DISQ: <u>D</u>ynamic <u>I</u>teration <u>S</u>kipping for Variational <u>Q</u>uantum Algorithms

Junyao Zhang[1], Hanrui Wang[2], Gokul Subramanian Ravi[3,4],
Frederic T. Chong[3], Song Han[2], Frank Mueller[5], Yiran Chen[1]

[1]Duke University, [2]MIT, [3]University of Chicago, [4]University of Michigan, [5]North Carolina State University

jz420@duke.com

Abstract—In the noisy intermediate scale quantum (NISQ) era, the Variational Quantum Algorithm (VQA) has emerged as one of the most promising approaches to harness the power of quantum computers. In VQA, a classical optimizer iteratively updates the parameters of a variational quantum circuit to minimize a cost objective obtained by executing the quantum circuit on real quantum hardware. However, the deployment of VQA applications on NISQ devices encounters substantial noise, which degrades training stability. Moreover, the drift of noise is particularly intractable due to its dynamic nature in duration and magnitude. Noise drift leads to significant deviations in VQA iteration's objective function estimation and shapes a dynamic noisy landscape, which poses a considerable challenge for stable VQA parameter training, thereby hampering the accurate convergence of VQA optimizations.

This paper proposes DISQ to craft a stable landscape for VQA training and tackle the noise drift challenge. DISQ adopts a "drift detector" with a reference circuit to identify and skip iterations that are severely affected by noise drift errors. Specifically, the circuits from the previous training iteration are re-executed as a reference circuit in the current iteration to estimate noise drift impacts. The iteration is deemed compromised by noise drift errors and thus skipped if noise drift flips the direction of the ideal optimization gradient. To enhance noise drift detection reliability, we further propose to leverage multiple reference circuits from previous iterations to provide a well-founded judge of current noise drift. Nevertheless, multiple reference circuits also introduce considerable execution overhead. To mitigate extra overhead, we propose Pauli-term subsetting (prime and minor subsets) to execute only observable circuits with large coefficient magnitudes (prime subset) during drift detection. Only this minor subset is executed when the current iteration is drift-free.

Evaluations across various applications and QPUs demonstrate that DISQ can mitigate a significant portion of the noise drift impact on VQAs and achieve 1.51-$2.24\times$ fidelity improvement over the traditional baseline. DISQ's benefit is 1.1-$1.9\times$ over the best alternative approach while boosting average noise detection speed by $2.07\times$.

Index Terms—Variational Quantum Algorithm, Variational Quantum Eigensolver, Quantum Computing, Noise Mitigation
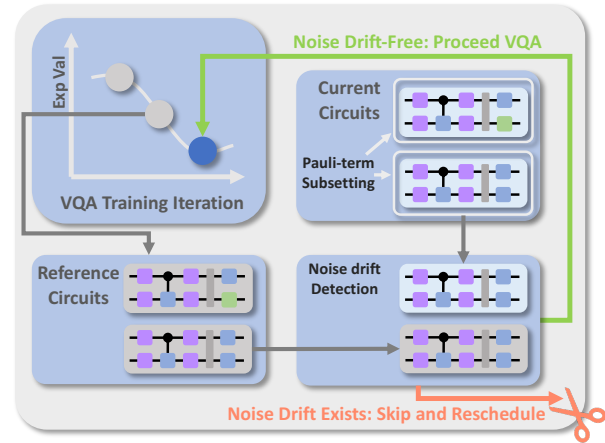
Fig. 1. Computing resource-efficient iteration skipping approach to filter out the noise drift impact. Previous iterations act as optimal reference circuits to detect the noise drift on the current VQA iteration. Pauli-term subsetting is utilized to proactively minimize computation during noise drift detection.

I. INTRODUCTION

Quantum computing is a revolutionary computational model that is poised to leverage substantial quantum mechanical phenomena to provide computing advantages in resolving some classically intractable problems in domains, such as chemistry [1], [2], biology [3], fundamental software algorithms [4], [5], and machine learning [6], [7].

One of the most promising noise intermediate scale quantum (NISQ) [8] algorithms that can provide a quantum advantage is the variational quantum algorithm (VQA) [9], [10], which has been widely applied to critical applications in chemistry [1], [10], approximation [11], and physics [12], among others. VQA is a long-running iterative algorithm that deploys a classical optimizer to train a parameterized quantum circuit on a quantum machine. The quantum circuit parameters are tuned in each iteration to approach the application's targets, which are usually minimization problems, such as estimating the ground state energy of molecules.

In spite of quantum supremacy having been theorized in the aforementioned domains [3]–[6], [10], quantum processing units (QPU) in the contemporary NISQ era are still vulnerable to various types of noise, such as decoherence errors, gate errors, state preparation and measurement (SPAM) errors, and crosstalk. These noise errors stem from a multitude of sources, such as device defects [13], thermal fluctuations [14], [15], magnetic flux [14], [16]–[18], qubit coupling destruction [19], insulation problems [13], and other external stimuli [20]–[22]. Due to the dynamic nature of quantum systems and the current limitations of quantum devices fabrication [23], these noise sources vary over time in intensity, thereby causing unpredictable fluctuations that deteriorate the quantum system.

With awareness of qubit restrictions in deploying full-scale quantum error correction (QEC) [24], [25] and the perniciousness of noise, multiple works have investigated mitigation

techniques for general quantum circuits [26]–[33]. To harvest quantum computing power while maintaining robustness in the face of different forms of noise errors. [34]–[36] have proposed noise-aware training frameworks to boost VQA training speed and accuracy of results. [37] have proposed noise-aware search for robust quantum ansatz. Although these works provide the potential to improve result fidelity or enhance training robustness, they primarily focus on static noise. Noise is assumed to be static or to remain stable for sufficient periods, such that these techniques can adequately capture and model the characteristics of static noise to provide mitigation methods. Unfortunately, focusing solely on static noise is insufficient for deploying VQA and harvesting its power on QPUs.

VQA iteratively tunes circuit parameters with the support of gradients [38]. The objective of the optimizer is to filter out the optimal gradient among disparate gradients to generate a corresponding set of parameters in each iteration, thereby achieving convergence towards the desired objective function. This convergence is expected to occur when the noise landscape of the quantum device is stable and consistent during the gradient estimation process, which allows the optimizer to accurately discern the most optimal circuit parameters. One of the fundamental challenges hampering the deployment of VQAs for practical utilization on QPUs is noise drift error originating from dynamic sources. The term "noise drift" refers to deviations in a quantum circuit's output distribution caused by shifts in the device characteristics of one or more qubits in the quantum circuit [13]–[16]. Noise drift, characterized by the unpredictable and time-varying nature of noise in quantum devices, poses a significant challenge to quantum algorithms, particularly for long-running iterative applications, such as VQA. Since noise drift alters the VQA tuning landscape over iterations, it leads to inconsistent gradient estimation each time and thus disrupts the convergence of VQA. This dynamic nature results in noise drift being hard to model and suppress with static noise mitigation techniques. The state-of-the-art method targeting dynamic noise, QISMET [39], predominantly addresses spike-like transient noise but performs inadequately in the presence of complex noise, such as noise drift, potentially misdirecting the VQA training.

In this paper, we propose DISQ, <u>D</u>ynamic <u>I</u>teration <u>S</u>kipping for Variational <u>Q</u>uantum Algorithms, a novel computing resource-efficient iteration skipping approach that crafts a reliable landscape for VQA applications by filtering out noise drift errors as in Fig.1. Drawing inspiration from QISMET in handling dynamic noise, (i) DISQ employs a previous iteration as an optimal reference to detect the noise drift for the current iteration by estimating the discrepancy between the reference output from the prior iteration and its result in the current iteration. Subsequently, DISQ adopts traditional per-iteration gradient calculation and combines it with noise drift error detection to estimate machine-obtained and drift-free gradients. A VQA iteration is accepted only if the direction of the machine-observed gradient loosely matches the direction of the drift-free gradient; Advancing beyond QISMET, (ii) DISQ augments multiple previous iterations as references

to further enhance noise drift detection reliability in facing intractable noise drift. Although multi-reference circuits aid in shaping the VQA training landscape, they also introduce extra execution costs; (iii) To minimize the extra execution cost, DISQ groups the Pauli terms of the target Hamiltonian by Pauli-term subsetting (prime and minor subsets) and partitions the execution. Prime subsets (Pauli terms with dominating coefficients) of references and prime subsets corresponding to the current iteration are executed during noise drift detection. The minor subset for the current iteration is executed only if noise drift is not detected. Subsetting conserves computation in noise drift detection, and execution partitioning eliminates unnecessary computation in skipping iterations. Evaluation demonstrates that DISQ achieves 1.1-1.9 fidelity improvements over the best alternative approach while boosting average noise detection speed by 2.07 . This paper thus makes the following contributions:

- **DISQ Framework**: We propose DISQ, a computing resource-efficient method to actively discover noise drift instances that severely impact VQA accurate convergence. DISQ controls VQA iterations to actively skip noise drift errors and maintain VQA optimization under reliable scenarios.

- **Noise Drift Detection with Multi-References**: We leverage multiple previous iterations as references to faithfully detect noise drift errors in VQA iterations, enabling precise estimation of noise drift impacts with DISQ.

- **Concept and Design of Pauli-term Subsetting**: We introduce a novel perspective on quantum circuits for noise detection: instead of employing all the observable circuits in a brute-force manner, Pauli-term subsetting groups the dominant observable circuits (prime subset) as effective substitutes for noise detection. Subsetting proactively enables execution acceleration in DISQ even with numerous references.

- **Robust Landscape with Reasonable Overhead**:
Through the aforementioned steps, DISQ effectively filters a substantial portion of noise drift's impact on VQAs, attaining 1.51-2.24 fidelity increase over traditional baseline. DISQ surpasses the best alternative approach with a 1.1-1.9 benefit while enhancing average noise detection speed by 2.07 .

II. BACKGROUND

A. Noise in NISQ QPUs

In the era of NISQ, two primary technologies in quantum architectures: superconducting transmon qubits [40] and trapped-ions [17], [41] are being pursued for universal quantum computing. Despite the fact that no definitive verdict has been reached in their performance comparison [42], both require precise control due to their sensitivity to various types of noise. The scalability of QPUs built using these two technologies is limited by noise errors, such as (i) Decoherence error: natural decay caused by the energy exchange between a qubit and its environment, which makes the qubit lose
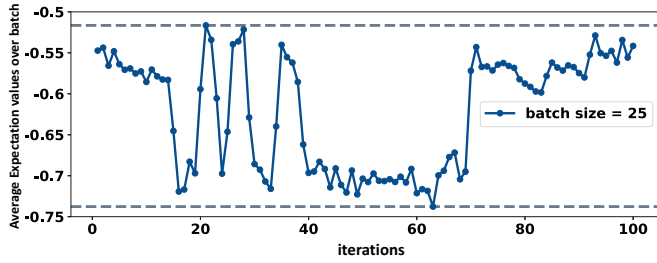
Fig. 2. Noise drift errors on circuits. Circuit data are collected by 100 continuous runs of a circuit batch from an experiment on IBMQ Belem. Each data point is the average expectation value of the circuit batch (25 identical circuits). The mean value is -0.62, but the range value is concerning 0.22.

its quantum properties (state) over time; (ii) State preparation and measurement (SPAM) errors: caused by imprecise state initialization and measurement [43]; (iii) Gate error: imperfect gate operations, such as depolarization, which are approximately 0.1% and 4% for 1-qubit gates and 2-qubit gates, respectively, for IBMQ [44]; (iv) Device-specific error: QPU-specific noise profiles that vary in spatial and temporal noise, arising from factors like topology cross-talk [23], [45], inhomogeneous qubits during manufacture [19], [46], [47], imperfections in gate implementation and control, or specific external interference [20], [21], where the latter is a significant source of error such as magnetic fields affecting Zeeman trapped-ion qubits [17], [18] and unstable near-resonant two-level systems affecting transmon qubits [15].

The presence of those stochastic errors and environmental variations renders the noise drift to be dynamic in nature [14], [16], [19], [48]. Further, the impacts of different noise drift errors compound with each other and accumulate over time, thereby increasing the overall probability of obtaining erroneous outcomes, particularly in long-running applications such as VQAs. Gradient evaluation discrepancies can mislead the tuning process in unfavorable directions, thereby impeding the convergence and accuracy of VQA results. Fig.2 demonstrates a severe case of noise drift affecting circuit fidelity, obtained from IBMQ Belem. Each data point represents the average expectation value obtained over a circuit batch (25 identical circuits). Although the mean of expectation values over 100 executions of this batch is -0.62, its range is large at 0.22. This proves that the statistical robustness offered by executing multiple circuit shots (for capturing probabilistic output distributions) cannot compensate for noise drift errors. While such severe noise drift might not always occur, milder noise drifts that continuously steer circuit results away from the correct objective ground state are frequently encountered and still perturb the accuracy of results.

B. Variational Quantum Algorithms

VQA, a hybrid quantum-classical algorithm, is widely used in chemistry and approximations such as Variational Quantum Eigensolver (VQE) [10] and Quantum Approximate Optimization Algorithm (QAOA) [11]. The VQA problem can be generalized to a Hamiltonian, $\hat{H}$, a linear combination of Pauli terms $\hat{P}_i$, and their numerical coefficients c_i, to describe the total energy of the system, as shown in Equation 1.

$$\hat{H} = \sum_i c_i \hat{P}_i \qquad (1)$$

The objective function is to find the ground state energy of the system, which corresponds to the lowest eigenvalue, λ_{min}, of $\hat{H}$ [10]. To approximate the expectation value for $\hat{P}_i$, the parameterized quantum circuit (ansatz) with parameters $\overrightarrow{\theta}, \theta \subseteq \mathbb{R}$, are iteratively executed on QPUs and tuned by a classical optimizer. The expectation value is derived from ansatz measurements over different observables (bases). The objective function is represented by

$$\min \lambda(\overrightarrow{\theta}) = \min \langle \psi(\overrightarrow{\theta})|\hat{H}|\psi(\overrightarrow{\theta})\rangle = \min \sum_i c_i \langle \psi(\overrightarrow{\theta})|\hat{P}_i|\psi(\overrightarrow{\theta})\rangle \quad (2)$$

$|\psi(\overrightarrow{\theta})\rangle$ is the eigenvector corresponding to $\lambda(\overrightarrow{\theta})$. Ideally, the application monotonically converges to the minimum. However, machine noise, such as various noise drifts, can taint the gradient computation and thus bias the VQA optimization process. Therefore, VQA relies on noise-robust optimizers to become noise resilience [9], [49].

III. MOTIVATION

A. Dynamic Noise Landscape Navigation

Classical gradient-based optimizers in VQA are implemented in multiple ways to calculate the gradient based on the prior gradients and then tune the parameters of the ansatz to drive the objective function into a ground state. The underlying assumption of the optimizer is the evaluation of the gradient under the same noise environment.

To better understand the impact of noise drift in VQA, Fig.3 visualizes the VQA landscape for a minimization problem under different scenarios. Each contour level represents a value of the objective function, with a darker blue color indicating a smaller value in the objective function estimation. The orange dot represents the objective function value for each parameter
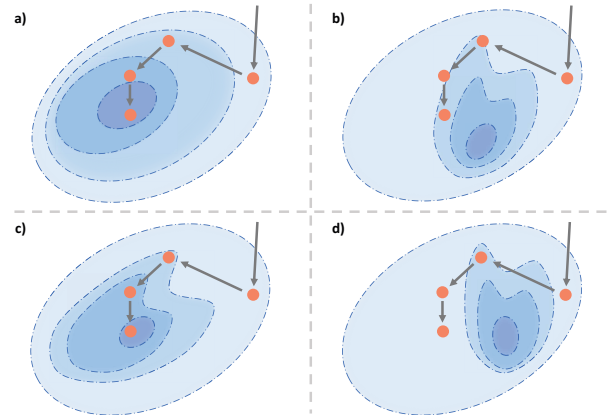

Fig. 3. VQA landscape navigation. Contour levels correspond to different objective function values, with darker colors indicating smaller values. The orange dot marks the objective function value for each parameter configuration chosen by the optimizer. The tuning process is represented by an arrow trace. a) represents the drift-free (ideal) scenario; b) and d) depict scenarios with out-of-range noise drift; c) illustrates a scenario where noise drift is acceptable.
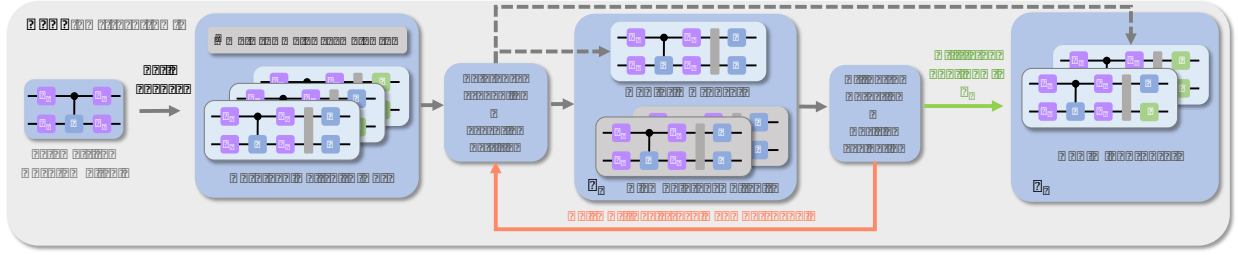
1064

Fig. 4. Overview of DISQ: a VQA iteration execution is partitioned into two stages (,); Prime subset corresponding to the current iteration (blue circuit in) and its references (gray circuits) are executed in to detect the noise drift; If noise drift is present, DISQ skips the results of the current job and reschedules all the circuits in via the next job (orange line). Otherwise, minor subset corresponding to the current iteration (blue circuits in) are executed to proceed with VQA (green line).

configuration selected by the optimizer. Ideally, the VQA optimizer steadily tunes the parameters towards a minimum value of the objective function, which is represented by moving the orange point from the contour level in light blue to the deeper blue contour level, as shown in Fig.3-a. However, in reality, the presence of noise drift alters contour levels' shape and position, leading to discrepancies between the machine-obtained and ideal values. Such discrepancies even vary with each iteration of the optimization process. In comparison with Fig.3-a (drift-free), Fig.3-b illustrates how optimizer tuning is disturbed by severe noise drift. Optimizers are incapable of finding the minimum value in such a landscape with substantial fluctuations in objective function estimation. Therefore, it becomes crucial to estimate noise drift errors and craft a stable environment for the gradient calculation of the optimizer. Specifically, optimizers only process VQA iterations under scenarios such as Fig.3-c, where the machine-obtained contour levels are loosely aligned with the ideal (drift-free) contour levels in Fig.3-a.

B. Limitations in Dynamic Noise Estimation

Traditional VQA optimizers calculate the tuning gradient for iteration from the discrepancy between the machine-obtained objective function estimations (or energy) of iterations and to select the next iteration parameter:

$$\_ \tag{3}$$

where denotes the energy obtained by the quantum machine corresponding to iteration .

Regarding the cutting-edge transient mitigation technique QISMET [39], it selects the adjacent previous iteration as the reference circuit and re-executes it in the current iteration to estimate noise errors for iteration . Noise impact is estimated by comparing the energy of the reference circuit repetition in the current iteration _ with its previous execution _ :

$$\_ \quad \_ \tag{4}$$

The noise-free energy is predicted by removing the noise error component from the energy estimation using Equation 5. The ideal gradient is then calculated with Equation 6.

$$\tag{5}$$

$$\_ \tag{6}$$

QISMET then employs gradients and to govern VQA progress. While this approach can address transient noise with short duration and spike-like magnitude, it proves insufficient for tackling prolonged and non-deterministic noise drift, as shown in Fig.2, leading to deceptive and detrimental iterations during tuning. Furthermore, the noise estimation process requires retrial, re-executing the reference circuit with current circuit configurations, which doubles the computational resources compared to traditional VQA iterations. Further details are discussed in Section IV.

IV. DISQ DESIGN

DISQ framework overview is depicted in Fig.4. It appends the observable circuits of prime subsets from multiple previous iterations, which serve as references to diligently detect the noise drift in VQA iterations. Noise drift errors are assessed by comparing the energy of the references in previous jobs with their repetitions in the current job. This process enables DISQ to decide whether to reschedule or accept a particular iteration, thus controlling VQA progression. The orange arrow indicates the action in case of the existence of noise drift: rescheduling of current circuits via the next job. The green line indicates the accepted scenario, proceeding to stage for total energy estimation. The design details of DISQ are discussed in the following sections, with a focus on three key insights: multi-reference enhancement, computing overhead minimization, and enhanced noise drift detection.

A. Multi-Reference Enhancement

The potential for detrimental impacts of noise drift errors is presented in Fig.2, which is seen beyond the statistical robustness offered by executing multiple circuit shots. In such cases, relying solely on a single adjacent iteration for dynamic noise detection, as with [39], can be deceptive to the optimizer and introduce fallacious iterations with inaccurate objective function (energy) estimations during the tuning process. Utilizing such fallacious references to estimated noise errors in subsequent iterations fails to establish a reliable environment for tuners. Instead, the bias accumulates, potentially causing VQA to be far from its target, as shown in Fig.3-d.

Additionally, deviations in objective function estimation also hinder accurate noise drift detection, particularly when relying on a single reference. Fig.5 presents the circuit expectation values collected over roughly a 24-hour period, featuring
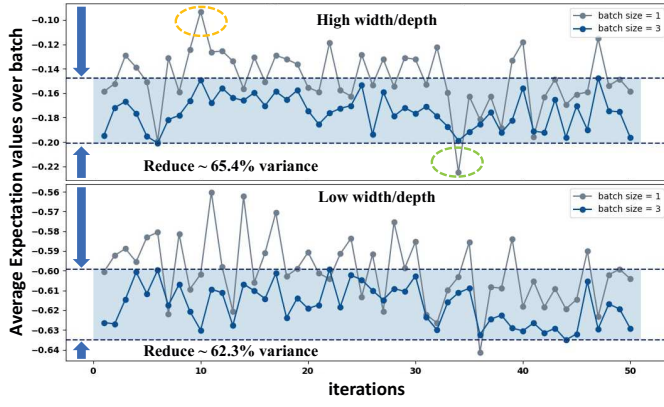
Fig. 5. The benefits of averaging in reducing deviations. Expectation value data over 50 executions of two circuit batches (blue line: three different circuits; gray line: one circuit) is compared in two different circuit features on IBMQ Lagos.

two qubit/gate-level characteristics and two circuit batch sizes. The gray line in each plot corresponds to a batch consisting of only one circuit configuration. In the bottom plot, the gray line has an average of about -0.59 and a range value of 0.08. Significant outliers (negative and positive, circled in green and orange, respectively) are observed in the top plot. The gray line in the top plot exhibits an average of around -0.14 and a large range value of 0.13, highlighting the challenges arising from fluctuating results in objective function estimations within noisy landscapes. Such deviations impede accurate detection of the noise drift and result in the termination of the VQA tuning process far from a minimum, such as Fig.3-d.

In contrast to the single-circuit configuration shown in Fig.5, the blue line in the same figure represents a batch consisting of three different circuits. This approach attenuates a significant portion of the fluctuations in both sub-figures, decreasing the variance by 65.4% and 62.3%, respectively. These results indicate that incorporating multiple references improves the reliability of noise detection, thereby mitigating the negative impacts of the aforementioned adverse factors, including fallacious iterations in the reference set and deviations in executing the references. Consequently, the use of multi-references is beneficial for VQA since reliable noise drift detection facilitates maintaining the gradient calculation of VQA in drift-free scenarios.

B. Computing Overhead Minimization

Dynamic noise estimation, discussed in Section III-B requires bundling the circuit configurations corresponding to the current iteration and reference circuit in each VQA iteration. This process doubles the required computing resources compared to a traditional VQA iteration. Section IV-A highlights the benefits of using multiple references, which further increases the computing overhead linearly with the number of reference circuits involved. Additionally, to align the machine-obtained gradient with the ideal gradient, several iterations are skipped when detrimental noise drift occurs. In extreme cases, it may even be necessary to execute the circuit bundle several times to accept a single iteration. Although it is
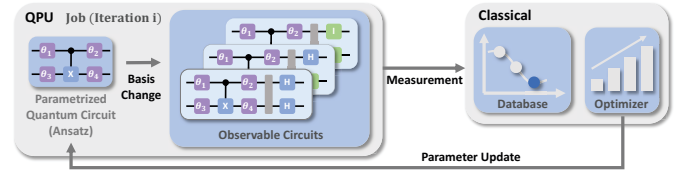

Fig. 6. Variational quantum algorithm (VQA), a hybrid quantum-classical algorithm

somewhat efficient in mitigating noise drifts, this level of computing overhead is unacceptable for a simple task and is even more severe for long-running applications such as VQA. To eliminate the aforementioned computing overhead, Pauli-term subsetting is proposed.

Equation 7 defines a Hamiltonian $\hat{H}$, which describes the total energy of a system and serves as the objective function of a problem.

$$\hat{H} = 1.4 \cdot XX + 0.05 \cdot ZI + 0.02 \cdot ZX \qquad (7)$$

To estimate the total energy (or objective function) for $\hat{H}$ with a parameterized quantum circuit (ansatz) via VQA, each Pauli term is converted into its corresponding basis to generate observable circuits (OC) for their energy estimation. Subsequently, all terms are aggregated for the total energy of $\hat{H}$. Fig.6 depicts a VQA iteration of $\hat{H}$, where the circuits with blue backgrounds are the OCs that are bundled into a "job" to be executed for energy estimation [50]. In other words, even a simple Hamiltonian like $\hat{H}$ necessitates QPU to execute multiple OCs for a single iteration.

In Equation 7, the Pauli term "XX" holds the largest absolute coefficient value, which accounts for over 95% of all the coefficients of Pauli terms. This indicates that the OC corresponding to the Pauli term "XX" (the circuit in Fig.6, with each qubit followed by a Hadamard gate) significantly impacts the energy estimation of $\hat{H}$, as shown in Fig.7. Hence, we introduce a new term "prime subset" to refer to such OCs that are deemed to dominate the energy estimation of a Hamiltonian. Specifically, the OC with an absolute coefficient exceeding a specified threshold value TH_p relative to the sum of absolute coefficients for all Pauli terms. The setting of threshold value TH_p is discussed further in Section V-C1. Note: The prime subset is not limited to a single circuit but rather is a collection of circuits. Additionally, we define a "minor subset", which are the remaining OCs.
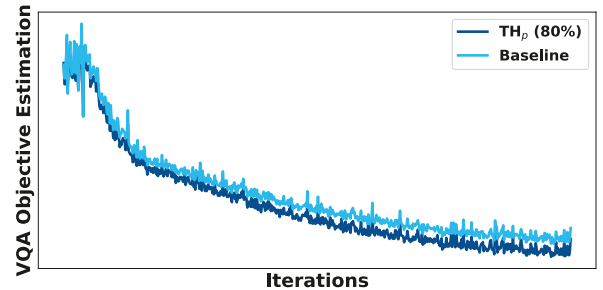

Fig. 7. Traditional VQA convergences. The light blue line represents the training trace of the Hamiltonian $\hat{H}$, while the dark blue line illustrates the convergence of the prime subset with a setting of $TH_p = 80\%$. Prime subset dominates the energy estimation of $\hat{H}$.

Based on the Pauli-subsetting design, DISQ actively takes the following steps in each iteration to alleviate computing overhead without sacrificing the reliability of noise drift detection: (i) rearranging the execution order of OCs corresponding to the current iteration by partitioning them into two stages (prime subset in S_1, minor subset in S_2); (ii) executing the prime subset of reference circuits for noise drift detection in S_1, instead of all the references' observable circuits; (iii) proceeding to S_2 only if the current iteration passes the noise drift detection.

C. DISQ Noise Drift Detection

As discussed in Section IV-A and Section IV-B, DISQ incorporates multi-references and deploys Pauli-term subsetting to improve the reliability of noise drift detection with less execution time in S_1 stage (Note: The adjacent previous iterations are usually suitable choices for reference since they typically have similar objective function estimations to the circuit configuration in the current iteration, and the adjacent iterations generally have closer noise landscapes compared to other iterations). The noise drift error estimation in Equation 4 is enhanced to incorporate K reference circuits from iteration $\{(i - n) \mid n \in [1..K]\}$:

$$D_i = \sum_{n=1}^{K} c_{i-n}[Er_{i-n}(P) - E_{i-n}(P)] \ s.t. \ \sum_{n=1}^{K} c_n = 1 \quad (8)$$

where (P) denotes the prime subset, K represents the incorporated reference number, which is discussed next in Section V-C2, and c_{i-n} is the proportion factor corresponding to the reference circuits from the previous iteration $(i - n)$. c_{i-n} is fixed at $\frac{1}{K}$ to ensure the sum of all factors equals 1 (Note: dynamic factors could potentially enhance benefits even further, but this is beyond the scope of our current study).

With the design of K reference circuits, DISQ provides a more faithful estimation of the drift-free energy $Ef_i(P)$ and the corresponding drift-free gradient Gf_i. The calculation process is enhanced using the updated Equations 9 and 10, respectively:

$$Ef_i(P) = E_i(P) - D_i \quad (9)$$

$$Gf_i = Ef_i(P) - \sum_{n=1}^{K} c_{i-n} \cdot E_{i-n}(P) \quad (10)$$

The multi-references based machine-obtained G_i in Equation 3 is described below:

$$G_i = E_i(P) - \sum_{n=1}^{K} c_{i-n} \cdot E_{i-n}(P) \quad (11)$$

To handle various intractable noise drift, DISQ employs an intelligent control policy to diligently guide VQA, ensuring its training remains in a mild landscape. The aforementioned gradients Gf_i and G_i are utilized to make informed decisions, which control VQA progress by accepting or rescheduling particular iterations. The underlying principle is to accept VQA iterations only if the direction of Gf_i (gradient obtained from DISQ noise drift detection) coincides with the direction of G_i (gradient observed by the VQA tuner based on machine energy estimations). Fig.8-a, c describe the acceptance
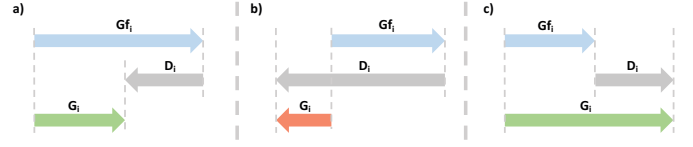

Fig. 8. DISQ control policy. a) and c): Machine-obtained and drift-free gradients have the same direction, thus acceptable; b): Noise drift flips the machine-obtained direction, thus iteration is rescheduled.

scenarios. This precautionary approach ensures that the entire VQA tuning process takes place under the same or similar landscapes, enabling the tuner to steadily and reliably approach its objective without deviating from the target due to the negative impact of noise drift.

For the scenario in Fig.8-b, in response, DISQ skips the result in the current iteration due to noise drift errors and reschedules the job. As a result, identical circuit configurations are repeated in the next iteration. Note that noise drift effects can persist for extended periods of time. Therefore, the skipping and repetition may span multiple jobs but are limited by a max-out σ. Once max-out is reached, DISQ deems that the landscape has completely changed, rendering the references' energy in previous jobs inapplicable to the current landscape. In this situation, DISQ updates the reference energy (used only for detection) and detects the noise drift based on the updated reference energy. Algorithm 1 outlines the process.

D. Functionality Across Iterations

DISQ workflow with three references over several continuous VQA iterations is illustrated in Fig.9. In S_1 stage of job α, QPU executes the circuits consisting of the prime subset corresponding to iteration (i), depicted as the inner box with blue background, and the configuration of prime subsets corresponding to iteration $\{(i - n) \mid n \in [1..3]\}$, depicted with light gray background. Then, the input of the noise

---

Algorithm 1 DISQ Workflow

Require: execute N iterations
  $i = 0$, $i_r = 0$
  while $i \neq N$ **do**
    $\{E_i(P), Er_{i-n}(P)\} \leftarrow QPU(P_i, P_{i-n}) \mid n \in [1..K]$
    for $n \in [1..K]$ **do**
      $E_{i-n}(P) \leftarrow$ Database
    end for
    $D_i \leftarrow \sum_{n=1}^{K} c_{i-n}[Er_{i-n}(P) - E_{i-n}(P)]$
    $Ef_i \leftarrow E_i(P) - D_i$
    $Gf_i \leftarrow Ef_i(P) - \sum_{n=1}^{K} c_{i-n} \cdot E_{i-n}(P)$
    $G_i \leftarrow E_i(P) - \sum_{n=1}^{K} c_{i-n} \cdot E_{i-n}(P)$
    if $G_i \cdot Gf_i > 0$ **then**
      $E_i(M) \leftarrow QPU(M_i)$
      $i \leftarrow i + 1$
      Database $\leftarrow E_i(P) + E_i(M)$
    else
      $i_r \leftarrow i_r + 1$
      if $i_r == \sigma$ **then**
        Database $\leftarrow \{Er_{i-n}(P)\} \mid n \in [1..K]$
        Database $\leftarrow E_i(P) + E_i(M)$
        $i_r \leftarrow 0$
      end if
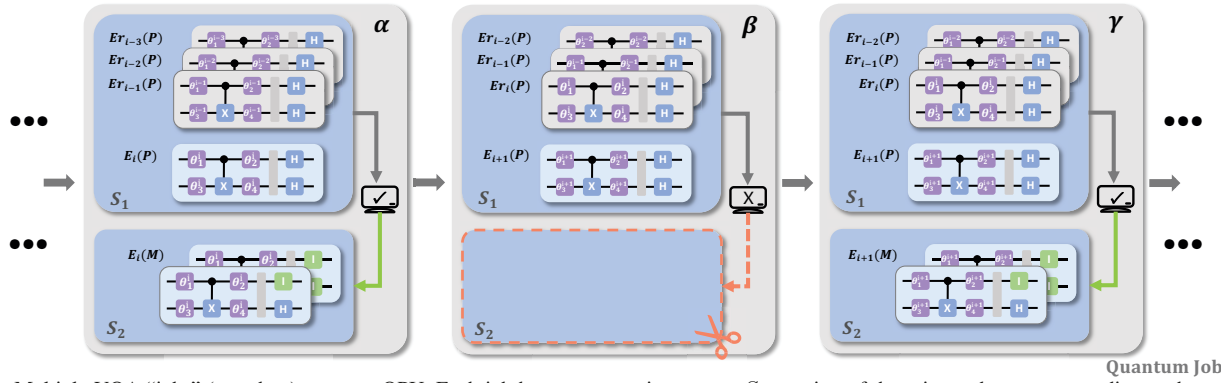   **end if**
  end while

Fig. 9. Multiple VQA "jobs" (gray box) run on a QPU. Each job has two execution stages. S_1 consists of the prime subset corresponding to the current VQA iteration (circuits with blue background) and the prime subset of reference circuits appended by DISQ (circuits with gray background). S_2 is determined by DISQ (as check mark ✓ or X) and executes the minor subset corresponding to the current VQA iteration. VQA progresses when both S_1 and S_2 have completed.

drift detection includes the above executed results $E_i(P)$ and $\{Er_{i-n}(P) \mid n \in [1..3]\}$ along with reference prime subset energy from previous jobs $\{E_{i-n}(P) \mid n \in [1..3]\}$ (not shown in the figure). In job α, DISQ indicates the impact of noise drift on iteration (i) is acceptable (check mark). Therefore, job α proceeds to S_2 stage and the minor subset corresponding to iteration (i) is scheduled to execute for the total energy estimation of iteration (i). VQA thus progresses to iteration $(i+1)$.

Job β is the adjacent job right after job α to compute the circuits of iteration $(i+1)$. The prime subset of iteration (i) replaces the prime subset of iteration $(i-3)$ as one of the reference circuits for this job. Other executed circuits in stage S_1 are shown in the figure. Once stage S_1 of β is executed, the following information is utilized for noise drift detection:
a) prime subset energy corresponding to iteration $(i+1)$;
b) $\{E_{i-n}(P) \mid n \in [0..2]\}$, prime subset energy of reference circuits from the previous jobs;
c) $\{Er_{i-n}(P) \mid n \in [0..2]\}$, the repetitions of reference circuits in the current job β.

In this example, DISQ detects that the noise drift error is out of range and derails VQA convergence (indicated by an X in the figure). Therefore, iteration $(i+1)$ is rescheduled, and the outcomes of job β are discarded. The figure shows the termination of job β (S_2 is not executed) and a subsequent repetition of circuits via job γ.

Once stage S_1 of job γ is completed, DISQ checks again and determines whether the noise drift present in job β that still exists in job γ. In this job, DISQ observes that the noise drift has passed or deems its impact to be within an acceptable range. The S_2 stage of job γ is processed and contributes further to the progression of VQA. By taking above steps in each VQA iteration, DISQ provides a stable landscape for the gradient computation of the tuning process.

V. EVALUATION

A. Evaluation Methodology

1) General Infrastructure: DISQ is a software optimization approach implemented using the Qiskit Runtime Python package [50], which allows for running quantum programs en-

tirely on the IBM Cloud and executing workloads on quantum systems at scale. DISQ is broadly applicable across all VQA applications and can be integrated into any VQA classical optimizer to enhance noise drift resilience. In this work, we restrict the evaluation to eight qubits considering the machine limitations on circuit metrics (depth, width). The evaluation deploys SPSA as the main classical optimizer and primarily focuses on VQE, one of the VQA applications (introduced in Section II-B). The evaluations encompass four Hamiltonians, four different ansatzes, and six QPUs from IBM.

2) Benchmarks: The primary Hamiltonian evaluated is the potential energy of the helium hydrogen ion over bond lengths of 1.7 $\mathring{A}$, with additional evaluations for the potential energy of the hydrogen fluoride molecule, the lithium hydride molecule, and the hydrogen molecule. The molecules are labeled with a superscript "c", indicating that the number of terms in their Hamiltonians has been reduced using the reduction method from [51]. This method is capable of compressing large fermionic Hamiltonians into several qubits, enabling more efficient calculations and analysis. The hardware-efficient SU2 [52] and RA [53] ansatz are used in the experiments, varying the block repetitions between 2, 6, and 10 to change the number of parameters in the ansatz. The selected QPUs are Kolkata (27 qubits), Toronto (27 qubits), Montreal (27 qubits), Perth (7 qubits), Jakarta (7 qubits), and Lagos (7 qubits). All the machine details can be found at IBMQ's website [54].

Despite the convenience of accessing IBM's quantum machines through the cloud and facilitating the workload with Qiskit Runtime, the limited access to quantum machines still prevents a holistic evaluation of our proposal. To enable fine-granularity noise drift evaluation, we combine the noise trace generated by the IBM machine Toronto with the enhanced noise traces upon the model from [39]. This enhanced model captures the noise drift effects in each iteration and normalizes these effects to the magnitude of VQA estimations, allowing the simulations to exactly mimic observed noise drift errors.

3) Baselines: The following schemes are assessed in several comparative evaluations:

- DISQ: Setting the number of optimal reference circuits to three (two for HeH$^+$ ion) and the prime threshold TH$_p$

to 80%.

- Baseline: Traditional variational quantum eigensolver with deploying SPSA for optimization.
- QISMET: Transient noise mitigation approach [39] with tuning the threshold to 10% to provide optimal performance.

B. Evaluation Results

1) Evaluation on QPUs: The evaluation of DISQ, QIS-MET, and Baseline for VQA energy estimation is conducted on IBMQ machines to solve the VQA energy estimation problem for a four qubits HeH^+ ion, with a fixed number of 350 iterations. The experiments are conducted synchronously to ensure temporal adjacency between the iterations of DISQ and other schemes.

The results in Fig.10 compare DISQ against other schemes on IBMQ Montreal device. Two periods of significant fluctuations, which heavily derail both QISMET and Baseline, are highlighted. QISMET only eliminates the noise instance in the second period and fails to avoid the first serious turbulence due to the variation in perceived gradient estimates. In contrast, benefiting from multi-reference circuits, DISQ predominantly bypasses both turbulent periods ensuring steady progress, ultimately achieving improvements of 59.8% and 44.9% over Baseline and QISMET, respectively.

Fig.11 shows a comparison of DISQ against Baseline and QISMET on IBMQ Toronto. The noise drift errors persist for an extended period (from 100 to 170 iterations), resulting in multiple instances of noise drift errors with moderate or tiny magnitudes. The tuning of QISMET and Baseline stagnates during this period. Although DISQ is also impacted for several early iterations, it quickly recovers and continues its navigation to the target, improving fidelity over Baseline and QISMET by 29.0% and 37.4%, respectively.

In the evaluation on IBMQ Jakarta shown in Fig.12, the machine behavior is mostly smooth except for one instance of an inconspicuous noise drift error, which is highlighted. While this period does not seem to cause any severely detrimental impact, QISMET is deceived by a mediocre estimation and languishes at a local optimum. DISQ heuristically identifies this camouflaged noise drift period and is able to further progress toward the global optimum target with benefits of 51.6% and 36.3% over Baseline and QISMET, respectively.
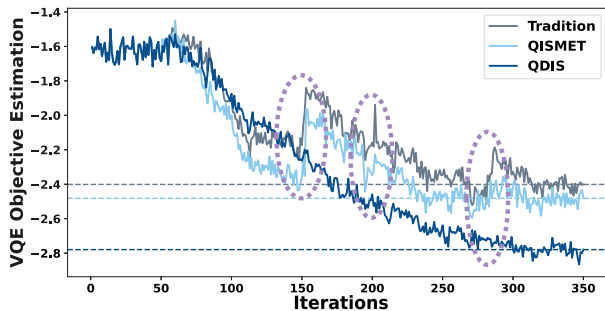

Fig. 10. DISQ benefits for a HeH^+ VQA application on IBMQ Montreal with several high fluctuating noise regions (circled). Sharp noise drift errors are avoided by DISQ.

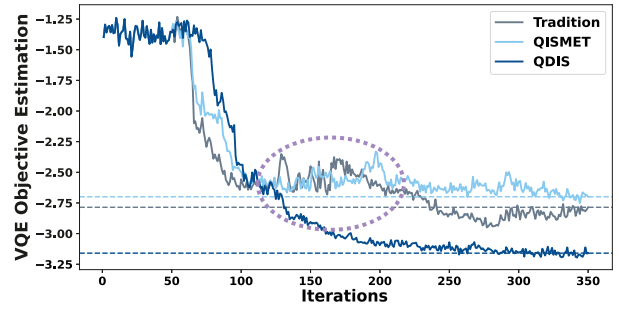Fig. 11. A HeH^+ VQA application on IBMQ Toronto with an instance of noise drift error spanning multiple iterations (circled). This phase of noise is bypassed by DISQ, thus improving convergence of application.
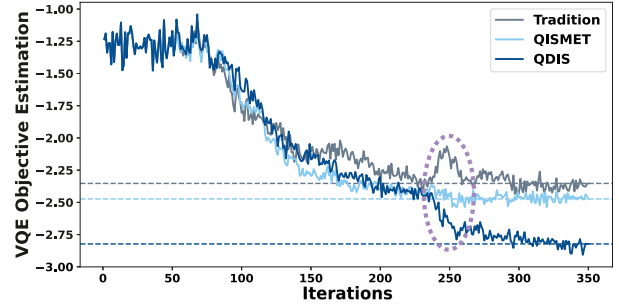

Fig. 12. A HeH^+ VQA application on IBMQ Jakarta with an inconspicuous instance of noise drift error (circled). The convergence of application benefits from DISQ by skipping this malicious instance of noise.

The benefits of DISQ over Baseline and QISMET on six different IBMQ machines are shown in Fig.13. The primary vertical axis shows the improvements in measured VQA expectations over Baseline, while the secondary vertical axis shows the improvement over QISMET. Across all the machines, DISQ consistently improves the expectation values over Baseline by 1.29-$2.057\times$, with a mean improvement of $1.67\times$. Furthermore, DISQ boosts the fidelity by an average of 45% (up to 62.1% and by at least 35.5%) over QISMET. The above improvements are achieved over 350 iterations due to access constraints. It is expected that benefits increase with more iterations since this would provide higher potential for more noise drift errors.

2) Evaluation with Multiple Benchmarks: A thorough comparison is conducted with QISMET and Baseline to comprehensively evaluate the performance of DISQ across the six different applications listed in Table I. The evaluation is conducted over 1000 simulation iterations using different molecules and ansatzes to explore the adaptability of DISQ. The total number of OC for each molecule is listed under
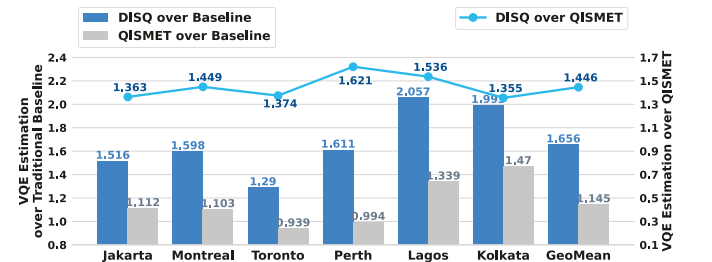

Fig. 13. DISQ's benefits for a HeH^+ VQA application on six IBMQ machines

TABLE I
APPLICATION INFORMATION SUMMARY

Application	Molecule	Qubits	Ansatz	Layer Reps	# Observable
HF	HF	8	RA	6	76 (8)
HF-SU2	HF	8	SU2	6	76 (8)
HF-RA-2	HF	8	RA	2	76 (8)
HF-RA-10	HF	8	RA	10	76 (8)
LiH	LiH	6	RA	6	13 (4)
HeH	HeH^+	4	RA	6	4 (1)

the "# Observable" column and the values in the parentheses are the number of OC in its prime subset. Expectation value comparisons are presented on the left side of Fig.14. DISQ consistently outperforms Baseline, achieving improvements ranging from 1.51 to 2.24. Moreover, DISQ improves the fidelity by up to 1.89 over QISMET. Multi-reference circuit scheduling employed in DISQ showcases its superior performance under convoluted, noisy traces across various molecules and ansatzes, which is consistent with the discussion presented in section IV-A.

The number of circuit executions in each scheme enables an intuitive comparison of computation overhead, depicted on the right side of Fig.14. The leverage of Pauli-term subsetting not only allows DISQ to schedule more reference circuits in noise detection, but also boosts its computing speed. DISQ actively reduces executed circuit cost by up to 39.2% with a mean of 23.5% across the application, and boosts noise drift detection speed by an average of 2.07 over QISMET.

3) Molecule Estimation Evaluation: VQE is extensively used in molecule chemistry to estimate the energy value of a molecule in a specific geometry, where energy variation indicates the chemical reaction rates for the molecule. The geometry of a molecule typically varies with different bond lengths, resulting in a multitude of Hamiltonians [10]. The energy of molecules is calculated from the expectation value of their Hamiltonians. However, noise drift errors affect the expectation values differently for different bond lengths, leading to skewed energy differences. Fig.15 depicts the potential expectation for the molecule over 10 different H-H bond lengths, each corresponding to a unique Hamiltonian and VQE experiment. The gray line represents the noise-free scenario, while DISQ closely models the noise-free trace by accurately estimating the expectation for each bond length. However, QISMET and Baseline deviate from the ideal scenario as the bond length descends, especially at lower bond lengths where the noise drift errors have a substantial impact. At a bond
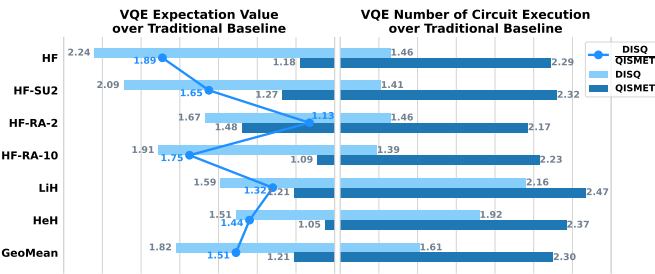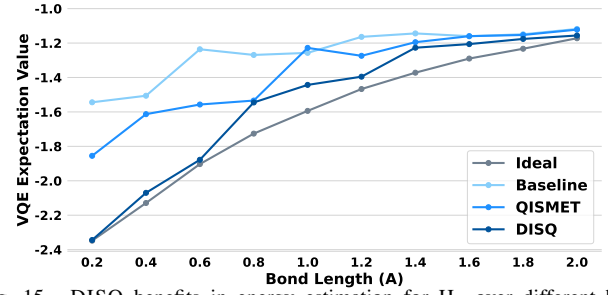

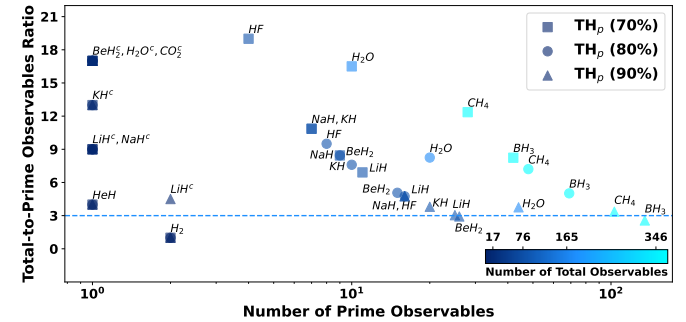Fig. 15. DISQ benefits in energy estimation for H over different bond lengths.

length of 0.8, the step-wise estimation difference relative to the previous length of 0.6 has opposite gradients when comparing Baseline and Ideal, which results in misleading indications in the chemical reaction rates for Baseline.

C. Analysis

1) Sweeping over DISQ Prime Threshold: In section IV-B, we discuss the use of a prime threshold TH to delineate the criteria for Pauli-term subsetting. To identify an appropriate threshold for the prime subset, we conduct a statistical analysis of the observable circuit (OC) number in different molecules with different TH. Our findings reveal that the energy of most molecules heavily relies on a few OCs, aligning with our prime subset design and highlighting the substantial potential of DISQ to conserve computational resources.

The statistical results from our analysis are summarized in Fig.16, which depicts evaluated molecule configurations in terms of their OC numbers in the prime subset and total-to-prime ratio (the ratio of the total number of OCs to the number of OCs in the prime subset). Each point is color-coded by OC number and shaped by the applied TH. The y-axis represents the total-to-prime ratio, with higher points indicating greater potential for DISQ to reduce computing overhead. The x-axis (log scale) represents the OC number in the prime subset. The molecules, denoted by the superscript, represent compressed molecules, as introduced in Section V-A.

To further investigate the optimal TH for molecules, four different thresholds (50, 70, 80, and 90) are analyzed along with Baseline in Fig.17. The evaluation is conducted on simulation across three molecules: HeH, LiH, and HF (Note: 60 has the same number of OC as 70 in HeH and LiH molecules, while HeH maintains a consistent OC count across these
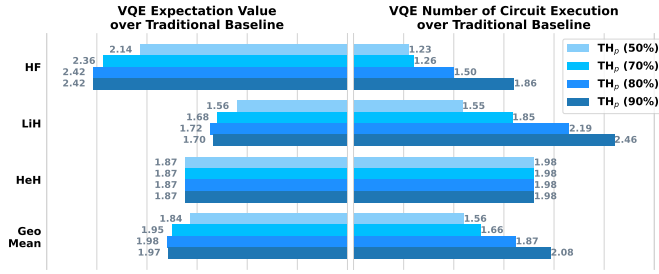

Fig. 14. DISQ provides consistent benefits outperforming the baseline in six different applications.

Fig. 16. Statistics of prime subset and the ratio of total-to-prime across different molecules with different TH values.

Fig. 17. Simulation evaluation with varying DISQ prime threshold values TH_p. HeH+ and LiH molecules exhibit the same number of OC at TH_p and TH_p, while HeH+ maintains a consistent OC count across these thresholds.

thresholds). The 50 and 70 percent thresholds skip fewer impactful noise drift instances, resulting in relatively poor results. Although the 90 threshold yields similar improvements as 80, it requires more circuits to be executed. Consequently, the optimal threshold for selected molecules is found to be 80, which achieves strong benefits over Baseline while minimizing computing resources. Note that intelligent dynamic thresholding for different molecules may further improve benefits, but that is beyond the scope of this study.

2) Sweeping over DISQ References Number: Section IV-A discusses using multi-references to improve the reliability of noise detection. Four different K values (number of incorporated references) are analyzed alongside Baseline in Fig.18 for three molecules HeH+, LiH, and HF.

For all the selected molecules, $K = 1$ is a conservative scheme, which accepts some detrimental noise drift instances, and thus performs negative impacts on the VQA trace. For HeH+ ion, incorporating three or four references for noise drift detection pushes DISQ to be worse than two references, since some iterations with tiny noise drift errors are avoided unnecessarily. Incorporating three references for molecules HF and LiH provides a good trade-off, achieving a 2.47× and 1.71× improvement over Baseline, respectively.

VI. RELATED WORK

To contribute to the success of quantum computing during the NISQ era, it is critical to comprehend and control sources of noise, which typically include thermal fluctuations [14] and magnetic flux [16]–[18]. Other sources, such as cosmic rays [55], device defects [13], quantum drift [48] and external stimuli [20]–[22] also cause the pernicious effects on qubits.

By understanding noise, multiple works from different hierarchies have investigated both error correction [56]–[59] and error mitigation techniques to efficiently drive the computing
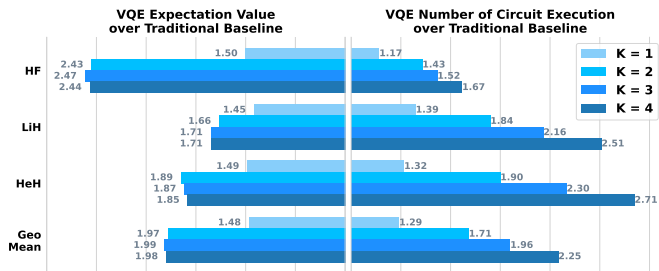

Fig. 18. Simulation evaluation of DISQ using K reference circuits.

power of quantum [28]–[33]. Some works target specific noise sources and propose corresponding solutions, such as crosstalk-oriented compilation methods [27] and experiment-based techniques for measurement error [26]. Beyond these works, mapping virtual to physical qubits is also a primary technique used to alleviate noise effects [37], [60]. The works in [61], [62] even ensemble diverse qubit mappings or multi-device mappings. Other proposed mitigation techniques include learning-based methods [63]–[65], co-design of training, and noise robustness [24], [35] and circuit structure analysis-based one [66] and pulse-level optimizations [67]. However, the preceding works all primarily focus on static noise.

Regarding dynamic noise, [39] profiles spike-like noise impacts and proposes a method with reference circuits. Comparatively, our overhead-aware scheduling method can provide significantly more accurate detection in facing various types of noise drift (not limited to transient spikes) while heuristically minimizing required computing resources.

VII. CONCLUSION

DISQ, as proposed in this paper, takes proactive steps to address the detrimental impact of noise drifts and to neutralize the required execution overhead. To achieve this, DISQ incorporates multi-reference circuits to faithfully detect noise drift errors in VQA iterations in order to maintain the fidelity of the iterations in drift-free scenarios. Furthermore, DISQ deploys Pauli-term subsetting to replace reference circuits with their prime subset circuits and partition the circuit execution to efficiently reduce the corresponding circuit execution overhead by detecting noise drift.

REFERENCES

[1] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets," *Nature*, vol. 549, no. 7671, pp. 242–246, 2017.

[2] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. Sawaya *et al.*, "Quantum chemistry in the age of quantum computing," *Chemical reviews*, vol. 119, no. 19, pp. 10 856–10 915, 2019.

[3] N. Lambert, Y.-N. Chen, Y.-C. Cheng, C.-M. Li, G.-Y. Chen, and F. Nori, "Quantum biology," *Nature Physics*, vol. 9, no. 1, pp. 10–18, 2013.

[4] L. K. Grover, "A fast quantum mechanical algorithm for database search," in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, 1996, pp. 212–219.

[5] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM review*, vol. 41, no. 2, pp. 303–332, 1999.

[6] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," *Nature*, vol. 549, no. 7671, pp. 195–202, 2017.

[7] P. Rebentrost, M. Mohseni, and S. Lloyd, "Quantum support vector machine for big data classification," *Physical review letters*, vol. 113, no. 13, p. 130503, 2014.

[8] J. Preskill, "Quantum computing in the nisq era and beyond," *Quantum*, vol. 2, p. 79, 2018.

[9] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn *et al.*, "Quantum optimization using variational algorithms on near-term quantum devices," *Quantum Science and Technology*, vol. 3, no. 3, p. 030503, 2018.

[10] 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," *Nature communications*, vol. 5, no. 1, p. 4213, 2014.

[11] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithm," *arXiv preprint arXiv:1411.4028*, 2014.

[12] A. Roggero, A. C. Li, J. Carlson, R. Gupta, and G. N. Perdue, "Quantum computing for neutrino-nucleus scattering," *Physical Review D*, vol. 101, no. 7, p. 074038, 2020.

[13] J. M. Martinis, K. B. Cooper, R. McDermott, M. Steffen, M. Ansmann, K. Osborn, K. Cicak, S. Oh, D. P. Pappas, R. W. Simmonds *et al.*, "Decoherence in josephson qubits from dielectric loss," *Physical review letters*, vol. 95, no. 21, p. 210503, 2005.

[14] P. Klimov, J. Kelly, Z. Chen, M. Neeley, A. Megrant, B. Burkett, R. Barends, K. Arya, B. Chiaro, Y. Chen *et al.*, "Fluctuations of energy-relaxation times in superconducting qubits," *Physical review letters*, vol. 121, no. 9, p. 090502, 2018.

[15] J. J. Burnett, A. Bengtsson, M. Scigliuzzo, D. Niepce, M. Kudra, P. Delsing, and J. Bylander, "Decoherence benchmarking of superconducting qubits," *npj Quantum Information*, vol. 5, no. 1, pp. 1–8, 2019.

[16] S. Schlör, J. Lisenfeld, C. Müller, A. Bilmes, A. Schneider, D. P. Pappas, A. V. Ustinov, and M. Weides, "Correlating decoherence in transmon qubits: Low frequency noise by single fluctuators," *Physical review letters*, vol. 123, no. 19, p. 190502, 2019.

[17] C. D. Bruzewicz, J. Chiaverini, R. McConnell, and J. M. Sage, "Trapped-ion quantum computing: Progress and challenges," *Applied Physics Reviews*, vol. 6, no. 2, p. 021314, 2019.

[18] T. Ruster, C. T. Schmiegelow, H. Kaufmann, C. Warschburger, F. Schmidt-Kaler, and U. G. Poschinger, "A long-lived zeeman trapped-ion qubit," *Applied Physics B*, vol. 122, no. 10, p. 254, 2016.

[19] C. Müller, J. H. Cole, and J. Lisenfeld, "Towards understanding two-level-systems in amorphous solids: insights from quantum circuits," *Reports on Progress in Physics*, vol. 82, no. 12, p. 124501, 2019.

[20] P. Murali, N. M. Linke, M. Martonosi, A. J. Abhari, N. H. Nguyen, and C. H. Alderete, "Full-stack, real-system quantum computer studies: Architectural comparisons and design insights," in *Proceedings of the 46th International Symposium on Computer Architecture*, 2019, pp. 527–540.

[21] J. L. Orrell and B. Loer, "Sensor-assisted fault mitigation in quantum computation," *Physical Review Applied*, vol. 16, no. 2, p. 024025, 2021.

[22] R. de Sousa, X. Hu, and S. D. Sarma, "Effect of an inhomogeneous external magnetic field on a quantum-dot quantum computer," *Physical Review A*, vol. 64, no. 4, p. 042307, 2001.

[23] 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*, 2020, pp. 1001–1016.

[24] D. A. Lidar and T. A. Brun, *Quantum error correction*. Cambridge university press, 2013.

[25] J. Chiaverini, D. Leibfried, T. Schaetz, M. D. Barrett, R. Blakestad, J. Britton, W. M. Itano, J. D. Jost, E. Knill, C. Langer *et al.*, "Realization of quantum error correction," *Nature*, vol. 432, no. 7017, pp. 602–605, 2004.

[26] S. Bravyi, S. Sheldon, A. Kandala, D. C. Mckay, and J. M. Gambetta, "Mitigating measurement errors in multiqubit experiments," *Physical Review A*, vol. 103, no. 4, p. 042605, 2021.

[27] Y. Ding, P. Gokhale, S. F. Lin, R. Rines, T. Propson, and F. T. Chong, "Systematic crosstalk mitigation for superconducting qubits via frequency-aware compilation," in *2020 53rd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*. IEEE, 2020, pp. 201–214.

[28] 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, 2020, pp. 306–316.

[29] G. S. Ravi, K. N. Smith, P. Gokhale, A. Mari, N. Earnest, A. Javadi-Abhari, and F. T. Chong, "Vaqem: A variational approach to quantum error mitigation," in *2022 IEEE International Symposium on High-Performance Computer Architecture (HPCA)*. IEEE, 2022, pp. 288–303.

[30] Y. Li and S. C. Benjamin, "Efficient variational quantum simulator incorporating active error minimization," *Physical Review X*, vol. 7, no. 2, p. 021050, 2017.

[31] B. Pokharel, N. Anand, B. Fortman, and D. A. Lidar, "Demonstration of fidelity improvement using dynamical decoupling with superconducting qubits," *Physical review letters*, vol. 121, no. 22, p. 220502, 2018.

[32] K. Temme, S. Bravyi, and J. M. Gambetta, "Error mitigation for short-depth quantum circuits," *Physical review letters*, vol. 119, no. 18, p. 180509, 2017.

[33] S. S. Tannu and M. K. Qureshi, "Mitigating measurement errors in quantum computers by exploiting state-dependent bias," in *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture*, 2019, pp. 279–290.

[34] S. Endo, Z. Cai, S. C. Benjamin, and X. Yuan, "Hybrid quantum-classical algorithms and quantum error mitigation," *Journal of the Physical Society of Japan*, vol. 90, no. 3, p. 032001, 2021.

[35] H. Wang, J. Gu, Y. Ding, Z. Li, F. T. Chong, D. Z. Pan, and S. Han, "Quantumnat: quantum noise-aware training with noise injection, quantization and normalization," in *Proceedings of the 59th ACM/IEEE Design Automation Conference*, 2022, pp. 1–6.

[36] H. Wang, Z. Li, J. Gu, Y. Ding, D. Z. Pan, and S. Han, "Qoc: quantum on-chip training with parameter shift and gradient pruning," in *Proceedings of the 59th ACM/IEEE Design Automation Conference*, 2022, pp. 655–660.

[37] H. Wang, Y. Ding, J. Gu, Y. Lin, D. Z. Pan, F. T. Chong, and S. Han, "Quantumnas: Noise-adaptive search for robust quantum circuits," in *2022 IEEE International Symposium on High-Performance Computer Architecture (HPCA)*. IEEE, 2022, pp. 692–708.

[38] W. Lavrijsen, A. Tudor, J. Müller, C. Iancu, and W. De Jong, "Classical optimizers for noisy intermediate-scale quantum devices," in *2020 IEEE international conference on quantum computing and engineering (QCE)*. IEEE, 2020, pp. 267–277.

[39] G. S. Ravi, K. N. Smith, J. M. Baker, T. Kannan, N. Earnest, A. Javadi-Abhari, H. Hoffmann, and F. T. Chong, "Navigating the dynamic noise landscape of variational quantum algorithms with qismet," 2022. [Online]. Available: https://arxiv.org/abs/2209.12280

[40] J. Clarke and F. K. Wilhelm, "Superconducting quantum bits," *Nature*, vol. 453, no. 7198, pp. 1031–1042, 2008.

[41] IonQ, "Trapped ion quantum computing," 2023. [Online]. Available: https://ionq.com/

[42] N. M. Linke, D. Maslov, M. Roetteler, S. Debnath, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe, "Experimental comparison of two quantum computing architectures," *Proceedings of the National Academy of Sciences*, vol. 114, no. 13, pp. 3305–3310, 2017.

[43] M. Sun and M. R. Geller, "Efficient characterization of correlated spam errors," *arXiv preprint arXiv:1810.10523*, 2018.

[44] S. S. Tannu and M. Qureshi, "Ensemble of diverse mappings: Improving reliability of quantum computers by orchestrating dissimilar mistakes," in *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture*, 2019, pp. 253–265.

[45] S. Sheldon, E. Magesan, J. M. Chow, and J. M. Gambetta, "Procedure for systematically tuning up cross-talk in the cross-resonance gate," *Physical Review A*, vol. 93, no. 6, p. 060302, 2016.

[46] W. Neuhauser, M. Hohenstatt, P. E. Toschek, and H. Dehmelt, "Localized visible mono-ion oscillator," *Phys. Rev. A*, vol. 22, pp. 1137–1140, Sep 1980. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.22.1137

[47] M. Raizen, J. Gilligan, J. C. Bergquist, W. M. Itano, and D. J. Wineland, "Ionic crystals in a linear paul trap," *Physical Review A*, vol. 45, no. 9, p. 6493, 1992.

[48] T. Patel and D. Tiwari, "Disq: a novel quantum output state classification method on ibm quantum computers using openpulse," in *Proceedings of the 39th International Conference on Computer-Aided Design*, 2020, pp. 1–9.

[49] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, "The theory of variational hybrid quantum-classical algorithms," *New Journal of Physics*, vol. 18, no. 2, p. 023023, 2016.

[50] A.-M. A-tA v, MD SAJID ANIS and et al., "Qiskit: An open-source framework for quantum computing," 2021.

[51] A. Zhao, A. Tranter, W. M. Kirby, S. F. Ung, A. Miyake, and P. J. Love, "Measurement reduction in variational quantum algorithms," *Physical Review A*, vol. 101, no. 6, p. 062322, 2020.

[52] IBM, "Efficientsu2," 2023. [Online]. Available: https://qiskit.org/documentation/stubs/qiskit.circuit.library.EfficientSU2.html

[53] ——, "Realamplitudes," 2023. [Online]. Available: https://qiskit.org/documentation/stubs/qiskit.circuit.library.RealAmplitudes.html

[54] ——, "Ibm quantum," 2023. [Online]. Available: https://quantum-computing.ibm.com/

[55] M. McEwen, L. Faoro, K. Arya, A. Dunsworth, T. Huang, S. Kim, B. Burkett, A. Fowler, F. Arute, J. C. Bardin *et al.*, "Resolving catastrophic error bursts from cosmic rays in large arrays of superconducting qubits," *Nature Physics*, vol. 18, no. 1, pp. 107–111, 2022.

[56] D. A. Lidar and T. A. Brun, *Quantum error correction*. Cambridge university press, 2013.

[57] Y. Ueno, M. Kondo, M. Tanaka, Y. Suzuki, and Y. Tabuchi, "Qulatis: A quantum error correction methodology toward lattice surgery," in *2022 IEEE International Symposium on High-Performance Computer Architecture (HPCA)*. IEEE, 2022, pp. 274–287.

[58] B. M. Terhal, "Quantum error correction for quantum memories," *Reviews of Modern Physics*, vol. 87, no. 2, p. 307, 2015.

[59] J. Chiaverini, D. Leibfried, T. Schaetz, M. D. Barrett, R. Blakestad, J. Britton, W. M. Itano, J. D. Jost, E. Knill, C. Langer *et al.*, "Realization of quantum error correction," *Nature*, vol. 432, no. 7017, pp. 602–605, 2004.

[60] T. Patel and D. Tiwari, "Veritas: accurately estimating the correct output on noisy intermediate-scale quantum computers," in *SC20: International Conference for High Performance Computing, Networking, Storage and Analysis*. IEEE, 2020, pp. 1–16.

[61] S. Stein, N. Wiebe, Y. Ding, P. Bo, K. Kowalski, N. Baker, J. Ang, and A. Li, "Eqc: ensembled quantum computing for variational quantum algorithms," in *Proceedings of the 49th Annual International Symposium on Computer Architecture*, 2022, pp. 59–71.

[62] S. S. Tannu and M. Qureshi, "Ensemble of diverse mappings: Improving reliability of quantum computers by orchestrating dissimilar mistakes," in *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture*, 2019, pp. 253–265.

[63] A. Strikis, D. Qin, Y. Chen, S. C. Benjamin, and Y. Li, "Learning-based quantum error mitigation," *PRX Quantum*, vol. 2, no. 4, p. 040330, 2021.

[64] P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, "Error mitigation with clifford quantum-circuit data," *Quantum*, vol. 5, p. 592, 2021.

[65] H. Wang, P. Liu, J. Cheng, Z. Liang, J. Gu, Z. Li, Y. Ding, W. Jiang, Y. Shi, X. Qian *et al.*, "Quest: Graph transformer for quantum circuit reliability estimation," *arXiv preprint arXiv:2210.16724*, 2022.

[66] S. Deshpande, C. Xu, T. Trochatos, H. Wang, F. Erata, S. Han, Y. Ding, and J. Szefer, "Design of quantum computer antivirus," in *2023 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*, 2023, pp. 260–270.

[67] Z. Liang, H. Wang, J. Cheng, Y. Ding, H. Ren, Z. Gao, Z. Hu, D. S. Boning, X. Qian, S. Han *et al.*, "Variational quantum pulse learning," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*. IEEE, 2022, pp. 556–565.