A QUANTUM APPROACH FOR STOCHASTIC CONSTRAINED BINARY OPTIMIZATION

Sarthak Gupta and Vassilis Kekatos

Bradley Dept. of ECE, Virginia Tech, Blacksburg, VA 24061, USA; {gsarthak,kekatos}@vt.edu

ABSTRACT

Analytical and practical evidence indicates the advantage of quantum computing solutions over classical alternatives. Quantum-based heuristics relying on the variational quantum eigensolver (VQE) and the quantum approximate optimization algorithm (QAOA) have been shown numerically to generate high-quality solutions to hard combinatorial problems, yet incorporating constraints to such problems has been elusive. To this end, this work puts forth a quantum heuristic to cope with stochastic binary quadratically constrained quadratic programs (QCQP). Identifying the strength of quantum circuits to efficiently generate samples from probability distributions that are otherwise hard to sample from, the variational quantum circuit is trained to generate binary-valued vectors to approximately solve the aforesaid stochastic program. The method builds upon dual decomposition and entails solving a sequence of judiciously modified standard VQE tasks. Tests on several synthetic problem instances using a quantum simulator corroborate the near-optimality and feasibility of the method, and its potential to generate feasible solutions for the deterministic QCQP too.

Index Terms— QAOA, VQE, dual decomposition, quantum unconstrained binary optimization (QUBO).

1. INTRODUCTION

Quantum computers exhibit an innate ability to handle exponentially large computations in a parallel fashion yet with a strong probabilistic flavor. Quantum algorithms such as Shor's integer factorization, Grover's search, and the linear system solver of Harrow-Hassidim-Lloyd (HHL) can attain polynomial or even exponential speedups over the best known algorithms on classical computers [1]. Nonetheless, some of these quantum algorithms presume a large number of qubits on fault-tolerant quantum computers. In the *near-term intermediate scale* (NISQ) era, quantum computers are noisy and thus oftentimes limited in terms of number of gates and/or qubits. With such limitations in mind, *variational* quantum algorithms have been suggested as promising tools to practically showcase quantum advantage in the NISQ setup [2].

Variational quantum computers involve a sequence of parameterized gates. Their parameters are updated externally by a classical computer in a closed-loop fashion to steer the quantum state towards a desirable direction. The variational quantum eigensolver (VQE) used to provide high-quality solutions to combinatorial problems is a representative example. The Quantum Approximate Optimization Algorithm (QAOA) [3] is a special instance of VQE. In QAOA, not only the parameters but also the architecture of the quantum circuit become problem-dependent. The quantum circuit trained by QAOA operates as a sampler to efficiently generate near-optimal solutions of binary quadratic problems (e.g., MAXCUT); see [4] for a summary of claims on QAOA.

While most VQE/QAOA schemes target unconstrained problems, dealing with constraints is crucial to several applications in machine learning, wireless communications, and financial (stock trading) optimization. Adding constraints to QAOA or adiabatic quantum computing [5] (the QAOA counterpart for non-gate-based quantum computers) has been pursued in two ways. One approach has been to convert the constrained problem into an unconstrained minimization of a Lagrangian-like function [6, 7]. However, the weights for constraint penalties can be safely selected only if constraints are expressed as Boolean functions or linear equalities. An alternative approach modifies the architecture of the quantum circuit (via the mixer Hamiltonian of QAOA) to confine quantum states on the subspace spanned by constraints [8, 9, 4, 10]. Nonetheless, constructing such 'driver' mixer Hamiltonians is again highly problem-dependent and often limited to equality constraints. Reference [11] develops a quantum adiabatic approach to tackle binary linearly-constrained quadratic programs. It targets the dual problem and interfaces the quantum computer with a branch-and-bound scheme ran classically. Reference [12] treats mixed-binary quadratic-plus-convex problems using the alternating direction method of multipliers (ADMM) to split binary and continuous variables into separate minimizations, solved by QAOA and classical convex optimizers respectively per ADMM iteration.

Relation to prior work. Addressing binary QCQPs by quantum heuristics has been largely unexplored to the authors' knowledge. We put forth a quantum-based heuristic to solve a stochastic binary QCQP. Harnessing the power of quantum circuits to sample from probability mass functions

(PMF) that are hard to sample classically, we devise a dual decomposition technique that solves a sequence of standard VQE tasks to systematically adjust Lagrangian multipliers. Numerical tests using quantum computer simulators provided by IBM evaluate this technique on randomly generated stochastic and deterministic binary QCQPs.

2. QUANTUM COMPUTING PRELIMINARIES

A quantum system consisting of n quantum bits (qubits) is described by an exponentially large state vector $|\mathbf{x}\rangle \in \mathbb{C}^N$ with $N = 2^n$ assuming the system is in a *pure state*. The Dirac notation $|\mathbf{x}\rangle$ named *ket* emphasizes that vector $\mathbf{x}$ is unit-norm or $\sum_{k=0}^{N-1} |x_k|^2 = 1$. If $\mathbf{e}_k$ is the k-th canonical vector of length N, we can write $|\mathbf{x}\rangle = \sum_{k=0}^{N-1} x_k |\mathbf{e}_k\rangle$. The vector $\mathbf{e}_k$ is oftentimes alternatively expressed as $|\mathbf{e}_k\rangle = |k\rangle$, where k is the binary representation of index k. For example, a system with $n = 2$ qubits has a state in $\mathbb{C}^4$, which is spanned by canonical vectors $\{\mathbf{e}_k\}_{k=0}^3$ and $\mathbf{e}_0 = [1\ 0\ 0\ 0]^\top = |00\rangle$. Vector $|\mathbf{x}\rangle$ provides a statistical characterization for the quantum state: If we measure the quantum system output, its qubits will be in configuration $|k\rangle$ with probability $|x_k|^2$ for all k. Symbol $\langle\mathbf{x}|$ termed *bra* denotes the conjugate transpose of $|\mathbf{x}\rangle$, while the *braket* $\langle\mathbf{x}|\mathbf{y}\rangle$ denotes the inner product between states.

The fundamental operations we can perform on a quantum system is evolution and measurement. The former can be described by the application of a unitary $\mathbf{U}$ on state $|\mathbf{x}\rangle$ to produce state $|\mathbf{y}\rangle = \mathbf{U}|\mathbf{x}\rangle$. Although $\mathbf{U}$ is exponentially large, it is usually implemented efficiently using quantum gates. Among various types of measurements, we focus on *projective measurements*. A projective measurement is associated with a Hermitian matrix (named *observable*) and its eigenvalue decomposition $\mathbf{H} = \sum_{m=1}^M \lambda_m \mathbf{v}_m \mathbf{v}_m^H$. If such measurement is performed on $|\mathbf{x}\rangle$, outcome m is observed with probability $p_m := |\langle\mathbf{x}|\mathbf{v}_m\rangle|^2$. Define a random variable taking value λ_m when outcome m is observed. The expected value of this variable is $\langle\mathbf{x}|\mathbf{H}|\mathbf{x}\rangle = \sum_{m=1}^M p_m \lambda_m$. If $\mathbf{H}$ is diagonal, the measurement is on the *computational basis*. This is practically important because now $\mathbf{v}_m = \mathbf{e}_m$, outcome m relates to $|m\rangle$, and each qubit can be measured individually.

If quantum system i has been prepared in state $|\mathbf{x}_i\rangle$ for $i = 1, 2$, their joint state would be $|\mathbf{x}_1\rangle \otimes |\mathbf{x}_2\rangle$, where $\otimes$ is the Kronecker product. This is oftentimes represented as $|\mathbf{x}_1\rangle |\mathbf{x}_2\rangle$ or $|\mathbf{x}_1, \mathbf{x}_2\rangle$. The Kronecker product rule generalizes to the composition of n systems. For example, $|1\rangle |1\rangle |0\rangle = \mathbf{e}_1 \otimes \mathbf{e}_1 \otimes \mathbf{e}_0 = \mathbf{e}_6 = |110\rangle$, where the canonical vectors shown in the middle are in $\mathbb{R}^2$ and those at the end are in $\mathbb{R}^8$.

3. VARIATIONAL QUANTUM EIGENSOLVER (VQE)

VQE is a heuristic approach to find near-optimal solutions for combinatorial problems of the general form

$$\min_{\mathbf{b} \in \{0,1\}^n} f(\mathbf{b}). \tag{1}$$

A particular example of interest is the quadratic unconstrained binary optimization (QUBO) problem with

$$f(\mathbf{b}) = \mathbf{b}^\top \mathbf{A} \mathbf{b} + \mathbf{b}^\top \mathbf{c} + d \tag{2}$$

which is known to be NP-hard. For later developments, it is convenient to reformulate QUBO in terms of the *spin* $\{\pm 1\}$ variables through the transformation

$$s_i = 1 - 2b_i = (-1)^{b_i} \text{ for } i = 0, \ldots, n-1. \tag{3}$$

Collecting the spin variables in vector $\mathbf{s} = \mathbf{1} - 2\mathbf{b}$, the quadratic objective can be equivalently expressed as

$$f(\mathbf{b}) = \bar{f}(\mathbf{s}) = \mathbf{s}^\top \bar{\mathbf{A}} \mathbf{s} + \mathbf{s}^\top \bar{\mathbf{c}} + \bar{d} \tag{4}$$

where $\bar{\mathbf{A}} := \frac{1}{4}\mathbf{A}$; $\bar{\mathbf{c}} := -\frac{1}{2}(\mathbf{A}\mathbf{1} + \mathbf{c})$; and $\bar{d} := \frac{1}{4}\mathbf{1}^\top \mathbf{A}\mathbf{1} + \frac{1}{2}\mathbf{1}^\top \mathbf{c} + d$. We next explain how VQE samples high-quality solutions of (1) by solving an eigenvalue minimization task.

The VQE method falls under the family of *variational* quantum algorithms. The term *variational* pertains to the fact that the quantum circuit is not fixed, but parameterized by relatively few parameters collected in vector $\boldsymbol{\theta} \in \mathbb{R}^P$. These parameters are iteratively adjusted by classical computer in a closed-loop fashion so that the quantum system eventually reaches a desirable state. The process resembles the training of a neural network whose weights are updated by an optimization algorithm. Similarly to neural networks where the learner has to select an architecture (e.g., network depth/width and type of activations), the parameterized form (also termed *ansatz*) of the variational quantum circuit is specified *a priori*. We will be using a 2-local ansatz where single-qubit R_Y gates are applied to all qubits, followed by a full entanglement circuit, all repeated for 3 layers (iterations) [2].

Given $\boldsymbol{\theta}$ and driven by input $|0\rangle^n$, the quantum circuit produces at its output the quantum state $|\mathbf{x}(\boldsymbol{\theta})\rangle = \mathbf{U}(\boldsymbol{\theta}) |0\rangle^n$ for a unitary $N \times N$ matrix $\mathbf{U}(\boldsymbol{\theta})$. To simplify notation, we will oftentimes write $|\mathbf{x}\rangle$ in lieu of $|\mathbf{x}(\boldsymbol{\theta})\rangle$. Albeit $|\mathbf{x}\rangle \in \mathbb{C}^N$ is exponentially long, it can be easily generated by the quantum circuit though it cannot be read out of the circuit as a vector in a computationally efficient manner. Instead, it is relatively easy to sample from it. Every time we run the quantum circuit driven by $|0\rangle^n$, we will be observing one of the binary outputs $|k\rangle = |\mathbf{e}_k\rangle$ with probability $p_k := |x_k|^2$ for $k = 0, \ldots, N-1$. The quantum circuit thus serves as an efficient sampler from the exponentially large probability mass function (PMF) $\{p_k\}_{k=0}^{N-1}$.

To exploit this sampling property, we next relate the cost $f(\mathbf{b})$ with a so-termed *Hamiltonian* matrix $\mathbf{H}$ so that

$$\mathbf{H} |\mathbf{e}_k\rangle = f(|k\rangle) |\mathbf{e}_k\rangle \quad \text{for all } k. \tag{5}$$

Matrix $\mathbf{H}$ is apparently diagonal and carries all N function evaluations $f(\mathbf{e}_k)$ on its diagonal. Moreover, the canonical vectors $\mathbf{e}_k$ constitute the eigenvectors of $\mathbf{H}$, each with corresponding eigenvalue $f(|k\rangle)$. Therefore, the minimization

in (1) can be reformulated as the problem of finding the eigenvector corresponding to the minimum eigenvalue of $\mathbf{H}$

$$\min_{|\mathbf{x}\rangle} \langle \mathbf{x}| \mathbf{H} |\mathbf{x}\rangle. \tag{6}$$

As long as $|\mathbf{x}\rangle$ is allowed to take any of the values $\{\mathbf{e}_k\}_{k=0}^{N-1}$, the minimizer of (6) corresponds to the minimizer of (1). For example, if a quantum system has $n = 3$ qubits, its state would be $|\mathbf{x}\rangle \in \mathbb{C}^8$. Here $\mathbf{e}_k$'s are the columns of the identity matrix $\mathbf{I}_8$. If the minimizer of (6) is $|\mathbf{e}_5\rangle = |b_1 b_2 b_3\rangle = |101\rangle$, then the minimizer of (1) is $\mathbf{b} = [1\ 0\ 1]^\top$; and vice versa.

Although $\mathbf{H}$ is exponentially large, it can be implemented using only $\mathcal{O}(n^2)$ quantum gates since it can be expressed as

$$\mathbf{H} = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \bar{A}_{ij}\mathbf{Z}_i\mathbf{Z}_j + \sum_{i=0}^{n-1} \bar{c}_i\mathbf{Z}_i + \bar{d}\mathbf{I}_N \tag{7}$$

where the $N \times N$ Hermitian matrix $\mathbf{Z}_i$ is defined as

$$\mathbf{Z}_i = \mathbf{I}_2 \otimes \cdots \otimes \mathbf{Z} \otimes \cdots \otimes \mathbf{I}_2 \ \text{ with } \ \mathbf{Z} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

This is a Kronecker product involving $(n-1)$ identity matrices $\mathbf{I}_2$ and one *Pauli-Z* operator $\mathbf{Z}$ applied to the i-th qubit. Matrix $\mathbf{H}$ as defined in (7) is obviously diagonal. To establish (5), note first that $\mathbf{Z}|0\rangle = |0\rangle$ and $\mathbf{Z}|1\rangle = -|1\rangle$, or more compactly, $\mathbf{Z}|b\rangle = (-1)^b |b\rangle$. Consequently, when $\mathbf{Z}_i$ is applied to a state $|\mathbf{b}\rangle = |b_1 b_2 \cdots b_n\rangle$, the effect is $\mathbf{Z}_i |\mathbf{b}\rangle = (-1)^{b_i} |\mathbf{b}\rangle = s_i |\mathbf{b}\rangle$ from (3). Similarly, it also holds that $\mathbf{Z}_i\mathbf{Z}_j |\mathbf{b}\rangle = s_i s_j |\mathbf{b}\rangle$. Property (5) now follows immediately by postmultiplying (7) by any $|\mathbf{e}_k\rangle$ and using $f(\mathbf{b}) = \bar{f}(\mathbf{s})$.

If $|\mathbf{x}\rangle$ in (6) is restricted to set $\mathcal{E} := \{\mathbf{e}_k\}_{k=0}^{N-1}$, problem (6) is as hard as (1). VQE relaxes (6) to the set of all quantum states $|\mathbf{x}(\boldsymbol{\theta})\rangle$ that can be parameterized by the chosen ansatz and via $\boldsymbol{\theta}$. Problem (6) is then solved over $\boldsymbol{\theta}$ rather than $|\mathbf{x}\rangle$

$$\min_{\boldsymbol{\theta}} F(\boldsymbol{\theta}) := \langle \mathbf{x}(\boldsymbol{\theta})|\mathbf{H}|\mathbf{x}(\boldsymbol{\theta})\rangle. \tag{8}$$

From the eigenvalue property (5), it follows $\langle \mathbf{e}_n| \mathbf{H} |\mathbf{e}_k\rangle = f(|k\rangle)$ for all k. How about $\langle \mathbf{x}| \mathbf{H} |\mathbf{x}\rangle$ for a general state $|\mathbf{x}\rangle$? Because $|\mathbf{x}\rangle = \sum_{k=0}^{N-1} x_k |\mathbf{e}_k\rangle$, it is easy to show that

$$\langle \mathbf{x}|\mathbf{H}|\mathbf{x}\rangle = \sum_{k=0}^{N-1} |x_k|^2 f(|k\rangle) = \sum_{k=0}^{N-1} p_k f(|k\rangle). \tag{9}$$

In other words, function $F(\boldsymbol{\theta})$ is the average of f under the PMF defined by $|\mathbf{x}\rangle$. For instance, the random outcome $|k\rangle = |101\rangle$ occurring with probability $|x_5|^2$ is assigned to the random variable f taking the value $f([1\ 0\ 1]^\top)$. Hence, function $F(\boldsymbol{\theta})$ is really an expectation (an *observable* in the quantum computation parlance) of function $f(\mathbf{b})$ when $\mathbf{b}$ is drawn from the PMF $\{|x_k(\theta)|^2\}_{k=0}^{N-1}$. Ideally, the global minimizer $\boldsymbol{\theta}$ of (8) defines a PMF via $|\mathbf{x}(\boldsymbol{\theta})\rangle$ that samples with non-zero probability only the canonical vectors $|\mathbf{e}_k\rangle$ associated with the smallest eigenvalue of $\mathbf{H}$.

Problem (8) is solved in a hybrid fashion: The quantum computer samples from $|\mathbf{x}(\boldsymbol{\theta})\rangle$ and estimates $F(\boldsymbol{\theta})$ and possibly its gradient $\nabla_{\boldsymbol{\theta}} F$. A classical computer uses the previous information and iteratively updates $\boldsymbol{\theta}$ based on a zero- or first-order optimization algorithm, such as gradient descent or Bayesian optimization. As with training neural networks, $F(\boldsymbol{\theta})$ is nonconvex due to the form of the ansatz. Moreover, the ensemble statistic $F(\boldsymbol{\theta})$ cannot be computed exactly, but estimated as the sample average $\hat{F}(\boldsymbol{\theta}) := \sum_{r=1}^R f(\mathbf{b}_r)/R$ over R runs, where $\mathbf{b}_r$ is the quantum output after run r.

4. CONSTRAINED VQE

As discussed earlier, VQE provides a successful heuristic for solving QUBO through the variational formulation of (8). Can VQE be generalized to deal with a binary QCQP of the ensuing form?

$$\min_{\mathbf{b} \in \{0,1\}^n} f_0(\mathbf{b}) \tag{10}$$
$$\text{s.to } f_m(\mathbf{b}) \le 0, \quad m = 1:M.$$

Here $f_m(\mathbf{b}) := \mathbf{b}^\top \mathbf{A}_m \mathbf{b} + \mathbf{b}^\top \mathbf{c}_m + d_m$ for $m = 0, \ldots, M$. Solving such problems is also known to be NP-hard. Providing a quantum heuristic to directly deal with (10) seems to be challenging. To this end, we relax expectations and aim at designing a quantum state $|\mathbf{x}\rangle$ from which we can draw binary-valued $\mathbf{b}$ that solve the *stochastic* binary QCQP:

$$\min_{|\mathbf{x}\rangle} \mathbb{E}_\mathbf{x}[f_0(\mathbf{b})] \tag{11}$$
$$\text{s.to } \mathbb{E}_\mathbf{x}[f_m(\mathbf{b})] \le 0, \quad m = 1:M.$$

As in the unconstrained setup, rather than minimizing over $|\mathbf{x}\rangle$, we propose optimizing over a PMF parameterized by $\boldsymbol{\theta}$ and captured by quantum state $|\mathbf{x}(\boldsymbol{\theta})\rangle$. Specifically, we suggest solving the constrained minimization

$$\min_{\boldsymbol{\theta}} F_0(\boldsymbol{\theta}) \tag{12}$$
$$\text{s.to } F_m(\boldsymbol{\theta}) \le 0: \quad \lambda_m, \quad m = 1:M$$

where each observable $F_m(\boldsymbol{\theta}) := \langle \mathbf{x}(\boldsymbol{\theta})|\mathbf{H}_m|\mathbf{x}(\boldsymbol{\theta})\rangle$ depends on the Hamiltonian $\mathbf{H}_m$ defined similar to $\mathbf{H}$ in (7) for all m. Heed that problem (12) can be reformulated and solved as a linear program (LP) over the PMF of $\mathbf{b}$. Nonetheless, that requires evaluating $\{f_m(\mathbf{b})\}_{m=0}^M$ for all 2^n values of $\mathbf{b}$. Moreover, the optimization variable of this LP is the vector of PMF values that is exponentially large too. That is also the case with standard VQE/QAOA.

Contrary to (10), problem (12) is over the continuous variable $\boldsymbol{\theta}$, and thus, we can associate a dual variable λ_m for each constraint and define its Lagrangian function

$$L(\boldsymbol{\theta}; \boldsymbol{\lambda}) := F_0(\boldsymbol{\theta}) + \sum_{m=1}^M \lambda_m F_m(\boldsymbol{\theta}) \tag{13}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^M$ collects all dual variables. Problem (12) could be solved via *dual decomposition*, according to which $\boldsymbol{\lambda}$ is updated iteratively via a subgradient ascent step on L as

$$\lambda_m^{t+1} := \max\left\{\lambda_m^t + \mu_t F_m(\boldsymbol{\theta}^t), 0\right\}, \quad m = 1 : M \quad (14)$$

for a positive step size $\mu_t = \mu_0/(t + \alpha)$ with $\alpha > 0$, and $\boldsymbol{\theta}^t$ is a minimizer of the Lagrangian $L(\boldsymbol{\theta}; \boldsymbol{\lambda}^t)$ evaluated at $\boldsymbol{\lambda}^t$:

$$\boldsymbol{\theta}^t \in \arg\min_{\boldsymbol{\theta}} \langle \mathbf{x}(\boldsymbol{\theta}) | \mathbf{H}_0 + \sum_{m=1}^{M} \lambda_m^t \mathbf{H}_m | \mathbf{x}(\boldsymbol{\theta}) \rangle. \quad (15)$$

Problem (15) takes the QUBO form of (8), and is therefore amenable to standard VQE or even the celebrated QAOA approach. Under the latter, the ansatz takes a particular form that depends on the problem Hamiltonian $\mathbf{H}_0 + \sum_{m=1}^{M} \lambda_m^t \mathbf{H}_m$. Here, we used a problem-independent ansatz under the general VQE framework and leave QAOA for future work.

5. NUMERICAL TESTS

The novel solver for (12) was implemented in Python using the Qiskit library [13]. The *VQE* class in Qiskit was used to solve the minimization for the primal update (15). In addition to providing the ansatz described in Section 3, the VQE class was configured with the 'SLSQP' optimizer. The maximum number of iterations was set to $1,000$, and we used the `aer_simulator_statevector` quantum simulation backend. For the dual update in (14), constraint violations were measured over the observables $\mathbf{H}_m$ using the minimum eigenstate returned by VQE. The stopping criteria $\|\boldsymbol{\lambda}^t - \boldsymbol{\lambda}^{t-1}\|_2 \leq 1 \cdot 10^{-5}$ was utilized to ascertain the convergence of the dual updates (14).

To illustrate the application of the proposed strategy to solving the stochastic binary QCQP in (11), several 2-bit problem instances were sampled randomly by drawing the entries of $\{\mathbf{A}_0, \mathbf{c}_0, \mathbf{d}_0\}$ and $\{\mathbf{A}_1, \mathbf{c}_1, \mathbf{d}_1\}$ from the standard normal distribution, while ensuring the resulting problem was feasible. The VQE approach was compared against a linear program that finds a PMF solving (12); this was possible due to the small value of 2^n. For the two approaches, the obtained PMFs along with the associated dual variables are reported in Table 1 for 4 randomly sampled problem instances.

To study the scalability of the approach and to verify the compatibility of the solutions with the deterministic QCQP in (10), we also sampled 30 feasible 5-bit problem instances with three constraints each. The quadratic cost and constraint functions were generated as in the previous test. To avoid instances with non-binding constraints, the constants d_m in the constraint functions were manually adjusted so that at least one of the constraints was active and yielded a non-zero dual variable. From the sampled problems, it was found that the dual decomposition involving VQE was able to produce the optimal solutions for 28 out of the 30 problem instances

Table 1. Comparing the exact solution of (12) obtained via a linear program and the proposed quantum-based approach.

#	Found PMF		Dual	
	Quantum	LP	Quantum	LP
1	$[0.44, 0, 0.56, 0]$	$[0.44, 0, 0.56, 0]$	0.854	0.851
2	$[0.71, 0, 0.29, 0]$	$[0.70, 0, 0.30, 0]$	0.337	0.337
3	$[0, 0.80, 0, 0.20]$	$[0, 0.80, 0, 0.20]$	0.459	0.459
4	$[0, 0, 0.61, 0.39]$	$[0, 0, 0.60, 0.40]$	0.566	0.566

Fig. 1. Convergence of dual variables under dual updates (14) for a stochastic binary QCQP with $M = 3$ constraints.

tested, whereas infeasible binary candidates were obtained for the remaining 2 instances. Figure 1 illustrates the convergence of the dual variables for one of the problem instances, where all three constraints were found to be active.

6. CONCLUSIONS

A novel generalization of VQE to address the need for dealing with stochastic binary QCQPs has been developed. Leveraging dual decomposition, the approach entails solving a sequence of judiciously modified VQE tasks. Numerical tests demonstrate that upon convergence of the constrained VQE algorithm, the variational quantum circuit is able to sample from a stochastic policy to generate binary-valued vectors that minimize the binary QCQP and satisfy its constraints in expectation. Some of these samples seem to be feasible for the deterministic binary QCQP too. This novel heuristic sets the foundation for further developments towards constrained discrete optimization. We are currently exploring several exciting directions: *i)* Coupling this approach with QAOA rather than VQE; *ii)* skipping the nested optimization in (15) through a primal-dual decomposition alternative as in [14, 15]; and *iii)* dealing with mixed-binary setups.

7. REFERENCES

[1] Michael A. Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.

[2] Osvaldo Simeone, "An introduction to quantum machine learning for engineers," *Foundations and Trends in Signal Processing*, vol. 16, no. 1–2, pp. 1–223, 2022.

[3] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann, "A quantum approximate optimization algorithm applied to a bounded occurrence constraint problem," *arXiv: Quantum Physics*, 2014.

[4] S. Hadfield, Z. Wang, B. O'Gorman, E. G. Rieffel, D. Venturelli, and R. Biswas, "From the quantum approximate optimization algorithm to a quantum alternating operator ansatz," *Algorithms*, vol. 12, no. 2, pp. 34, 2019.

[5] C. C. McGeoch, *Adiabatic quantum computation and quantum annealing: Theory and practice*, vol. 5, Springer, Switzerland, 2014.

[6] A. Lucas, "Ising formulations of many NP problems," *Frontiers in Physics*, vol. 2, no. 5, pp. 1–15, 2014.

[7] M. Ohzeki, "Breaking limitation of quantum annealer in solving optimization problems under constraints," *Scientific reports*, vol. 10, no. 1, pp. 1–12, 2020.

[8] I. Hen and M. S. Sarandy, "Driver Hamiltonians for constrained optimization in quantum annealing," *Phys. Rev. A*, vol. 93, no. 6, pp. 062312, 2016.

[9] I. Hen and F. M. Spedalieri, "Quantum annealing for constrained optimization," *Phys. Rev. Appl.*, vol. 5, no. 63, pp. 034007, 2016.

[10] S. Hadfield, Z.Wang, E. G. Rieffel, B. O'Gorman, D. Venturelli, and R. Biswas, "Quantum approximate optimization with hard and soft constraints," in *ACM Intl. Workshop on Post Moore's Era Supercomputing*, New York, NY, 2017, pp. 15–21.

[11] Pooya Ronagh, Brad Woods, and Ehsan Iranmanesh, "Solving constrained quadratic binary problems via quantum adiabatic evolution," *Quantum Info. Comput.*, vol. 16, no. 11–12, pp. 1029–1047, Sept. 2016.

[12] Claudio Gambella and Andrea Simonetto, "Multi-block ADMM heuristics for mixed-binary optimization on classical and quantum computers," *IEEE Trans. on Quantum Engineering*, vol. 1, pp. 1–22, 10 2020.

[13] "Qiskit: An open-source framework for quantum computing," 2021.

[14] S. Gupta, S. Misra, D. Deka, and V. Kekatos, "DNN-based policies for stochastic AC-OPF," in *Proc. Power Syst. Comput. Conf.*, Porto, Portugal, June 2021, (to appear also in the Elsevier Electric Power Systems Research).

[15] S. Gupta, V. Kekatos, and M. Jin, "Controlling smart inverters using proxies: A chance-constrained DNN-based approach," *IEEE Trans. Smart Grid*, vol. 13, no. 2, pp. 1310–1321, Mar. 2022.