Hamiltonian:

$$C = \sum_{\nu=1}^{n} C_\nu = \sum_{\nu=1}^{n} (1 - 2y_\nu) \prod_{\kappa \in \mathcal{I}_\nu^{\mathrm{c}}} \mathbf{Z}_\kappa, \tag{7}$$

where $\mathcal{I}_\nu^{\mathrm{c}}$ is a set of nonzero-element indices in the νth column of a generator matrix $\mathbf{G}$, i.e., $\mathcal{I}_\nu^{\mathrm{c}} = \{\kappa : [\mathbf{G}]_{\kappa,\nu} = 1\}$ where $[\cdot]_{i,j}$ denotes the element of an argument matrix at the ith row and jth column. Since the Z-gate performs as $+|\phi\rangle$ or $-|\phi\rangle$ for the qubit state of $|\phi\rangle = |0\rangle$ or $|1\rangle$, respectively, maximizing the cost Hamiltonian (7) is equivalent to ML decoding (6). The QAOA decoding has a linear complexity with respect to n.

C. Degree Optimization of Generator Matrix

The proposed cost Hamiltonian in (7) is a generalized version used for MaxCut problem [12], in which case the column degree is identical to be two, i.e., $d_\nu^{\mathrm{c}} = |\mathcal{I}_\nu^{\mathrm{c}}| = 2$, regardless of column ν. Even for such regular degree-2 cost Hamiltonian, it was shown in [15] that the quality of QAOA approximation is highly dependent on the graph connectivity, specifically, the number of girth-6 (i.e., triangles in graph) and row degrees $d_\kappa^{\mathrm{r}} = |\mathcal{I}_\kappa^{\mathrm{r}}|$ for a row-wise nonzero-entry index set of $\mathcal{I}_\kappa^{\mathrm{r}} = \{\nu : [\mathbf{G}]_{\kappa,\nu} = 1\}$.

In order to obtain an insight to optimize generator matrix $\mathbf{G}$ suited for our QAOA decoding, we consider to create different degree distributions by applying the following basis transform to the original generator matrix $\mathbf{G}$: $\mathbf{G}' = \mathbf{P}\mathbf{G}$, where $\mathbf{P} \in \mathbb{F}_2^{k\times k}$ is a full-rank matrix that performs basic row operations. It should be noticed that the performance of linear block codes over symmetric channels is invariant with respect to such a basis transform for the classic ML decoder because the Hamming weight spectrum remains unchanged. In the case of QAOA decoder, however, the decoder performance depends on the specific structure of generator matrices.

For example, the generator matrix of $[7, 4]$ systematic Hamming codes is given by

$$\mathbf{G} = \begin{bmatrix} 1&0&0&0&1&1&0 \\ 0&1&0&0&1&0&1 \\ 0&0&1&0&0&1&1 \\ 0&0&0&1&1&1&1 \end{bmatrix}. \tag{8}$$

The column-degree distribution of this matrix is $[1, 1, 1, 1, 3, 3, 3]$, whose average degree is $\bar{d}^{\mathrm{c}} = \mathbb{E}[d_\nu^{\mathrm{c}}] = 1.86$ with $\mathbb{E}[\cdot]$ denoting an expectation. Suppose the following transform matrix for instance:

$$\mathbf{P} = \begin{bmatrix} 1&1&1&1 \\ 0&1&1&1 \\ 1&0&1&0 \\ 0&0&0&1 \end{bmatrix}, \tag{9}$$

then the new generator matrix will be written as

$$\mathbf{G}' = \begin{bmatrix} 1&1&1&1&1&1&1 \\ 0&1&1&1&0&0&1 \\ 1&0&1&0&1&0&1 \\ 0&0&0&1&1&1&1 \end{bmatrix}, \tag{10}$$

whose average degree increases to $\bar{d}^{\mathrm{c}} = 2.71$ without changing Hamming weight spectrum of the linear codes. In this way, we can evaluate the performance of QAOA decoder with various generator matrices to optimize degree distributions.

IV. Performance Analysis of QAOA Decoding

We describe how to systematically analyze the performance of QAOA channel decoding. We generalize the analytical method [15] investigated for MaxCut problem, by considering irregular degree distributions. We focus on level-1 QAOA having single driver B and cost C Hamiltonians.

A. Theoretical Analysis Method for Level-1 QAOA

To analyze idealistic QAOA behavior as performance limit, we here assume zero-word transmission over error-free channels $\mathbf{y} = \mathbf{x} = \mathbf{0}$ as generalization is straightforward. Since the cost expectation can be decomposed as $F_p(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \sum_\nu \langle C_\nu \rangle$, we focus the νth cost Hamiltonian whose degree is $d_\nu^{\mathrm{c}} = |\mathcal{I}_\nu^{\mathrm{c}}|$. For example in Hamming code generator (8), the 5th column $[\mathbf{G}]_{:,5} = [1, 1, 0, 1]^{\mathrm{T}}$ corresponds to the Hamiltonian $C_5 = (1 - 2y_5)\mathbf{Z}_1\mathbf{Z}_2\mathbf{Z}_4$ whose degree is $d_5^{\mathrm{c}} = 3$.

In $\langle C_\nu \rangle = \langle \phi | U^\dagger(C, \gamma_1) U^\dagger(B, \beta_1) C_\nu U(B, \beta_1) U(C, \gamma_1) | \phi \rangle$, observe that most terms in the operator $U(B, \beta_1) = \prod \exp(-\jmath\beta_1 \mathbf{X}_i)$ will commute and result in

$$U(B, \beta_1)^\dagger \big(\prod \mathbf{Z}_\kappa\big) U(B, \beta_1) = \prod (c'\mathbf{Z}_\kappa + s'\mathbf{Y}_\kappa), \tag{11}$$

where $c' = \cos(2\beta_1)$ and $s' = \sin(2\beta_1)$, by recalling that Pauli matrix $\boldsymbol{\Sigma} \in \{\mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$ satisfies $\exp(\jmath\beta\boldsymbol{\Sigma}) = \cos(\beta)\mathbf{I} + \jmath\sin(\beta)\boldsymbol{\Sigma}$ and circulation rule such as $\mathbf{X}\mathbf{Z} = -\jmath\mathbf{Y}$. The d_ν^{c}-ary product of binary additions in (11) can be expanded in the $2^{d_\nu^{\mathrm{c}}}$-ary sum of d_ν^{c}-ary products. Let $\mathbf{b} \in \mathbb{F}_2^k$ represent such terms to indicate whether $s'\mathbf{Y}_\kappa$ is used if $[\mathbf{b}]_\kappa = 1$ otherwise $[\mathbf{b}]_\kappa = 0$. For example, the Hamiltonian C_5 in (8) needs to account for 2^3-terms of $c'^3\mathbf{Z}_1\mathbf{Z}_2\mathbf{Z}_4$, $c'^2 s'\mathbf{Z}_1\mathbf{Z}_2\mathbf{Y}_4$, …, $s'^3\mathbf{Y}_1\mathbf{Y}_2\mathbf{Y}_4$ by associative binary vector of $\mathbf{b} = [0, 0, 0, 0]$, $[0, 0, 0, 1]$, …, $[1, 1, 0, 1]$, respectively. Letting ϖ be the weight of binary vector $\mathbf{b}$, the cost expectation will be proportional to $(-\jmath s')^\varpi c'^{(d_\nu^{\mathrm{c}} - \varpi)}$ due to $\mathbf{Z}\mathbf{Y} = -\jmath\mathbf{X}$ and $\langle + | \mathbf{X} | + \rangle = 1$.

Next, we consider cost operator $U(C, \gamma_1)$ on each decomposed Pauli terms $\mathbf{W}^{\mathbf{b}}$. Selecting only non-commutable cost Hamiltonians $C^{\mathbf{b}}$, we can write

$$U(C, \gamma_1)^\dagger \mathbf{W}^{\mathbf{b}} U(C, \gamma_1) = U(C^{\mathbf{b}}, 2\gamma_1)^\dagger \mathbf{W}^{\mathbf{b}}. \tag{12}$$

We refer to the number of such non-commutable Hamiltonians as the rank ρ. It can be obtained by selecting columns of $\mathbf{G}$ having non-zero element after modulo-2 product of $\mathbf{b}$. For example, $\mathbf{b} = [1, 1, 0, 1]$ corresponding to $\mathbf{W}^{\mathbf{b}} = s'^3\mathbf{Y}_1\mathbf{Y}_2\mathbf{Y}_4$ yields $\mathbf{b}\mathbf{G} = [1, 1, 0, 1, 1, 0, 0]$, and thus non-commutable submatrix is $\mathbf{G}^{\mathbf{b}} = [\mathbf{G}]_{:,\{1,2,4,5\}}$ whose rank is $\rho = 4$.

We then consider binary representation of combinatorials in

$$U(C^{\mathbf{b}}, 2\gamma_1)^\dagger = \prod_\nu^\rho \mathrm{e}^{2\jmath\gamma_1 C_\nu} = \prod_\nu^\rho \big(c\mathbf{I} + \jmath s \prod_\kappa \mathbf{Z}_\kappa\big), \tag{13}$$

where $c = \cos(2(-1)^y\gamma_1)$ and $s = \sin(2(-1)^y\gamma_1)$. Specifically, letting $\mathbf{a} \in \mathbb{F}_2^\rho$ indicate binary choice of either $c\mathbf{I}$ or $\jmath s\prod \mathbf{Z}_\kappa$, the above ρ-ary products of binary additions can be expressed by 2^ρ-ary sums of ρ-ary multiplications. For example, $(c\mathbf{I})(c\mathbf{I})(\jmath s\mathbf{Z}_1\mathbf{Z}_2\mathbf{Z}_4)(\jmath s\mathbf{Z}_1\mathbf{Z}_3\mathbf{Z}_4) = -s^2c^2\mathbf{Z}_2\mathbf{Z}_3$ corresponds to $\mathbf{a} = [0, 0, 1, 1]$ for sub-generator $[\mathbf{G}]_{:,\{1,2,4,5\}}$. Letting ω be the weight of binary vector $\mathbf{a}$, the cost function

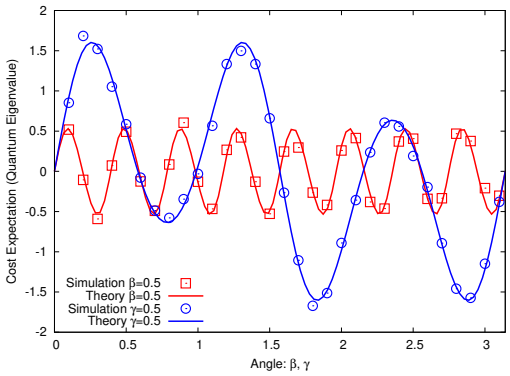


Fig. 1: Cost expectation $F_1(\gamma_1, \beta_1)$ for level-1 QAOA decoding of $[16, 5]$ systematic Reed–Muller codes.

will be proportional to $(\jmath s)^\omega c^{(\rho-\omega)}$ if it is non-commutable to cost Hamiltonian associated with $\mathbf{b}$. Specifically, such $\mathbf{a}$ must satisfy $\mathbf{b}^{\mathrm{T}} = \mathbf{G}^{\mathbf{b}}\mathbf{a}^{\mathrm{T}}$. There may exist plural of such binary vector pairs $\mathbf{a}$ and $\mathbf{b}$. Define $A_\nu^{\mathbf{a},\mathbf{b}}$ as the number of such pairs.

In consequence, the cost expectation can be obtained by counting the number of combinatorials subject to $\mathbf{b}^{\mathrm{T}} = \mathbf{G}^{\mathbf{b}}\mathbf{a}^{\mathrm{T}}$ for each column ν, as follows:

$$F_1(\gamma_1, \beta_1) = \sum_{\nu=1}^{n}(1-2y_\nu) \sum_{\mathbf{b}\in\mathbb{F}_2^k} \sum_{\mathbf{a}\in\mathbb{F}_2^\rho\,:\,\mathbf{b}^{\mathrm{T}}=\mathbf{G}^{\mathbf{b}}\mathbf{a}^{\mathrm{T}}} A_\nu^{\mathbf{a},\mathbf{b}} (\jmath s)^\omega c^{(\rho-\omega)} (-\jmath s')^{\varpi} c'^{(d_\nu^{\mathrm{c}}-\varpi)}. \quad (14)$$

B. Numerical Validation in Quantum Simulations

Using the above-described method, we can systematically derive the analytic expression of $F_1(\gamma_1, \beta_1)$ for level-1 QAOA decoder of any arbitrary binary linear codes given its generator matrix. For instance, we obtain the following quantity for $[16, 5]$ systematic Reed–Muller codes:

$$F_1(\gamma_1, \beta_1) = \tfrac{1}{32}\sin(4\gamma_1)\sin(2\beta_1) \big(4(\cos(4\gamma_1) + \cos(12\gamma_1) + \cos(20\gamma_1) + \cos(24\gamma_1))\sin^4(2\beta_1) + 5(\cos(4\gamma_1) + \cos(12\gamma_1))(25 + 36\cos(4\beta_1) + 3\cos(8\beta_1))\big).$$

Fig. 1 shows the cost expectation $\langle C\rangle$ with sweeping angles of β_1 or γ_1, for the $[16, 5]$ systematic Reed–Muller code. For quantum simulations, we use qiskit to obtain the averaged cost function over 8192-shot measurements. It was verified that our theoretical analysis agrees well the simulation results.

C. Degree Distribution Optimization

Figs. 2(a)–(c) show landscape of analytic cost expectation $F_1(\gamma_1, \beta_1)$ for Hamming codes with different generator matrix having an average degree $\bar{d}^{\mathrm{c}}$ of 1.71, 1.86, and 2.29, respectively. It was shown that the quality of QAOA decoding highly depends on degree distributions even though those basis-transformed codes have identical Hamming weight spectrum.

Table I lists the theoretical cost functions derived by our method for eight different Hamming codes with an average degree from 1.71 to 2.71. From the analytical expression, we can obtain optimal angle parameters $\gamma_1^\star$ and $\beta_1^\star$ to maximize the cost expectation. It can be seen that the maximum cost expectation $F_1^\star$ tends to improve as the average degree decreases. For example, lowest-degree non-symmetric Hamming code achieves $F_1^\star = 2.409$ which is larger than naïve random sampling, i.e., $\langle C\rangle = 0$, whereas a smaller cost of $F_1^\star = 1.790$ is achieved by higher-degree systematic Hamming code. This trend suggests that low-density generator-matrix (LDGM) codes can be a good candidate for level-1 QAOA decoder. This is intuitive because there exist fewer qubit interactions in cost Hamiltonian operator $U(C, \gamma)$.

D. Higher-Level QAOA Decoder

Fig. 3 shows the cross-entropy loss [16] as a function of average degree for level-p QAOA decoding of Hamming codes. The angle parameters were optimized by VQE employing NM method. One can see that higher-level QAOA offers significant gain in decoding accuracy, approaching to error-free decision, i.e., zero cross-entropy. Interestingly, systematic Hamming code achieves the best performance for $p \geq 2$ unlike level-1 QAOA. A non-heuristic design of generator matrix for high-level QAOA is an open question to pursue in the future.

E. Success Rate of ML Decision

Fig. 4 plots the accumulated success probability that QAOA measurement gives the ML decision. For real quantum processor, we use *ibmq_14_melbourne*. The success rate increases with the number of quantum shots (trial measurements). Although the real quantum processor has a degraded success rate, it is still much better than naïve random decision.

V. Conclusion

We proposed to make use of QAOA algorithm for classical channel decoding. Theoretical analysis was investigated and insight to optimize generator matrix was provided by discussing degree distributions. Since this is the first proof-of-principle study, there remain many challengs including integration with quantum error correction, proving quantum advantage over classical methods, consideration of quantum coupling maps, how to design long codes, and how to deal with soft information for high-level QAOA as future work.

References

[1] E. Berlekamp, R. McEliece, and H. Van Tilborg, "On the inherent intractability of certain coding problems (corresp.)," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 384–386, 1978.

[2] A. R. Calderbank and P. W. Shor, "Good quantum error-correcting codes exist," *Physical Review A*, vol. 54, no. 2, p. 1098, 1996.

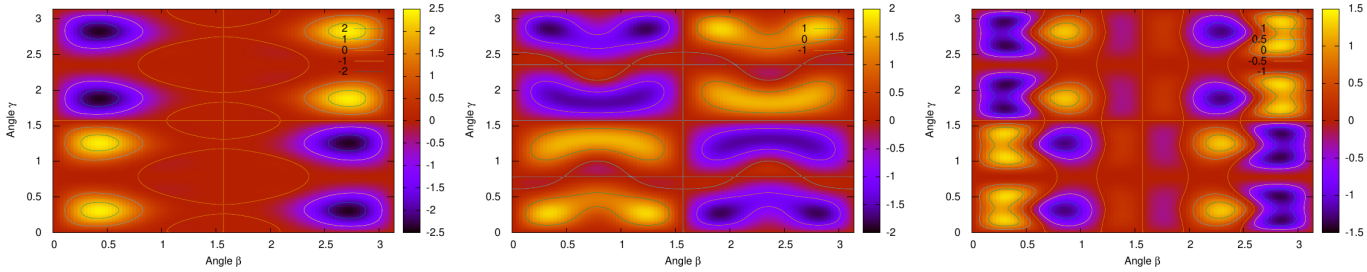
[3] A. R. Calderbank, E. M. Rains, P. Shor, and N. J. Sloane, "Quantum error correction via codes over gf (4)," *IEEE Trans. Inf. Theory*, vol. 44, no. 4, pp. 1369–1387, 1998.

[4] D. Poulin, J.-P. Tillich, and H. Ollivier, "Quantum serial turbo codes," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2776–2798, 2009.

[5] D. J. MacKay, G. Mitchison, and P. L. McFadden, "Sparse-graph codes for quantum error correction," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2315–2330, 2004.

[6] M. M. Wilde and S. Guha, "Polar codes for classical-quantum channels." *IEEE Trans. Inf. Theory*, vol. 59, no. 2, pp. 1175–1187, 2013.

[7] Z. Babar, S. X. Ng, and L. Hanzo, "EXIT-chart-aided near-capacity quantum turbo code design," *IEEE Trans. Veh. Technol.*, vol. 64, no. 3, pp. 866–875, 2015.



(a) $\bar{d}^{\mathrm{c}} = 1.71$: $F_1^\star = 2.409$ (b) $\bar{d}^{\mathrm{c}} = 1.86$: $F_1^\star = 1.790$ (c) $\bar{d}^{\mathrm{c}} = 2.29$: $F_1^\star = 1.367$

Fig. 2: Landscape of cost expectation $F_1(\gamma_1, \beta_1)$ for level-1 QAOA decoding of $[7, 4]$ Hamming codes.

TABLE I: Theoretical Expression of Level-1 QAOA Cost Expectation for Decoding $[7, 4]$ Hamming Codes

$\bar{d}^{\mathrm{c}}$	$\mathbf{P}$	$F_1(\gamma_1, \beta_1)$	$F_1^\star$	$\beta_1^\star$	$\gamma_1^\star$
1.71	$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\1&0&0&1\end{bmatrix}$	$3sc^2s'(1+c')^2 - sc^2s'^3(c^2-3s^2)(c^2-s^2)$	2.409	0.424	0.311
1.86	$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$	$-2sc(c^2-s^2)s'(1-3c'^2) + 3sc^2s'(1+2c'^2)$	1.790	0.345	0.277
2.00	$\begin{bmatrix}1&0&0&0\\1&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$	$sc^2s'(1+c'+c'^2+3c'^3) + 2sc(c^2-s^2)s'(1+c'^2+2c'^3) - sc^2(c^2-3s^2)(c^2-s^2)s'^3(1+c')$	1.606	0.329	0.239
2.14	$\begin{bmatrix}1&1&0&0\\0&1&0&0\\0&1&1&0\\0&0&0&1\end{bmatrix}$	$3sc^2s'(c'^2-s'^2) + 2sc(c^2-s^2)s'(1+5c'^2)$	1.562	0.785	1.820
2.29	$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\1&1&1&1\end{bmatrix}$	$-3sc^2s'(1-c'-3c'^2) + sc^2(c^2-3s^2)(c^2-s^2)s'(1+3c'+3c'^2)$	1.367	0.310	0.512
2.43	$\begin{bmatrix}1&1&1&1\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$	$sc^2s'c'(1+2c'+3c'^2) - 2sc(c^2-s^2)s'(1+c'-2c'^2-2c'^3) + sc^2(c^2-3s^2)(c^2-s^2)s'(1+3c'+2c'^2+c'^3)$	1.308	0.283	1.034
2.57	$\begin{bmatrix}1&1&1&1\\0&1&1&1\\0&0&1&0\\0&0&0&1\end{bmatrix}$	$-sc^2s'(1+2c'-3c'^2-3c'^3) - 2sc(c^2-s^2)s'(1-3c'^2-2c'^3) + sc^2(c^2-3s^2)(c^2-s^2)s'(1+2c'+3c'^2+c'^3)$	1.420	0.275	1.005
2.71	$\begin{bmatrix}1&1&1&1\\0&1&1&1\\1&0&1&0\\0&0&0&1\end{bmatrix}$	$-3sc^2s'^3(1+c') + 2sc(c^2-s^2)s'c'(1+2c')(1+c') + sc^2(c^2-3s^2)(c^2-s^2)(3+3c'+c'^2)s'c'$	1.671	0.506	1.846

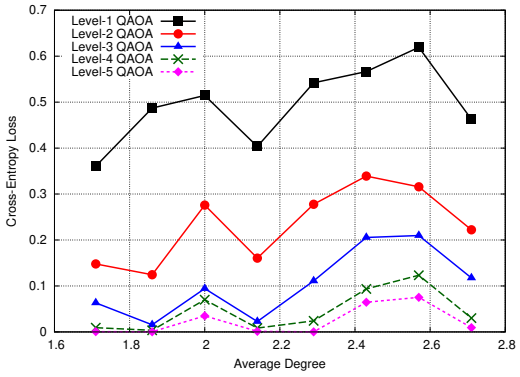


Fig. 3: Cross-entropy loss for level-p QAOA decoding of Hamming codes. Angles vectors are optimized by NM method.

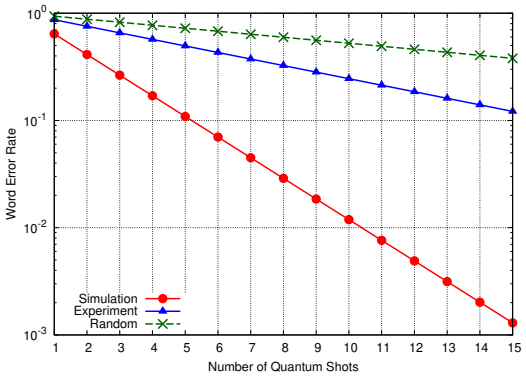


Fig. 4: ML-decision error rate of level-1 QAOA over quantum measurements for Hamming code at 1% BSC ($\bar{d}^{\mathrm{c}} = 1.71$).

[8] P. Botsinis, D. Alanis, Z. Babar, S. X. Ng, and L. Hanzo, "Iterative quantum-assisted multi-user detection for multi-carrier interleave division multiple access systems," *IEEE Trans. Commun.*, vol. 63, no. 10, pp. 3713–3727, 2015.

[9] P. Botsinis, D. Alanis, Z. Babar, H. Nguyen, D. Chandra, S. X. Ng, and L. Hanzo, "Quantum algorithms for wireless communications," *IEEE Communications Surveys & Tutorials*, 2018.

[10] P. Botsinis, D. Alanis, Z. Babar, S. X. Ng, and L. Hanzo, "Joint quantum-assisted channel estimation and data detection," *IEEE Access*, vol. 4, pp. 7658–7681, 2016.

[11] P. Botsinis, D. Alanis, Z. Babar, H. V. Nguyen, D. Chandra, S. X. Ng, and L. Hanzo, "Quantum-aided multi-user transmission in non-orthogonal multiple access systems," *IEEE Access*, vol. 4, pp. 7402–7424, 2016.

[12] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithm," *arXiv:1411.4028*, 2014.

[13] Patrick J. Coles et al, "Quantum algorithm implementations for beginners," *arXiv:1804.03719*, 2018.

[14] E. Farhi and A. W. Harrow, "Quantum supremacy through the quantum approximate optimization algorithm," *arXiv:1602.07674*, 2016.

[15] S. Hadfield, "Quantum algorithms for scientific computing and approximate optimization," *arXiv preprint arXiv:1805.03265*, May 2018.

[16] C. S. Calude and E. Calude, "The road to quantum computational supremacy," *arXiv preprint arXiv:1712.01356*, 2017.