# Improved Success Probability with Greater Circuit Depth for the Quantum Approximate Optimization Algorithm

(article starts on next page)

# Improved Success Probability with Greater Circuit Depth for the Quantum Approximate Optimization Algorithm

Andreas Bengtsson[1], Pontus Vikstål[1], Christopher Warren,[1] Marika Svensson,[2,3] Xiu Gu[1],
Anton Frisk Kockum[1], Philip Krantz[1], Christian Križan[1], Daryoush Shiri,[1] Ida-Maria Svensson,[1]
Giovanna Tancredi,[1] Göran Johansson,[1] Per Delsing,[1] Giulia Ferrini,[1] and Jonas Bylander[1,*]

[1]*Microtechnology and Nanoscience, Chalmers University of Technology, Göteborg SE-412 96, Sweden*

[2]*Computer Science and Engineering, Chalmers University of Technology, Göteborg SE-412 96, Sweden*

[3]*Jeppesen, Göteborg SE-411 03, Sweden*

Present-day, noisy, small or intermediate-scale quantum processors—although far from fault tolerant—support the execution of heuristic quantum algorithms, which might enable a quantum advantage, for example, when applied to combinatorial optimization problems. On small-scale quantum processors, validations of such algorithms serve as important technology demonstrators. We implement the quantum approximate optimization algorithm on our hardware platform, consisting of two superconducting transmon qubits and one parametrically modulated coupler. We solve small instances of the NP (nondeterministic polynomial time)-complete exact-cover problem, with 96.6% success probability, by iterating the algorithm up to level two.

## I. INTRODUCTION

Quantum computing promises exponential computational speedup in a number of fields, such as cryptography, quantum simulation, and linear algebra [1]. Even though a large, fault-tolerant quantum computer is still many years away, impressive progress has been made over the last decade using superconducting circuits [2–4], leading to the noisy intermediate-scale quantum (NISQ) era [5]. It was predicted that NISQ devices should allow for "quantum supremacy" [6], that is, solving a problem that is intractable on a classical computer in a reasonable time. This was recently demonstrated on a 53-qubit processor by sampling the output distributions of random circuits [7].

Two of the most prominent NISQ algorithms are the quantum approximate optimization algorithm (QAOA) for combinatorial optimization problems [8–10] and the variational quantum eigensolver (VQE) for the calculation of molecular energies [11–13]. The QAOA is a heuristic algorithm that could bring a polynomial speedup to the solution of specific problems encoded in a quantum Hamiltonian [14,15]. Moreover, the QAOA should produce output distributions that cannot be efficiently calculated on a classical computer [16].

The QAOA is a hybrid algorithm, as it is executed on both a classical and a quantum computer. The quantum part consists of a circuit with $p$ levels, where better approximations to the solution of the encoded problem are generally achieved with higher $p$. In this work, we report on using our superconducting quantum processor to demonstrate the QAOA with up to $p = 2$, enabled by adequately high gate fidelities. We solve small toy instances of the NP (nondeterministic polynomial time) complete exact-cover problem with 96.6% success probability. For $p > 1$, the QAOA solution cannot be efficiently calculated on a classical computer, as the computational complexity scales doubly exponentially in $p$ [8].

Our interest in solving the exact-cover problem originates from its use in many real-world applications, for instance, the exact-cover problem can provide feasible solutions to airline planning problems such as tail assignment [17]. Currently, this is solved by well-developed optimization techniques in combination with heuristics. By leveraging heuristic quantum algorithms such as the QAOA, the current approach can be augmented and might provide high-quality solutions while reducing the running time. Applying the QAOA to instances of the exact-cover problem extracted from real-world data in the context of tail assignment has been numerically studied with 25 qubits, corresponding to 25 routes and 278 flights [18].

*bylander@chalmers.se

Other quantum algorithms for solving the exact-cover problem, specifically quantum annealing, have been considered in Refs. [19–21].

## II. QAOA

All NP-complete problems can be formulated in terms of finding the ground state of an Ising Hamiltonian [22]. The QAOA aims at finding this state by applying two noncommuting Hamiltonians, $\hat{B}$ and $\hat{C}$, in an alternating sequence (with length $p$) to an equal superposition state of $n$ qubits [visualized in Fig. 1(a)],

$$|\vec{\gamma}, \vec{\beta}\rangle = \prod_{i=1}^{p} [e^{-i\beta_i \hat{B}} e^{-i\gamma_i \hat{C}}] \left( \frac{|0\rangle + |1\rangle}{2} \right)^{\otimes n}, \quad (1)$$

where $\gamma_i$ and $\beta_i$ are (real) variational angles. The first Hamiltonian in the sequence is the Ising (cost) Hamiltonian specifying the problem,

$$\hat{C} = \sum_{i=1}^{n} h_i \hat{\sigma}_i^z + \sum_{i<j} J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z, \quad (2)$$

and the second is a transverse field (mixing) Hamiltonian defined by

$$\hat{B} = \sum_{i=1}^{n} \hat{\sigma}_i^x, \quad (3)$$

where $h_i$ and $J_{ij}$ are real coefficients, and the $\hat{\sigma}_i^{x(z)}$ are the Pauli X (Z) operators applied to the $i$th qubit.

The ground state of Eq. (2) corresponds to the lowest-energy state. We therefore define the energy expectation value of Eq. (1) as a cost function

$$F(\vec{\gamma}, \vec{\beta}) = \langle \vec{\gamma}, \vec{\beta} | \hat{C} | \vec{\gamma}, \vec{\beta} \rangle = \sum_{i=1}^{n} h_i \langle \hat{\sigma}_i^z \rangle + \sum_{i<j} J_{ij} \langle \hat{\sigma}_i^z \hat{\sigma}_j^z \rangle. \quad (4)$$

This cost function is evaluated by repeatedly preparing and measuring $|\vec{\gamma}, \vec{\beta}\rangle$ on a quantum processor. To find the state that minimizes Eq. (4), a classical optimizer is used to find the optimal variational angles $\vec{\gamma}^*, \vec{\beta}^*$. For a high enough $p$, $|\vec{\gamma}^*, \vec{\beta}^*\rangle$ is equal to the ground state of $\hat{C}$ and hence yields the answer to the optimization problem [8]. However, for algorithms executed on real hardware without error correction, noise will inevitably limit the circuit depth, implying that there is a trade-off between algorithmic errors (too low $p$) and gate errors (too high $p$). Note that, in order to find the solution to the optimization problem, it is not necessary for $|\vec{\gamma}^*, \vec{\beta}^*\rangle$ to be equal to the ground state: as long as the ground-state probability is high enough, the quantum processor can be used to generate a shortlist of

potential solutions that can be checked efficiently (in polynomial time) on a classical computer. For instance, even if the success probability of measuring the ground state is only 5%, we could measure 100 instances and still attain a probability greater than 99% of finding the correct state. Moreover, the angles $\vec{\gamma}^*, \vec{\beta}^*$ themselves are not interesting, as long as they yield the lowest-energy state. This gives some robustness against coherent gate errors, since any over or under rotations can be compensated for by a change in the variational angles [12].

We apply the QAOA to the exact-cover problem, which reads: given a set $X$ and several subsets $S_i$ containing parts of $X$, which combination of subsets include all elements of $X$ just once? Mathematically speaking, this combination of subsets should be disjoint, and their union should be $X$. This problem can be mapped onto an Ising Hamiltonian, where the number of spins equals the number of subsets, while the size of $X$ can be arbitrary.

Let us consider $n = 2$, for which the two-spin Ising Hamiltonian is

$$\hat{C} = h_1 \hat{\sigma}_1^z + h_2 \hat{\sigma}_2^z + J \hat{\sigma}_1^z \hat{\sigma}_2^z. \quad (5)$$

The exact-cover problem is mapped onto this Hamiltonian by choosing $h_i$ and $J$ as [23]

$$J > \min(c_1, c_2), \quad (6)$$
$$h_1 = J - 2c_1, \qquad h_2 = J - 2c_2,$$

where $c_i$ is the number of elements in subset $i$, and $J > 0$ if the two subsets share at least one element. We are free to choose $J$, as long as it fulfills the criterion in Eq. (6).
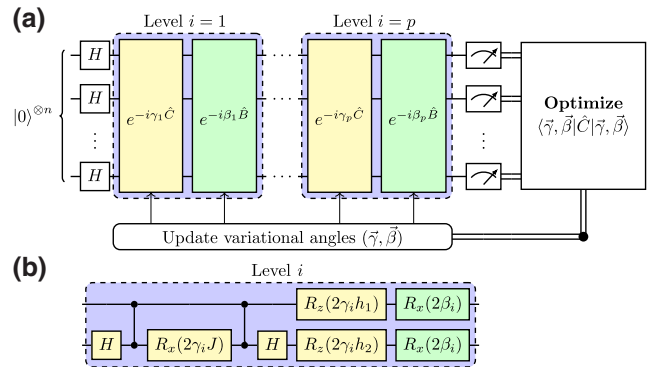


FIG. 1.   (a) The QAOA for a problem specified by the Ising Hamiltonian $\hat{C}$. An alternating sequence of two Hamiltonians ($\hat{C}$ and $\hat{B}$) is applied to an equal superposition of $n$ qubits. After measurement of the qubit states, a cost is calculated, which a classical optimization algorithm minimizes by varying the angles $\vec{\gamma}, \vec{\beta}$. (b) Our implementation of one QAOA level with $n = 2$ using controlled-phase, Hadamards ($H$), and single-qubit $x$ and $y$ rotations ($R_x$ and $R_y$). The background color of each gate identifies which part in (a) it implements.

TABLE I. The four different exact-cover problems available with two subsets, and their solutions and respective sets of coefficients in the Ising Hamiltonian $\hat{C} = h_1\hat{\sigma}_1^z + h_2\hat{\sigma}_1^z + J\hat{\sigma}_1^z\hat{\sigma}_2^z$.

| Problem | Subsets | $h_1$ | $h_2$ | $J$ | Solution |
|---|---|---|---|---|---|
| A | $S_1 = \{x_1, x_2\}$, $S_2 = \{x_1\}$ | $-1/2$ | 0 | $1/2$ | $|10\rangle$ |
| B | $S_1 = \{x_1, x_2\}$, $S_2 = \{\}$ | $-1$ | 0 | 0 | $|10\rangle$ or $|11\rangle$ |
| C | $S_1 = \{x_1\}$, $S_2 = \{x_2\}$ | $-1/2$ | $-1/2$ | 0 | $|11\rangle$ |
| D | $S_1 = \{x_1, x_2\}$, $S_2 = \{x_1, x_2\}$ | 0 | 0 | 1 | $|10\rangle$ or $|01\rangle$ |

For example, consider $X = \{x_1, x_2\}$ and two subsets $S_1 = \{x_1, x_2\}$ and $S_2 = \{x_1\}$. This gives $c_1 = 2$ and $c_2 = 1$, and we could choose $J = 2$, yielding $h_1 = -2$ and $h_2 = 0$. It is easy to check that the corresponding ground state is $|10\rangle$ (i.e., $S_1$ is the solution). Finally, we normalize $J$ and $h_i$ such that the Ising Hamiltonian has integer eigenvalues, allowing us to restrict $\gamma_i$ and $\beta_i$ to the interval $[0, \pi[$.

For two subsets, four different problems exist, which all yield different sets of $h_i$ and $J$. These are summarized in Table I. Problem A is the example given above; it is the most interesting, as the other problems are trivial. Problems B and C are trivial since they do not contain any qubit-qubit interaction ($J = 0$). Problem D is also trivial since both subsets are equal. Additionally, the ground states are degenerate for problems B and D.

## III. REALIZATION ON QUANTUM HARDWARE

We implement Eq. (1) on our quantum processor using the circuit in Fig. 1(b). The circuit can be somewhat compiled by simple identities (e.g., two Hadamard gates act as an identity gate). We stress that our implementation of the QAOA is scalable in that we do not use any exponentially costly precompilation (e.g., calculating the final circuit unitary and using Cartan decomposition to minimize the number of two-qubit gates).

Our quantum processor is fabricated using the same processes as in Ref. [24] and consists of two fixed-frequency transmon qubits with individual control and readout. Both qubits are coupled to a common frequency-tunable coupler used to mediate a controlled-phase (CZ) gate between the qubits. The CZ gate is realized by a full coherent oscillation between the $|11\rangle$ and $|02\rangle$ states. The interaction is achieved by parametrically modulating the resonant frequency of the coupler at a frequency close to the difference frequency between the $|0\rangle - |1\rangle$ and $|1\rangle - |2\rangle$ transitions of qubit 1 and 2, respectively [25,26]. We have benchmarked such a gate on the same device during the same cooldown to above 99%; however, the benchmark performed closest in time to the experiments presented here

showed a fidelity of 98.6%. These kinds of fidelity fluctuations might be related to fluctuations in the qubits' coherence times [24]. Single-qubit X rotations are driven by microwave pulses at the qubit transition frequencies with fidelities of 99.86% and 99.93% for the respective qubits. Z rotations are implemented in software as a shift in drive phase and thus have unity fidelity [27]. All the reported gate fidelities are measured by (interleaved) randomized benchmarking [28]. More experimental details, a measurement setup along with a device schematic, and benchmarking results are given in Appendices A and B.

## IV. APPLYING THE QAOA TO FOUR PROBLEMS

For $p = 1$, we apply a simple grid ($61 \times 61$) search of $\beta_1, \gamma_1 \in [0, \pi[$ while recording 5000 measurements of each qubit. From these, we calculate $\langle\sigma_i^z\rangle$, $\langle\sigma_1^z\sigma_2^z\rangle$, the cost function $F$, and the occupation probability for each of the four possible states, while accounting for the limited, but calibrated, readout fidelity (86% and 95% for the two qubits). By collecting sufficiently many samples, the statistical error on the estimated quantities can be made small.

The grid search allows us to explore the shape of the optimization landscape, which may bring important understanding in the difficulty of finding global minima for black-box optimizers. In Fig. 2, we show measured cost functions for the four problems in Table I. Because of the normalization of $h_i$ and $J$, the ground state for each problem corresponds to $F = -1$. In Fig. 2(a), the cost function for problem A never reaches below $-0.5$. To achieve costs approaching $-1$, additional levels ($p > 1$) are needed. Moreover, the existence of a local minimum
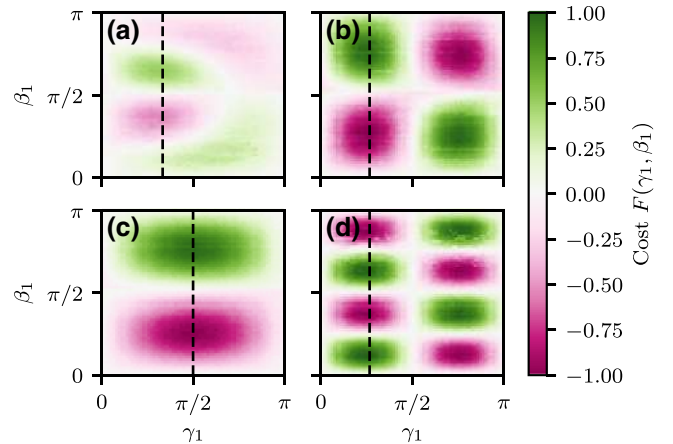


FIG. 2. Cost functions $F(\vec{\gamma}, \vec{\beta})$ for the QAOA applied to four instances of the exact-cover problem with $p = 1$ and $n = 2$. Panels (a)–(d) correspond to problems A–D in Table I. Each experimental data point is evaluated from the average of 5000 measurements on our quantum processor. The dashed lines indicate the positions of the linecuts in Fig. 3.

around $\gamma_1 \approx \beta_1 \approx 3\pi/4$ could cause difficulties for optimizers trying to find the global minimum. For problems B–D [Figs. 2(b)–2(d)], we see clear minima where $F \approx -1$, indicating that we have found the optimal variational angles $|\vec{\gamma}^*, \vec{\beta}^*\rangle$ corresponding to the ground state.

In Fig. 3, we take linecuts along the dashed lines in Fig. 2 and benchmark our measured cost functions and state probabilities against those of an ideal quantum computer without any noise. We see excellent agreement between measurement and theory: the measured positions of each minimum and maximum are aligned with those of the theory, consistent with low coherent-error rates. In addition, we observe excellent agreement between the absolute values at the minima and maxima, indicating low incoherent-error rates as well. Even with high gate fidelities, a high algorithmic fidelity is not guaranteed. Randomized benchmarking gives the average fidelity over a large number of random gates, which transforms any coherent errors into incoherent errors. For real quantum algorithm circuits, the gates are generally not random. Therefore, any coherent errors can quickly add up and yield algorithmic performance far lower than expected from randomized benchmarking fidelities alone [29,30].

To quantify the performance of the QAOA with $p = 1$, we compare the highest-probability state at the minima of $F$ with the solutions in Table I. Problem A [Fig. 3(a)] does not reach its ground state ($F \approx -0.5$); however, the
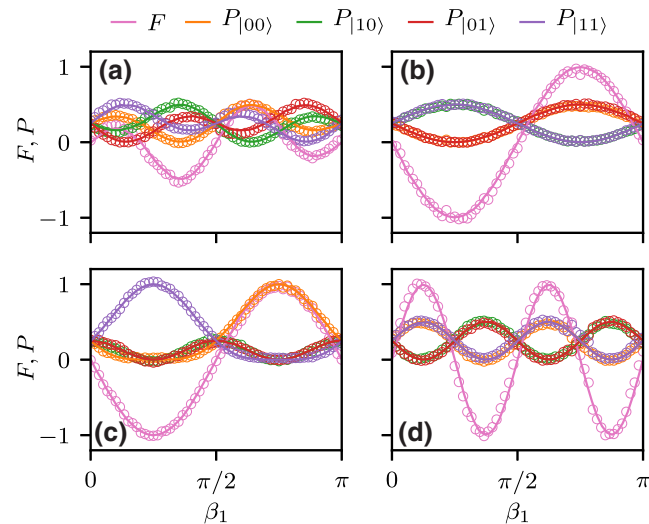


FIG. 3. A comparison between experiment (open circles) and theory (solid lines) for four exact-cover problems using the QAOA with $p = 1$. Each color (given at the top) corresponds to either a state probability or the value of the cost function $F$. The four panels (a)–(d) correspond to the four problems (A–D) in Table I. The linecuts are taken at the vertical dashed lines in Fig. 2. The theory curves are calculated assuming an ideal quantum processor, whereas each experimental data point is derived from the average of 5000 measurements on our quantum processor.

probability of measuring the correct state ($|10\rangle$) is approximately 50%, which is still better than random guessing. For problem C [Fig. 3(c)], we see that $F \approx -1$ does indeed correspond to a probability close to unity of measuring the ground state ($|11\rangle$). Problems B and D [Figs. 3(b) and 3(d)] have degenerate ground states, indicated by two state probabilities close to 50% each at $F \approx -1$.

## V. INCREASING THE SUCCESS PROBABILITY

To increase the success probability for problem A, we add an additional level ($p = 2$). For $p > 1$, a grid search to map out the full landscape becomes unfeasible due to the many parameters (equal to $2p$). Therefore, we instead use black-box optimizers to find the optimal variational angles. We try three different gradient-free optimizers: Bayesian optimization with Gaussian processes (BGPs), Nelder-Mead, and covariance matrix adaptation evolution strategy (CMA ES). We choose BGPs due to its ability to find global minima, Nelder-Mead due to it being common and simple, and CMA ES due to its favorable scaling with the number of optimization parameters.

We evaluate the optimizer performances by running 200 independent optimizations with random starting values ($\vec{\gamma}, \vec{\beta} \in [0, \pi[$) for each optimizer. For each set of variational angles, we repeat the circuit and measure 5000 samples to accurately estimate the expectation values. We set a threshold for convergence at $F < -0.95$ and count the number of converged optimization runs as well as the number of calls to the quantum processor (function calls) required to converge. We also record the success probability of measuring the problem solution ($P_{|10\rangle}$). The results are summarized in Table II.

We observe that the success probabilities after convergence are similar for all three optimizers. However, there is a difference in convergence probability, of which BGPs has the highest and Nelder-Mead has the lowest. The lower performance of Nelder-Mead is most likely due to its sensitivity to local minima, a well-known problem for most optimizers. In contrast, one of the strengths of Bayesian optimization is its ability to find global minima,

TABLE II. Comparison between different optimizers. We run the QAOA for problem A over 200 iterations with random starting parameters. We extract the convergence probability for reaching a cost below $-0.95$, the average number of function calls required to reach that level, and the highest achieved probability of measuring the problem solution ($P_{|10\rangle}$).

| Optimizer | Convergence (%) | Function calls | $P_{|10\rangle}$ (%) |
|---|---|---|---|
| BGPs | 61.5 | $44 \pm 16$ | 96.5 |
| Nelder-Mead | 20.0 | $38 \pm 13$ | 96.3 |
| CMA ES | 49.5 | $121 \pm 46$ | 96.6 |

which could explain why it performs better than Nelder-Mead and CMA ES. Additionally, Bayesian optimization is designed to handle optimization where the time of each function call is high (costly), such that the number of calls is kept low. However, for more optimization parameters (higher $p$), the performance of BGPs is generally decreased due to an increasing need for classical computation. CMA ES, on the other hand, excels when the number of parameters is high, and thus might be a good optimizer for the QAOA with tens or hundreds of parameters. Here, with just four parameters, CMA ES has a convergence
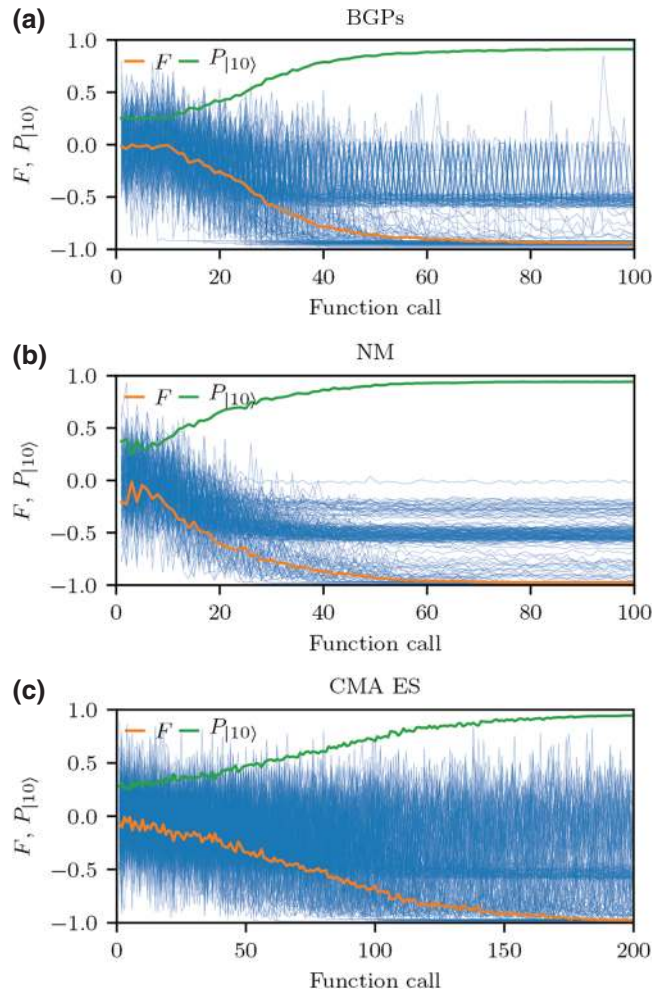


**(a)**

**(b)**

**(c)**

FIG. 4.   Optimization of variational angles for problem A using $p = 2$ iterations of the algorithm and three different black-box and gradient-free optimizers: (a) Bayesian optimization with Gaussian processes, (b) Nelder-Mead (NM), and (c) covariance matrix adaptation evolution strategy. We run the optimization 200 times with random starting parameters. Plotted as blue lines are the individual optimization trajectories for $F$, where each data point is the average of 5000 measurements. In orange and green are the costs ($F$) and success probabilities ($P_{|10\rangle}$) averaged over the converged runs.

probability similar to that of BGPs, although with a greater number of function calls on average.

To quantify the optimization further, we study the trajectories of each optimization run (Fig. 4). For each run, we plot the costs $F$. The trajectories for BGPs and Nelder-Mead [Figs. 4(a)–4(b)] corroborate the indications about local minima. We see groups of horizontal lines corresponding to different local minima, especially clear at $F \approx -0.55$ for both BGPs and Nelder-Mead. We also see that BGPs tries, and sometimes succeeds, to escape these local minima, which is one of the advantages of Bayesian optimization. In comparison, Nelder-Mead rarely gets out of a local minimum once it has found it. For the third optimizer, CMA ES [Fig. 4(c)], it is hard to draw any conclusions from the trajectories other than that the convergence is slower than for the other optimizers. However, we include the CMA ES trajectories for completeness. For each optimizer, we also plot the averaged (over all the converged) trajectories for $F$ and the probability of finding the solution state $P_{|10\rangle}$.

At the end of the optimization, the highest recorded probability of generating the correct state is 96.6%. The success probability is limited by imperfect gates (we have verified that an ideal quantum computer and $p = 2$ can achieve $P_{|10\rangle} = 1$). We compare our measured success probability to what we would expect from the randomized-benchmarking fidelities. The quantum circuit for $p = 2$ consists of 6 X, 4 Hadamard, 4 Z, and 3 CZ gates, which, when multiplied together with the fidelities for each gate, predicts a total fidelity of 96.3%, in good agreement with the measured fidelity considering experimental uncertainties (e.g., fluctuations in qubit coherence and gate fidelities). Note that $p = 3$ would not yield a higher success probability, since adding more gates would lower the total fidelity further (predicted to be 94.2%).

Finally, we examine histograms over the success probabilities at the end of each optimization run for the three different optimizers; see Fig. 5. Again, we observe that
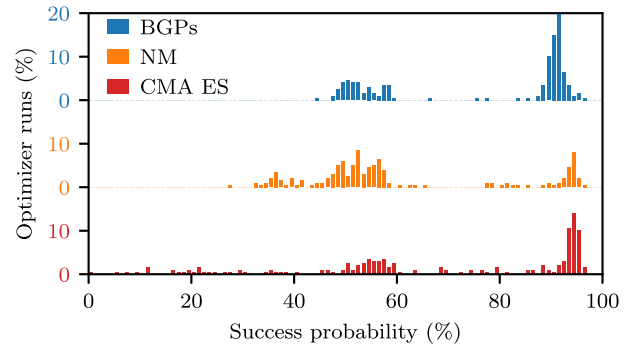


FIG. 5.   Histogram of the final success probabilities ($P_{|10\rangle}$) for three different optimizers on problem A using $p = 2$ iterations of the algorithm. Each optimizer is run 200 times. The bars are vertically offset for clarity.

BGPs has the most converged runs out of the three. We see clusters around 55% and 95% success probabilities for all three optimizers, possibly corresponding to one local and the global minima. For CMA ES, the success probabilities are more scattered, where some runs even have below 40% success. All in all, Bayesian optimization performs the best; however, further studies will be needed to determine which classical optimizer is the most suitable for variational quantum algorithms, such as the QAOA and VQE.

## VI. CONCLUSION

In conclusion, we implement the quantum approximate optimization algorithm with up to $p = 2$ levels. Using a superconducting quantum processor with state-of-the-art performance, we successfully optimize four instances of the exact-cover problem. For the nontrivial instance (problem A), we use $p = 2$ and black-box optimization to reach a success probability of 96.6% (up from 50% with $p = 1$), in good agreement with a prediction from our gate fidelities. Even if many more qubits are needed to solve problems that are intractable for classical computers, algorithmic performance serves as a critical quantum-processor benchmark since performance can be much lower than what individual gate fidelities predict. Although further experiments with larger devices are needed to explore whether the QAOA can have an advantage over classical algorithms, our results show that the QAOA can be used to solve the exact-cover problem.

## ACKNOWLEDGMENTS

## APPENDIX A: MEASUREMENT SETUP

The experimental measurement setup used here is a standard circuit quantum electrodynamics setup; see the schematic in Fig. 6. The quantum processor consists of two xmon-style transmon qubits coupled via a frequency-tunable anharmonic oscillator. The tunability is provided by two Josephson junctions in a superconducting quantum interference device (SQUID) configuration. The two qubits are capacitively coupled to individual control lines and quarter-wavelength resonators for readout. There is
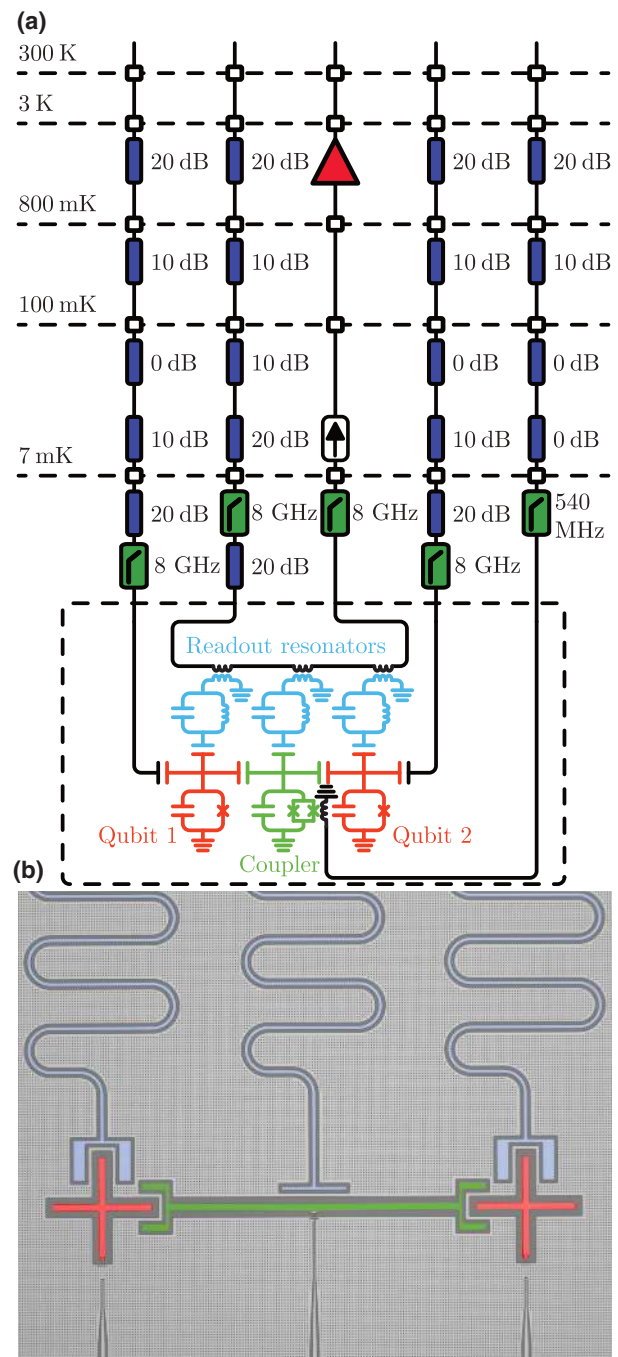


FIG. 6. (a) Cryogenic setup and electrical circuit of the quantum processor. All lines are attenuated and filtered to minimize the amount of noise reaching the qubits. The readout output contains cryogenic isolators and a high electron mobility transistor amplifier. (b) False-colored micrograph of the processor. The colors match the circuit elements in (a). The three waveguides at the bottom are for control over the qubits and the coupler.

also a readout resonator for the coupler, which is only used as a debugging tool (i.e., it is not involved during any algorithm execution). The SQUID for the tunable coupler is inductively coupled to a waveguide to allow for both static and fast modulation of the resonant frequency.

The processor is fabricated on a high-resistivity intrinsic silicon substrate. After initial chemical cleaning, an aluminium film is evaporated. All features except the Josephson junctions are patterned by direct-write laser lithography and etched with a warm mixture of acids. The Josephson junctions are patterned by electron-beam lithography, and evaporated from the same target as previously. A third lithography and evaporation step (with *in-situ* ion milling) is performed to connect the Josephson junctions to the rest of the circuit. Finally, the wafer is diced into individual dies and subsequently cleaned by a combination of wet and dry chemistry.

A die is then selected and packaged in a copper box and wire bonded to a palladium- and gold-plated printed circuit board with 16 nonmagnetic coaxial connectors. For the present device, we use five of these connectors, two for readout, two for single-qubit control, and one for control of the magnetic flux through the SQUID loop of the coupler. These are connected to filtered and attenuated coaxial lines leading up to room temperature. We point out that the dc current for the static flux bias is also provided through the coaxial line. Finally, the processor is attached to the mixing chamber of a Bluefors LD250 cryo-free dilution refrigerator. There, it is shielded from stray magnetic fields by two high permeability shields and two superconducting shields.

We perform multiplexed readout by using the Zurich Instruments UHFQA for generating and detecting the readout signals, together with a Rohde & Schwarz SGS100A continuous-wave signal generator and two Marki IQ mixers for up- and down-conversion. The single-qubit pulses are synthesized by the Zurich Instruments HDAWG and up-converted using Rohde & Schwarz SGS100A vector signal generators. The flux drive is generated directly by the HDAWG since the modulation frequency is within the bandwidth of the instrument. Finally, all instruments are controlled and orchestrated by the measurement and automation software Labber. Labber also does cost-function evaluations and calls external Python packages for the three different optimizers. All three optimizers are run using publicly available packages: Scikit-Optimize for BGPs, scipy for Nelder-Mead, and pycma for CMA ES.

## APPENDIX B: CHARACTERIZATION AND TUNE-UP

Initially, we perform basic spectroscopy and decoherence benchmarking of each qubit individually. This allows us to extract readout frequencies, qubit frequencies and anharmonicities, relaxation and dephasing times, and static couplings between qubit and resonator, as well as between qubit and coupler. The extracted parameters are found in Table III.

After the initial characterization, we tune up high-fidelity single-qubit gates. The drive pulses have cosine envelopes together with first-order derivative removal by

TABLE III. Device parameters. Readout-resonator frequency $f_R$ and qubit transition frequencies $f_{ij}$. Here $g$ is the coupling between qubit and resonator, and $j$ is the coupling between qubit and coupler; $T_1$ and $T_2^*$ are the relaxation and free induction decay times measured over 14 h; $F_{1q}$, $F_m$, and $F_{CZ}$ are the single-qubit, measurement, and CZ fidelities, respectively.

| Parameter | Qubit 1 | Qubit 2 |
|---|---|---|
| $f_R$ | 6.17 GHz | 6.04 GHz |
| $f_{01}$ | 3.82 GHz | 4.30 GHz |
| $f_{12} - f_{01}$ | −229 MHz | −225 MHz |
| $j$ | 29.1 MHz | 33.0 MHz |
| $g$ | 53.2 MHz | 56.9 MHz |
| $T_1$ | 77 $\mu$s | 55 $\mu$s |
| $T_2^*$ | 49 $\mu$s | 82 $\mu$s |
| $F_{1q}$ | 0.9986 | 0.9993 |
| $F_m$ | 0.86 | 0.95 |
| $F_{CZ}$ | 0.986 | |

adiabatic gate (DRAG) components to compensate for the qubit frequency shift due to the driving. Our rather long (50 ns) pulses makes leakage from $|1\rangle$ to $|2\rangle$ minimal even without DRAG. To find optimal pulse amplitudes and DRAG coefficients, we use error amplification by applying varying lengths of trains of $\pi$ pulses. Qubit drive frequencies are measured accurately by detuned Ramsey fringes.

Next, we calibrate our readout fidelities. By collecting raw voltages of the readout signals (as measured by the digitizer in the UHFQA), with and without a calibrated $\pi$ pulse applied to the qubit ($|0\rangle$ and $|1\rangle$ states, respectively), and as a function of readout frequency and amplitude, we can find the optimal readout parameters. Because of our rather low coupling strengths, we cannot achieve short readout times in this device. However, the QAOA does not require any measurement feedback, so a long readout time is not an issue as long as the time is shorter than the relaxation times of the qubits. Also, longer readout times give greater signal-to-noise ratios, which allow us to achieve high readout fidelities even in the absence of a quantum-limited amplifier. Here, the readout is 2.3 $\mu$s long, well below our relaxation times (several tens of microseconds). After finding the optimal readout parameters, a voltage threshold is used to differentiate between $|0\rangle$ or $|1\rangle$ of the measured qubit.

To accurately extract state probabilities in the presence of limited readout fidelity, we collect statistics of the measured qubit population as a function of qubit drive amplitude (Rabi oscillations). Since the measured population increases monotonically with the expected population, we can renormalize the populations, similarly to Ref. [31]. This calibration allows us to accurately measure the average quantities $\langle \sigma_i^z \rangle$, $\langle \sigma_i^1 \sigma_i^2 \rangle$ and state probabilities even in the presence of limited readout fidelities.

Our two-qubit gate of choice is the CZ. This interaction is induced by parametrically modulating the resonant frequency of the coupler at a frequency close to the difference of $|11\rangle$ and $|02\rangle$. For our device, this frequency is 255 MHz. However, due to the frequency modulation and the nonlinear relationship between flux and frequency, the transition frequencies are slightly lowered. This frequency shift will also induce deterministic phase shifts on the individual qubits, which we compensate for by applying $Z$ gates on both qubits after each CZ gate. We choose a static bias point and a modulation amplitude that yield a moderate effective coupling strength of 5 MHz between the two states. From here, we find the modulation frequency and time that yield a full oscillation between the $|11\rangle$ and $|02\rangle$ states. We then fine tune the frequency and time such that the controlled phase is $\pi$ and the leakage to $|02\rangle$ is minimal. Here, the final gate frequency and duration are 253 MHz and 271 ns.

We benchmark our single- and two-qubit fidelities using randomized benchmarking. A sequence of random gates drawn from the Clifford group is applied together with a final recovery gate that should take the system back to the ground state. The number of random gates is varied and the probability of measuring the ground state is recorded. In Fig. 7, we plot these probabilities for each qubit individually and for the two-qubit case. In the single-qubit case, it is important to note that it is done simultaneously for both qubits. Generally, the gate fidelities are higher if they are done in isolation. However, to reduce the total run time of algorithms, we usually run single-qubit gates in parallel. Therefore, simultaneous randomized benchmarking fidelities are more relevant metrics than isolated ones.
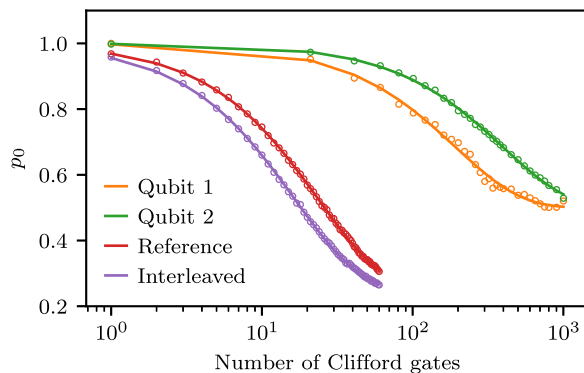


FIG. 7. Randomized benchmarking of single- and two-qubit gates. Plotted are the probability of measuring the ground state as a function of the number of Clifford gates applied. Circles are data, and lines are fits to extract the gate fidelity. For qubits 1 and 2, the extracted single-qubit fidelities (averaged over all possible single-qubit Clifford gates) are 0.9986 and 0.9993. For benchmarking of the two-qubit gate, we take a reference (random Clifford gates) and an interleaved (a CZ gate between each Clifford gate) trace to extract the CZ fidelity (0.986).

[1] A. Montanaro, Quantum algorithms: An overview, npj Quantum Inf. **2**, 15023 (2016).

[2] G. Wendin, Quantum information processing with superconducting circuits: A review, Rep. Prog. Phys. **80**, 106001 (2017).

[3] X. Gu, A. F. Kockum, A. Miranowicz, Y. X. Liu, and F. Nori, Microwave photonics with superconducting quantum circuits, Phys. Rep. **718–719**, 1 (2017).

[4] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, Superconducting qubits: Current state of play, Annu. Rev. Condens. Matter Phys. **11**, 369 (2020).

[5] J. Preskill, Quantum computing in the nisq era and beyond, Quantum **2**, 79 (2018).

[6] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Characterizing quantum supremacy in near-term devices, Nat. Phys. **14**, 595 (2018).

[7] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[8] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, arXiv preprint arXiv:1411.4028 (2014).

[9] J. Otterbach, R. Manenti, N. Alidoust, A. Bestwick, M. Block, B. Bloom, S. Caldwell, N. Didier, E. S. Fried, S. Hong, et al., Unsupervised machine learning on a hybrid quantum computer, arXiv preprint arXiv:1712.05771 (2017).

[10] G. Pagano, A. Bapat, P. Becker, K. Collins, A. De, P. Hess, H. Kaplan, A. Kyprianidis, W. Tan, C. Baldwin, et al., Quantum approximate optimization with a trapped-ion quantum simulator, arXiv preprint arXiv:1906.02700 (2019).

[11] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).

[12] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, *et al.*, Scalable Quantum Simulation of Molecular Energies, Phys. Rev. X **6**, 031007 (2016).

[13] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, Nature **549**, 242 (2017).

[14] Z. Jiang, E. G. Rieffel, and Z. Wang, Near-optimal quantum circuit for grover's unstructured search using a transverse field, Phys. Rev. A **95**, 062317 (2017).

[15] M. Y. Niu, S. Lu, and I. L. Chuang, Optimizing qaoa: Success probability and runtime dependence on circuit depth, arXiv preprint arXiv:1905.12134 (2019).

[16] E. Farhi and A. W. Harrow, Quantum supremacy through the quantum approximate optimization algorithm, arXiv preprint arXiv:1602.07674 (2016).

[17] M. Grönkvist, Ph.D. thesis, School Chalmers University of Technology and Göteborg University, 2005.

[18] P. Vikstål, M. Grönkvist, M. Svensson, M. Andersson, G. Johansson, and G. Ferrini, Applying the quantum

approximate optimization algorithm to the tail assignment problem, arXiv preprint arXiv:1912.10499 (2019).

[19] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, A quantum adiabatic evolution algorithm applied to random instances of an np-complete problem, Science **292**, 472 (2001).

[20] H. Wang and L.-A. Wu, Ultrafast adiabatic quantum algorithm for the np-complete exact cover problem, Sci. Rep. **6**, 22307 (2016).

[21] T. Graß, Quantum Annealing with Longitudinal Bias Fields, Phys. Rev. Lett. **123**, 120501 (2019).

[22] A. Lucas, Ising formulations of many np problems, Front. Phys. **2**, 5 (2014).

[23] V. Choi, Adiabatic quantum algorithms for the np-complete maximum-weight independent set, exact cover and 3sat problems, arXiv preprint (2010).

[24] J. J. Burnett, A. Bengtsson, M. Scigliuzzo, D. Niepce, M. Kudra, P. Delsing, and J. Bylander, Decoherence benchmarking of superconducting qubits, npj Quantum Inf. **5**, 9 (2019).

[25] D. C. McKay, S. Filipp, A. Mezzacapo, E. Magesan, J. M. Chow, and J. M. Gambetta, Universal Gate for Fixed-Frequency Qubits via a Tunable bus, Phys. Rev. Appl. **6**, 064007 (2016).

[26] S. Caldwell, N. Didier, C. Ryan, E. Sete, A. Hudson, P. Karalekas, R. Manenti, M. da Silva, R. Sinclair, E. Acala, *et al.*, Parametrically Activated Entangling Gates Using Transmon Qubits, Phys. Rev. Appl. **10**, 034050 (2018).

[27] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, Efficient z gates for quantum computing, Phys. Rev. A **96**, 022330 (2017).

[28] A. D. Córcoles, J. M. Gambetta, J. M. Chow, J. A. Smolin, M. Ware, J. Strand, B. L. T. Plourde, and M. Steffen, Process verification of two-qubit quantum gates by randomized benchmarking, Phys. Rev. A **87**, 030301(R) (2013).

[29] K. Michielsen, M. Nocon, D. Willsch, F. Jin, T. Lippert, and H. De Raedt, Benchmarking gate-based quantum computers, Comput. Phys. Commun. **220**, 44 (2017).

[30] M. Kjaergaard, M. Schwartz, A. Greene, G. Samach, A. Bengtsson, M. O'Keeffe, C. McNally, J. Braumüller, D. Kim, P. Krantz, *et al.*, A quantum instruction set implemented on a superconducting quantum processor, arXiv preprint (2020).

[31] J. M. Chow, L. DiCarlo, J. M. Gambetta, A. Nunnenkamp, L. S. Bishop, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Detecting highly entangled states with a joint qubit readout, Phys. Rev. A **81**, 062325 (2010).