

Probing the ground state of the antiferromagnetic Heisenberg model on the kagome lattice using geometrically informed variational quantum eigensolver

Abdellah Tounsi ^{1,2} Nacer Eddine Belaloui ^{1,2} Abdelmouheymen Rabah Khamadja ^{1,2} Takei Eddine Fadi Lalaoui ^{1,2}
Mohamed Messaoud Louamri ^{1,3} David E. Bernal Neira ⁴ and Mohamed Taha Rouabah ^{1,2,5,*}

¹Constantine Quantum Technologies, *Frères Mentouri University Constantine 1*, Ain El Bey Road, Constantine 25017, Algeria

²Laboratoire de Physique Mathématique et Subatomique, *Frères Mentouri University Constantine 1*,
Ain El Bey Road, Constantine 25017, Algeria

³Theoretical Physics Laboratory, *University of Science and Technology Houari Boumediene*, BP 32 Bab Ezzouar, Algiers 16111, Algeria

⁴Davidson School of Chemical Engineering, *Purdue University*, 480 Stadium Road, West Lafayette, Indiana 47907, USA

⁵Department of Electrical and Computer Engineering, *North Carolina State University*, Raleigh, North Carolina 27606, USA



(Received 17 December 2025; accepted 8 September 2026; published 21 October 2026)

This work investigates the nature of the ground state of the antiferromagnetic Heisenberg model on fundamental kagome cells, a triangle and a star, using the variational quantum eigensolver algorithm on real quantum hardware. We demonstrate that the ground-state preparation is achievable using a shallow hardware-efficient quantum circuit with a naturally Euclidean parameter space. Our custom ansatz is capable of accurately recovering meaningful properties of the ground state such as the spin-spin correlation terms and static structure factor without explicit error mitigation. These features are found to be resilient to noise. We exploited the Fubini-Study metric in constructing the ansatz, ensuring a singularity-free parameter space. With this ansatz design, the adaptive gradient descent optimizer achieves a faster convergence in terms of the number of iterations compared to simultaneous perturbation stochastic approximation. We further apply error mitigation techniques, including zero-noise extrapolation (ZNE) and qubitwise readout-error mitigation (REM). While ZNE does not obey the Rayleigh-Ritz variational principle, the conditions under which REM preserves it are discussed.

DOI: [10.1103/82g5-bsvp](https://doi.org/10.1103/82g5-bsvp)

I. INTRODUCTION

Finding the ground-state energy of a k -local Hamiltonian is known to be a quantum Merlin-Arthur hard (QMA-hard) problem [1], the quantum analog of a classical NP-hard problem. This is mainly due to the exponential growth of the Hilbert space in interacting quantum many-body systems, rendering the problem generally intractable for classical computers. Despite QMA-hard problems being challenging to solve even on a perfect quantum computer, the latter remains the most suitable tool to crack this class of problems because it follows the same laws of quantum mechanics [2–4]. However, despite recent advancements in the technology [5,6], current quantum computers are yet to reach their promised potential to efficiently simulate complex quantum systems. Short coherence times, high levels of noise, and limited connectivity, all characteristics of noisy intermediate-scale quantum (NISQ) era computers [7,8], still pose significant challenges to quantum advantage. Furthermore, identifying and preparing the ground state is crucial for studying many-body systems,

enabling the understanding of their properties in the low-temperature regime. For this task, the variational quantum eigensolver (VQE) was introduced [9,10]. It is a hybrid quantum-classical algorithm specifically designed to approximate the ground-state energy using parametrized quantum circuits that are classically optimized. Thanks to its flexibility and hybrid nature, this algorithm offers several potential advantages. First, it can be scalable with a suitable choice of ansatz, thus handling larger systems with a lower classical computational cost, effectively reducing the computational cost associated with the growth of the Hilbert space. Second, the VQE is designed to be compatible with current NISQ devices [9,11,12] and their low quantum volume [8]. This is because it relies on relatively shallow quantum circuits to prepare trial states, and the process of observable measurements is based on primitive sampling. Nevertheless, estimating the energy with this algorithm comes at the cost of lower accuracy compared to quantum phase estimation (QPE) [13], a trade-off that is necessary in the absence of the fault-tolerant quantum computers that are required to deal with the deep circuits of QPE. Different methods have been developed to address this accuracy gap with NISQ computers. Multiple error mitigation [14] and measurement optimization schemes [15] have been proposed to improve the energy estimates during the VQE and its scalability, while alternative methods, such as quantum sampling-based algorithms [16,17], have shown better scalability and a higher resilience to noise.

*Contact author: m.taha.rouabah@umc.edu.dz

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

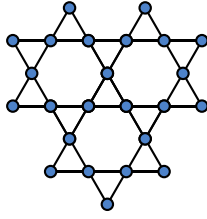


FIG. 1. The geometry of the kagome lattice.

This paper is interested in the spin- $\frac{1}{2}$ isotropic Heisenberg antiferromagnetic model on a kagome lattice, a geometric arrangement of spins that has attracted increasing interest among researchers [18–25] for its exotic magnetic properties and potential to manifest quantum spin liquids [18], or topological phases [18,19] that are still promising candidates to perform quantum computation [26–28]. Due to the frustrated nature of this system, preparing its ground state is nontrivial [18,29].

The kagome lattice is made up of corner-sharing triangles as shown in Fig. 1. When spins are placed on each node, the interaction between spins can be modeled quantum mechanically by the Hamiltonian of the Heisenberg interaction, $\sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j$, where the total spin is a conserved quantity [30]. Although the ferromagnetic regime ($J_{ij} < 0$) is well understood in the literature, the antiferromagnetic regime ($J_{ij} > 0$) is very intriguing since the triangular connections make it impossible to satisfy all constraints simultaneously.

Specifically, we are interested in an isotropic antiferromagnetic spin-1/2 Heisenberg model:

$$H = \sum_{(i,j)} X_i X_j + Y_i Y_j + Z_i Z_j, \quad (1)$$

where (i, j) is a pair of coupled sites i and j . The Hamiltonian forms a commuting set of operators with $\mathbf{S}^2$ and S_z , such that $\mathbf{S} = \sum_i \mathbf{S}_i$, where $\mathbf{S}_i = (S_{x_i}, S_{y_i}, S_{z_i})^T$ is the total spin vector operator and $S_z = \sum_i S_{z_i} = \frac{\hbar}{2} \sum_i Z_i$ is the z component of the total spin vector. That is, $H = \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j$, such that $J_{ij} = 4/\hbar^2$ if i and j are nearest neighbors to each other and $J_{ij} = 0$ otherwise. Moreover, the Hamiltonian along with the total spin operator commutes with the lattice symmetry operations. Generally, the Hamiltonian commutes with $\mathbf{S}$, as shown in Appendix C. Therefore, the Hamiltonian commutes with the $S_{\pm}$ ladder operators, and it is SU(2) invariant.

For a Heisenberg model in a lattice of simply connected triangles, where it is possible to form a singlet state in each triangle, we may show, following the same arguments as Shastri and Sutherland [31], that a dimer state $|\Psi_{\text{dimer}}^n\rangle$ is a ground state. The latter is defined as follows:

$$|\Psi_{\text{dimer}}^n\rangle = \bigotimes_{(i,j) \in \mathcal{D}_n} [i, j], \quad (2)$$

where $[i, j] = \frac{1}{\sqrt{2}}(|\uparrow_i \downarrow_j\rangle - |\downarrow_i \uparrow_j\rangle)$ for each neighbor pair (i, j) in a dimer configuration $\mathcal{D}_n$ from the set of all possible dimer configurations $\mathcal{D}$.

Examples of simply connected triangles are the simple one triangle, two connected triangles, one kagome star, and a chain of kagome stars. However, the argument fails for an arbitrary

kagome lattice, such as three overlaid kagome stars, as shown in Fig. 1, because the configuration becomes geometrically frustrated and the condition of forming a singlet state in each triangle is violated. For that reason, the resonating valence bond (RVB) spin liquid state is proposed, which features short-range correlation and gapped excitations [32]. However, numerical studies showed signs of long-range correlations with algebraic decay and gapless spectrum [33]. For instance, using the Gutzwiller projected wave function, it has been shown that U(1) Dirac spin liquid has better agreement with exact diagonalization for 48 sites [34].

This system's general ground state is not fully understood yet, especially in the thermodynamic limit, i.e., an infinite lattice. Interestingly, numerical studies show very distinct results; initial numerical approaches of the density-matrix renormalization group (DMRG) showed a gapped ground state of $\mathbb{Z}_2$ spin liquid with short-range entanglement [35], while recent DMRG studies hint at algebraic states such as the U(1) Dirac spin liquid [36] or chiral spin liquid [37]. Moreover, Liao *et al.* [38] used the projected entangled simplex states to show that the ground state exhibits gapless quantum spin liquid in the infinite lattice limit. Exact diagonalization methods showed that the ground state differs from the theoretical candidates, at least in the small finite sizes of 36 and 48 spin sites [39]. Recently, a physically oriented neural network ansatz has been employed for lattices of 48 sites and showed signs of a spinon pair density wave that does not break time reversal symmetry or any of the lattice symmetries [40].

Quantum experiments on the kagome lattice are still very limited. Theoretical proposals and classical simulations have been carried out. Some studies are devoted to optimizing the implementation of the Hamiltonian variational (HV) ansatz [20,21], while some studies like Guo *et al.* [23] utilized the HV ansatz and showed phase transitions when considering different ratios of second-nearest-neighbor interactions to first-nearest-neighbor ones. Although the HV ansatz is physically inspired, it is still too deep for NISQ devices.

To motivate our approach, we compare it to recent state-of-the-art efforts that explore similar directions from different starting points. Sajjan *et al.* [22] follow a hybrid classical-quantum pipeline: They begin with DMRG computations on a reduced Hamiltonian to extract correlation data, then design an ansatz to replicate those features, and finally simulate and implement the VQE circuit, complemented by error mitigation techniques like the twirled readout-error extinction and the zero-noise extrapolation (ZNE). However, their implementation ultimately reduces the whole 12-qubit problem to an effective two-qubit ansatz. Although this reduction is well motivated, it does not offer a useful benchmark for our setting, where our objective is to evaluate the VQE procedure under more problem-agnostic and less informed conditions. Conversely, we design a generic ansatz based on the hardware connections and native gates, then adapt it to give at least one ground state, and analyze it with the Fubini-Study metric to avoid zero metric or undefined geometry in the energy landscape. Despite the presence of noise, we are able to characterize the resulting quantum state and extract structural information. Notably, our optimization starts from product states of high energy and still converges, highlighting the robustness of the variational landscape and the optimization

strategy. Weaving *et al.* [24] use a symmetry-based qubit reduction strategy, identifying exact and approximate $\mathbb{Z}_2$ symmetries to taper the Hamiltonian down to five qubits. After validating the spectrum via DMRG, they apply VQE with extensive error mitigation, including the readout-error mitigation (REM), multifold ZNE, and symmetry verification of measurement outcomes. While their method excels at preserving energy estimates under strong qubit compression, its scope is restricted to the energy alone. Our aim, in contrast, focuses on the structure and correlations of the quantum state itself, which is essential for understanding ordering, entanglement, and potential quantum phases beyond the spectral analysis. Semeghini *et al.* [19] follow a different experimental direction by preparing the RVB states using Rydberg atom arrays. They encode each edge of the kagome lattice as an atom and use adiabatic evolution with the blockade mechanism's constraints to prepare dimer coverings, enabling the study of topological properties and quasiparticle excitations. Despite the scalability of their platform and its suitability to probe RVB physics, it is confined to a specific candidate state and does not address the general solution of the Heisenberg model.

In this work, we focus on the ability to characterize the ground state in current quantum devices. We will show that it is possible to probe the spin-spin correlations with sufficient accuracy despite noise. To prepare the ground state, we design a custom shallow hardware-efficient ansatz targeting two systems: a single triangle and six corner-sharing triangles making one kagome star—geometries whose ground states correspond to simple dimer coverings and are thus product states. We employ the quantum natural gradient descent (QNGD), an optimization technique designed to linearly transform gradient updates using the Fubini-Study metric tensor [41], thereby accounting for the local geometry of the parameter space. Recently, it has been shown that the QNGD method is able to outperform the vanilla gradient descent method on Rydberg atoms and superconducting circuit platforms using realistic simulations [42]. To further enhance convergence, Atif *et al.* [43] developed the adaptive QNGD (AQNGD) that combines QNGD with a backtracking line search to dynamically adapt the step size. Instead of explicitly implementing the QNGD optimization, we construct the ansatz such that the Fubini-Study metric is implicitly diagonal and constant. Because of this, the AQNGD is reduced to the adaptive gradient descent (AGD) and any optimizers used with this ansatz would be equivalent to its quantum natural implementation.

In addition to this setting, error mitigation with low post-processing cost is applied at the end of the VQE to improve the final ground energy estimate. We implement the REM to reverse the state preparation and measurement (SPAM) errors using a partitioned response matrix approach [44]. Rather than reconstructing the full response matrix, we apply the partitioned inverse response matrices directly to the measured bitstrings. For sparse VQE measurements where $M \ll 2^n$ unique bitstrings are observed (with n qubits), this approach scales as $O(M \times 2^n)$ in the worst case, with the possibility of truncation-based enhancement, making an improvement over the work of Geller and Sun [44] that requires $O(2^n \times 2^n)$ full matrix multiplication for postprocessing. This is detailed in Appendix E.

Beyond the energy estimation, we are able to characterize the dimer state by measuring the spin-spin correlations and the static structure factor, which exhibits resilience to noise. We provide the static structure factor map with and without REM and show a plausible agreement with the ground-state static structure factor.

In the following Sec. II, we outline the quantum natural gradient-based optimization method and its implementation in Sec. II A, and the design of the hardware-efficient Euclidean ansatz circuit in Sec. II B. The grouping of Pauli terms for efficient energy measurement along with measurement variance and standard error expectation is detailed in Appendix D. Readout-error mitigation and the zero-noise extrapolation methods are detailed in Appendixes E and F, respectively. The results of experiments on one triangle and a single kagome star are detailed in Sec. III, while the main discussion of findings can be found in Sec. IV, and the conclusion in Sec. V.

II. METHODS

In general, the parameters of the quantum circuit in variational quantum algorithms are optimized similarly to other classical optimization problems in machine learning models. Nevertheless, the quantum nature of the circuit imposes some genuine challenges. To begin with, the quantum measurement outcome is intrinsically probabilistic even in the presence of a noise-free ideal quantum computer. Moreover, current quantum devices suffer from additional noise sources coming from coherent and incoherent errors, state preparation errors, and measurement errors. In addition, the ansatz requires a significant number of parameters to cover the Hilbert space region that includes the ground state.

The cost to estimate the gradient in gradient-based optimizers is proportional to the number of parameters d ; it is still possible, though, to optimize the number of shots for gradient estimation using the parameter-shift rule [45,46]. Finally, the cost function landscape is complex. Besides local minima, variational quantum algorithms can also face the problem of barren plateaus, when the gradient magnitude vanishes exponentially with the system's size [47].

Given these properties, the VQE algorithm must be designed with careful considerations of the hardware capabilities and the problem's characteristics to design a well-suited ansatz-optimizer pair.

A. Adaptive quantum natural gradient descent

Searching for the correct parameters that represent the ground state is the central challenge for the classical part of VQE. For this task, various optimizers have been used, such as the normal gradient descent and stochastic perturbative ones. However, they do not take into account the geometry of the space of parametrized quantum states, called the Kähler manifold or the projective quantum space, considering it to be represented by a Euclidean space that is free of singularities. In quantum applications, the space of parametrized quantum states is rarely Euclidean [48,49]; it usually has a curvature featured by the Fubini-Study metric and hosts singularities and covariant parameters, which make the search

for the optimal parameters very challenging [50]. Following the natural gradient approach from classical machine learning [41,51], the QNGD optimizer was proposed as its quantum analog. This optimizer uses the steepest gradient descent in the parameter space with an induced Fubini-Study metric, successfully avoiding singularities and considering parameters' covariance. The transformation from the Euclidean space to the Riemannian parameter space is done by taking the inverse of the Fubini-Study metric tensor, which is the real part of the quantum Fisher information matrix. The Fubini-Study metric can also be used, as we will show below, to decide on an ansatz design with a naturally Euclidean parameter space that is more easily trainable at a lower cost than more complex ones. For more details on the QNGD, see Appendix C and the works of Yamamoto [50] and Stokes *et al.* [41].

Although QNGD trains parameters faster and in the correct convergence direction, it still faces the problem of choosing the ideal learning rate. The algorithm can get stuck around the ground state and never reach the minimum if the geometry becomes tighter than the step size, which is common in entangled systems with a large learning rate [50]. On the other hand, choosing an extremely small step size will significantly slow the algorithm since it will be making really small steps when the geometry does not necessitate it. Choosing a step size that monotonically decreases is not a reliable solution either as it is not geometrically informed and will be heavily problem dependent.

A way around this was presented in Atif *et al.* [43], who used QNGD with a variable, geometrically informed, learning rate following the Armijo rule, which is an efficient backtracking rule with proven convergence. This allows us to adaptively choose a large step size when needed to converge faster without sacrificing the ability to navigate narrow parts of the geometry, thus avoiding stagnation by adapting the step size to the local geometry at each iteration. With this addition, the dependence of the QNGD's performance on the chosen step size is mitigated. Although this method introduces new hyperparameters to set, these are less problematic to choose than the step size. This AQNGD algorithm is detailed in Appendix A.

Notably, the Fubini-Study metric does not mitigate barren plateaus (BP). Because BPs are defined as the exponential decay of the gradient amplitude, merely rescaling the gradient with a metric cannot recover the lost signal. In fact, the gradient amplitude is dictated by the observable's locality, circuit depth, and initial state complexity [47]. Additionally, in the case of noise-induced barren plateaus the information is completely lost; hence, the metric is completely ineffective.

B. Euclidean ansatz

Ideally, the ansatz for a specific physical problem should satisfy three conditions: (1) Expressivity, i.e., the ability of the ansatz to prepare the system's ground state. (2) Hardware efficiency, ensuring that meaningful information survives the inherent quantum processing unit (QPU) noise. (3) Trainability, meaning the parameter space is well behaved so that the classical optimizer can efficiently converge to the optimal parameters. The best physically inspired ansätze designs for

the kagome lattice so far have been based on the Hamiltonian variational ansatz (HVA), which can be viewed as a digital approximation of adiabatic evolution. Park and Kildoran [52] proved that, in the absence of incoherent noise and under specific constraints, the HVA does not suffer from barren plateaus. Bosse and Montanaro [20] optimized the number of layers required for the HVA on QPUs with a square-lattice architecture. Kattemölle and van Wezel [21] proposed an HVA-inspired symmetry-preserving ansatz using dense layers of parametrized Heisenberg interaction gates starting from a dimer state, and further predicted the error rate and circuit depth required to achieve a given fidelity for arbitrary system sizes. A more hardware-friendly ansatz targeting valence-bond states composed of dimer bonds was introduced by Sajjan *et al.* [22]. In this work, our goal is to develop an expressive, trainable, hardware-efficient ansatz that can be implemented on current QPUs without relying on heavy problem-specific reductions, but only hardware constraints. To achieve this, we construct an ansatz based on a simple pattern of parametrized single-qubit rotations combined with entangling gates between physically connected qubits to avoid SWAP gates. Our ansatz, shown in Fig. 2, is inspired by the well-known real amplitude ansatz, but reduces the number of R_y gates. Specifically, we found that placing R_y rotations on every qubit both before and after the entangling layer was unnecessary. Instead, we alternate their placement, splitting them between the two sides of the entangling layer so that each qubit carries only a single R_y rotation. This choice serves to reduce the number of parameters, making the ansatz more trainable, while also mitigating some of the coherent noise coming from the entangling layer through the parametrized gates after it. Although the choice of a ladder CNOT structure increases depth, a more compact brick-wall structure will increase coherent crosstalk noise [53,54].

Beyond this structural simplification, a central feature of our ansatz is that it naturally has a parameter space with a Euclidean geometry. By analytically evaluating the Fubini-Study metric (see Appendix C), we show that it is diagonal and constant for our construction. This means that the parameters vary independently (i.e., no overparametrization), the parameter space is free of singularities, and the gradients obtained via the parameter-shift rule coincide with the true natural gradients. In other words, our ansatz is not only hardware efficient, but also intrinsically trainable. Since the Fubini-Study metric is diagonal and constant, the number condition (i.e., the ratio of the largest eigenvalue to the smallest eigenvalue) is minimum. This implies that the metric is stable as recently shown by Haddou and Bennai [55]. In this setting, AQNGD is equivalent to just adaptive gradient descent and will be referred to as AGD for the rest of the paper. Altogether, this ansatz fulfills the three conditions outlined above: It is expressive, hardware efficient, and intrinsically trainable, thanks to its Euclidean parameter space. Our results confirm its expressivity, as it successfully prepares the ground state and reproduces physical observables such as the static spin structure factor $S(\mathbf{q})$. Looking forward, this construction opens the way for further investigations into scaling this family of Euclidean hardware-efficient ansätze to larger kagome lattices and, more broadly, to other frustrated quantum magnets.

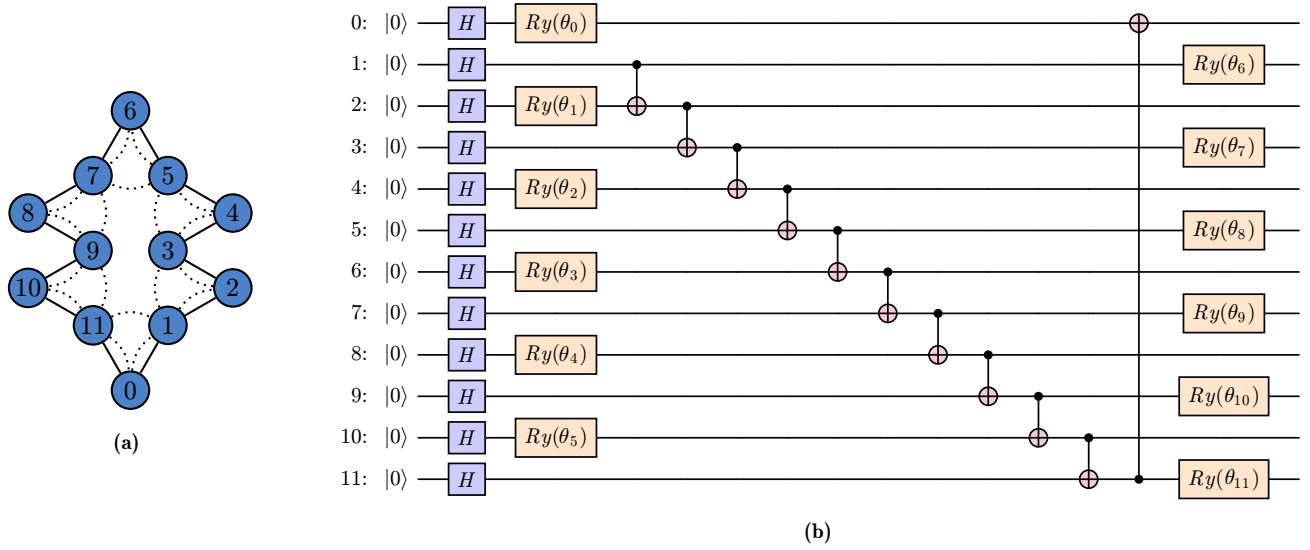


FIG. 2. (a) Hardware qubit connection is shown by solid lines, while Heisenberg interactions are shown in dotted lines. (b) The ansatz circuit used to prepare a trial wave function for the kagome star case. This ansatz is hardware compatible: It is linearly entangling, restricted to the hardware connectivity, and it is low depth with only one entangling layer and 12 parametrized rotation gates.

III. EXPERIMENTAL RESULTS

In this section, we expose the variational quantum eigen-solving for the ground-state experiments on quantum devices. The purpose is to test the performance of optimizers and quantum error mitigation techniques in real situations. First, we show the experiments on three qubits for a system of three spin-1/2 sites. Then, we show experiments for 12 Heisenberg interacting spin-1/2 sites connected in one kagome star.

A. Antiferromagnetic Heisenberg interacting 1/2-spin triangle

Antiferromagnetic Heisenberg interacting spins in one triangle have three degenerate ground states whose energy value is $E_{\text{GS}}^{\Delta} = -3$. We show in the upper plot of Fig. 4, the variational energy as evaluated on IBM Oslo (7 qubits) and IBM Kyoto (127 qubits) using the AGD optimization that is able to smoothly converge on three linearly connected qubits. The VQE algorithms have been implemented for 19 iterations on IBM Oslo. Moreover, we performed an additional experiment on a newer quantum hardware platform, IBM Kyoto, using a limited number of iterations sufficient to achieve satisfactory convergence. The native gates of both QPUs are detailed in Appendix G. The two plots show a very close behavior, even

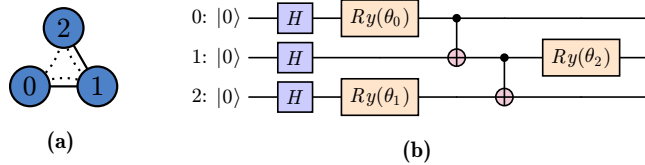


FIG. 3. (a) Hardware qubit connection is shown by solid lines, while Heisenberg interactions are shown in dotted lines. (b) The ansatz circuit used as trial wave function in the triangle case. This ansatz is hardware compatible: First, it is linearly entangling, following the native hardware connectivity. Second, it is low depth with only one entangling layer and three parametrized rotation gates.

though the initial ansatz parameter vector $\theta_{\text{init}}^{\Delta}$, in the same index order as Fig. 3, is randomly picked. Namely,

$$\theta_{\text{init}}^{\Delta} = (0.2427\pi, \quad 0.2510\pi, \quad 0.1233\pi)$$

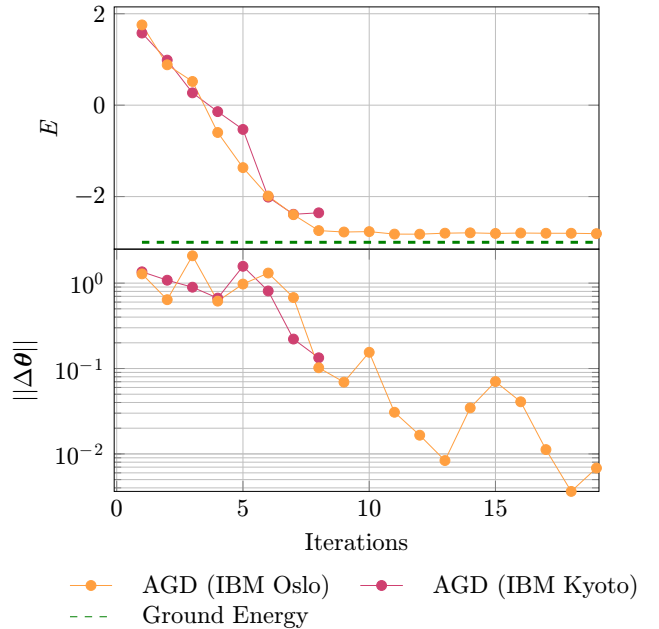


FIG. 4. The upper plot displays the VQE convergence graph by iterations for the Heisenberg model on a triangle lattice, run on the IBM Kyoto 127-qubit and IBM Oslo 7-qubit QPUs, using the AGD optimization method. The REM and ZNE quantum error-mitigation techniques are not applied during the VQE optimization. The lower plot represents the step-size magnitudes $\|\Delta\theta_i^{\Delta}\| = \|\theta_{i+1}^{\Delta} - \theta_i^{\Delta}\| = (\beta/2^{k_i})\|\nabla\theta_i^{\Delta}\|$, which are determined by the product of the gradient-vector norm and the learning rate $\beta/2^k$ set by the Armijo adaptive number k .

are given for IBM Kyoto experiment, while

$$\theta_{\text{init}}^{\Delta} = (0.2914\pi, \quad 0.2812\pi, \quad 0.1266\pi)$$

for IBM Oslo experiment. It is worth recalling that the Fubini-Study metric is constant due to the structure of the quantum circuit, such that $g = 0.25 \cdot I_{3 \times 3}$, as detailed in Appendix C. The step size is adapted using line search backtracking as described in Sec. II A. The number of executions k required to find the best step size in each iteration is plotted in Fig. 8(a). The lower graph in Fig. 4 shows clearly the working mechanism of the learning rate adaptation due to the Armijo's rule fulfillment. In the beginning, the initial large step size is enough to start exploration, and then it tends to slow down reciprocally with the increasing Armijo's adaptive number k value to make the step size adaptive to the cost function landscape. We notice that the step size continues decreasing in IBM Oslo experiment due to diminishing quantum natural gradient norm and the flexible learning rate. Overall, the convergence of the AGD optimized VQE energy curves in Fig. 4 is remarkable. Even though the experiment is small and the quantum circuit is shallow, it showcases the low noise rate of the real quantum devices and the trainability of the ansatz using quantum natural gradient descent and Armijo's adaptive method.

We note that the final trained ansatz parameter vector $\theta_{\text{final}}^{\Delta}$ is given by

$$\theta_{\text{final}}^{\Delta} = (0.2306\pi, \quad 1.5005\pi, \quad 1.0031\pi) \quad (3)$$

for IBM Kyoto, in the same index order as in Fig. 3, and

$$\theta_{\text{final}}^{\Delta} = (0.3111\pi, \quad 1.5240\pi, \quad 1.0144\pi) \quad (4)$$

for IBM Oslo experiment, which are very close to the exact parameter values $\theta_1^{\Delta} = 3\pi/2$, $\theta_2^{\Delta} = \pi$, while θ_0^{Δ} is free to have any value. Typically, when $\theta_0^{\Delta} = \pi/2$, we can reproduce one of the theoretical ground states $|\psi_{\text{GS}}^{\Delta}\rangle$, which is

$$|+\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle) = |+\rangle \otimes [1, 2]. \quad (5)$$

That is nothing but a dimer state, with the fidelity value $|\langle \psi_{\text{SVS}}(\theta_{\text{final}}^{\Delta}) | \psi_{\text{GS}}^{\Delta} \rangle|^2 = 99.89\%$ on the state-vector simulation (SVS) and an error of $\Delta E = E_{\text{SVS}}(\theta_{\text{final}}^{\Delta}) - E_{\text{GS}}^{\Delta} = 9 \times 10^{-5}$ using the VQE-optimized parameters in Eq. (3) from IBM Kyoto.

Nevertheless, the exact parameters are not enough if the quantum state cannot be experimentally characterized. To enhance the accuracy of the energy measurement of the last state, error mitigation (EM) techniques are explored. We recall that EM was not used during the VQE procedure. The values of the energy prior to and subsequent to EM implemented on the IBM Kyoto machine are plotted in the upper plot of Fig. 5. The unmitigated value is found to be -2.3634 ± 0.0110 in the normalized units, while the readout-error mitigation using the full calibration matrix gives -2.9423 ± 0.0034 with straightforward REM, and -2.8917 ± 0.0046 with positive probabilities preserving REM, both showing better accuracy. Applying ZNE error mitigation on top of REM using quadratic extrapolation exceeds the ground energy, as shown in Fig. 5. On the other hand, applying ZNE on the raw energy values in the first folding points yields the values

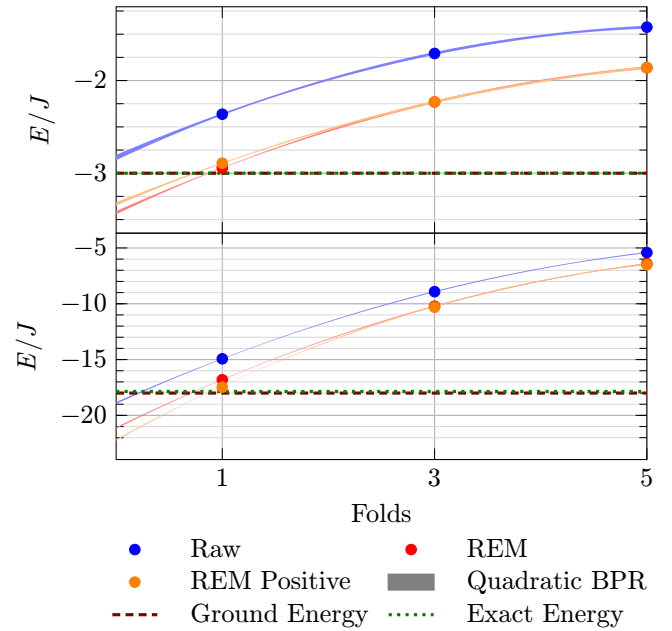


FIG. 5. Benchmarking quantum error mitigation using ZNE and REM techniques as described in Appendixes F and E, where Bayesian polynomial regression (BPR) is used. Readout and zero-noise error mitigation evaluated on the IBM Kyoto of the ground-state energy measurement of antiferromagnetic spin-1/2 Heisenberg model on one triangle (top), and on one kagome star in IBM Torino (bottom).

-2.8306 ± 0.0286 for quadratic extrapolation, which is the minimal polynomial that fits the three points. It is observed that the REM and the ZNE work better separately than in combination. These results will be discussed in Sec. IV.

Remarkably, the detection of local dimers can be enhanced with the REM method, as shown in Table I. The error mitigation makes the gap clear between spin correlations of different pairs when the dimer bond is present or absent. This observation fits qualitatively the theoretical expectation from the exact ground state, while quantitative closeness is enhanced through REM.

B. Antiferromagnetic Heisenberg interacting 1/2-spin kagome star

The antiferromagnetic Heisenberg model of 1/2 spins on a kagome star has two possible dimer ground states,

$$|\psi_{\text{GS}}^1\rangle = [1, 2] \otimes [3, 4] \otimes \cdots \otimes [11, 0], \quad (6)$$

TABLE I. Correlation terms between spins of each edge pair in the triangle.

Edges (i, j)	Correlation $\langle S_i \cdot S_j \rangle$				Exact
	IBM Oslo		IBM Kyoto		
	Raw	REM	Raw	REM	
(0, 1)	−0.0756	−0.0636	0.1476	0.0701	0
(1, 2)	−2.7828	−2.9902	−1.7740	−2.8604	−3
(2, 0)	0.0472	0.0440	−0.7288	−0.1434	0

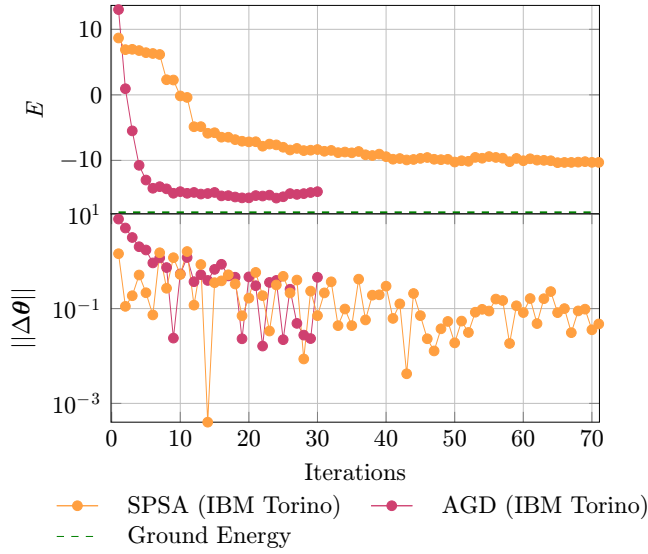


FIG. 6. The upper plot shows the VQE convergence graph per iteration as illustrated in terms of energy variation for the Heisenberg model on a single kagome star lattice run on IBM Torino 133-qubit QPU with the AGD and SPSA optimizers. Notice that AGD requires more quantum circuit executions per iteration than SPSA to evaluate the gradient (one execution for each the 12 parameters) and to adapt the learning rate, which is reflected in the number of executions k in Fig. 8(b). However, most of those executions are independent of each other and hence parallelizable. The lower plot represents the step-size magnitudes $\|\Delta\theta_i\| = \|\theta_{i+1} - \theta_i\|$, which are determined by the product of the gradient-vector norm and the learning rate $\beta/2^k$ set by the Armijo adaptive number k in the AGD case.

which is the unique dimer state that is expected from the ansatz in Fig. 2. The other state should be

$$|\psi_{GS}^2\rangle = [0, 1] \otimes [2, 3] \otimes \cdots \otimes [10, 11],$$

where the qubits are labeled as shown in Fig. 2. Both states correspond to the ground energy $E_{GS} = -18$, since the kagome star is constructed by simply connected triangles, as discussed in Sec. I. The ansatz depicted in the quantum circuit in Fig. 2 targets only the first state. Although this introduces a physical bias in the ansatz, it serves as a benchmark to the quantum device, allowing us to assess the performance of a quantum computer on such problems in the NISQ regime. To assess the benefits of AGD's adaptivity and its exact gradient, we compare it with simultaneous perturbation stochastic approximation (SPSA).

The VQE convergence graph is depicted in Fig. 6 for both SPSA and AGD optimizers. The final components of the trained ansatz parameter vector θ_{final} , in the same index order as Fig. 2, for the AGD optimization experiment are given, up to a global phase, by

$$\begin{pmatrix} \theta_0 = -0.3962\pi & \theta_6 = +1.0056\pi \\ \theta_1 = -0.5324\pi & \theta_7 = +1.0190\pi \\ \theta_2 = +1.4827\pi & \theta_8 = +0.9836\pi \\ \theta_3 = -0.4658\pi & \theta_9 = -0.9768\pi \\ \theta_4 = -0.4879\pi & \theta_{10} = +1.0007\pi \\ \theta_5 = +1.4993\pi & \theta_{11} = -0.9830\pi \end{pmatrix} \approx \begin{pmatrix} -\pi/2 & \pi \\ -\pi/2 & \pi \\ -\pi/2 & \pi \\ -\pi/2 & \pi \\ -\pi/2 & \pi \\ -\pi/2 & \pi \end{pmatrix}, \quad (7)$$

where the right side defines the converged exact ansatz parameters of the ground state. Namely, according to the ansatz in Fig. 2, the parametrized state $|\psi(\theta_{\text{final}})\rangle$ should yield the following state $|\psi_{GS}^1\rangle$.

Using the parameters in Eq. (7), we get the fidelity value $|\langle\psi_{SVS}(\theta_{\text{final}})|\psi_{GS}\rangle|^2 = 96.37\%$ and $\Delta E = E_{SVS}(\theta_{\text{final}}) - E_{GS} = 0.1463$ in the state-vector simulator when the result is compared with the ground state and the ground energy, respectively.

To retrieve more accurate information from the quantum results, we exploit the REM and the ZNE techniques. First, we show in Fig. 5, the results of error mitigation techniques. The raw energy value is -14.9316 ± 0.0265 , which is around three units away from the theoretical value. However, with REM only, we get -16.8028 ± 0.0195 and -17.3675 ± 0.0123 for straightforward REM and positivity preserving REM, respectively. Using ZNE only with Bayesian quadratic extrapolation gives -18.8826 ± 0.0790 . On the other hand, the combination of REM and ZNE undershoots beyond -19 , as detailed in Table II. These results are consistent with the similar behavior of REM and ZNE in the triangle case in Sec. III A.

For more insight, we run two different error mitigation experiments on the IBM Torino and IBM Algiers quantum computers. The results are detailed in Table II. In these experiments, we use parameters from a classically simulated VQE. Even though REM is not consistently well performing as in IBM Algiers, it provides an upper bound especially when the positivity of sampling counts is retained. On the other hand, we observe again the matching between the ZNE with Bayesian quadratic extrapolation and the theoretical value. This observation is also consistent with the ZNE experiment in the triangle case.

Noticeably, the spin correlation structure can be effectively characterized even with the noisy backend, as shown in Fig. 7. Here, the static spin structure factor is computed based on the measurement of each spin-to-spin correlation, and which is defined as

$$S(\mathbf{q}) = \frac{1}{N} \sum_{ij} e^{i\mathbf{q} \cdot (\mathbf{r}_i - \mathbf{r}_j)} \langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle, \quad (8)$$

where the summation goes over all spins in sites $\mathbf{r}_i$ for $i = 1, \dots, N$, and $\mathbf{q}$ is the momentum vector in the reciprocal lattice. The static structure factor reveals the signature of the magnetization order of the system. This is the Fourier transform of the equal-time spin-spin correlation function, not merely the Fourier transform of the lattice geometry, and thus reflects the correlation patterns between spins in the quantum state. The comparison between the exact and experimental spin structure factor shows that errors in the individual spin correlations are not enough to erase the correlation order signature. Moreover, we exploit REM method to enhance the landscape of the spin structure factor. The Pearson correlation coefficient indicates similarities of 99.97%, 99.97%, and 99.94% with the exact $S(\mathbf{q})$ for the simulated noisy results, the QPU results, and the REM-mitigated QPU results, respectively. The corresponding mean-square errors are 0.013, 0.042, and 0.007 in the same order. Hence, REM balances between accuracy and qualitative pattern similarity.

TABLE II. (a) Results for the AGD-driven VQE for one kagome star performed on IBM Torino QPU as shown in Fig. 6 and mitigated using REM and ZNE as shown in Fig. 5. (b) IBM Algiers ground-state measurement and mitigation for one kagome star. (c) IBM Torino ground-state energy measurement and mitigation for one kagome star. The parameters of (b) and (c) experiments are obtained from classically simulated VQE.

Method	Energy value $\pm \epsilon$ (E/J)		
	IBM Torino (a)	IBM Algiers (b)	IBM Torino (c)
Unmitigated	-14.9316 ± 0.0265	-10.9748 ± 0.0343	-14.6346 ± 0.0260
REM	-16.8028 ± 0.0195	-12.4031 ± 0.0306	-16.2172 ± 0.0199
REM + positivity	-17.3675 ± 0.0123	-12.1712 ± 0.0314	-16.0097 ± 0.0202
Polynomial deg = 1 ZNE	-17.2644 ± 0.0355	-12.8618 ± 0.0441	-16.3731 ± 0.0348
Polynomial deg = 2 ZNE	-18.8826 ± 0.0790	-18.0060 ± 0.0922	-17.9740 ± 0.0780
REM + polynomial deg = 2 ZNE	-21.1330 ± 0.0705	-20.3412 ± 0.0877	-19.8208 ± 0.0705
REM + positivity + polynomial deg = 2 ZNE	-22.2032 ± 0.0640	-19.9411 ± 0.0888	-19.4483 ± 0.0707
Exact energy	-17.8533	-17.9998	-17.9998
Ground energy		-18.0000	

IV. DISCUSSION

Here, we will discuss the ability of AGD to prepare the ground state, the need to measure noise resilient properties such as the static structure factor, and we discuss further the enhanced energy values using EM.

In Fig. 6, we can clearly see the superior optimization performance of AGD over SPSA in real, noise-affected, quantum

devices. AGD requires fewer iterations to converge compared to the inherently noise-adapted SPSA method. Even though the number of circuit executions is larger than what is required by an SPSA implementation, scaling linearly in the number of parameters, AGD can be boosted by distributed computation in each iteration and is guaranteed to converge and thus more scalable than zeroth-order stochastic optimization methods [41,43]. To assess the role of adaptation in the AGD performance, it is sufficient to check the profile of the step size $\|\Delta\theta\|$ during the optimization. The lower plot of Fig. 6 showcases the comparison between the magnitudes of the step sizes taken by AGD and SPSA. First, we notice that the profile of Armijo's adapted step size shows qualitatively similar evolution pattern as SPSA with exponentially decaying learning rate [10]. Second, the Armijo rule allows to take arbitrary larger initial steps to converge faster as shown in the same figure, while adapting the learning rate to fine-tune when needed. These observations are supported by simulation experiments in Figs. 8 and 9.

In addition, as shown in Secs. III A and III B, despite the significant noise-induced gap in the energy of the ground state, the optimized parameters show impressive accordance with the exact ground-state parameters without any reliance on error mitigation techniques. This observation can be quantified by evaluating the fidelity and energy values using the state-vector simulation of the ansätze. Namely, the state-vector fidelities are 99.89% in the case of one triangle and 96.37% in the case of one kagome star, while energy gaps are $\Delta E = 9 \times 10^{-5}$ and $\Delta E = 0.1463$ for the triangle and the kagome star, respectively.

Although this observation highlights the promising capabilities of the VQE protocol to prepare accurate states despite errors [10,56], it becomes impractical and not scalable if such accurate states are inaccessible by quantum computing means. Our results show that characterizing the dimer state using spin correlation and spin structure factor is possible even without error mitigation. Moreover, it is possible to enhance the state's spin structure resolution through error mitigation means such as REM, as shown in Fig. 7. This result is in agreement with the recent work of Lively *et al.* [57] who showed that the prediction of the energy derivative, two-site spin correlation

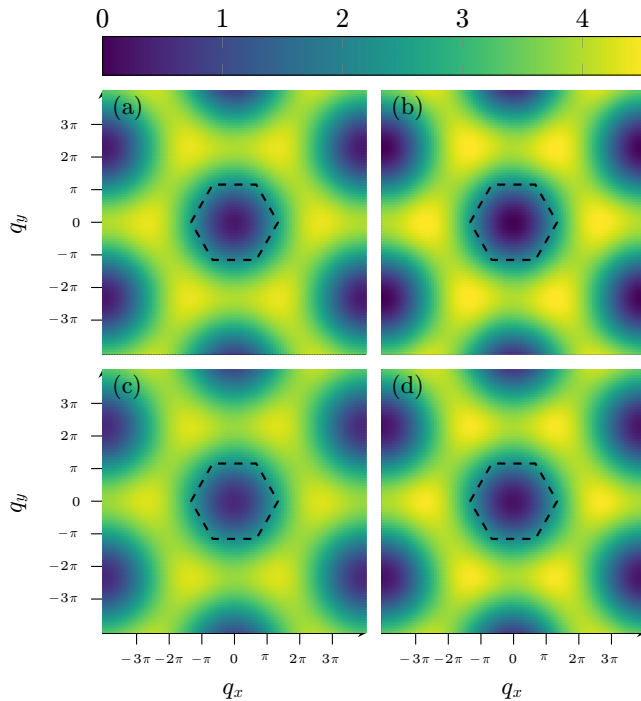


FIG. 7. Spin-spin correlation static spin structure factor $S(\mathbf{q})$ (the Fourier transform of the equal-time spin-spin correlation function) as computed from the measurement results of noisy simulation of IBM Torino device [panel (a)], ideal simulator with shot noise [panel (b)], real IBM Torino QPU without error mitigation [panel (c)], and with REM [panel (d)]. The dashed hexagon draws the border of the Brillouin zone. The suppression at zero momentum indicates the absence of ferromagnetic order.

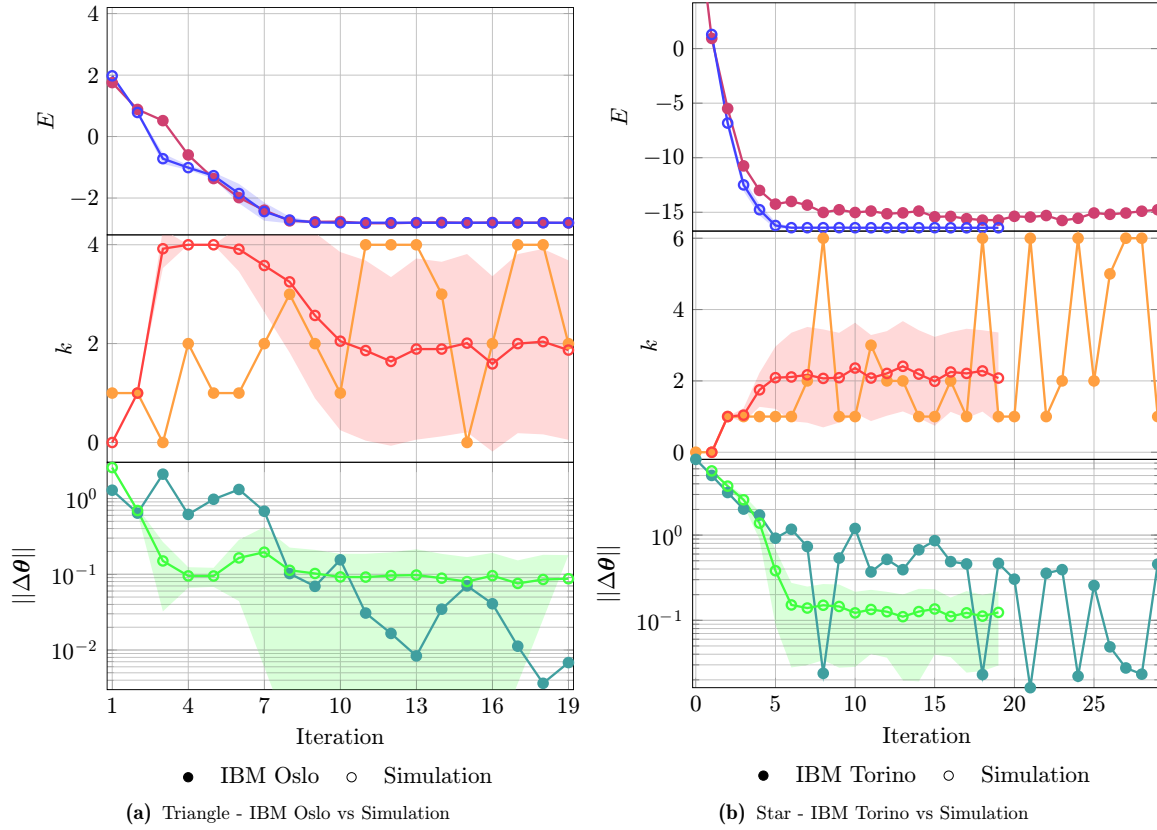


FIG. 8. AGD optimization results for the Heisenberg model on the triangle lattice comparing IBM Oslo QPU and simulation of the VQE experiment on IBM Oslo noise model over 100 runs [panel (a)]; and on one kagome star lattice comparing IBM Torino QPU and simulation of the same VQE experiment on IBM Torino noise model over 100 runs, starting from the same initial parameters [panel (b)]. The step norm $||\Delta\theta|| = \beta/2^k ||\nabla\theta||$ in both cases.

functions, and the fidelity susceptibility from VQE results can be accurate even without error mitigation. For instance, local observables are less sensitive to noise than nonlocal ones [58].

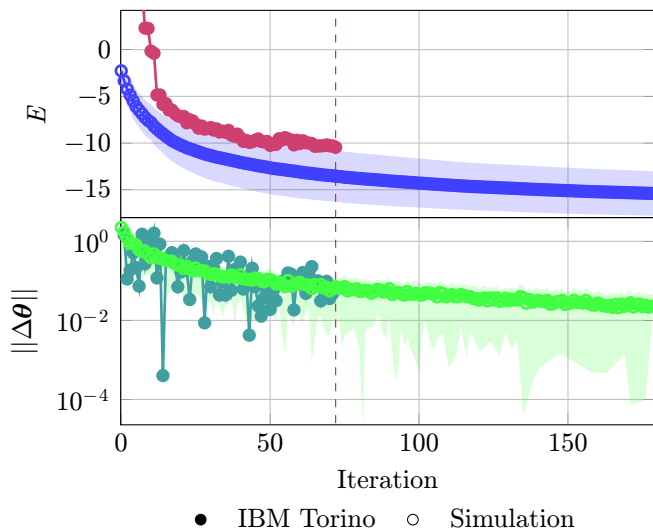


FIG. 9. SPSA optimization results for the Heisenberg model on one star kagome lattice comparing IBM Torino QPU and simulation of the same VQE experiment on IBM Torino noise model over 100 runs, starting from the same initial parameters. The step norm $||\Delta\theta_i|| = ||\theta_{i+1} - \theta_i||$.

One of the main results of this work is that the spin-spin correlation can be noise resilient. Noise resilience is in the sense that it is steady under some decoherence noise threshold. To demonstrate that, we conduct many simulation experiments where we measure the spin-spin static structure factor (SSF) in different noise levels, as shown in Fig. 10. As the noise level increases, the SSF similarity error does not exhibit linear or polynomial scaling. Instead, it displays a steplike, sigmoid-shaped dependence, characterized by two quasi-plateau regimes separated by a sharp crossover region. In the low-noise regime, the SSF remains nearly constant and close to the ideal value, indicating robustness of the correlation structure. Beyond a critical noise threshold, the SSF rapidly transitions to a second plateau where correlations are substantially degraded. Although the transition is continuous, its sharp crossover resembles an approximate phase transition between two distinct correlation regimes. Notably, the width of the crossover coincides with the region where the SSF remains physically meaningful. In contrast to the SSF behavior under noise, the energy exhibits a monotonic decay toward the mixed-state value as noise increases. This inverse trend highlights that while energy degrades gradually under noise, the SSF retains structural integrity up to a threshold and then undergoes a rapid qualitative change, emphasizing the nontrivial sensitivity of correlation-based observables to noise amplification. We should direct the attention of the reader to the fact that the noise strength is arbitrary and

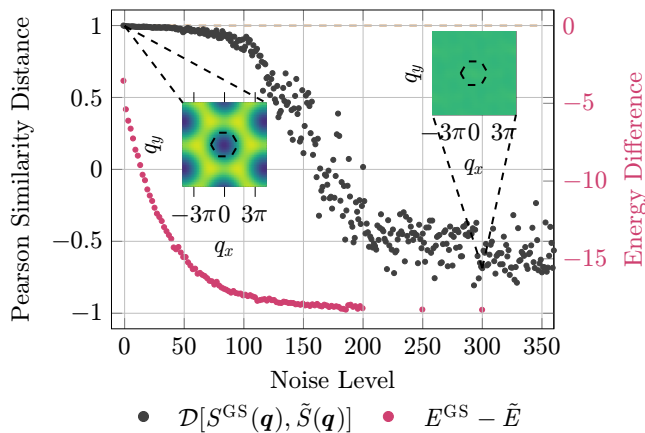


FIG. 10. Pearson similarity distance $\mathcal{D}[S^{\text{GS}}(\mathbf{q}), \tilde{S}(\mathbf{q})]$ between the spin-spin correlation SSF of the prepared state under noise $[\tilde{S}(\mathbf{q})]$ and the ideal spin-spin correlation SSF $[S^{\text{GS}}(\mathbf{q})]$ as a function of noise level. The noise strength is introduced by controlled folding of arbitrary one- and two-qubit gates, in this case, the sequence $(H \cdot CX \cdot CX \cdot H)$. For comparison, the same plot includes the energy deviation, defined as $E^{\text{GS}} - \tilde{E}$, where $E^{\text{GS}} = -18$ is the ground-state energy and $\tilde{E}$ is the energy of the same state under incoherent noise.

unitless. It is introduced by controlled folding of one- and two-qubit gates following the sequence $(H \cdot CX \cdot CX \cdot H)$, effectively amplifying the underlying hardware noise without altering the target unitary. It is important to say that such noise resilience is expected since the spin-spin correlation is featured by a local dimer formation energy gap that protects the SSF.

On the other hand, we should emphasize that a mere two-site descriptor is not sufficient to distinguish between arbitrary two pure quantum states. Namely, Bosse and Montanaro [20] have shown that spin-spin correlation structure can be very close to the ground state while the state fidelity is very low. Nevertheless, they showed also that VQE states can represent these local observables well even when they have not yet reached high global fidelity.

Finally, EM techniques are of great importance in this context. The results of error mitigation implemented on the last ansatz obtained from the VQE for one kagome star are detailed in Table II. Although the results are strongly stochastic, the accuracy is enhanced using different EM techniques. First of all, we notice that REM alone fails in the IBM Algiers' case to reach a significant accuracy despite its success in other cases. The combination of REM and ZNE overshoots in most cases. Impressively, ZNE with quadratic extrapolation obtained repeatedly the best accuracy across different experiments, as shown in Fig. 5 and Table II. The combination of REM and ZNE overshoots the ground energy, which is consistent with the results of Kurita *et al.* [59] that show that ZNE yields a better accuracy when combined with randomized compilation, to suppress coherent noise, in the absence of SPAM errors. In addition, ZNE without REM may go below the exact energy, but to a lesser extent. This is in part due to coherent noise channels shifting the measured energy values in a nonstochastic way, which can potentially be captured by the ZNE fitting functions. A careful study, currently the focus

of an ongoing work by the authors, is necessary to provide a more solid explanation to this mechanism. In the end, it is important to recall that REM does not break the variational principle as explained in Appendix E, while there is no known guarantee for this with ZNE.

V. CONCLUSION

In this work, we focused on the preparation of the ground state of kagome lattice fragments on real quantum devices, showing that this can be achieved with shallow circuits using the VQE. Remarkably, this holds even in the presence of significant noise and without applying error mitigation. While noise limits the quantitative accuracy of energy estimation, essential qualitative features—such as the spin correlation structure of the dimer state—remain accessible, owing to the energy gap that protects certain observables from noise effects. We exploited the Fubini-Study metric as a tool to analyze the geometry of the ansatz and designed it to have a Euclidean parameter space (i.e., constant diagonal Fubini-Study metric), which allows for a smooth and easy optimization without facing singularities. Building upon this, we believe that this approach of carefully designing the ansatz's geometry holds significant potential. We posit that such a design methodology can be exploited to design problem-tailored ansätze with analytically computable Fubini-Study metrics, which leads to expressive, hardware-efficient, and intrinsically trainable ansätze at low costs. This suggests that this kind of ansätze architecture could enable increasingly accurate and robust preparation of complex ground states for frustrated systems as quantum hardware continues to progress, including larger kagome lattices and quantum spin liquids. We analyzed the performance of the AGD, where backtracking search is used, in comparison with SPSA, showing that AGD achieves convergence in fewer iterations while maintaining competitive run-time, with implicit Euclidean parametrized space of ansätze. We showed that the adaptive feature through backtracking search enables faster convergence and less dependency on the starting point. Our results further reveal that error mitigation techniques can substantially improve the accuracy of observable estimates: REM preserves the variational principle, providing reliable upper bounds, while ZNE does not respect this bound but often achieves the highest accuracy, suggesting potential for future studies. We also showed that certain observables, such as the static spin structure factor, are less sensitive to noise than the Hamiltonian, suggesting a good characterization of some of the ground-state properties despite the challenges of the quantum hardware.

Finally, we highlight that solving device-native problems such as the one kagome star ground state on heavy-hex architectures not only enables realistic benchmarking but also paves the way for scaling to larger kagome lattices. Future work will focus on extending these methods to characterize the ground state of larger quantum spin liquid systems.

ACKNOWLEDGMENTS

This document has been produced with the financial assistance of the European Union, Directorate-General for

International Partnerships (Grant No. DCI-PANAF/2020/420-028), through the African Research Initiative for Scientific Excellence (ARISE), pilot program. ARISE is implemented by the African Academy of Sciences with support from the European Commission and the African Union Commission. We are grateful to the Algerian Ministry of Higher Education and Scientific Research and DGRST for the financial support. We thank the Quantum Collaborative for their support and access to IBM Quantum Resources. We acknowledge the use of IBM Quantum services for this work. The authors gratefully acknowledge fruitful discussions with Dr. Manas Sajjan.

The contents of this document are the sole responsibility of the authors and can under no circumstances be regarded as reflecting the position of the European Union, the African Academy of Sciences, and the African Union Commission. The views expressed are those of the authors and do not reflect the official policy or position of IBM or the IBM Quantum team.

DATA AVAILABILITY

The QPU data and postprocessing code for ZNE and REM that support the findings of this article are openly available on Zenodo [60].

APPENDIX A: ADAPTIVE QUANTUM NATURAL GRADIENT DESCENT

In the quantum natural gradient descent, the parameters are updated according to the gradient value, which is measured using the parameter-shift rule method that gives the Euclidean gradient, multiplied by the inverse of the Fubini-Study block diagonal metric, which squeezes and stretches the gradient according to the geometry of the parameter space. The latter is a practical approximation of the full Fubini-Study metric that would require exhaustive operations and techniques that are near impossible to implement on NISQ devices [41]. Another near-term alternative is the efficient protocol by Gómez-Lurbe [61] for lowering the number of executed circuits from $O(\text{number of parameters})$ to $O(\text{number of layers})$. The learning rate is adapted using the Armijo rule [62], as detailed in Algorithm 1. This subroutine searches for the largest learning rate that verifies the Armijo condition. However, AQNGD, as detailed in Algorithm 2, requires the careful tuning of the following hyperparameters that can affect the performance of optimization:

(1) $\alpha \in [0, 1]$: the constant to tune the sensitivity of Armijo's line search principle. Lower values of α skew the rule toward the preference of higher step sizes, and higher values skew it toward the preference of lower step sizes. In this work, we used $\alpha = 0.01$.

(2) β : the maximum learning rate (step size) used by the algorithm. It is fixed in this work to $\beta = 0.5$.

(3) k_m : the maximum number of step sizes searched by the line search algorithm to determine the best step size. In this work, $k_m = 6$.

(4) ϵ : the tolerance parameter used in the PseudoINVERT function, which is available as a Numpy function [63]. In our case, ϵ is set to 10^{-15} .

ALGORITHM 1. Apply Armijo rule.

Require: f : Cost function, α, β, θ_i

Ensure: k

```

1: Assert  $\alpha \in [0, 1], \beta > 0$ 
2:  $k \leftarrow 0$ 
3:  $\theta_{i+1} \leftarrow \theta_i - \frac{\beta}{2^k} \nabla f(\theta_i)$ 
4: ArmijoCondition  $\leftarrow f(\theta_i) - f(\theta_{i+1}) \geq \alpha \frac{\beta}{2^k} \|\nabla f(\theta_i)\|^2$ 
5: while not ArmijoCondition do
6:    $k \leftarrow k + 1$ 
7:    $\theta_{i+1} \leftarrow \theta_i - \frac{\beta}{2^k} \nabla f(\theta_i)$ 
8:   ArmijoCondition  $\leftarrow f(\theta_i) - f(\theta_{i+1}) \geq \alpha \frac{\beta}{2^k} \|\nabla f(\theta_i)\|^2$ 
9: end while

```

APPENDIX B: HEISENBERG HAMILTONIAN PROPERTIES

The Heisenberg Hamiltonian is defined as

$$H = \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j. \quad (\text{B1})$$

Let us recall the algebraic properties of the spin operators:

$$[S_i^\alpha, S_j^\beta] = i\delta_{ij}\epsilon^{\alpha\beta\gamma}\hbar S_i^\gamma, \quad (\text{B2})$$

$$[S_i^\pm, S_j^z] = \mp\delta_{ij}\hbar S_i^\pm, \quad (\text{B3})$$

$$[S_i^+, S_j^-] = 2\delta_{ij}\hbar S_i^z. \quad (\text{B4})$$

The total spin operators are defined as

$$\mathbf{S} = \sum_i \mathbf{S}_i = \sum_i (S_i^x, S_i^y, S_i^z)^T, \quad (\text{B5})$$

$$\mathbf{S}^2 = \mathbf{S} \cdot \mathbf{S} = \left(\sum_i S_i^x \right)^2 + \left(\sum_i S_i^y \right)^2 + \left(\sum_i S_i^z \right)^2. \quad (\text{B6})$$

ALGORITHM 2. AQNGD algorithm: It follows typically the QNGD algorithm, but it is featured by an additional step to adapt the step size rate λ_i using the line search backtracking that involves the optimization of the Armijo-Goldstein rule, as detailed in Algorithm 1.

Require: f : Cost function, $\alpha, \beta, \theta_0, \epsilon$

Ensure: $f(\theta_i)$

```

1: Assert  $\alpha \in [0, 1], \beta > 0, \epsilon > 0$ 
2:  $i \leftarrow 0$ 
3:  $\theta_i \leftarrow \theta_0$ 
4: repeat
5:    $\nabla_{\text{Euclid}} f(\theta_i) \leftarrow \text{ParameterShift}(\theta_i);$ 
6:    $g(\theta_i) \leftarrow \text{FubiniStudyMetric}(\theta_i)$ 
7:    $g^{-1}(\theta_i) \leftarrow \text{PseudoInvert}(g(\theta_i), \epsilon)$ 
8:    $\nabla f(\theta_i) \leftarrow g^{-1}(\theta_i) \nabla_{\text{Euclid}} f(\theta_i)$ 
9:    $k \leftarrow \text{ApplyArmijoRule}(\theta_i, \alpha, \beta)$ 
10:   $\lambda_i \leftarrow \beta/2^k$ 
11:   $\theta_i \leftarrow \theta_i - \lambda_i \nabla f(\theta_i)$ 
12:   $i \leftarrow i + 1$ 
13: until:  $f(\theta_{i-1}) - f(\theta_i) < \epsilon$ 

```

The raising and lowering operators are defined as

$$S_i^\pm = S_i^x \pm iS_i^y. \quad (\text{B7})$$

The total raising and lowering operators are defined as

$$S^\pm = \sum_i S_i^\pm. \quad (\text{B8})$$

The commutation relation between H and S^z is given by

$$\begin{aligned} [H, S^z] &= \sum_{ij} J_{ij} [\mathbf{S}_i \cdot \mathbf{S}_j, S^z] \\ &= \sum_{ij} J_{ij} \sum_k [\mathbf{S}_i \cdot \mathbf{S}_j, S_k^z]. \end{aligned} \quad (\text{B9})$$

Since

$$[S_i^x S_j^x, S_k^z] = -\delta_{ik} S_i^y S_j^x - \delta_{jk} S_i^x S_j^y, \quad (\text{B10})$$

$$[S_i^y S_j^y, S_k^z] = \delta_{ik} S_i^x S_j^y + \delta_{jk} S_i^y S_j^x, \quad (\text{B11})$$

$$[S_i^z S_j^z, S_k^z] = 0, \quad (\text{B12})$$

then, $[H, S^z] = 0$. Generally,

$$[H, S^z] = [H, S^x] = [H, S^y] = 0. \quad (\text{B13})$$

Therefore, H is $SU(2)$ invariant. Moreover,

$$[H, S^2] = [H, S^{x2}] + [H, S^{y2}] + [H, S^{z2}] = 0. \quad (\text{B14})$$

APPENDIX C: FUBINI-STUDY METRIC

The infinitesimal Euclidean distance between two parametrized quantum states $\rho_\theta = |\psi(\theta)\rangle\langle\psi(\theta)|$ and $\rho_{\theta+d\theta} = |\psi(\theta+d\theta)\rangle\langle\psi(\theta+d\theta)|$, such that $\theta \in \mathbb{R}^n$ can be expressed as

$$d^2(\rho_\theta, \rho_{\theta+d\theta}) = \sum_{(i,j) \in [n]^2} g_{ij}(\theta) d\theta_i d\theta_j, \quad (\text{C1})$$

where $g_{ij}(\theta) = \text{Re } G_{ij}(\theta)$, and the quantum geometric tensor $G_{ij}(\theta)$ is defined as [41]

$$G_{ij}(\theta) = \left\langle \frac{\partial \psi(\theta)}{\partial \theta_i} \frac{\partial \psi(\theta)}{\partial \theta_j} \right\rangle - \left\langle \frac{\partial \psi(\theta)}{\partial \theta_i} \psi(\theta) \right\rangle \left\langle \psi(\theta) \frac{\partial \psi(\theta)}{\partial \theta_j} \right\rangle. \quad (\text{C2})$$

This tensor is computed based on the Fubini-Study distance.

In the case of ansatz involving layers of parametrized Pauli rotation gates, the quantum geometric tensor can be simplified analytically since the parametrized ansatz form is given by

$$U(\theta) |0\rangle = V_L U_L(\theta_L) V_{L-1} U_{L-1}(\theta_{L-1}) \cdots V_1 U_1(\theta_1) V_0 |0\rangle, \quad (\text{C3})$$

such that

$$U_l(\theta_l) = \prod_{j=1}^N \exp(-i\theta_l^j P_l^j)$$

is the unitary operator of the l th layer involving Pauli rotations on N qubits, and V_l is a fixed unitary operator that does not depend on the parameters θ_l . Therefore, if the Fubini-Study metric is approximated by block matrices for each layer, the

quantum geometric tensor can be approximated as $G_{ij}(\theta) \approx \bigoplus_l G_{ij}^l(\theta_l)$ [41], where

$$G_{ij}^l(\theta_l) = \langle \psi_l | P_l^i P_l^j | \psi_l \rangle - \langle \psi_l | P_l^i | \psi_l \rangle \langle \psi_l | P_l^j | \psi_l \rangle. \quad (\text{C4})$$

This quantity reflects the covariance of the two observables P_l^i and P_l^j . According to this approximation, the blocks are symmetric matrices, and they are real valued. Moreover, each block requires at most $O(N^2)$ circuit executions since the second term in Eq. (C4) can be computed with single circuit in the $P_l^1 \cdots P_l^N$ basis, while the first term can be computed at most with $O(N^2)$ circuits or with $O(1)$ circuits if the Pauli operations are the same for all gates.

However, the block matrices need to be invertible and significantly larger than zero to be meaningful. Actually, a small eigenvalued block matrix means a flat landscape. Therefore, it is good to design the ansatz in such a way that the blocks of the metric are not singular.

In case of the block corresponding to the first layer, generally, the qubits are initialized in tensor product state. Hence, the first block matrix is diagonal. To avoid singular first block, we can ensure that

$$\langle \psi_1 | P_1^N \otimes \cdots \otimes P_1^1 | \psi_1 \rangle \neq 0.$$

This can be achieved by using suitable Pauli rotation generator according to the choice of the first layer V_0 .

In our case, our ansatz was specifically designed to have a diagonal constant Fubini-Study metric. Namely,

$$G_{ij}(\theta_l) = \langle \psi_l | P_l^i U_{l,k-1}^\dagger P_l^j | \psi_k \rangle - \langle \psi_l | P_l^i | \psi_l \rangle \langle \psi_k | P_l^j | \psi_k \rangle, \quad (\text{C5})$$

where

$$U_{l,k-1} = U_l^\dagger(\theta_l) V_l^\dagger \cdots U_{k-1}^\dagger(\theta_{k-1}) V_{k-1}^\dagger,$$

and (l, k) are the layer indices where (i, j) parameter indices are located. We can exploit the following identities to analytically prove that the full metric is constant and diagonal:

$$\begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \boxed{Y} \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} = \begin{array}{c} \boxed{Z} \\ \boxed{Y} \end{array} \quad (\text{C6})$$

$$\begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \boxed{Y} \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} = \begin{array}{c} \boxed{Y} \\ \boxed{X} \end{array} \quad (\text{C7})$$

$$\begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \boxed{Z} \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} = \begin{array}{c} \boxed{Z} \\ \boxed{Z} \end{array} \quad (\text{C8})$$

$$\begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \boxed{Y} \\ \boxed{Z} \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} = \begin{array}{c} \boxed{X} \\ \boxed{Y} \end{array} \quad (\text{C9})$$

$$\begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \boxed{X} \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} \begin{array}{c} \bullet \\ \oplus \end{array} = \begin{array}{c} \boxed{X} \end{array} \quad (\text{C10})$$

In the end, the expectation values that involve single Y_i gates or double $Y_i Y_j$ gates in separate qubits can be simplified with these identities to remove the entangling gates leaving us with a term of the form $\lambda \times \langle 0 | Y H Y | 0 \rangle$, which is equal to zero no matter what λ is. Only the expectation values of Y_i^2 give unities.

APPENDIX D: QUBITWISE COMMUTING PAULI TERMS GROUPING MEASUREMENT

The Hamiltonian in Eq. (1) for one kagome star (12 qubits) yields 54 terms to be measured on a quantum computer. However, it is possible to optimize this number of evaluations by grouping qubitwise commuting Pauli terms, since all commuting operators are simultaneously measurable. Namely,

$$[X_i X_j, X_k X_l] = [Y_i Y_j, Y_k Y_l] = [Z_i Z_j, Z_k Z_l] = 0 \quad (\text{D1})$$

for any qubits i, j, k, l . Hence, it is possible to compute the 54 terms—or any number of Heisenberg interaction terms—by evaluating three quantum circuits only, namely,

$$\langle \psi | \sum_{(i,j)} \sigma_i \sigma_j | \psi \rangle$$

for $\sigma = X, Y, Z$ such that ψ is the trial wave function. The effect of grouping commuting Pauli terms on the measurement accuracy has been studied in the literature [64–67]. Interestingly, Crawford *et al.* [67] showed that splitting a commuting group to two sets does not improve the accuracy under the condition that the number of measurements is optimally distributed among the groups. In this study, the standard error calculation, based on the variance of energy obtained from sampling the grouped commuting terms, is discussed in the following section.

Measured energy variance

We show here how we compute the empirical variance and the standard deviation of energy value obtained from qubitwise commuting Pauli words grouping. In the experiment, we showed that sampling three circuits in the X, Y , and Z basis is enough to estimate the energy as follows:

$$\langle H \rangle = \langle H_X \rangle + \langle H_Y \rangle + \langle H_Z \rangle, \quad (\text{D2})$$

such that

$$\langle H_X \rangle \approx \frac{1}{N} \sum_{b \in S_X} \langle b | \sum_{(i,j)} Z_i Z_j | b \rangle, \quad (\text{D3})$$

$$\langle H_Y \rangle \approx \frac{1}{N} \sum_{b \in S_Y} \langle b | \sum_{(i,j)} Z_i Z_j | b \rangle, \quad (\text{D4})$$

$$\langle H_Z \rangle \approx \frac{1}{N} \sum_{b \in S_Z} \langle b | \sum_{(i,j)} Z_i Z_j | b \rangle, \quad (\text{D5})$$

where S_X, S_Y , and S_Z are independent sets of bitstrings that are measured in the $X^{\otimes n}, Y^{\otimes n}$, and $Z^{\otimes n}$ eigenbases, respectively. Statistically, since the sets of samples are independent, the covariance of the variables H_X, H_Y , and H_Z is zero. Hence, the variance of H is nothing but

$$\text{Var}(H) = \text{Var}(H_X) + \text{Var}(H_Y) + \text{Var}(H_Z). \quad (\text{D6})$$

The variance of each variable H_P can be computed as follows:

$$\text{Var}(H_P) = \frac{1}{N} \sum_{b \in S_P} \langle b | \left(\sum_{(i,j)} Z_i Z_j - \langle H_P \rangle \right)^2 | b \rangle, \quad (\text{D7})$$

which leads to the reduced form

$$\text{Var}(H_P) = \langle H_P^2 \rangle - \langle H_P \rangle^2. \quad (\text{D8})$$

We can show using the identity projection $I = \sum_{b \in \mathbb{Z}_2^n} |b\rangle \langle b|$ that

$$\text{Var}(H_P) = \frac{1}{N} \sum_{b \in S_P} \left(\langle b | \sum_{(i,j)} Z_i Z_j | b \rangle - \langle H_P \rangle \right)^2, \quad (\text{D9})$$

which is nothing but the empirical variance of the energy value obtained from the P basis measurements. The standard deviation is then given by

$$\sigma(H) = \sqrt{\text{Var}(H)}. \quad (\text{D10})$$

The standard error is then

$$\epsilon = \sqrt{\frac{\text{Var}(H)}{N}}. \quad (\text{D11})$$

APPENDIX E: READOUT-ERROR MITIGATION

One of the main sources of noise in quantum devices is the SPAM errors. Specifically, the measurement or readout errors are well known in many areas of physics such as astronomy and high-energy physics. The main assumption is that such readout errors are governed by the response matrix R , which is a stochastic matrix that stores the probability of measuring a specific state $|f\rangle = |f_0 f_1 \dots f_N\rangle$ given that the N qubits are prepared in a prior state $|i\rangle = |i_0 i_1 \dots i_N\rangle$. Here, $i, f \in \{0, 1\}^{\otimes N}$. Namely, the $2^N \times 2^N$ components of the R matrix are

$$R_{fi} = P(f|i), \quad (\text{E1})$$

which is interpreted as the probability of measuring $|f\rangle \langle f|$ state given that the true state is $|i\rangle \langle i|$. Therefore, the relation between the true and the noisy vectors of bitstrings $\mathbf{t} = \sum_k t_k |k\rangle \langle k|$ and $\mathbf{n} = \sum_k n_k |k\rangle \langle k|$, respectively, are related as follows:

$$n_k = \sum_l R_{kl} t_l, \quad (\text{E2})$$

which suggests that the true readout vector is nothing but $\mathbf{t} = R^{-1} \mathbf{n}$. Traditionally, Moore-Penrose pseudoinverse is taken since the stochastic matrices are not generally invertible. Unfortunately, taking the inverse of an exponential matrix is a hard computation overhead. Nachman *et al.* [68] suggested unfolding techniques borrowed from high-energy physics. However, the characterization of the full R matrix by itself is not efficient since it requires 2^N state preparations, $\mathcal{O}(2^n)$ measurements on the QPU, and $2^n \times 2^n$ amount of memory for storage. Nevertheless, the full characterization of the R -matrix assumes correlated all-to-all interaction between the qubits. Geller and Sun [44] proposed an efficient measurement method based on the characterization of the k nearest-neighbor qubits, requiring $\mathcal{O}(4^k N^2)$ where N is the number of qubits, but that leaves the burden of classically postprocessing a $2^N \times 2^N$ matrix that is still not scalable. In practice, we can assume that correlated readout errors occur just locally between nearest neighbors in QPUs of restricted connectivity [44], so

$$P(a_1 a_2 \dots a_N | b_1 b_2 \dots b_N) = \bigotimes_{(jk \dots l) \in \mathcal{P}} P_{jk \dots l}(a|b), \quad (\text{E3})$$

such that $P_{jk\dots l}$ is the probability of transition of the specific qubits labeled by $jk\dots l$ sharing the same neighborhood, and $\mathcal{P}$ is the partition of the qubit register. The immediate advantage of this model is the efficiency of taking the inverse of the response matrix since $P^{-1} = \bigotimes_{(jk\dots l) \in \mathcal{P}} P_{(jk\dots l)}^{-1}(a|b)$. Moreover, the correction process for a given measurement result $\mathbf{m} = \sum_k m_k |k\rangle \langle k|$ involves only the application of the inverse partitioned response matrix on the relevant sections of the qubit register. That is, the recovered measurement result is

$$\mathbf{t} = R^{-1} \sum_p m_p |p\rangle \langle p| = \sum_p m_p R^{-1} |p\rangle \langle p|. \quad (\text{E4})$$

We use the partitioned response matrix. Hence,

$$\mathbf{t} = \sum_p m_p \bigotimes_{(jk\dots l)} R_{(jk\dots l)}^{-1} |p_{(jk\dots l)}\rangle \langle p_{(jk\dots l)}|. \quad (\text{E5})$$

We apply each R^{-1} matrix partition to the corresponding $|p\rangle \langle p|$ partition as a probability vector using the inverse of Eq. (E2). That is,

$$\mathbf{t} = \sum_p m_p \bigotimes_{(jk\dots l)} \sum_{q_{(jk\dots l)}} P_{(jk\dots l)}^{-1}(p|q) |q_{(jk\dots l)}\rangle \langle q_{(jk\dots l)}|. \quad (\text{E6})$$

Since the number of measured samples of bitstrings in $\mathbf{m}$ should be small in practice, the global sum is still practical. However, the tensor products can yield exponentially large samples. Let $N_p = N/P$ be the number of qubits, such that P is the number of partitions. Then, the inner sum spans 2^{N_p} states, while the tensor product will yield $2^{N_p P} = 2^N$ states in the computational basis. Nevertheless, it is possible to truncate results with the largest probabilities. It is worth noting that the REM method faces the issue of negative probabilities since the inverse does not preserve the stochasticity of the response matrix. Even though the negative pseudoprobabilities are absorbed in the estimation of any physical observable, it is important to note that the positivity of the pseudoprobabilities is crucial for maintaining the variational principle of Rayleigh and Ritz to be valid. Actually, positivity can be retained by involving more efforts to minimize $\|\mathbf{R}\mathbf{t} - \mathbf{m}\|$ [44] or minimize $\|\mathbf{t} - R^{-1}\mathbf{m}\|$. The latter is computationally less demanding. In this work, we show the REM results with both scenarios: the positivity preserving and nonpreserving REM since the REM with positivity recovery is still valid given that, in practice, the negative pseudoprobability values have small magnitudes.

APPENDIX F: ZERO-NOISE EXTRAPOLATION ERROR MITIGATION

One of the efficient error mitigation techniques is the zero-noise extrapolation, which was initially developed by Li and Benjamin [69] and Temme *et al.* [70]. The basic idea behind this technique is the possibility to tune the quantum machine's noise with a single parameter λ in such a way that the expectation value $\langle O \rangle$ depends on this error parameter as follows:

$$\langle O \rangle(\lambda) = E_0 + \sum_{k=1}^n a_k \lambda^k + O(\lambda^{n+1}), \quad (\text{F1})$$

where E_0 is the expected value without noise. Li and Benjamin [69] showed that this is true when the evolution of the quantum circuit is modeled with the master equation

$\frac{\partial \rho}{\partial t} = -[H, \rho] + \lambda \mathcal{L}[\rho]$, where $\mathcal{L}$ is the Lindblad superoperator. Although originally tuning λ was performed by rescaling execution time in quantum computers [58,69–71], it is practically preferred to apply a unitary circuit or gate folding on digital quantum devices [72–74]. Folding technique, despite its practicality, scales only the circuit internal decoherence noise, while in principle keeping the SPAM errors fixed. In this work, unitary global folding is performed considering the folding points 1, 3, and 5. Previously, different extrapolation methods have been used; Li and Benjamin [69] used the Richardson's deferred approach, while Giurgica-Tiron *et al.* [73] benchmarked different methods on IBMQ London five-qubit chip, showing the instability of Richardson's extrapolation and better accuracy with exponential extrapolation. The choice of the extrapolation method is still an ambiguous feature of ZNE method. Nevertheless, in this work, we choose the second-degree polynomial model since it fits the points minimally to avoid overfitting. Moreover, to infer the standard error of the extrapolated zero-noise energy, we use the BPR with heteroscedastic noise technique assuming a Gaussian process [75,76]. Finally, it is worth noting that ZNE does not guarantee an upper bound of the exact energy and allows the violation of Rayleigh-Ritz variational principle since the method is not based on Hamiltonian measurement.

APPENDIX G: QPU DEVICES TECHNICAL DETAILS

The native gates of Oslo IBM quantum device—which is featured by seven qubits—are CNOT, RZ, $\sqrt{X}$, X, and the identity I. Consequently, the required gates are implemented as follows:

$$H \equiv \text{RZ}(\pi/2) \cdot \sqrt{X} \cdot \text{RZ}(\pi/2), \quad (\text{G1})$$

$$\text{Ry}(\theta) \equiv \text{RZ}(3\pi) \cdot \sqrt{X} \cdot \text{RZ}(\theta + \pi) \cdot \sqrt{X}. \quad (\text{G2})$$

The native gates of Kyoto IBM quantum device—which is from Eagle processor generation with 127 qubits—are ECR = $1/\sqrt{2}(\text{IX} - \text{XY}) = \text{XI} \cdot \text{CNOT} \cdot \text{S}\sqrt{X}$, RZ, $\sqrt{X}$, X, and the identity I. Therefore, the required gates are implemented as follows:

$$H \equiv \text{RZ}(\pi/2) \cdot \sqrt{X} \cdot \text{RZ}(\pi/2), \quad (\text{G3})$$

$$\text{Ry}(\theta) \equiv \text{RZ}(3\pi) \cdot \sqrt{X} \cdot \text{RZ}(\theta + \pi) \cdot \sqrt{X}, \quad (\text{G4})$$

$$\text{CNOT} \equiv \text{IX} \cdot \text{ECR} \cdot$$

$$\text{RZ}(-\pi) \cdot \sqrt{X} \cdot \text{RZ}(-\pi) \otimes \text{RZ}(-\pi/2). \quad (\text{G5})$$

The native gates of Torino IBM quantum device with the Heron architecture and 133 qubits are CZ, RZ, X, and $\sqrt{X}$. The compiled gates in the Torino IBM quantum device with the Heron architecture are decomposed as follows:

$$H \equiv \text{RZ}(\pi/2) \cdot \sqrt{X} \cdot \text{RZ}(\pi/2), \quad (\text{G6})$$

$$\text{Ry}(\theta) \equiv \text{RZ}(3\pi) \cdot \sqrt{X} \cdot \text{RZ}(\theta + \pi) \cdot \sqrt{X}, \quad (\text{G7})$$

$$\text{CNOT} \equiv (I \otimes \text{RZ}(\pi/2) \cdot \sqrt{X}) \cdot \text{CZ} \cdot$$

$$I \otimes \text{RZ}(\pi) \sqrt{X} \cdot \text{RZ}(\pi/2). \quad (\text{G8})$$

- [1] J. Kempe, A. Kitaev, and O. Regev, The complexity of the local Hamiltonian problem, *SIAM J. Comput.* **35**, 1070 (2006).
- [2] R. P. Feynman, Simulating physics with computers, *Int. J. Theor. Phys.* **21**, 467 (1982).
- [3] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, *Rev. Mod. Phys.* **92**, 015003 (2020).
- [4] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, 10th ed. (Cambridge University Press, Cambridge, 2010).
- [5] A. Somoroff, Q. Ficheux, R. A. Mencia, H. Xiong, R. Kuzmin, and V. E. Manucharyan, Millisecond coherence in a superconducting qubit, *Phys. Rev. Lett.* **130**, 267001 (2023).
- [6] S. Ganjam, Y. Wang, Y. Lu, A. Banerjee, C. U. Lei, L. Krayzman, K. Kisslinger, C. Zhou, R. Li, Y. Jia, M. Liu, L. Frunzio, and R. J. Schoelkopf, Surpassing millisecond coherence in on chip superconducting quantum memories by optimizing materials and circuit design, *Nat. Commun.* **15**, 3687 (2024).
- [7] J. Preskill, Quantum Computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
- [8] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, Noisy intermediate-scale quantum algorithms, *Rev. Mod. Phys.* **94**, 015004 (2022).
- [9] 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).
- [10] N. E. Belaloui, A. Tounsi, A. R. Khamadja, M. M. Louamri, A. Benslama, D. E. Bernal Neira, and M. T. Rouabah, Ground-state energy estimation on current quantum hardware through the variational quantum eigensolver: A practical study, *J. Chem. Theory Comput.* **21**, 6777 (2025).
- [11] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik, Quantum chemistry in the age of quantum computing, *Chem. Rev.* **119**, 10856 (2019).
- [12] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, *Nat. Rev. Phys.* **3**, 625 (2021).
- [13] A. Y. Kitaev, Quantum measurements and the Abelian stabilizer problem, *arXiv:quant-ph/9511026*.
- [14] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, Quantum error mitigation, *Rev. Mod. Phys.* **95**, 045005 (2023).
- [15] C. Yi, C. Zhou, and J. Takahashi, Quantum phase estimation by compressed sensing, *Quantum* **8**, 1579 (2024).
- [16] J. Robledo-Moreno, M. Motta, H. Haas, A. Javadi-Abhari, P. Jurcevic, W. Kirby, S. Martiel, K. Sharma, S. Sharma, T. Shirakawa, I. Sitdikov, R.-Y. Sun, K. J. Sung, M. Takita, M. C. Tran, S. Yunoki, and A. Mezzacapo, Chemistry beyond the scale of exact diagonalization on a quantum-centric supercomputer, *Sci. Adv.* **11**, eadu9991 (2025).
- [17] M. Mikkelsen and Y. O. Nakagawa, Quantum-selected configuration interaction with time-evolved state *Phys. Rev. Res.* **7**, 043043 (2025).
- [18] L. Balents, Spin liquids in frustrated magnets, *Nature (London)* **464**, 199 (2010).
- [19] G. Semeghini, H. Levine, A. Keesling, S. Ebadi, T. T. Wang, D. Bluvstein, R. Verresen, H. Pichler, M. Kalinowski, R. Samajdar, A. Omran, S. Sachdev, A. Vishwanath, M. Greiner, V. Vuletić, and M. D. Lukin, Probing topological spin liquids on a programmable quantum simulator, *Science* **374**, 1242 (2021).
- [20] J. L. Bosse and A. Montanaro, Probing ground-state properties of the kagome antiferromagnetic Heisenberg model using the variational quantum eigensolver, *Phys. Rev. B* **105**, 094409 (2022).
- [21] J. Kattemölle and J. van Wezel, Variational quantum eigensolver for the Heisenberg antiferromagnet on the kagome lattice, *Phys. Rev. B* **106**, 214429 (2022).
- [22] M. Sajjan, R. Gupta, S. S. Kale, V. Singh, K. Kumaran, and S. Kais, Physics-inspired quantum simulation of resonating valence bond states: A prototypical template for a spin-liquid ground state, *J. Phys. Chem. A* **127**, 8751 (2023).
- [23] Y. Guo, F. Guo, B. Zhong, X. Chen, X. Yuan, X. Jin, and H. Tang, Quantum-enhanced computing for the antiferromagnetic $j_1 - j_2$ Heisenberg model, *Adv. Quantum Technol.* **8**, 2300240 (2025).
- [24] T. Weaving, A. Ralli, V. Wimalaweera, P. J. Love, and P. V. Coveney, Simulating the antiferromagnetic Heisenberg model on a spin-frustrated kagome lattice with the contextual subspace variational quantum eigensolver, *arXiv:2506.12391*.
- [25] Q. Wang, H. Lei, Y. Qi, and C. Felser, Intriguing kagome topological materials, *npj Quantum Mater.* **10**, 72 (2025).
- [26] B. Bauer, L. Cincio, B. Keller, M. Dolfi, G. Vidal, S. Trebst, and A. Ludwig, Chiral spin liquid and emergent anyons in a kagome lattice Mott insulator, *Nat. Commun.* **5**, 5137 (2014).
- [27] J.-W. Mei, J.-Y. Chen, H. He, and X.-G. Wen, Gapped spin liquid with $\mathbb{Z}_2$ topological order for the kagome Heisenberg model, *Phys. Rev. B* **95**, 235107 (2017).
- [28] A. Tounsi, N. E. Belaloui, M. M. Louamri, A. Benslama, and M. T. Rouabah, Optimized topological quantum compilation of three-qubit controlled gates in the Fibonacci anyon model: A controlled-injection approach, *Phys. Rev. A* **110**, 012603 (2024).
- [29] P. Mendels and F. Bert, Quantum kagome antiferromagnet: $\text{ZnCu}_3(\text{OH})_6\text{Cl}_2$, *J. Phys.: Conf. Ser.* **320**, 012004 (2011).
- [30] A. S. Teixeira Pires, *Theoretical Tools for Spin Models in Magnetic Systems*, IOP ebooks (Institute of Physics Publishing, London, England, 2021).
- [31] B. Sriram Shastry and B. Sutherland, Exact ground state of a quantum mechanical antiferromagnet, *Physica B+C* **108**, 1069 (1981).
- [32] S. Sachdev, Kagome' and triangular-lattice Heisenberg antiferromagnets: Ordering from quantum fluctuations and quantum-disordered ground states with unconfined bosonic spinons, *Phys. Rev. B* **45**, 12377 (1992).
- [33] C. Waldtmann, H.-U. Everts, B. Bernu, C. Lhuillier, P. Sindzingre, P. Lecheminant, and L. Pierre, First excitations of the spin $\frac{1}{2}$ Heisenberg antiferromagnet on the kagomé lattice, *Eur. Phys. J. B* **2**, 501 (1998).
- [34] Y. Ran, M. Hermele, P. A. Lee, and X.-G. Wen, Projected-wavefunction study of the spin- $\frac{1}{2}$ Heisenberg model on the kagomé lattice, *Phys. Rev. Lett.* **98**, 117205 (2007).

- [35] S. Yan, D. A. Huse, and S. R. White, Spin-liquid ground state of the $S = \frac{1}{2}$ kagome Heisenberg antiferromagnet, *Science* **332**, 1173 (2011).
- [36] Y.-C. He, M. P. Zaletel, M. Oshikawa, and F. Pollmann, Signatures of Dirac cones in a DMRG study of the kagome Heisenberg model, *Phys. Rev. X* **7**, 031020 (2017).
- [37] R.-Y. Sun, H.-K. Jin, H.-H. Tu, and Y. Zhou, Possible chiral spin liquid state in the $S = \frac{1}{2}$ kagome Heisenberg model, *npj Quantum Mater.* **9**, 16 (2024).
- [38] H. J. Liao, Z. Y. Xie, J. Chen, Z. Y. Liu, H. D. Xie, R. Z. Huang, B. Normand, and T. Xiang, Gapless spin-liquid ground state in the $S = \frac{1}{2}$ kagome antiferromagnet, *Phys. Rev. Lett.* **118**, 137202 (2017).
- [39] A. M. Läuchli, J. Sudan, and R. Moessner, $S = \frac{1}{2}$ kagome Heisenberg antiferromagnet revisited, *Phys. Rev. B* **100**, 155142 (2019).
- [40] T. Đurić, J. H. Chung, B. Yang, and P. Sengupta, Spin-1/2 kagome Heisenberg antiferromagnet: Machine learning discovery of the spinon pair-density-wave ground state, *Phys. Rev. X* **15**, 011047 (2025).
- [41] J. Stokes, J. Izaac, N. Killoran, and G. Carleo, Quantum natural gradient, *Quantum* **4**, 269 (2020).
- [42] F. Dell’Anna, R. Gómez-Lurbe, A. Pérez, and E. Ercolessi, Quantum natural gradient optimizer on noisy platforms: Quantum approximate optimization algorithm as a case study, *Phys. Rev. A* **112**, 022612 (2025).
- [43] T. A. Atif, U. Chukwu, J. Berwald, and R. Dridi, Accelerating NISQ variational methods using geometry, in *Quantum Computing, Communication, and Simulation III*, edited by P. R. Hemmer and A. L. Migdall (SPIE, San Francisco, California, 2023).
- [44] M. R. Geller and M. Sun, Toward efficient correction of multiqubit measurement errors: Pair correlation method, *Quantum Sci. Technol.* **6**, 025009 (2021).
- [45] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, Quantum circuit learning, *Phys. Rev. A* **98**, 032309 (2018).
- [46] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, Evaluating analytic gradients on quantum hardware, *Phys. Rev. A* **99**, 032331 (2019).
- [47] M. Larocca, S. Thanasilp, S. Wang, K. Sharma, J. Biamonte, P. J. Coles, L. Cincio, J. R. McClean, Z. Holmes, and M. Cerezo, Barren plateaus in variational quantum computing, *Nat. Rev. Phys.* **7**, 174 (2025).
- [48] M. Kolodrubetz, D. Sels, P. Mehta, and A. Polkovnikov, Geometry and non-adiabatic response in quantum and classical systems, *Phys. Rep.* **697**, 1 (2017).
- [49] M. Bukov, D. Sels, and A. Polkovnikov, Geometric speed limit of accessible many-body state preparation, *Phys. Rev. X* **9**, 011034 (2019).
- [50] N. Yamamoto, On the natural gradient for variational quantum eigensolver, *arXiv:1909.05074*.
- [51] S.-I. Amari, Natural gradient works efficiently in learning, *Neural Comput.* **10**, 251 (1998).
- [52] C.-Y. Park and N. Killoran, Hamiltonian variational ansatz without barren plateaus, *Quantum* **8**, 1239 (2024).
- [53] P. Murali, D. C. McKay, M. Martonosi, and A. Javadi-Abhari, Software mitigation of crosstalk on noisy intermediate-scale quantum computers, in *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems*, ASPLOS ’20 (Association for Computing Machinery, New York, NY, 2020), pp. 1001–1016.
- [54] D. C. McKay, I. Hincks, E. J. Pritchett, M. Carroll, L. C. G. Govia, and S. T. Merkel, Benchmarking quantum processor performance at scale, *arXiv:2311.05933*.
- [55] M. A. Haddou and M. Bennai, Sculpting quantum landscapes: Fubini-study metric conditioning for geometry-aware learning in parameterized quantum circuits, *arXiv:2506.21940*.
- [56] F. Sorourifar, M. T. Rouabah, N. E. Belaloui, M. M. Louamri, D. Chamaki, E. J. Gustafson, N. M. Tubman, J. A. Paulson, and D. E. Bernal Neira, Toward efficient quantum computation of molecular ground-state energies, *AIChE J.* **71**, e18887 (2025).
- [57] K. Lively, T. Bode, J. Szangolies, J.-X. Zhu, and B. Fauseweh, Noise robust detection of quantum phase transitions, *Phys. Rev. Res.* **6**, 043254 (2024).
- [58] Y. Kim, C. J. Wood, T. J. Yoder, S. T. Merkel, J. M. Gambetta, K. Temme, and A. Kandala, Scalable error mitigation for noisy quantum circuits produces competitive expectation values, *Nat. Phys.* **19**, 752 (2023).
- [59] T. Kurita, H. Qassim, M. Ishii, H. Oshima, S. Sato, and J. Emerson, Synergetic quantum error mitigation by randomized compiling and zero-noise extrapolation for the variational quantum eigensolver, *Quantum* **7**, 1184 (2023).
- [60] A. Tounsi, N. E. Belaloui, A. R. Khamadja, T. E. F. Lalaoui, M. M. Louamri, D. E. Bernal Neira, and M. T. Rouabah, Supplementary data and code for: “Probing the ground state of the antiferromagnetic Heisenberg model on the kagome lattice using geometrically informed variational quantum eigensolver”, Zenodo, 2026, <http://doi.org/10.5281/zenodo.21264601>.
- [61] R. Gómez-Lurbe, Efficient protocol to estimate the quantum Fisher information matrix for commuting-block circuits, *Quantum Inf. Process.* **25**, 278 (2026).
- [62] L. Armijo, Minimization of functions having Lipschitz continuous first partial derivatives, *Pac. J. Math.* **16**, 1 (1966).
- [63] C. R. Harris, K. Jarrod Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, *et al.*, Array programming with NumPy, *Nature (London)* **585**, 357 (2020).
- [64] D. Wecker, M. B. Hastings, and M. Troyer, Progress towards practical quantum variational algorithms, *Phys. Rev. A* **92**, 042303 (2015).
- [65] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, *New J. Phys.* **18**, 023023 (2016).
- [66] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz, *Quantum Sci. Technol.* **4**, 014008 (2019).
- [67] O. Crawford, B. v. Straaten, D. Wang, T. Parks, E. Campbell, and S. Brierley, Efficient quantum measurement of Pauli operators in the presence of finite sampling error, *Quantum* **5**, 385 (2021).
- [68] B. Nachman, M. Urbanek, W. A. de Jong, and C. W. Bauer, Unfolding quantum computer readout noise, *npj Quantum Inf.* **6**, 84 (2020).
- [69] Y. Li and S. C. Benjamin, Efficient variational quantum simulator incorporating active error minimization, *Phys. Rev. X* **7**, 021050 (2017).

- [70] K. Temme, S. Bravyi, and J. M. Gambetta, Error mitigation for short-depth quantum circuits, *Phys. Rev. Lett.* **119**, 180509 (2017).
- [71] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Error mitigation extends the computational reach of a noisy quantum processor, *Nature (London)* **567**, 491 (2019).
- [72] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski, Cloud quantum computing of an atomic nucleus, *Phys. Rev. Lett.* **120**, 210501 (2018).
- [73] T. Giurgica-Tiron, Y. Hindy, R. LaRose, A. Mari, and W. J. Zeng, Digital zero noise extrapolation for quantum error mitigation, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, Denver, CO, 2020).
- [74] A. He, B. Nachman, W. A. de Jong, and C. W. Bauer, Zero-noise extrapolation for quantum-gate error mitigation with identity insertions, *Phys. Rev. A* **102**, 012426 (2020).
- [75] C. M. Bishop, *Pattern Recognition and Machine Learning*, 1st ed., Information Science and Statistics (Springer, New York, NY, 2006).
- [76] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*, 3rd ed. (CRC Press, Boca Raton, FL, 2025), Chap. IV.14.7, electronic edition with errors fixed as of 20 February 2025, For non-commercial purposes only.