Letter

Variational quantum eigensolvers in the era of distributed quantum computers

Ilia Khait,[1,2,*] Edwin Tham,[1] Dvira Segal,[2,3] and Aharon Brodutch[1]

[1]*Entangled Networks Ltd., Toronto, Ontario, Canada M4R 2E4*

[2]*Department of Physics and Centre for Quantum Information and Quantum Control, University of Toronto, Toronto, Ontario, Canada M5S 1A7*

[3]*Department of Chemistry, University of Toronto, 80 Saint George Street, Toronto, Ontario, Canada M5S 3H6*

(Received 7 March 2023; accepted 6 September 2023; published 20 November 2023)

The computational power of a quantum computer is limited by the number of qubits available for information processing. Increasing this number within a single device is difficult; it is widely accepted that distributed modular architectures are the solution to large-scale quantum computing. The major challenge in implementing such architectures is the need to exchange quantum information between modules. In this work, we show that a distributed quantum computing architecture with *limited* capacity to exchange information between modules can accurately solve quantum computational problems. Using the example of a variational quantum eigensolver with an ansatz designed for a two-module (dual-core) architecture, we show that three intermodule operations provide a significant advantage over no intermodule (or serially executed) operations. These results provide a strong indication that near-term *modular* quantum processors can be an effective alternative to their monolithic counterparts.

DOI: 10.1103/PhysRevA.108.L050401

Introduction. Quantum computers promise significant speed-up for a diverse set of problems [1–4]. However, the quantum advantage over classical computation only becomes appreciable when the problem size (i.e., the number of qubits required to solve the problem) is sufficiently large. Yet in practice, increasing the number of useful qubits on a quantum processing unit (QPU) is challenging: Generally, there is a trade-off between qubit count and qubit quality [5–8]. Modular architectures, where small high-quality QPUs are interconnected, offer a more sustainable solution to the scaling problem than a monolithic approach [9–14]. In small devices, high-fidelity qubit operations are easier to engineer, and corresponding verification and validation are more tractable. Modular approaches, however, require transmission of quantum information between QPUs. This information exchange can be used to create effective interactions between qubits residing on different QPUs. In general, information transfer between different modules is significantly slower and less reliable than that between qubits assigned to the same module. We call this the *quantum interconnect bottleneck* (QIB).

An increasingly salient architectural question for quantum computers concerns trade-offs in using an interconnected multimodule quantum device: Do the overheads associated with the QIB outweigh the price of adding qubits to a monolithic device?

Suppose one aims to run a circuit that requires N qubits but only has access to M-qubit devices with $M < N$. Assuming that these devices can exchange quantum information using a quantum interconnect, it is possible to recompile the circuit [15] such that it uses the interconnect n_i times. Alternatively, methods such as entanglement forging and circuit knitting [16–19] attempt to solve this problem at the cost of running the smaller M-qubit circuit more times, by a number which is exponential in n_i. If the QIB is ignored, an interconnected approach is favorable since it requires exponentially fewer shots. However, with QIB overheads, the operation of an interconnect may extend run times and reduce quality of results. To quantify the benefits of a quantum interconnect we compare a dual-core solution to a naïve approach with comparable running time—solving different parts of the problem on separate QPUs and relying solely on classical communication; we refer to this as the *separable* (or $n_i = 0$) solution. The dual-core solution consists of two interconnected QPUs with $\frac{N}{2}$ qubits each, while assuming that each QPU individually has an all-to-all connectivity map; i.e., within each module, qubits can interact directly with every other qubit.

If one allows $\mathcal{O}(N)$ interconnect uses, the aforementioned architecture becomes equivalent to an all-to-all N-qubit device. However, the QIB combined with practical considerations, such as decoherence, requires limiting n_i. As described below, $n_i = 3$ is not only sufficient for the problems we consider but it also shows a significant improvement over the separable solution. Specifically, we show that for a dual-core architecture, the estimation error arising from the expressibility of a limited-connectivity ansatz is exponentially suppressed with n_i.

In Fig. 1(a) we show our variational ansatz, which is composed of single-qubit operations along with the ZZ gate, $ZZ(\phi) = \exp(i\frac{\phi}{2}\sigma_z \otimes \sigma_z)$, a common entangling operation in trapped-ion devices [20,21]. We treat ZZ gates that straddle two clusters of $N/2$ qubits as interconnect-mediated remote gates. We compare performances of the separable and modular architectures for a common algorithm, the variational

*ilia@entanglednetworks.com

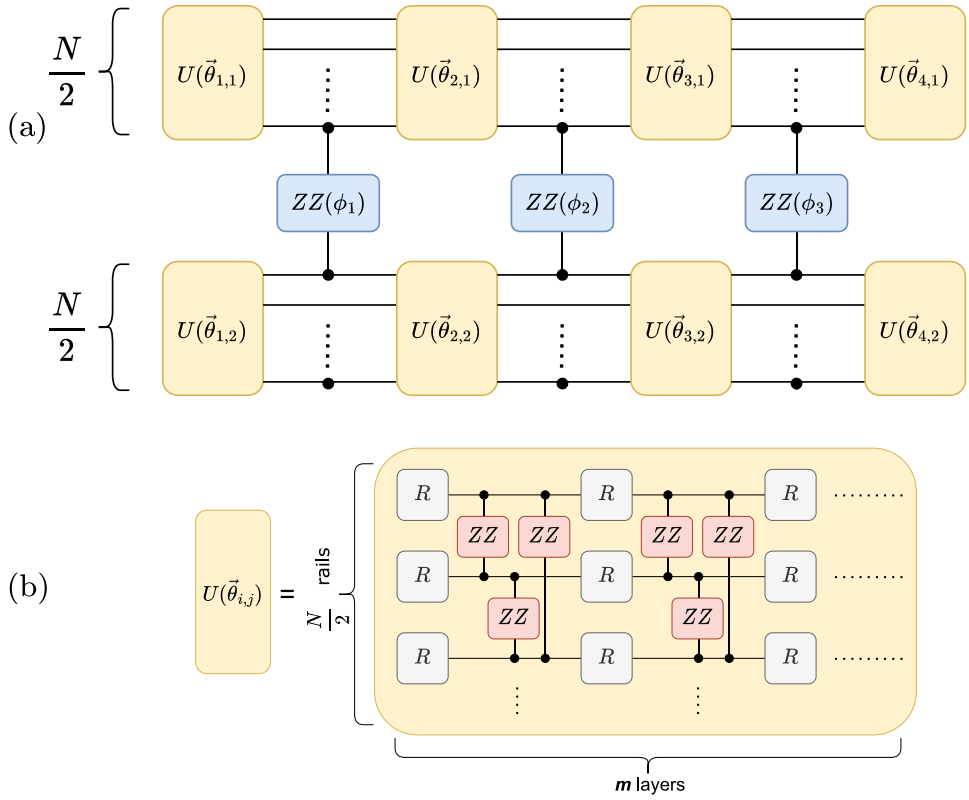

FIG. 1. A VQE ansatz circuit for a dual-core quantum architecture. (a) VQE ansatz's structure: Each QPU contains half of the available qubits, $\frac{N}{2}$. The unitary $U(\vec{\theta}_{i,j})$ (yellow block) acts on the qubits of the jth QPU at the ith stage, followed by a parametric remote gate operation, ZZ (blue block), allowing entanglement to be shared between QPUs. Throughout this Letter, the number of intercore operations is $n_i = 3$. (b) The unitary $U(\vec{\theta}_{i,j})$ contains m layers; gates (R) are implicitly parametrized by elements of $\vec{\theta}_{i,j}$. Throughout this Letter, $m = 3$ for a total 12 layers executed on each QPU.

quantum eigensolver (VQE) [3,22], which estimates the ground-state energy of a Hamiltonian. To make a simple comparison between the separable solution and the interconnected one, we only consider the precision of the result on the final circuit, and thus, we avoid issues related to the performance of the optimization stage.

Interconnect advantage. Decomposing a state into its principal components [23] is a core technique in many numerical recipes such as the density matrix remormalization group (DMRG) algorithm [24], where a truncation is performed based on the diminishing return on the fidelity of storing more basis components (at a high cost). Such a rationale is used for understanding the power of interconnects. Suppose one has two QPUs, each capable of preparing any state in its M-qubit Hilbert space. The state of that dual-QPU system is a product state, $|\psi\rangle = |\psi_1\rangle|\psi_2\rangle$, where $|\psi_j\rangle$ is the state of the jth QPU. Every application of a remote operation between QPUs increases inter-QPU entanglement, as expressed by the rank (d) of the Schmidt decomposition of $|\psi\rangle$, cut along the two QPUs: $\psi^{(d)} = \sum_{i=1}^{d} c_i |\psi_1^{(i)}\rangle|\psi_2^{(i)}\rangle$. Note that with a sufficiently expressive *intra*-QPU ansatz [$U(\vec{\theta})$ in Fig. 1] the entanglement rank d can rise quickly, up to exponential in number of inter-QPU operations n_i (i.e., $d \leqslant 2^{n_i}$).

Procedure. VQE is an iterative classical-quantum hybrid algorithm, which estimates the ground-state energy of a given Hamiltonian [3,25]. A quantum computer produces an approximation of the Hamiltonian's ground-state based on a parametrized ansatz; in turn, a classical strategy for converting a Hamiltonian into a series of compactly implementable observables estimates an energy eigenvalue from that approximate ground state [26–28]. This process is repeated, with varied ansatz parameters chosen by an optimization strategy, until a sufficiently refined ground state is reached [29]. In what follows, we demonstrate that the $n_i = 3$ dual-core parametrized circuit suggested in Fig. 1 provides an excellent approximation to the exact ground state of interacting systems.

Apart from ansatz expressibility (how closely the ansatz can approximate an arbitrary quantum state), VQE results also depend on classical factors like the Hamiltonian-to-observable map and parameter optimization routines. Since we intend to test how expressibility is augmented by interconnects in a multi-QPU setup, we avoid confounding classical issues by maximizing the fidelity between the variational state and the exact ground state, instead of minimizing the expectation value of the Hamiltonian, E_{var}. The following summarizes our procedure (for details, see Ref. [30]). First, we diagonalize the Hamiltonian obtaining the exact ground state, $|\psi_{\text{GS}}\rangle$. We perform a singular value decomposition (SVD) (where the system is divided into two units) and retain the eight most

significant contributions,

$$|\psi_{\rm GS}\rangle \to \left|\psi_{\rm GS}^{(8)}\right\rangle = \sum_{i=1}^{8} \lambda_i \left|\phi_1^{(i)}\right\rangle \left|\phi_2^{(i)}\right\rangle. \tag{1}$$

Here, λ_i and $|\phi_1^{(i)}\rangle|\phi_2^{(i)}\rangle$ are the ith Schmidt eigenvalue and eigenvector, respectively. We iteratively build the variational solution (see Fig. 1). We start by optimizing the fidelity towards a product state, our crudest approximation to the ground state, $|\psi_{\rm GS}^{(1)}\rangle$, defined by the largest Schmidt eigenvalue λ_1. We construct a variational approximation of this target state using the first set of unitaries $U(\vec{\theta}_{1,1})$ and $U(\vec{\theta}_{1,2})$, by applying them on an all-polarized state, $|0\ldots0\rangle_1|0\ldots0\rangle_2$. Next, we add another Schmidt coefficient and target the state $|\psi_{\rm GS}^{(2)}\rangle$ by enlarging the set of variational parameters: a first remote operation $ZZ(\phi_1)$ and another set of unitaries, $U(\vec{\theta}_{2,1})$ and $U(\vec{\theta}_{2,2})$. After optimization, we add another Schmidt coefficient, with another interconnected-remote operation, and we repeat till $|\psi_{\rm GS}^{(8)}\rangle$ [31].

Models. We test three models to benchmark the interconnect-mediated ansatz: the transverse field Ising model (TFIM), the spin-1/2 anisotropic Heisenberg model, and the spin-1 Heisenberg model. The first two models are paradigmatic examples in benchmarking performances of novel methods [32,33]; the spin-1 Heisenberg model enables exploration of the impact of less local interactions (due to the casting onto spin-1/2 operators) on the quality of the interconnected QPU solution.

The TFIM is defined as

$$H_{\rm TFIM} = -J \sum_{i=1}^{N-1} \sigma_i^z \sigma_{i+1}^z - h_x \sum_{i=1}^{N} \sigma_i^x, \tag{2}$$

where σ^α ($\alpha = x$, y, and z) are the Pauli matrices and N is the number of spins (qubits). The phase diagram of the TFIM consists of a (anti)ferromagnetic ordered product state for positive (negative) J spin-spin interaction and vanishing transverse field h_x, and a disordered state at strong transverse magnetic field. In the thermodynamic limit, a quantum phase transition to a gapless phase occurs at $h_x = J$.

The anisotropic Heisenberg model, which is used in studies of magnetic systems is given by

$$H_{\rm XYZ} = \sum_{i=1}^{N-1} \left(J_x S_i^x S_{i+1}^x + J_y S_i^y S_{i+1}^y + J_z S_i^z S_{i+1}^z\right) + h_x \sum_{i=1}^{N} S_i^x, \tag{3}$$

where S^α ($\alpha = x$, y, z) are spin-1/2 operators.

Lastly, the $S = 1$ Heisenberg model resembles the Affleck-Kennedy-Lieb-Tasaki model [34,35], and it is of great interest due to its topological properties. Here, it allows us to probe and benchmark the interconnected ansatz for a Hamiltonian with fewer local operators,

$$H_{\rm Heis} = J \sum_{\alpha,i=1}^{N-1} S_i^\alpha S_{i+1}^\alpha \to \frac{J}{4} \sum_{\alpha,i=1}^{N-1} \left(\sigma_{2i-1}^\alpha + \sigma_{2i}^\alpha\right)\left(\sigma_{2i+1}^\alpha + \sigma_{2i+2}^\alpha\right) + J_{\rm FM} \sum_{\alpha,i=1}^{N} \sigma_{2i-1}^\alpha \sigma_{2i}^\alpha. \tag{4}$$

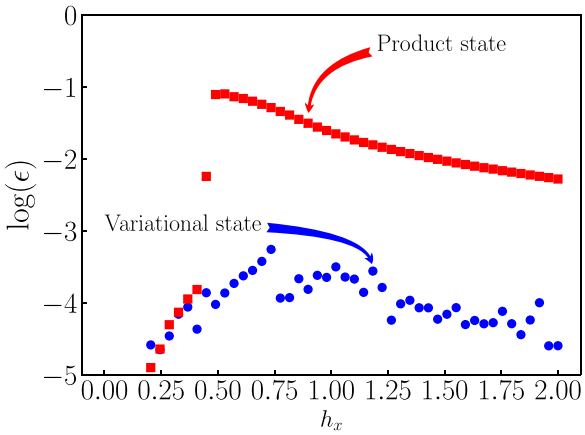

FIG. 2. TFIM model with 12 spins at $J = 1$, Eq. (2), studied as a function of the transverse magnetic field h_x. We depict the logarithm of the relative energy difference between the exact ground state and the variational ansatz, ϵ. The two variational ansätze are a product state (red squares), in which a separable solution is forced, and the interconnected solution (blue circles). For presentation purposes, the error of the separable solution is decreased by an order of magnitude. For either vanishing or strong magnetic fields, the exact solution is a product state. However, for intermediate h_x values, the separable solution is far inferior.

Here S^α ($\alpha = x$, y, and z) are spin-1 operators, and the mapping marked by the arrow splits every $S = 1$ operator into a pair of spin-1/2 operators. $J_{\rm FM} \gg J$ is chosen such that a triplet is selected for every *other* bond (originating from the $S = 1$ operators), with the resulting interactions having a distance of 4 units.

Results. We study systems with up to 12 qubits and demonstrate next an immense advantage for $n_i = 3$ over $n_i = 0$. In Fig. 2, we compare the ground-state approximation for the TFIM given the two distinct architectures. We denote the VQE solution by $E_{\rm var}$, and we compare it to the exact ground-state energy $E_{\rm GS}$. Specifically, throughout the Letter, we analyze the error measure $\epsilon = \frac{E_{\rm var}}{E_{\rm GS}} - 1$.

An indication of a phase transition in the thermodynamic limit is detectable even in this small system, as can be seen in Fig. 2 by examining the product state (red squares). While for extreme field values ($h_x \to 0$ and $\frac{h_x}{J} \gg 1$) the product state solution well approximates the exact ground state, around $\frac{h_x}{J} \approx 0.5$ this approximation completely fails. In contrast, the $n_i = 3$ ansatz maintains a lower error throughout. For presentation purposes, we display the product state error scaled down by a factor of 10. Overall, the interconnected solution performs well, and its error does not exceed 0.07% throughout the phase diagram, compared to an error of up to 8% in the separable solution. The error in the interconnected solution is comparable in magnitude to the exact (nonvariational) wave function truncated to eight Schmidt terms, where infidelity is 0.01%.

In Fig. 3, we examine a portion of the phase diagram of the anisotropic Heisenberg model at the fixed value $J_x = 1.0$ and magnetic field values $h_x \in \{0, 0.5, 1.0\}$, thus including the symmetric Heisenberg point $(J_x, J_y, J_z, h_x) = J(1, 1, 1, 0)$. We plot ϵ for a 12-qubit problem. Excluding extreme malper-

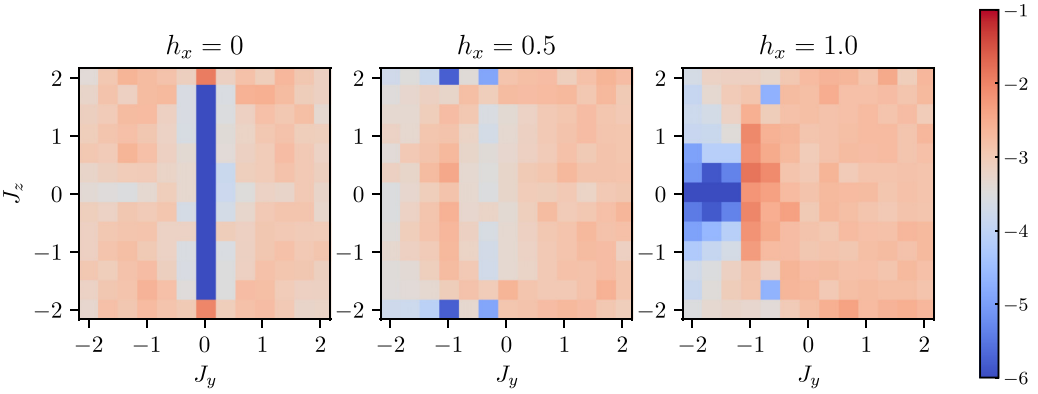

FIG. 3. VQE results for the one-dimensional XYZ model with 12 spins [see Eq. (3)]. We present a color map of the logarithm of the relative energy differences between the exact ground state and the variational ansatz, ϵ. The panels depict different magnetic field strengths: $h_x = 0$, 0.5, and 1.0, and $J_x = 1$. Fidelities are usually above 99.9%; the highest reported variation is 1.4%.

forming data points, the energy convergence ratio stays well below 10^{-3} [36]: The worst performing data point, occurring in the vicinity of $(J_x, J_y, J_z, h_x) = (1, -1, 0.5, 1)$ along the J_z direction, appears to be an outlier with an error $\epsilon \approx 1.4\%$. We are unable to pinpoint the reason for this failing; neighboring data points in the J_y direction show significantly better convergence.

In Fig. 4, we study the $S = 1$ Heisenberg model at $J_{\mathrm{FM}} = 10$ [37] [see Eq. (4)]. The main panel presents the log-infidelity of the result, $\ln\{1 - |\langle\psi_{\mathrm{var}}|\psi_{\mathrm{GS}}\rangle|^2\}$. The $n_i = 3$ solution (blue circles) displays significantly better fidelities compared to the separable one ($n_i = 0$, red squares), with factor of 20 *decrease* in infidelity at $N = 12$, and significantly better results for smaller systems. The inset shows the relative energy estimation error ϵ as a function of the number of qubits. Besides a single outstanding point ($N = 10$), the upward trend reflects the increasing complexity of the solution as the system size grows. The outlier at $N = 10$ was further examined by introducing another layer, after the third remote operation (with the VQE ansatz including 13 layers in total). This brought the relative energy error to 0.2%, consistent with the linear trend seen in Fig. 4. Introducing the same change to other values of N showed no significant change. We attribute this deviation to the optimization process as elaborated on next.

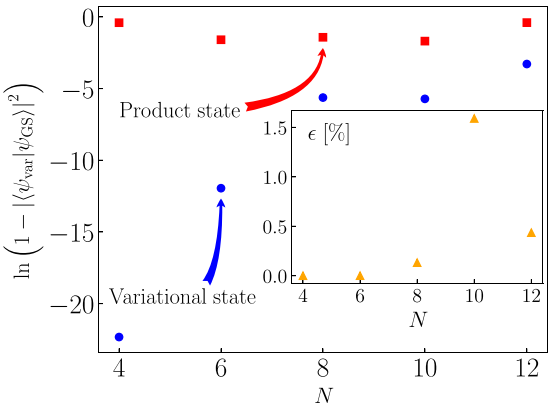

FIG. 4. VQE results for the $S = 1$ Heisenberg model, Eq. (4), with up to a six-spin chain represented by $\times 2$ qubits, N. We display the log infidelity of the variational state compared to the exact ground state for the separable product states of two $\frac{N}{2}$ qubits (red squares) and the interconnected ansätze (blue circles). The inset shows the relative energy error ϵ, in percent. While the error of the interconnected solution increases with N, it still delivers fidelity orders of magnitude better than that of the separable solution.

Discussion. Two issues limit convergence to the exact solution.

(i) Classical optimization. As described under *Procedure*, we are using differentiable programming [30] to find the optimal variational parameters for the VQE ansatz, a task of growing complexity when increasing the qubit count. The number of parameters increases depending on the ansatz structure (in our ansatz for $N = 12$ we have 753 variational parameters).

(ii) Expressibility of the ansatz. [38] As discussed above, one can separate the effect of the interconnect (the remote gate) from the "local" layers [the unitaries $U(\vec{\theta}_{i,j})$]. Each interconnect operation doubles the potential Schmidt rank of the state, and the role of subsequent layers is to facilitate quantum information spreading. Whether information spreads far enough depends on the number of layers and their inner structure (Fig. 1). To assess the expressibility of the ansatz without the effect of the interconnect, we have examined in Ref. [30] a single QPU architecture with all-to-all connected qubits. While an all-to-all architecture performs slightly better than the interconnected one, it comes at a greater cost as increasing the number of qubits on a QPU is a nontrivial task, which a multi-QPU modular architecture aims to avoid.

The SVD eigenvalues of the TFIM decay faster than those of the $S = 1$ Heisenberg model due the topological nature of the latter's ground state [39]. The *theoretical* lower bound on the infidelity is the sum of the discarded SVD eigenvalues squared. Considering that only three remote operations were allowed here, the discarded weight in the TFIM model was found to be 4×10^{-10}. Hence, the resulting infidelity, 10^{-6} is not limited by the interconnect. Similarly, in the $S = 1$ Heisenberg model with $N = 12$, the discarded weight is 2×10^{-3}, and the reported infidelity of 3×10^{-3} is close to this bound. In conclusion, the limiting factor in solving the VQE on an interconnected hardware is classical optimization

combined with the limited expressibility of the ansatz, rather than the introduction of remote operations. Interestingly, we note that comparing the TFIM to the $S = 1$ Heisenberg model, the TFIM is converging much better than the latter, both in the interconnected case and in the all-to-all connected case [30]. This could be explained by the suitability of the ansatz to the specific model, though in-depth consideration of this aspect is outside the scope of this Letter.

Conclusions. In this work, we demonstrated that a distributed quantum architecture with only modest inter-QPU capacity provides a dramatic advantage in VQE computations over serial architectures with no interconnects. In all cases studied, we found that three judiciously placed inter-QPU gates were sufficient to produce a significantly better approximation to the ground-state energy for Hamiltonians of interest compared to an architecture with no quantum interconnects. For the Hamiltonians studied here we find that increasing n_i allows for an exponential improvement in the fidelity [30]. Our comparison is based on simulations, and it is therefore limited to a small number of qubits. The main conclusion from this work is that an exponential increase in the Schmidt rank with respect to n_i (and subsequently in the dimension of the effective Hilbert space) manifests itself when solving a practical algorithm.

The relatively small instances explored here allowed us to overcome some aspects of classical optimization, yet prevented us from showing advantages over other methods such as circuit knitting and entanglement forging. These techniques scale exponentially worse with increased number of interconnect uses n_i, but are generally expected to perform well for $n_i = 3$. As such, we expect that in future work these methods could be combined with interconnects to increase the effective Hilbert space.

A number of important questions remain open, including the impact of noise and the imperfect nature of interconnects. Slow interconnects would increase run time and make the computation more susceptible to decoherence; inter-QPU operations are generally expected to have lower fidelity [14,40]; and the use of fixed resource states creates overheads in gate counts and the actual implementation of the interconnect would impact the other qubits. These limitations need to be weighed against the downside of increasing qubit count in a monolithic architecture, as well as artificially increasing qubit size using classical resources. Identifying algorithms where a limited number of interconnect uses can be proved advantageous will be an incentive for the implementation of multi-QPU architectures. We hope that our study would stimulate further work in this direction.

Acknowledgments. The authors acknowledge fruitful discussions with Finn Lasse Buessen and Kevin Smith. The work of I.K. was supported by the Centre for Quantum Information and Quantum Control (CQIQC) at the the University of Toronto. D.S. acknowledges support from an NSERC Discovery Grant and the Canada Research Chair program.

[1] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp, Contemp. Math. **305**, 53 (2002).
[2] L. K. Grover, in *Proceedings of the 28th ACM Theory of Computing* (ACM, New York, 1996), pp. 212–219.
[3] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. Obrien, Nat. Commun. **5**, 4213 (2014).
[4] E. Farhi, J. Goldstone, and S. Gutmann, arXiv:1411.4028.
[5] R. J. Hughes, D. F. V. James, E. H. Knill, R. Laflamme, and A. G. Petschek, Phys. Rev. Lett. **77**, 3240 (1996).
[6] A. Steane, C. F. Roos, D. Stevens, A. Mundt, D. Leibfried, F. Schmidt-Kaler, and R. Blatt, Phys. Rev. A **62**, 042305 (2000).
[7] C. Monroe and J. Kim, Science **339**, 1164 (2013).
[8] P. Murali, D. M. Debroy, K. R. Brown, and M. Martonosi, in *2020 ACM/IEEE 47th Annual International Symposium on Computer Architecture (ISCA)* (ACM/IEEE, New York, 2020), pp. 529–542.
[9] C. Monroe, R. Raussendorf, A. Ruthven, K. R. Brown, P. Maunz, L.-M. Duan, and J. Kim, Phys. Rev. A **89**, 022317 (2014).
[10] J. M. Pino, J. M. Dreiling, C. Figgatt, J. P. Gaebler, S. A. Moses, M. S. Allman, C. H. Baldwin, M. Foss-Feig, D. Hayes, K. Mayer, C. Ryan-Anderson, and B. Neyenhuis, Nature (London) **592**, 209 (2021).
[11] N. H. Nickerson, J. F. Fitzsimons, and S. C. Benjamin, Phys. Rev. X **4**, 041041 (2014).
[12] K. R. Brown, J. Kim, and C. Monroe, npj Quantum Inf. **2**, 16034 (2016).
[13] S. Bravyi, O. Dial, J. M. Gambetta, D. Gil, and Z. Nazario, J. Appl. Phys. **132**, 160902 (2022).
[14] D. Awschalom, K. K. Berggren, H. Bernien, S. Bhave, L. D. Carr, P. Davids, S. E. Economou, D. Englund, A. Faraon, M. Fejer, S. Guha, M. V. Gustafsson, E. Hu, L. Jiang, J. Kim, B. Korzh, P. Kumar, P. G. Kwiat, M. Lončar, M. D. Lukin *et al.*, PRX Quantum **2**, 017002 (2021).
[15] E. Tham, I. Khait, and A. Brodutch arXiv:2206.09938.
[16] T. Peng, A. W. Harrow, M. Ozols, and X. Wu, Phys. Rev. Lett. **125**, 150504 (2020).
[17] C. Piveteau and D. Sutter, arXiv:2205.00016.
[18] A. Eddins, M. Motta, T. P. Gujarati, S. Bravyi, A. Mezzacapo, C. Hadfield, and S. Sheldon, PRX Quantum **3**, 010309 (2022).
[19] Y. Zhang, L. Cincio, C. F. A. Negre, P. Czarnik, P. J. Coles, P. M. Anisimov, S. M. Mniszewski, S. Tretiak, and P. A. Dub, npj Quantum Inf. **8**, 96 (2022).
[20] A. Sørensen and K. Mølmer, Phys. Rev. Lett. **82**, 1971 (1999).
[21] A. Sørensen and K. Mølmer, Phys. Rev. A **62**, 022311 (2000).
[22] D. A. Fedorov, B. Peng, N. Govind, and Y. Alexeev, Mater. Theory **6**, 2 (2022).
[23] This decomposition can be achieved via SVD, or the Schmidt decomposition, which is used here.
[24] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[25] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature (London) **549**, 242 (2017).

[26] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, Quantum Sci. Technol. **4**, 014008 (2018).

[27] I. G. Ryabinkin, T.-C. Yen, S. N. Genin, and A. F. Izmaylov, J. Chem. Theory Comput. **14**, 6317 (2018).

[28] I. G. Ryabinkin, R. A. Lang, S. N. Genin, and A. F. Izmaylov, J. Chem. Theory Comput. **16**, 1055 (2020).

[29] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn, A. Kandala, A. Mezzacapo, P. Müller, W. Riess, G. Salis, J. Smolin, I. Tavernelli, and K. Temme, Quantum Sci. Technol. **3**, 030503 (2018).

[30] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevA.108.L050401 for additional details on (i) the optimization procedure that builds the interconnected solution, (ii) the interconnect advantage, (iii) ansatz expressibility and (iv) a comparison between the single-core and the interconnected architectures.

[31] Note that this is not a viable VQE optimization method in a "production" use since knowledge of the exact ground-state cannot be assumed; we employ it here only to test expressiblity.

[32] G. Carleo and M. Troyer, Science **355**, 602 (2017).

[33] F. L. Buessen, D. Segal, and I. Khait, Phys. Rev. Res. **5**, L022003 (2023).

[34] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, Phys. Rev. Lett. **59**, 799 (1987).

[35] K. C. Smith, E. Crane, N. Wiebe, and S. M. Girvin, arXiv:2210.17548.

[36] See Ref. [30], Fig. S2, for the comparison with a separable solution.

[37] In choosing the value J_{FM}, we tested convergence of the resulting ground state with DMRG calculations and found complete agreement for $J_{\mathrm{FM}} > 6$.

[38] For a deeper discussion about the expressibility of parametrized quantum circuits, we refer the reader to Ref. [41].

[39] S. R. White and D. A. Huse, Phys. Rev. B **48**, 3844 (1993).

[40] L. J. Stephenson, D. P. Nadlinger, B. C. Nichol, S. An, P. Drmota, T. G. Ballance, K. Thirumalai, J. F. Goodwin, D. M. Lucas, and C. J. Ballance, Phys. Rev. Lett. **124**, 110501 (2020).

[41] S. Sim, P. D. Johnson, and A. Aspuru-Guzik, Adv. Quantum Technol. **2**, 1900070 (2019).