npj | Quantum Information

www.nature.com/npjqi

ARTICLE OPEN

Check for updates

Analyzing the performance of variational quantum factoring on a superconducting quantum processor

Amir H. Karamlou[1,2,4]✉, William A. Simon[1,4], Amara Katabarwa[1], Travis L. Scholten[3], Borja Peropadre[1] and Yudong Cao[1]✉

In the near-term, hybrid quantum-classical algorithms hold great potential for outperforming classical approaches. Understanding how these two computing paradigms work in tandem is critical for identifying areas where such hybrid algorithms could provide a quantum advantage. In this work, we study a QAOA-based quantum optimization approach by implementing the Variational Quantum Factoring (VQF) algorithm. We execute experimental demonstrations using a superconducting quantum processor, and investigate the trade off between quantum resources (number of qubits and circuit depth) and the probability that a given biprime is successfully factored. In our experiments, the integers 1099551473989, 3127, and 6557 are factored with 3, 4, and 5 qubits, respectively, using a QAOA ansatz with up to 8 layers and we are able to identify the optimal number of circuit layers for a given instance to maximize success probability. Furthermore, we demonstrate the impact of different noise sources on the performance of QAOA, and reveal the coherent error caused by the residual *ZZ*-coupling between qubits as a dominant source of error in a near-term superconducting quantum processor.

npj Quantum Information (2021)7:156 ; https://doi.org/10.1038/s41534-021-00478-z

INTRODUCTION

While near-term quantum devices are approaching the limits of classical tractability[1,2], they are limited in the number of physical qubits, and can only execute finite-depth circuits with sufficient fidelity to be useful in applications[3]. Hybrid quantum-classical algorithms hold great promise for achieving a meaningful quantum advantage in the near-term[4–6] by combining the quantum resources offered by near-term devices with the computational power of classical processors. In a hybrid algorithm, a classical computer preprocesses a problem instance to expose the core components which capture its essential computational hardness, and is also in a form compatible with a quantum algorithm. The classical computer then utilizes a near-term quantum computer to finish solving the problem. The output of the quantum computer is a set of classical bitstrings which are post-processed classically. This interaction may be iterative: the classical computer sends communication and control signals to and from the quantum device, for example, by proposing new parameter values to be used in a parameterized quantum circuit.

Quantum algorithms developed for near-term quantum hardware will have to contend with several hardware restrictions, such as limited number of qubits, coupling connectivity, gate fidelities, and various sources of noise, which are all highly dependent on the particular quantum processor that the algorithm is being run on. Trade offs between different resources have recently been demonstrated in the context of running quantum circuits[7]. However, even as fault-tolerant quantum processors are built, classical computing will still be required to perform preprocessing and post-processing, as well as for quantum error-correction[8]. Therefore, understanding how quantum and classical computing resources can be leveraged together is imperative for extracting maximal utility from quantum computers.

An illustration of this hybrid scheme, in the context of this work, is shown in Fig. 1. In this work, we implement the variational quantum factoring (VQF) algorithm[9] on a superconducting quantum processor[10]. VQF is an algorithm for tackling an NP problem with a tunable trade off between classical and quantum computing resources, and is feasible for near-term devices. VQF uses classical preprocessing to reduce integer factoring to a combinatorial optimization problem which can be solved using the quantum approximate optimization algorithm (QAOA)[11].

VQF is a useful algorithm to study for several reasons. First, an inverse relationship exists between the amount of classical preprocessing and the number of required qubits needed for solving the problem, resulting in a tunable trade off between the amount of classical computation and qubits used. Second, the complexity of the quantum circuit acting on those qubits is tunable by adjusting the number of layers used in the QAOA ansatz. Finally, VQF has the advantage that its success is quickly verifiable on a classical computer by checking if the proposed factors multiply to the input biprime.

RESULTS

Variational quantum factoring

The VQF algorithm maps the problem of factoring into an optimization problem[12]. Given an n-bit biprime number

$$N = \sum_{k=0}^{n-1} 2^k N_k, \tag{1}$$

factoring involves finding the two prime factors p and q satisfying

[1]Zapata Computing, Boston, MA 02110, USA. [2]Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]IBM Quantum, IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, USA. [4]These authors contributed equally. Amir H. Karamlou, William A. Simon. ✉email: karamlou@mit.edu; yudong@zapatacomputing.com

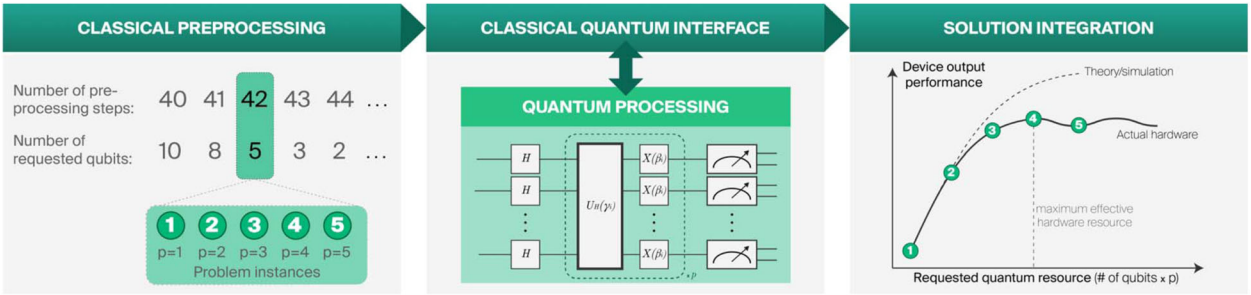

Fig. 1 Tunable resource tradeoffs with variational quantum factoring (VQF). Left: Given an integer N to be factored, varying amounts of classical preprocessing steps result in a different number of qubits required for the optimization problem, and defines a problem Hamiltonian. Middle: The optimization problem is solved on quantum hardware using the QAOA with p layers. Using classical optimization, the algorithm finds parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ that prepare a trial state which approximates the ground state of the problem Hamiltonian. Right: Classical post-processing combines the measurement results on the quantum device with classical preprocessing and evaluates the algorithm success.

$N = pq$, where

$$p = \sum_{k=0}^{n_p-1} 2^k p_k, \tag{2}$$

$$q = \sum_{k=0}^{n_q-1} 2^k q_k. \tag{3}$$

That is, p and q can be represented with n_q and n_p bits, respectively.

Factoring in this way can be thought of as the inverse problem to the longhand binary multiplication: the value of the i^{th} bit of the result, N_i, is known and the task is to solve for the bits of the prime factors $\{p_i\}$ and $\{q_i\}$. An explicit binary multiplication of p and q yields a series of equations that have to be satisfied by $\{p_i\}$ and $\{q_i\}$, along with carry bits $\{z_{i,j}\}$ which denote a bit carry from the i^{th} to the j^{th} position. Re-writing each equation in the series allows it to be associated to a particular *clause* in an optimization problem relating to the bit N_i:

$$C_i = N_i - \sum_{j=0}^{i} q_j p_{i-j} - \sum_{j=0}^{i} z_{i,j} + \sum_{j=1}^{n_p+n_q-1} 2^j z_{i,i+j}. \tag{4}$$

In order for each clause, C_i, to be 0, the values of $\{p_i\}$, $\{q_i\}$ and the carry bits $\{z_{i,j}\}$ in the clause must be correct. By satisfying the constraint for all clauses, the factors p and q can be retrieved.

Thus, the problem of factoring is reduced to a combinatorial optimization problem. Such problems can be solved using quantum computers by associating a qubit with each bit in the cost function. The number of qubits needed depends on the number of bits in the clauses. As discussed in[9], some number of classical preprocessing heuristics can be used to simplify the clauses (that is, assign values to some of the bits $\{p_i\}$ and $\{q_i\}$). As classical preprocessing removes variables from the optimization problem (by explicitly assigning bit values), the number of qubits needed to complete the solution to the problem is reduced. The new set of clauses will be denoted as $\{C_i'\}$ (see Supplementary Material for details of the classical preprocessing).

Each clause C_i' can be mapped to a term in an Ising Hamiltonian $\hat{C}_i$ by associating each bit value ($b_k \in \{p_i, q_i, z_{i,j}\}$) to a corresponding qubit operator:

$$b_k \mapsto \frac{1}{2}(1 - Z_k). \tag{5}$$

The solution to the factoring problem corresponds to finding the ground state of the Hamiltonian

$$\hat{H}_{C'} = \sum_{i=0}^{n} \hat{C}_i^2, \tag{6}$$

with a well defined ground state energy $E_0 = 0$. Because each clause C_i in equation (4) contains quadratic terms in the bits and the Hamiltonian $\hat{H}_{C'}$ is a sum of squares of $\hat{C}_i$, $\hat{H}_{C'}$ includes up to 4-local terms of the form $Z_i \otimes Z_j \otimes Z_k \otimes Z_l$. (An operator is k-local if it acts non-trivially on at most k qubits.) Much of the literature on quantum optimization has considered solving MAXCUT problems on d-regular graphs[11,13], which can be mapped to problems of finding the ground states of 2-local Hamiltonians. Other problems, such as MAX-3-LIN-2[14,15], can be directly mapped to ground state problems of 3-local Hamiltonians. The 4-local Ising Hamiltonian problems produced for VQF motivates a new entry to the classes of Ising Hamiltonian problems currently studied in the context of QAOA. While any k-local Hamiltonian can be reduced to a 2-local Hamiltonian by well-known techniques[16,17], the interactions between the qubits in the resulting 2-local Hamiltonian is not guaranteed to correspond to a d-regular graph, which again falls outside the scope of existing considerations in the literature.

The ground state of the Hamiltonian in equation (6) can be approximated on near-term, digital quantum computers using QAOA[11]. QAOA is a promising candidate for demonstrating quantum advantage through combinatorial optimization[18–20]. Each layer of a QAOA ansatz consists of two unitary operators, each with a tunable parameter. The first unitary operator is

$$U_H(\gamma) = e^{-i\gamma \hat{H}_{C'}} \text{ where } \gamma \in [0, 2\pi), \tag{7}$$

which applies an entangling phase according to the cost Hamiltonian $\hat{H}_{C'}$. The second unitary is the admixing operation

$$U_a(\beta) = \prod_i^n e^{-i\beta X_i} \text{ where } \beta \in [0, \pi), \tag{8}$$

which applies a single-qubit rotation around the X-axis with angle 2β to each qubit.

For a given number of layers p, the combination of $U_H(\gamma)$ and $U_a(\beta)$ are repeated sequentially with different parameters, generating the ansatz state

$$|\boldsymbol{\gamma}, \boldsymbol{\beta}\rangle = U_a(\beta_{p-1}) U_H(\gamma_{p-1}) \cdots U_a(\beta_0) U_H(\gamma_0) |+\rangle^{\otimes n}, \tag{9}$$

parameterized by $\boldsymbol{\gamma} = (\gamma_0, \cdots, \gamma_{p-1})$ and $\boldsymbol{\beta} = (\beta_0, \cdots, \beta_{p-1})$. The approximate ground state for the cost Hamiltonian can reached by tuning these $2p$ parameters. Given this ansatz state, a classical optimizer is used to find the optimal parameters $\boldsymbol{\gamma}_{\text{opt}}$ and $\boldsymbol{\beta}_{\text{opt}}$ that minimize the expected value of the cost Hamiltonian

$$E(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \langle \boldsymbol{\gamma}, \boldsymbol{\beta} | \hat{H}_{C'} | \boldsymbol{\gamma}, \boldsymbol{\beta} \rangle, \tag{10}$$

where for different parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ the value $E(\boldsymbol{\gamma}, \boldsymbol{\beta})$ is estimated on a quantum computer.

The circuit generating the state $|\boldsymbol{\gamma}_{\text{opt}}, \boldsymbol{\beta}_{\text{opt}}\rangle$ is then prepared on the quantum computer and measured. If the outcome of the measurement satisfies all the clauses, it can be mapped to the

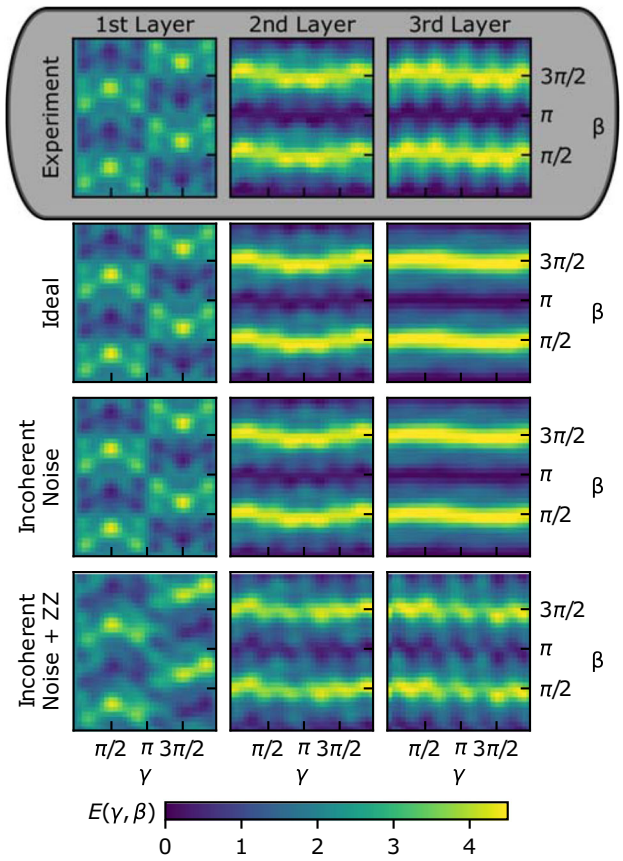

Fig. 2 QAOA optimization landscapes. For a fixed VQF instance on 3 qubits, we study the impact of noise (rows) on the optimization landscape for different number of QAOA layers in the ansatz (columns). To produce these landscapes, we use a resolution of $\pi/16$, resulting in 1024 grid points. For the second and third layers, the ansatz parameters for the previous layers are fixed to the optimal parameters found through ideal simulation.

remaining unsolved binary variables in $\{p_i\}$, $\{q_i\}$, and $\{z_{i,j}\}$, resulting in the prime factors p and q.

The success of VQF is measured by the probability that a measured bitstring encodes the correct factors. We define the set $M_s = \{m_j\}$ consisting of all bitstrings sampled from the quantum computer that satisfy all the clauses in the Hamiltonian, $\hat{H}_{C'}$. We can therefore define the success rate, $s(\boldsymbol{\gamma}, \boldsymbol{\beta})$, as the proportion of the bitstrings sampled from $|\boldsymbol{\gamma}, \boldsymbol{\beta}\rangle$ that satisfy all clauses in $\{C_i'\}$:

$$s(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \frac{|M_s|}{|M|}, \; M_s = \{m_j \in M | \sum C_i = 0\} \tag{11}$$

where $|M_s|$ is the number of samples satisfying all the clauses and $|M|$ is the total number of measurements sampled.

In order to successfully factor the input biprime, only one such satisfactory bitstring needs to be observed. However, if that bitstring occurs very rarely, then many repeated preparations and measurements of the trial state are necessary. Therefore, a higher value of $s(\boldsymbol{\gamma}, \boldsymbol{\beta})$ is generally preferred. $E(\boldsymbol{\gamma}, \boldsymbol{\beta}) \geq 0$, with the equality condition satisfied if $s(\boldsymbol{\gamma}, \boldsymbol{\beta}) = 1$.

Experimental analysis

In this work, we implement the VQF algorithm using the QAOA ansatz on the *ibmq_boeblingen* superconducting quantum processor (see Supplementary Material) to factor three biprime integers: 1099551473989, 3127, and 6557 which are classically preprocessed to instances with 3, 4, and 5 qubits respectively. The number of qubits required for each instance varies based the amount of classical preprocessing performed; 24 iterations of classical preprocessing were used for 1099551473989, 8 iterations for 3127, and 9 iterations for 6557.

In order to optimize the QAOA circuit and find $\boldsymbol{\gamma}_{\text{opt}}$ and $\boldsymbol{\beta}_{\text{opt}}$ we employ a layer-by-layer approach[21]. This approach has two phases: In the first phase, for each layer k, we evaluate a two-dimensional energy landscape by sweeping through different γ_k and β_k values while fixing the parameters for the previous layers (1 through $k-1$) to the optimal values found in previous iterations. In our experiments, for 1099551473989 we evaluate the energy for discrete values of γ_k and β_k ranging from 0 to 2π with a resolution of $\pi/6$, which yields 144 circuit evaluations for each layer (Fig. 2). For 3127 and 6557, we use a resolution of $2\pi/23$, which yields 529 circuit evaluations for each layer (see Supplementary Material). We use the degeneracy in the energy landscape as a result of evaluating β up to 2π instead of π as a feature to better understand the pattern of the energy landscape. We then select the optimal values for γ_k and β_k from this grid, and combine them with the optimal values from the previous layer. We increment k until it reaches the last layer, in which case the completion of the final round of the above steps marks the end of the first phase of the algorithm. In the second phase, we use the values $\{(\gamma_k, \beta_k)\}$ obtained from the first phase as the initial point for a gradient-based optimization using the L-BFGS-B method over all $2k$ parameters. Although there are alternative strategies for training QAOA circuits[21,22], this approach has been shown to require a grid resolution that scales polynomially with respect to the number of qubits required[9].

In order to measure the gradient of $E(\boldsymbol{\gamma}, \boldsymbol{\beta})$ with respect to the individual variational parameters, we use analytical circuit gradients using the parameter-shift rule[23]. We use the optimal parameters found for a circuit with p layers using this method to prepare and measure $|\boldsymbol{\gamma}_{\text{opt}}, \boldsymbol{\beta}_{\text{opt}}\rangle$ in order to calculate the algorithm success rate (Fig. 3).

The performance of our algorithm relies on the optimization over an energy surface spanned by the variational parameters γ and β for each layer. Consequently, by studying this energy surface, we can understand the performance of the optimizer in tuning the parameters of the ansatz state, which in turn impacts the success rate. We do so by visualizing the energy landscape for the k^{th} layer of the ansatz as a function of the two variational parameters, β_k and γ_k, fixing $(\gamma_0, .., \gamma_{k-1})$ and $(\beta_0, .., \beta_{k-1})$ to the optimal values obtained through ideal simulation. Comparing these landscapes between experimental results, ideal simulation, and noisy simulation provides us with a valuable insight into the performance of the algorithm on quantum hardware.

The main sources of incoherent noise present in the quantum processor are qubit relaxation and decoherence. In order to capture the effects of these sources of noise on our quantum circuit in simulation, we apply a relaxation channel with relaxation parameter ϵ_r and a dephasing channel with dephasing parameter ϵ_d after each gate, described by:

$$\mathcal{E}(\rho) = \sum_{m=1}^{3} E_m \rho E_m^{\dagger}, \tag{12}$$

with Kraus operators

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\epsilon_r-\epsilon_d} \end{pmatrix}, \quad E_2 = \begin{pmatrix} 1 & \sqrt{\epsilon_d} \\ 0 & 0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{\epsilon_r} \end{pmatrix}. \tag{13}$$

Furthermore, a dominant source of coherent error is a residual ZZ-coupling between the transmon qubits[24]. This interaction is caused by the coupling between the higher energy levels of the

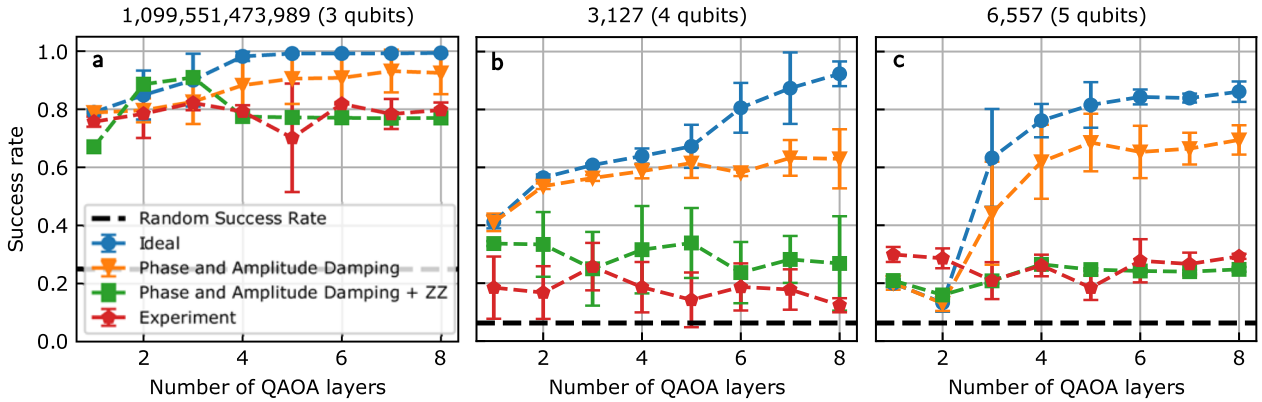

Fig. 3 VQF algorithm success rate. Success rates for running VQF on the integers (**a**) 1099551473989 (**b**) 3127, and (**c**) 6557 using 3, 4, and 5 qubits, respectively. We compare the results obtained from experiments (red line) to ideal simulation (blue line), simulation containing phase and amplitude damping noise (orange line), and the residual *ZZ*-coupling (green line). The amplitude damping rate ($T1 \approx 64\,\mu s$), dephasing rate ($T2 \approx 82\,\mu s$), and residual *ZZ*-coupling (see Supplementary Material) reflect the values measured on the quantum processor. We observe that our experimental performance is limited by the residual *ZZ*-coupling present in the quantum processor. The error bars indicate one standard deviation of uncertainty.

qubits, and is especially pronounced in transmons due to their weak anharmonicity. In our experiments, we measure the average residual *ZZ*-coupling strength to be $\xi/2\pi = 102 \pm 124$kHz (see Supplementary Material). While the additional *ZZ* rotation caused by this interaction when the qubits are idle between gates can be compensated by the tunable parameters $\boldsymbol{\gamma}$, this interaction has a severe impact on the two-qubit gates. In our experiments the Ising terms $e^{i\gamma ZZ} = \mathrm{CNOT}\circ I \otimes Z(\gamma)\circ\mathrm{CNOT}$ are realized using a single-qubit *Z* rotation conjugated by CNOT gates. Each CNOT is comprised of a *ZX* term generated by a two-pulse echo cross-resonance gate[25] and single-qubit rotations:

$$U_{\mathrm{CNOT}} = e^{-i\frac{\pi}{4}}[Z I]^{-1/2}[ZX]^{1/2}[IX]^{-1/2} \tag{14}$$

However, in the presence of residual *ZZ*-coupling, the axis of rotation of the cross-resonance (CR) gate is altered:

$$\mathrm{CR}(t) = e^{i\Gamma ZXt} \mapsto e^{i(\Gamma ZX + \Xi)t} \tag{15}$$

where Γ is the *ZX* coupling strength between the qubits as the result of the cross-resonance drive and Ξ is the overall effect of the residual *ZZ*-coupling between the qubits involved in the drive and their neighbors (see Supplementary Material). As a result, the two-qubit interactions are fundamentally altered, which leads to a source of error that cannot be corrected for through the variational parameters.

Figure 2 shows the energy landscapes from factoring 1099551473989 across 3 layers. Comparing the energy landscape produced in experiment with the ideal simulation and noisy simulations indicates that the energy landscapes produced in experiment is in close agreement with those produced by the ideal simulation and the simulation with incoherent noise for $p = 1$ and $p = 2$. However, for $p = 3$ the energy landscapes produced in experiment have considerable deviations from the ideal simulations due to coherent errors. The presence of incoherent sources of noise do not change the energy landscape pattern, and only reduce the contrast at every layer. On the other hand, the residual *ZZ*-coupling between qubits alters the pattern of the energy landscape. At lower depths, the impact of this noise source is minor, yet as we add more layers of the ansatz the effects accumulate, constructively interfere, and amplify due to the coherent nature of the error. To a first-order approximation, the coherent error contribution leads to changes in the weights of various frequency components of the optimization landscape (see Supplementary Material). In our energy landscapes these errors amplify the weights of high frequency components of the optimization landscape, leading to additional fluctuations compared to the ideal simulation.

We report the average success rate of VQF as a function of the number of layers in the ansatz for experiment and simulation (Fig. 3). As would be expected, in the ideal case the average success rate approaches the ideal value of 1 with increasing layers of the ansatz for each of the instances studied. Comparing experimental results to those obtained from ideal simulation shows a large discrepancy in the success rate. This discrepancy is not explained by the incoherent noise caused by qubit relaxation and dephasing; the coherent error caused by the residual *ZZ*-coupling is the dominant effect negatively impacting the success rate. Therefore, the performance of VQF, and other QAOA-based hybrid quantum-classical algorithms by extension, is limited by the residual *ZZ*-coupling between the qubits on transmon-based superconducting quantum processors.

We investigate the impact of different noise channels on the success rate scaling of VQF (Fig. 4). In the ideal scenario, with every added layer of QAOA, which increases the ansatz expressibility[26], the algorithm success rate increases. However, in the presence of phase damping, after an initial increase in the success rate we observe a plateau after which the success rate neither increases nor decreases as more layers of the ansatz circuit are added. We expect that this behavior is caused by a loss of quantum coherence at higher layers of circuit, which hinders the quantum interference required for the algorithm. In contrast, while undergoing amplitude damping, after an initial increase in the success rate, due to increased ansatz expressibility, we observe a steady decay in the algorithm's performance. This decay is a result of qubit relaxation back to the $|0\rangle$ state. When examining the effect of the coherent error induced by the residual *ZZ*-coupling, we see a more dramatic effect than that induced by either phase damping or amplitude damping. In the presence of the coherent error induced by the residual *ZZ*-coupling, we observe a plateau in the success rate for lower frequencies (10 and 25 kHz) and a decline in success rate for higher frequencies (50 and 100 kHz) within the first twenty layers of the ansatz.

While each of the sources of noise impact the algorithm's performance, the residual *ZZ*-coupling between the qubits in the quantum processor has the most dominant impact on the success rate; Our analysis indicates that this coherent error is the main source of error impeding the effectiveness of the algorithm. The accumulation of coherent noise from this source significantly alters the ansatz at higher layers, and cannot be corrected using the variational parameters. The results shown in Fig. 4 indicate that the performance of VQF can be significantly improved by

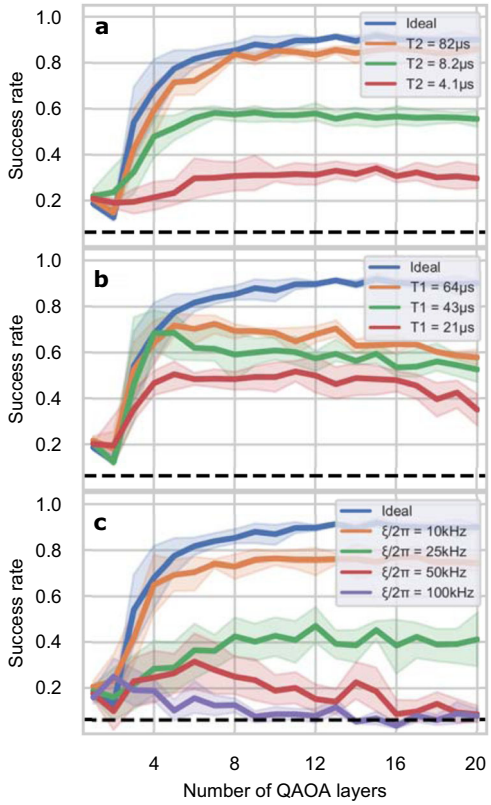

Fig. 4 Impact of different sources of noise on success rate when factoring 6557 using 5 qubits. a Phase damping, **b** amplitude damping, and **c** residual *ZZ*-coupling between qubits with a two-qubit gate time of $t_g = 315$ ns. The average T2, T1, and frequency of the residual *ZZ*-coupling for the given device are 82 μs, 64 μs, and 102 kHz respectively. The shaded regions indicate one standard deviation of uncertainty.

engineering *ZZ* suppression[27,28] and circuit compilation methods designed for mitigating coherent errors such as randomized compiling[29]. Since VQF is a non-trivial instance of QAOA we expect this performance analysis to also hold for other quantum algorithms based upon QAOA.

DISCUSSION

In this work, we analyzed the performance of VQF on a fixed-frequency transmon superconducting quantum processor and investigated several kinds of classical and quantum resource tradeoffs. We map the problem of factoring to an optimization problem and use variable amounts of classical preprocessing to adjust the number of qubits required. We then use the QAOA ansatz with a variable number of layers to find the solution to the optimization problem. We find that the success rate of the algorithm saturates as p increases, instead of decreasing to the accuracy of random guessing. While more layers increases the expressibility of the ansatz, at higher depths the circuit suffer from the impacts of noise.

Our analysis indicate that the residual *ZZ*-coupling between the transmon qubits significantly impacts performance. While relaxation and decoherence have an impact on the quantum coherence, especially at the deeper layers of QAOA, the effect of coherent noise can quickly accumulate and limit performance. By developing a noise model incorporating the coherent sources of noise, we are able to much more accurately predict our experimental results.

There have been various techniques[7,30] proposed for benchmarking the capability of a quantum device for performing arbitrary unitary operations. However, recent findings on experimental systems[3,31] suggest that existing benchmarking methods may not be effective for predicting the capability of a quantum device for specific applications. This has motivated recent works that focus on benchmarking the performance of a quantum device for applications such as generative modeling[32] and fermionic simulation[33]. Our study of VQF on a superconducting quantum processor has the potential to become another entry in the collection of such application-based benchmarks for quantum computers.

DATA AVAILABILITY

The raw data used for generating the plots in this paper are available upon request.

Received: 5 March 2021; Accepted: 12 August 2021;
Published online: 28 October 2021

REFERENCES

1. Arute, F. et al. Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).
2. Zhong, H.-S. et al. Quantum computational advantage using photons Downloaded from. *Science* (2020). https://doi.org/10.1126/science.abe8770
3. Kjaergaard, M. et al. Programming a quantum computer with quantum instructions. (2020). https://arxiv.org/abs/2001.08838
4. Google, AI. Quantum and Collaborators. Hartree-Fock on a superconducting qubit quantum computer. *Science* **369**, 1084–1089 (2020).
5. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 1–7 (2014).
6. Havlíček, V. et al. Supervised learning with quantum-enhanced feature spaces. *Nature* **567**, 209–212 (2019).
7. Proctor, T., Rudinger, K., Young, K., Nielsen, E. & Blume-Kohout, R. Measuring the Capabilities of Quantum Computers. *arXiv* (2020). https://arxiv.org/abs/2008.11294
8. Gottesman, D. An Introduction to Quantum Error Correction and Fault-Tolerant Quantum Computation. (2009). https://arxiv.org/abs/0904.2557
9. Anschuetz, E. R., Olson, J. P., Aspuru-Guzik, A. & Cao, Y. Variational Quantum Factoring. (2018). https://arxiv.org/abs/1808.08927v1
10. Krantz, P. et al. A quantum engineer's guide to superconducting qubits. *Appl. Phys. Rev.* **6**, 21318 (2019).
11. Farhi, E., Goldstone, J. & Gutmann, S. A Quantum Approximate Optimization Algorithm. (2014). https://arxiv.org/abs/1411.4028v1
12. Burges, C. J. C. Factoring as Optimization. *Tech. Rep*. MSR-TR-2002-83 (2002).
13. Guerreschi, G. G. & Matsuura, A. Y. QAOA for max-cut requires hundreds of qubits for quantum speed-up. *Scientific Reports* **9**, https://doi.org/10.1038/s41598-019-43176-9 (2019).
14. Farhi, E., Goldstone, J. & Gutmann, S. A quantum approximate optimization algorithm applied to a bounded occurrence constraint problem. (2014). https://arxiv.org/abs/1412.6062
15. Hastings, M. B. Classical and quantum bounded depth approximation algorithms. (2019). https://arxiv.org/abs/1905.07047
16. Biamonte, J. D. Nonperturbative k-body to two-body commuting conversion hamiltonians and embedding problem instances into ising spins. *Phys. Rev. A* **77**, 052331 (2008).
17. Dattani, N. Quadratization in discrete optimization and quantum mechanics. (2019). https://arxiv.org/abs/1901.04405
18. Harrigan, M. P. et al. Quantum approximate optimization of non-planar graph problems on a planar superconducting processor. *Nat. Phys.* **17**, 332–336 (2021).
19. Bengtsson, A. et al. Improved success probability with greater circuit depth for the quantum approximate optimization algorithm. *Phys. Rev. Appl.* **14**, 34010 (2020). https://doi.org/10.1103/PhysRevApplied.14.034010; https://arxiv.org/abs/1912.10495
20. Lacroix, N. et al. Improving the Performance of Deep Quantum Optimization Algorithms with Continuous Gate Sets. *PRX QUANTUM* **1**, 110304 (2020).
21. Zhou, L., Wang, S.-T., Choi, S., Pichler, H. & Lukin, M. D. Quantum Approximate Optimization Algorithm: Performance, Mechanism, and Implementation on Near-Term Devices. *Phys. Rev. X* **10** (2018). https://doi.org/10.1103/PhysRevX.10.021067; https://arxiv.org/abs/1812.01041
22. McClean, J. R. et al. Low depth mechanisms for quantum optimization. (2020). https://arxiv.org/abs/2008.08615

23. Schuld, M., Bergholm, V., Gogolin, C., Izaac, J. & Killoran, N. Evaluating analytic gradients on quantum hardware. *Phys. Rev. A* **99**, 032331 (2019).
24. Koch, J. et al. Charge-insensitive qubit design derived from the cooper pair box. *Phys. Rev. A* **76**, 042319 (2007).
25. Sundaresan, N. et al. Reducing unitary and spectator errors in cross resonance with optimized rotary echoes. *PRX Quantum* **1**, 020318 (2020).
26. Sim, S., Johnson, P. D. & Aspuru-Guzik, A. Expressibility and Entangling Capability of Parameterized Quantum Circuits for Hybrid Quantum-Classical Algorithms. *Adv. Quantum Technol.* **2**, 1900070 (2019).
27. Sung, Y. et al. Realization of high-fidelity CZ and ZZ-free iSWAP gates with a tunable coupler. *Phys. Rev. X* **11**, 021058 (2020).
28. Kandala, A. et al. Demonstration of a High-Fidelity CNOT for Fixed-Frequency Transmons with Engineered ZZ Suppression. (2020). https://arxiv.org/abs/2011.07050
29. Wallman, J. J. & Emerson, J. Noise tailoring for scalable quantum computation via randomized compiling. *Phys. Rev. A* **94**, 052325 (2016).
30. Cross, A. W., Bishop, L. S., Sheldon, S., Nation, P. D. & Gambetta, J. M. Validating quantum computers using randomized model circuits. *Phys. Rev. A* **100**, https://doi.org/10.1103/physreva.100.032328 (2019).
31. Braumüller, J. et al. Probing quantum information propagation with out-of-time-ordered correlators. (2021). https://arxiv.org/abs/2102.11751
32. Benedetti, M. et al. A generative modeling approach for benchmarking and training shallow quantum circuits. *npj Quantum Information* **5** (2019). https://doi.org/10.1038/s41534-019-0157-8
33. Dallaire-Demers, P.-L. et al. An application benchmark for fermionic quantum simulations. (2020). https://arxiv.org/abs/2003.01862

ACKNOWLEDGEMENTS

The authors would like to thank Eric Anschuetz, Morten Kjaergaard, Sukin Sim, Peter Johnson, and Jonathan Olson for fruitful discussion. We thank Maria Genckel for helping with the graphics. We would also like to thank the IBM Quantum team for providing access to the *ibmq_boeblingen* system, and for help debugging jobs. We acknowledge the use of IBM Quantum services for this work. The views expressed are those of the authors, and do not reflect the official policy or position of IBM.

AUTHOR CONTRIBUTIONS

Both A.H.K. and W.S. contributed equally. Y.C. conceived the project, A.H.K. and W.S. developed and executed the quantum workflow for measurements and simulations, A.H.K. and B.P. developed the noise model, T.S. assisted with Qiskit and IBM Q quantum backend support as relates to VQF. All authors contributed to writing the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41534-021-00478-z.

Correspondence and requests for materials should be addressed to Amir H. Karamlou or Yudong Cao.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.