Exploiting Contextuality in Variational Quantum Eigensolvers

William M. Kirby,[1] Andrew Tranter,[1, 2] and Peter J. Love[1, 3]

[1] *Department of Physics and Astronomy, Tufts University, Medford, MA 02155*
[2] *Cambridge Quantum Computing, 9a Bridge Street Cambridge, CB2 1UB United Kingdom*
[3] *Computational Science Initiative, Brookhaven National Laboratory, Upton, NY 11973*

In this work, we describe how contextuality may be identified and used to advantage in variational quantum eigensolvers (VQEs). Contextuality, which can appear in preparations, transformations, or measurements, is a feature of quantum mechanics that distinguishes it from classical physics [1–7]. In the context of VQEs, it is most natural to consider measurement contextuality, since the quantum component of a variational eigensolver consists of a series of ansatz preparations and measurements [8].

We focus on the most typical variant of VQE, in which the goal is to estimate the ground state energy of some Hamiltonian H [9]. This is accomplished by decomposing H into a linear combination of easily-measurable terms, preparing an ansatz, and estimating the expectation value of each term separately on the quantum computer. Then, classically, the expectation values of the terms are combined to obtain an expectation value for H, which is used to update the ansatz parameters. The quantum part of the procedure is therefore characterized by the set of measurements performed, i.e., the set of terms in the Hamiltonian. Hence, if we wish to identify contextuality in a VQE instance, the natural choice is to consider contextuality of the terms in the Hamiltonian.

A set of observables is *contextual* if it does not admit any self-consistent joint value assignments [8]. Inconsistencies in joint value assignments may arise due to *inference*, which refers to the following relation: if two observables A and B commute, then the value assigned to their product must be the product of their assigned values. This follows because an observer could measure A and B simultaneously, and thus infer the value of their product. The classic example of contextuality arising from such inference relations is the Peres-Mermin square [4–6], shown in Table I. For example, the first column of the square illustrates the fact that the value of XZ must be the product of the values of XI and IZ, while the last column of the table illustrates the fact that the value of $-YY$ must be the product of the values of XX and ZZ. However, there is no assignment of values to the complete square that satisfies all such inference relations, so the set of observables in the square is contextual.

Implementations of VQE typically use the Pauli decomposition of H:

$$H = \sum_{P \in \mathcal{S}} h_P P, \qquad (1)$$

where $\mathcal{S}$ is some set of Pauli operators P, and h_P are real coefficients. We wish to evaluate contextuality of the set $\mathcal{S}$, the measurements to be performed in the procedure. The first part

XI	IX	XX	$+1$
IZ	ZI	ZZ	$+1$
XZ	ZX	YY	$+1$
$+1$	$+1$	-1	

TABLE I. The Peres-Mermin square [4–6]. The observables in the square are two-qubit Pauli operators (tensor product symbols omitted). The observables commute within each row and column, and the product of each row or column is ± 1 (shown outside the lines). There is no assignment of values (± 1) to the observables that satisfies all six commuting products.

of this work [8] consists of proving a necessary and sufficient condition for noncontextuality of a set $\mathcal{S}$ of Pauli operators: for $\mathcal{Z} \subseteq \mathcal{S}$ defined to contain all operators that commute with all other operators in $\mathcal{S}$, commutation is an equivalence relation on $\mathcal{S} \setminus \mathcal{Z}$ if and only if $\mathcal{S}$ is noncontextual.

Given a noncontextual set of Pauli operators, there exist joint value assignments, so we should be able to construct probability distributions (corresponding to quantum states) over these joint value assignments. These probability distributions provide joint sets of expectation values for the Pauli operators, which are exactly what we need in order to reproduce the original VQE procedure. We can think of this as a classical phase-space description of a noncontextual Hamiltonian, where an uncertainty relation is imposed in the form of a restriction on the allowed probability distributions. This is demonstrated in the second part of this work [10]. The joint sets of expectation values are parametrized by at most $2n + 1$ real numbers for a Hamiltonian on n qubits, giving an expression for the expectation value of the Hamiltonian as a classical objective function:

$$\langle H \rangle = E(\vec{q}, \vec{r}). \tag{2}$$

Here E is a polynomial (with polynomially-many terms) of the real parameters $(\vec{q}, \vec{r})$, which satisfy $q_i = \pm 1$ and $|\vec{r}| = 1$.

One immediate consequence of the classical expression (2) for the energy is that we can perform a classical gradient search over the parameter space $(\vec{q}, \vec{r})$ using this objective function rather than estimating expectation values on a quantum computer. Thus, we have "dequantized" the VQE procedure for a noncontextual Hamiltonian, although it is not impossible that the quantum version could still afford an advantage due to the difference in how the ansatz parameters enter the objective function.

The other immediate consequence is that the NONCONTEXTUAL HAMILTONIAN PROBLEM — the decision problem of whether the ground state energy of a noncontextual Hamiltonian lies below a given energy gap $[a, b]$ such that $b - a \geq 1/\text{poly}(n)$ — is in NP. If the ground state energy is less than a, then there exists a parameter setting $(\vec{q}_0, \vec{r}_0)$ that witnesses this, i.e., such that $\langle H \rangle_0 = E(\vec{q}_0, \vec{r}_0) < a$. Since the function E is a polynomial with polynomially-many terms, it can be evaluated classically efficiently, providing a classical proof that $\langle H \rangle_0 < a$.

In the third and final part of this work [11], we show how to partition an arbitrary Hamiltonian into a noncontextual part and its complement, which will in general be contextual. By choosing the noncontextual part to be as large as possible, we can take the ground state energy of the noncontextual part (found via classical simulation) to be an initial approximation to the ground state energy of the full Hamiltonian. We then identify the subspace of quantum states that are consistent with the noncontextual ground state. By minimizing the expectation value of the remaining terms within this subspace (using VQE on a quantum computer), we can obtain a quantum correction to the initial noncontextual approximation.

The quantum correction subspace corresponding to a noncontextual state with parameters $(\vec{q}, \vec{r})$ is the joint eigenspace of a set of commuting observables $G \cup \{A\}$, determined during construction of the noncontextual objective function (2). G is an independent set of Pauli operators whose eigenvalues are the entries in $\vec{q}$, while A is a rotated Pauli operator determined by $\vec{r}$. The rotation that maps A to a single Pauli operator may be taken to leave G invariant, so the quantum correction subspace is a stabilizer subspace to which a rotation has been applied that preserves all but one of the stabilizers. We call this subspace the *contextual subspace*, because it encodes the degrees of freedom that are left when the noncontextual degrees of freedom have been factored out.

The remaining terms form a new Hamiltonian when restricted to the contextual subspace. This restricted Hamiltonian is smaller than the original Hamiltonian by one qubit per stabilizer (in the set $G \cup \{A\}$) of the contextual subspace, so the quality of the quantum correction obtained by minimizing the energy of the restricted Hamiltonian also depends on the dimension of the subspace. If we discard some of the stabilizers, the noncontextual approximation becomes worse,

but the quantum correction becomes better as the contextual subspace grows, and we prove that the overall approximation either stays the same or improves. Eventually, if we throw away all of the stabilizers, we recover full VQE. However, the more stabilizers we throw away, the larger the quantum correction Hamiltonian becomes in both qubits and terms. Thus, we have a parameter that we can adjust to use more quantum resources to get better accuracy. We call this method *contextual subspace VQE*, or CS-VQE.
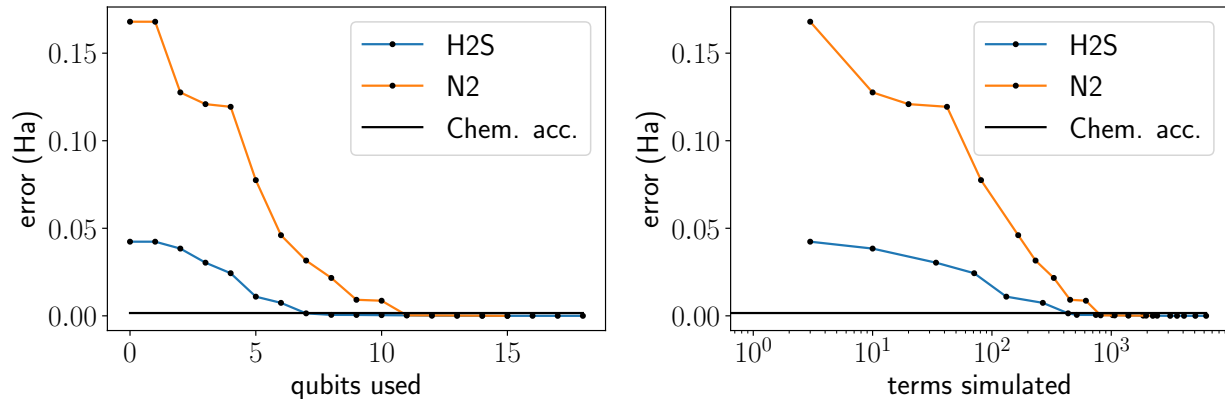

FIG. 1. The left plot shows CS-VQE approximation errors versus number of qubits used on the quantum computer, for tapered molecular Hamiltonians for H_2S and N_2. The right plot shows the same errors versus the number of terms simulated on the quantum computer. The solid black lines indicate chemical accuracy.

We simulated applying CS-VQE to electronic-structure Hamiltonians, the most common current application for VQE. In Fig. 1, we show as examples the results for H_2S and N_2. Fig. 1 plots the overall approximation error against the number of qubits used on the quantum computer, and against the number of terms simulated on the quantum computer. These examples illustrate that we can achieve chemical accuracy using fewer qubits and terms than required for standard VQE on tapered Hamiltonians. We have simulated CS-VQE on twenty-five molecular and atomic Hamiltonians on up to eighteen qubits (after tapering off qubits using symmetries) and found reductions in the resources required to reach chemical accuracy for all but three, all of which were among the smaller instances (fewer than ten qubits). The number of qubits can be chosen to match the available quantum processor, and reducing the number of terms can make the difference between the computation being worthwhile and being too expensive. Based on these results, we hope that CS-VQE can make it possible for smaller quantum computers to access larger applications.

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[3] S. Kochen and E. Specker, J. Math. Mech. **17**, 59 (1967).
[4] A. Peres, J. Phys. A: Math. Theor. **24**, L175 (1991).

[5] N. D. Mermin, Phys. Rev. Lett. **65**, 3373 (1990).

[6] N. D. Mermin, Rev. Mod. Phys. **65**, 803 (1993).

[7] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005).

[8] W. M. Kirby and P. J. Love, Phys. Rev. Lett. **123**, 200501 (2019).

[9] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nature Communications **5**, 4213 EP (2014).

[10] W. M. Kirby and P. J. Love, Phys. Rev. A **102**, 032418 (2020).

[11] W. M. Kirby, A. Tranter, and P. J. Love, arXiv preprint (2020), arXiv:2011.10027 [quant-ph].