arXiv:2504.01077v3 [quant-ph] 18 Dec 2025

Double-bracket algorithm for quantum signal processing without post-selection

Yudai Suzuki[1,2], Bi Hong Tiang[3], Jeongrak Son[3], Nelly H. Y. Ng [3,4], Zoë Holmes[1], and Marek Gluza[3]

[1]Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
[2]Quantum Computing Center, Keio University, Hiyoshi 3-14-1, Kohoku-ku, Yokohama 223-8522, Japan
[3]School of Physical and Mathematical Sciences, Nanyang Technological University, 637371, Singapore
[4]Centre for Quantum Technologies, Nanyang Technological University, 637371, Singapore

Quantum Signal Processing (QSP), a framework for implementing matrix-valued polynomials, is a fundamental primitive in various quantum algorithms. Despite its versatility, a potentially underappreciated challenge is that all systematic protocols for implementing QSP rely on post-selection. This can impose prohibitive costs for tasks when amplitude amplification cannot sufficiently improve the success probability. For example, in the context of ground-state preparation, this occurs when using a too poor initial state. In this work, we introduce a new formula for implementing QSP transformations of Hermitian matrices, which requires neither auxiliary qubits nor post-selection. Rather, using approximation to the exact unitary synthesis, we leverage the theory of the double-bracket quantum algorithms to provide a new quantum algorithm for QSP, termed Double-Bracket QSP (DB-QSP). The algorithm requires the energy and energetic variance of the state to be measured at each step and has a recursive structure, which leads to circuit depths that can grow super exponentially with the degree of the polynomial. With these strengths and caveats in mind, DB-QSP should be viewed as complementing the established QSP toolkit. In particular, DB-QSP can deterministically implement low-degree polynomials to "warm start" QSP methods involving post-selection.

Marek Gluza: marekludwik.gluza@ntu.edu.sg

1 Introduction

The efficient implementation of matrix-valued functions plays a central role in the design of modern quantum algorithms [1]. That is, the essence of many quantum algorithms boils down to constructing a polynomial function $p(H)$ of a given Hermitian matrix H and applying it to an input state $|\Psi\rangle$ to obtain a normalized state

$$|\Psi'\rangle = \frac{p(H)\,|\Psi\rangle}{\|p(H)\,|\Psi\rangle\,\|} \tag{1}$$

with $\|\,|\psi\rangle\,\| = \sqrt{\langle\psi|\psi\rangle}$ for any vector $|\psi\rangle$. For example, real and imaginary time evolution correspond to the transformations $p(H) \approx \exp(iHt)$ and $p(H) \approx \exp(-\tau H)$, while matrix inversion implements the transformation $p(H) \approx H^{-1}$. Quantum Signal Processing (QSP) is an algorithmic framework for realizing such polynomial transformations on quantum computers. QSP has enabled the development of advanced quantum algorithms for solving linear systems of equations [2, 3, 4], Hamiltonian simulation [5, 6, 7], and ground state preparation [8, 9].

Despite its versatility, an underappreciated challenge in QSP is the cost of post-selection [10]. For QSP implementation methods such as qubitization [5] and Linear Combination of Unitaries (LCU) [11, 12, 13] to be practically viable, the success probability for post-selection must be sufficiently high to avoid excessive resource overhead. While amplitude amplification techniques can improve success probabilities [14, 15], they may be insufficient when the success probability is exponentially small in the number of

qubits [16]. For instance, ground-state preparation algorithms with nearly optimal resource scaling may still incur exponential costs if the initial overlap between the input state and ground state is exponentially small [16, 9].

In this work, we propose a new QSP implementation that eliminates the need for auxiliary qubits and post-selection. Since the state after normalization in Eq. (1) is a genuine quantum state, there must exist a unitary operator U_Ψ such that $U_\Psi |\Psi\rangle = p(H)|\Psi\rangle / \|p(H)|\Psi\rangle\|$. This work identifies how to systematically perform the unitary synthesis of U_Ψ. The key insight is that any linear polynomial, $aI + bH$ with real coefficients $a, b \in \mathbb{R}$ and the identity operator I, can be *exactly* represented by a unitary operator in form of

$$\frac{(aI + bH)|\Psi\rangle}{\|(aI + bH)|\Psi\rangle\|} = e^{s[|\Psi\rangle\langle\Psi|, H]}|\Psi\rangle \ , \qquad (2)$$

where s is determined by the energy mean and variance. Using this building block, we prove that a recursion involving unitaries in Eq. (2) together with state-dependent reflection gates can realize *arbitrary* polynomial functions (see Fig. 1). We then utilize the recently-established theory of Double-Bracket Quantum Algorithms (DBQA) [17, 18] to derive a unitary synthesis that can be compiled into primitive gates using standard quantum computing methods. This leads to a new quantum algorithm which we call the Double-Bracket QSP (DB-QSP).

The advantages of DB-QSP come with two challenges. First, its recursive form means the depth of circuit required to converge with arbitrary precision grows super exponentially with the degree of the target polynomial functions. However, low-degree approximation techniques [3, 2] can be applied to keep the circuits depths efficient in certain cases. The second limitation is the need to estimate the energy and variance in energy of the state at each iteration in order to compute the step size used in the circuit at the next iteration. However, when the degree of polynomials scales logarithmically in the inverse of the desired precision, the corresponding sampling overhead should only be polynomial.

DB-QSP can be used both as a standalone method and as a tool in conjunction with other QSP methods [5, 12, 13]. In particular, it can be viewed as a (partial) alternative when the post-selection overhead of other QSP methods are prohibitively large. Namely, DB-QSP provides a deterministic approach to drive a state closer to a target state, such as an approximate ground state, regardless of the quality of the initial state. Conversely, in conventional QSP methods [5, 12, 13], a low post-selection success probability could prevent systematic improvements. Thus DB-QSP can provide a warm-starting procedure, i.e., a means of preparing approximate initial states, for existing methods.

2 Preliminaries

2.1 Overview of Quantum Signal Processing (QSP)

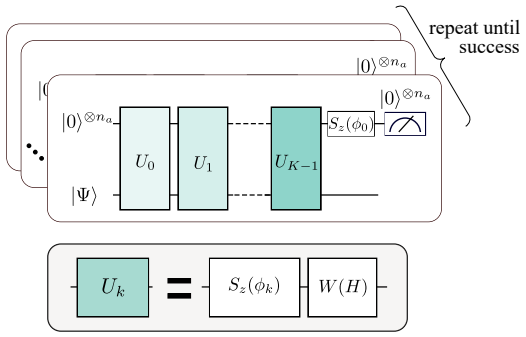
QSP is a framework for systematically constructing matrix-valued functions on quantum computers. The goal of QSP is to perform degree-K polynomial transformation $p(H)$ of a Hermitian matrix H to a n-qubit input state $|\Psi\rangle$ up to normalization (Eq. (1)). Sometimes, the implementation methodology proposed in Ref. [1] itself is referred to as "QSP". However, Eq. (1) can be achieved also via alternative techniques, e.g., Linear Combination of Unitaries (LCU) [12, 13]; see App. A for a detailed overview. In this manuscript, we use "QSP" to refer to the concept of implementing the polynomial functions, and distinguish it from the methodology in Ref. [1, 5] by referring to the latter as "qubitization".

Qubitization uses a circuit U_Q comprised of two types of operators: *signal operators* W and *signal processing operators* $S(\phi)$, where the phase ϕ is drawn from a set $\{\phi_k\}$. The desired polynomial transformation is obtained by performing a measurement in the so-called *signal basis*. Concretely, given the signal operator $W(H)$ of a Hermitian matrix H with $\|H\| \leq 1$ and the signal processing operator $S_z(\phi)$, there exists a sequence of QSP phase $\{\phi_k\}$ such that the following circuit

$$U_Q = S_z(\phi_0) \prod_{k=1}^{K} W(H) S_z(\phi_k) \ , \qquad (3)$$

followed by measurement in the basis $M = \{|+\rangle, |-\rangle\}$ can realize a degree-K real polynomial $p(H)$. The signal operator $W(H)$ can be constructed using block-encoding [19], which embeds a Hermitian matrix H into the top-left block

(a) Original QSP with Post-Selection [1]



(b) Our Proposal without Post-Selection (Thm. 2)

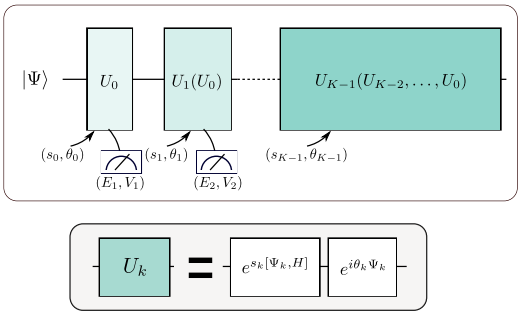


Figure 1: **Quantum Signal Processing (QSP) without auxiliary qubits and post-selection.** We introduce a new formula for implementing QSP of Hermitian matrices (Thm. 2). (a) To realize a degree-K polynomial of a Hermitian matrix H, original QSP performs measurement on auxiliary qubits so that the desired transformation is realized, as shown in Eq. (3). (b) In contrast, our formula does not require auxiliary qubits and accordingly the post-selection. Instead, we recursively apply the state-dependent unitary operators $e^{i\theta_k\Psi_k}e^{s_k[\Psi_k,H]}$ with $|\Psi_{k+1}\rangle = e^{i\theta_k\Psi_k}e^{s_k[\Psi_k,H]}|\Psi_k\rangle$, resulting in the circuit depth that grow significantly in the degree of polynomials K. Furthermore, to determine the time duration s_k and phase θ_k, energy $E_k = \langle\Psi_k|H|\Psi_k\rangle$ and variance in energy $V_k = \langle\Psi_k|H^2|\Psi_k\rangle - E_k^2$ must be known at each step.

of a larger unitary matrix as

$$W(H) = \begin{bmatrix} H & i\sqrt{I-H^2} \\ i\sqrt{I-H^2} & H \end{bmatrix}.$$

The signal processing operator $S_z(\phi)$,

$$S_z(\phi) = e^{i\phi Z} = \begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{bmatrix},$$

then acts on an auxiliary qubit. We provide details of the achievable functions via this technique in App. A.

QSP has led to asymptotically optimal Hamiltonian simulation algorithms [5] and a near-optimal method for ground-state preparation [9]. Furthermore, it serves as a fundamental tool for constructing primitive quantum algorithms that exhibit quantum advantages [2, 3]. Therefore, its efficiency in implementing linear algebraic operations and its role as a key building block for quantum algorithms have made QSP a subject of significant interest.

2.2 The Role of Post-Selection in Existing QSP Methods

Despite their versatility, existing QSP implementations face several challenges such as difficulty in finding angles [20] and demanding implementation costs for block-encodings [21]. As shown above, qubitization performs the measurement in the signal basis to post-select for the desired transformation. When this post-selection in qubitization is unsuccessful, it is possible to simply repeat the experiment until a successful implementation eventually appears. Amplitude amplification techniques [14] can often enhance success probabilities. For instance, Hamiltonian simulation benefits from this combination of techniques [15]. However, in some cases, the success probability for QSP could be exponentially small in the number of qubits [16]. For example, the successful probability of ground-state preparation can be prohibitively small if not initialized with a sufficiently good input state.

We illustrate the issue using an example of general qubitization. Given an input state $|\Psi\rangle$, the number of auxiliary qubits n_a and $\alpha \in \mathbb{R}$, applying U_Q in Eq. (3) to $|\Psi\rangle$ yields

$$|0\rangle^{\otimes n_a} \otimes p(H/\alpha)|\Psi\rangle + |\text{garbage}^{\perp}\rangle, \qquad (4)$$

where $|\text{garbage}^{\perp}\rangle$ is an orthogonal state, i.e., $|\text{garbage}^{\perp}\rangle \perp |0\rangle^{\otimes n_a} \otimes \frac{H}{\alpha}|\Psi\rangle$. The probability of projecting onto $|0\rangle^{\otimes n_a}$ is given by

$$p_{\text{succ}} = \| p(H/\alpha)|\Psi\rangle \|^2, \qquad (5)$$

which can be exponentially small. For example, in the case of Imaginary-Time Evolution (ITE), where $p(H) \approx e^{-\tau H}$, the success probability scales with the overlap of the initial state and the corresponding thermal state, which can

decay exponentially [22, 16, 23]. More generally, this dependence on state fidelity persists across various scenarios. For instance, the block-encoding query complexity for nearly-optimal ground-state preparation algorithm in Ref. [9] scales as $O(\alpha/\gamma)$, where $\gamma = |\langle\lambda_0|\Psi\rangle|^2$ is the fidelity of the input state $|\Psi\rangle$ with the ground state $|\lambda_0\rangle$ of H. The query scaling $O(\alpha/\gamma)$ corresponds to the inverse success probability and thus requires repeated trials for obtaining a successful outcome. This indicates that the success of the block-encoding depends on the input state. Additionally, since the number of queries to the block-encoding unitary scales with the degree of polynomials as shown in Eq. (3), the degree K needs to be sufficiently low to ensure successful post-selection each time. For more details and a discussion of similar challenges when using LCU for implementing QSP, see App. A.

3 Main Results

3.1 Overview of Analysis

In this section, we present an algorithm for QSP that requires neither auxiliary qubits nor post-selection. Our key insight, captured in Lem. 1 in Sec. 3.2, is that there exists a unitary that *exactly* implements the normalized action of the linear polynomial $H - \alpha I$ on an input state $|\Psi\rangle$ for any real α. We then show how repeated applications of this circuit to apply the linear polynomial can be used to implement any polynomial with real roots.

Sec. 3.3 tackles the extension to polynomials with complex roots. This leads to our main result, Thm. 2, which demonstrates that interleaving the unitary sequence U_Ψ from Lem. 1 with state-dependent reflection gates enables the realization of arbitrary polynomials.

Sec. 3.4 introduces a method to implement the unitary sequence in Thm. 2 called Double-Bracket QSP (DB-QSP), which performs general QSP without post-selection. Namely, we show that the recently-developed DBQA framework provides a means to efficiently implement the exponentials of commutators that appear in Thm. 2. Leveraging DBQA, we formulate DB-QSP outlined in Alg. 1. We analyze the errors introduced by this implementation compared to the idealized scenario in Thm. 2 and show that circuit depths of DB-QSP scale super exponentially with the degree of the polynomial to be implemented.

The DB-QSP algorithm (Alg. 1) also requires the energy and energy variance of the state at each iteration to be estimated in order to compute the step size for the next iteration. On quantum hardware, statistical noise is inevitable due to the finite number of measurement shots. In Sec. 3.5, we analyze how this noise affects the accuracy of the constructed state.

To further examine the practical implications of these challenges, Sec. 3.6 investigates the impact of circuit depth on applicability. Since the required depth depends on the polynomial degree, DB-QSP is limited to low-degree polynomials. We identify approximate ground-state preparation as a use case where DB-QSP can be practically useful.

Finally, Sec. 3.7 discusses a hybrid strategy that integrates DB-QSP with existing methods such as variational quantum algorithms, quantum dynamic programming, qubitization and LCU. The circuit depth scaling of DB-QSP suggests that available experimental resources may be insufficient for certain tasks. However, even qubitization with amplitude amplification sometimes demands exponential costs. In such difficult cases, combining qubitization or LCUs with DB-QSP could reduce resource requirements.

3.2 Main Tool: Unitary Synthesis for Polynomials with Real Roots without Post-Selection

In Sec. 2, we reviewed a QSP implementation relying on post-selection. An alternative is to find a unitary U_Ψ satisfying

$$U_\Psi |\Psi\rangle = \frac{p(H)|\Psi\rangle}{\|p(H)|\Psi\rangle\|} . \tag{6}$$

The following Lemma constructs a new tool that provides an explicit and exact construction of U_Ψ through an exponential of a specific commutator for linear polynomials. For simplicity, we hereafter use Ψ as a shorthand for the density matrix representation of a pure state, i.e., $\Psi = |\Psi\rangle\langle\Psi|$.

Lemma 1 (Unitary synthesis for linear polynomials without post-selection)**.** *Suppose $p(H) = H - \alpha I$ is any linear polynomial of a Hermitian matrix H with $\alpha \in \mathbb{R}$. Given an input state $|\Psi\rangle$ with energy mean $E_\Psi = \langle\Psi|H|\Psi\rangle$ and variance*

$V_\Psi = \langle\Psi| H^2 |\Psi\rangle - E_\Psi^2$, *the unitary synthesis for* $p(H)$ *in Eq.* (6) *can be achieved by*

$$U_\Psi = e^{s_\Psi [\Psi, H]}, \tag{7}$$

with

$$s_\Psi = \frac{-1}{\sqrt{V_\Psi}} \arccos\left(\frac{E_\Psi - \alpha}{\sqrt{V_\Psi + (E_\Psi - \alpha)^2}}\right) . \tag{8}$$

A rigorous proof of Lem. 1 is provided in App. B. Here, we present a proof sketch to clarify the derivation of the unitary operator in Eq. (7). First, we can see immediately that U_Ψ is indeed unitary as claimed because the commutator $[\Psi, H]$ in its exponent is anti-Hermitian, i.e., $[\Psi, H] = -([\Psi, H])^\dagger$.

Next, we derive Eq. (7), which establishes the equivalence between e^{sW_H} with $W_H = [\Psi, H]$ and a linear polynomial applied to $|\Psi\rangle$ for $s \in \mathbb{R}$. By definition, the unitary operator can be expressed as $e^{sW_H} = \sum_{k=0}^{\infty} \frac{s^k}{k!} W_H^k$ using all powers of W_H. However, when acting on $|\Psi\rangle$, we get

$$W_H |\Psi\rangle = -(H - E_\Psi I) |\Psi\rangle \ , \tag{9}$$

while for the second power

$$W_H^2 |\Psi\rangle = -(\langle\Psi| H^2 |\Psi\rangle - E_\Psi^2) |\Psi\rangle = -V_\Psi |\Psi\rangle \ . \tag{10}$$

This shows that the square of W_H leaves $|\Psi\rangle$ unchanged up to a rescaling prefactor. Thus, by substituting Eqs. (9), (10) into the series expansion, the resulting state can be simplified to

$$e^{s_\Psi W_H} |\Psi\rangle = (a(s_\Psi) I + b(s_\Psi) H) |\Psi\rangle , \tag{11}$$

with real-valued coefficients $a(s_\Psi), b(s_\Psi)$ corresponding to any duration $s_\Psi \in \mathbb{R}$ given by

$$a(s_\Psi) = \frac{E_\Psi}{\sqrt{V_\Psi}} \sin\left(s_\Psi \sqrt{V_\Psi}\right) + \cos\left(s_\Psi \sqrt{V_\Psi}\right) , \tag{12}$$

$$b(s_\Psi) = -\frac{1}{\sqrt{V_\Psi}} \sin\left(s_\Psi \sqrt{V_\Psi}\right) . \tag{13}$$

Here, the derivation exploits the Taylor series of trigonometric functions. Finally, by solving the equations $a(s_\Psi) = -\alpha / \| p(H) |\Psi\rangle \|$ and $b(s_\Psi) = 1 / \| p(H) |\Psi\rangle \|$, we obtain Eq. (8), the time duration s_Ψ to realize Eq. (6) for any linear polynomial.

Lem. 1 indicates that there exists a duration s such that the exponential of the commutator e^{sW_H} with $W_H = [\Psi, H]$ can realize any linear real polynomial. Importantly, the duration s can be found by precise measurements of the energy and variance of the state Ψ.

Higher order polynomials can then be realised by repeated applications of Lem. 1. The fundamental theorem of algebra shows that a polynomial of degree K with real roots can be represented as $p(H) = a_K \prod_{k=1}^{K} (H - \alpha_k I)$ with $\alpha_k \in \mathbb{R}$. This implies that such polynomials can be obtained by implementing Eq. (7) with the corresponding factors K times,

$$\frac{p(H) |\Psi_0\rangle}{\| p(H) |\Psi_0\rangle \|} = \prod_{k=0}^{K-1} e^{s_k [\Psi_k, H]} |\Psi_0\rangle \ , \tag{14}$$

where we start with an input state $|\Psi_0\rangle$ and define $|\Psi_{k+1}\rangle = e^{s_k [\Psi_k, H]} |\Psi_k\rangle$ using s_k in Eq. (8).

We stress that Eq. (14) only implements functions with real roots. Nonetheless, many functions, such as Chebyshev polynomials, have only real roots. Hence Eq. (14) can be used for applications including approximations of ITE; see App. D for the detail. However, Eq. (14) alone cannot construct *arbitrary* polynomial functions, as the roots can be complex in general. We will now proceed to discuss how to extend Eq. (14) to implement polynomials with complex roots.

3.3 Main Result: Unitary Synthesis for Arbitrary Polynomials without Post-Selection

In this section we show how Eq. (14) can be generalized to implement any *arbitrary* polynomial of the form

$$p(H) = a_K \prod_{k=1}^{K} (H - z_k I) \ , \tag{15}$$

where the roots can be complex, i.e., $z_k \in \mathbb{C}$. A core idea is that introducing a state-dependent reflection gate $e^{i\theta_\Psi \Psi}$ right after U_Ψ in Eq. (7) can realize any complex number z. That is, for any $z \in \mathbb{C}$, we obtain

$$\frac{(H - zI) |\Psi\rangle}{\| (H - zI) |\Psi\rangle \|} = e^{i\theta_\Psi \Psi} e^{s_\Psi [\Psi, H]} |\Psi\rangle . \tag{16}$$

Using this technique, we derive a unitary synthesis formula for QSP without the need for the auxiliary qubits and post-selection, which is the main result of this work. The proof is provided in App. B.

Theorem 2 (Unitary synthesis for QSP without post-selection)**.** *Consider an input state* $|\Psi_0\rangle$ *and any polynomial* $p(H)$ *of degree* K *for a given Hermitian matrix* H *in the form of Eq. (15). Given energy mean* $E_k = \langle\Psi_k| H |\Psi_k\rangle$ *and variance* $V_k = \langle\Psi_k| H^2 |\Psi_k\rangle - E_k^2$*, the unitary synthesis in Eq. (6) can be achieved by*

$$\frac{p(H)|\Psi_0\rangle}{\|p(H)|\Psi_0\rangle\|} = \prod_{k=0}^{K-1} e^{i\theta_k \Psi_k} e^{s_k[\Psi_k, H]} |\Psi_0\rangle , \quad (17)$$

with $s_k = \frac{-1}{\sqrt{V_k}} \arccos\left(\frac{|E_k - z_k|}{\sqrt{V_k + |E_k - z_k|^2}}\right)$ *and* $\theta_k = \arg\left(\frac{E_k - z_k}{|E_k - z_k|}\right)$*. Here, we recursively define the state* $|\Psi_k\rangle$ *by*

$$|\Psi_{k+1}\rangle = e^{i\theta_k \Psi_k} e^{s_k[\Psi_k, H]} |\Psi_k\rangle . \quad (18)$$

Thm. 2 establishes a recursive method for constructing any QSP polynomial through a sequence of unitary operators. Next, we explicitly demonstrate how this formulation can be implemented as a quantum algorithm.

3.4 Implementation: Double-Bracket QSP algorithm (DB-QSP)

Building upon Thm. 2, we present a unitary synthesis approach termed the Double-Bracket QSP algorithm (DB-QSP). A key challenge in implementing Eq. (17) lies in realizing the unitary operator $e^{s_\Psi[\Psi, H]}$. Here, we adopt the approach of DBQAs and utilize the group commutator formula [17, 18, 24, 25] given by [26, 17, 27]:

$$e^{s_\Psi[\Psi,H]} = \left(e^{is_\Psi^{(N)}\Psi} e^{is_\Psi^{(N)}H} e^{-is_\Psi^{(N)}\Psi} e^{-is_\Psi^{(N)}H}\right)^N + O(s_\Psi^{3/2}/\sqrt{N}) , \quad (19)$$

where $s_\Psi^{(N)} = \sqrt{|s_\Psi|/N}$ for $s_\Psi \leq 0$. Note that, since the range of the arccos function is $[0, \pi]$, the time duration s_k in Thm. 2 always takes a non-positive value. Based on this approximation, DB-QSP implements QSP using the Hamiltonian evolution $e^{is_\Psi^{(N)}H}$ and the state-dependent reflection gates $e^{is_\Psi^{(N)}\Psi}$. Specifically, the state-dependent reflection gate is implemented using the reflection about the initial state $|\Psi_0\rangle$ and a unitary operator U satisfying $|\Psi\rangle = U|\Psi_0\rangle$, i.e.,

$$e^{is_\Psi^{(N)}\Psi} = U e^{is_\Psi^{(N)}\Psi_0} U^\dagger. \quad (20)$$

Algorithm 1: DB-QSP

Input: Hermitian operator H, initial state $|\Psi_0\rangle$, degree K, parameters $\{z_k\}_{k=0}^{K-1}$, number of group commutator repetitions N.
Output: State $|\Psi_K\rangle = \frac{p(H)|\Psi_0\rangle}{\|p(H)|\Psi_0\rangle\|}$.
Initialize: $|\Psi\rangle \leftarrow |\Psi_0\rangle$.
for $k = 0$ to $K-1$ **do**
  Compute energy moment E_k and variance V_k for $|\Psi\rangle$.
  Use Thm. 2 to determine parameters s_k and θ_k for

$$e^{i\theta_k\Psi} e^{s_k[\Psi,H]} |\Psi\rangle = \frac{(H - z_k I)|\Psi\rangle}{\|(H - z_k I)|\Psi\rangle\|}.$$

  Set $s_k^{(N)} = \sqrt{|s_k|/N}$ and the group commutator unitary

$$G_k = e^{is_k^{(N)}\Psi} e^{is_k^{(N)}H} e^{-is_k^{(N)}\Psi} e^{-is_k^{(N)}H}.$$

  Update state by applying $|\Psi\rangle \leftarrow e^{i\theta_k\Psi} G_k^N |\Psi\rangle$.
end for
Return: $|\Psi\rangle$.

Alg. 1 summarizes the procedure of DB-QSP algorithm.

We note that DB-QSP assumes that both the state-dependent reflection gate with respect to the initial state and Hamiltonian evolution can be generated efficiently. Nevertheless, this assumption is not particularly restrictive. For the reflection gates, a straightforward approach is to perform the density matrix exponentiation of the initial state [28, 29]. Yet, if the input state $|\Psi_0\rangle$ is a computational basis state, the operation reduces to a multi-qubit controlled unitary, which can be implemented efficiently with cost scaling linearly in the number of qubits [30, 31, 32]. More concretely, when $|\Psi_0\rangle = |0\rangle$, the reflection gate takes the form of

$$e^{i\theta|0\rangle\langle 0|} = I + (e^{i\theta} - 1)|0\rangle\langle 0| = \begin{bmatrix} e^{i\theta} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (21)$$

which corresponds to a multi-qubit controlled parameterized phase gate. Even when $|\Psi_0\rangle$ is not a computational basis state, if a unitary U exists such that $|\Psi_0\rangle$ can be efficiently prepared from

a basis state, e.g., $|0\rangle$, the reflection gate can be realized as $e^{i\theta\Psi_0} = Ue^{i\theta|0\rangle\langle 0|}U^\dagger$. Similarly, when H is a local Hamiltonian, efficient compilation is feasible using established Hamiltonian simulation methods [33, 34, 12, 35, 36]. Thus, in many practical scenarios where such compilation subroutines are available, Eq. (7) serves as a unitary synthesis method of QSP without post-selection.

A key question is how efficiently DB-QSP can implement polynomials with small error. Eq. (19) indicates that the approximation error is governed by $|s_\Psi|^{3/2}/\sqrt{N}$, suggesting that the total number of group commutator repetiotions N may need to increase for higher precision. Thus, elucidating how N (or equivalently, the circuit depth) scales to achieve a fixed precision is crucial for evaluating the practicality of DB-QSP. In the following, we analytically estimate the circuit depth needed to accurately realize a polynomial $p(H)$ of degree K using DB-QSP. Before diving into this, we begin by analyzing the potential cost for implementing one step of DB-QSP with respect to the total discretization steps N.

Implementation cost for a single step of DB-QSP. We begin by analyzing the total number of group commutator repetitions N necessary to approximate $e^{s_\Psi[\Psi,H]}$ to ϵ_0-precision via the group commutator formula. That is, we compute the required N such that

$$\|e^{s_\Psi[\Psi,H]} - (e^{is_\Psi^{(N)}\Psi}e^{is_\Psi^{(N)}H}e^{-is_\Psi^{(N)}\Psi}e^{-is_\Psi^{(N)}H})^N\| \leq \epsilon_0. \tag{22}$$

From Eq. (19), we can immediately see that the relative size of s_Ψ and N determines the error ϵ_0. We further recall that from Thm. 2 we have

$$\begin{aligned} |s_\Psi| &= \frac{1}{\sqrt{V_\Psi}} \arccos\left(\frac{|E_\Psi - z|}{\sqrt{V_\Psi + |E_\Psi - z|^2}}\right) \\ &\leq \frac{1}{|E_\Psi - z|}, \end{aligned} \tag{23}$$

where the inequality is obtained by exploiting the fact that s_Ψ is monotonically decreasing in V_Ψ (as shown explicitly in App. B). Combining Eq. (19) and Eq. (23), we want $1/(|E_\Psi - z|)^{3/2}\sqrt{N} \leq \epsilon_0$, and so we find that there exists an N such that

$$N \in \mathcal{O}\left(\frac{1}{|E_\Psi - z|^3\epsilon_0^2}\right) \tag{24}$$

suffices to ensure Eq. (22) holds.

We thus see that a large gap $|E_\Psi - z|$ reduces the required number of steps N. Conversely, N diverges when $V_k = 0$ and $z_k = E_k$. This can intuitively be understood as arising from the fact that the operation $H - E_k I$ acts as an "annihilation operator". If $V_k = 0$, then the state is an eigenstate and E_k corresponds to its eigenvalue, meaning $(H - E_k I)|\Psi\rangle = 0$ and so the method breaks down. We note that a similar breakdown for eigenstates was observed for a quantum algorithm for ITE using the group commutator unitary in Ref. [18].

Circuit Depth of DB-QSP. We now proceed to analyze the circuit depth to realize a DB-QSP state that is ϵ-close to the ideal state for a degree-K polynomial. We define the circuit depth as the number of Hamiltonian evolution gates and reflection gates to construct quantum circuits for DB-QSP. For the analysis, consider the following state constructed by DB-QSP:

$$\begin{aligned} &|\omega_K\rangle \\ &= \prod_{k=0}^{K-1} e^{i\theta_k\omega_k}\left(e^{is_k^{(N)}\omega_k}e^{is_k^{(N)}H}e^{-is_k^{(N)}\omega_k}e^{-is_k^{(N)}H}\right)^N |\omega_0\rangle \end{aligned} \tag{25}$$

where the intermediate state is recursively constructed as

$$\begin{aligned} &|\omega_{k+1}\rangle \\ &= e^{i\theta_k\omega_k}\left(e^{is_k^{(N)}\omega_k}e^{is_k^{(N)}H}e^{-is_k^{(N)}\omega_k}e^{-is_k^{(N)}H}\right)^N |\omega_k\rangle \end{aligned} \tag{26}$$

with $s_k^{(N)} = \sqrt{|s_k|/N}$. We also define the exact QSP state derived from Thm. 2 as

$$|\Psi(\boldsymbol{\theta}, \boldsymbol{s})\rangle = \prod_{k=0}^{K-1} e^{i\theta_k\Psi_k}e^{s_k[\Psi_k,H]}|\Psi_0\rangle \tag{27}$$

with $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_{K-1})$ and $\boldsymbol{s} = (|s_0|, \ldots, |s_{K-1}|)$. The following Theorem captures the circuit depths required to ensure that the DB-QSP state in Eq. (25), agrees with the true circuit up to ϵ precision. The key assumption here is that the parameters (θ_k, s_k) are known exactly. In practice, the parameters will be computed with a finite number of measurement shots, requiring an additional sampling overhead and introducing additional errors. We will address this aspect in Section 3.5.

Theorem 3 (DB-QSP circuit depth)**.** *Suppose H is a Hermitian matrix whose spectral radius*

does not exceed unity, i.e., $\|H\| \leq 1$. *Let* $\zeta = \max(\boldsymbol{\theta}, \boldsymbol{s})$ *be the maximum value of all elements in* $\boldsymbol{\theta}$ *and* $\boldsymbol{s}$. *Also, consider* $|\omega_K\rangle$ *given by DB-QSP from Alg. 1 in Eq. (25) and the state* $|\Psi(\boldsymbol{\theta}, \boldsymbol{s})\rangle$ *from Thm. 2 in Eq. (27) for degree-*K *polynomials. Then there exists a circuit depth* $\mathcal{N}_K$ *such that*

$$\mathcal{N}_K \in \mathcal{O}\left(\left(\frac{8}{3}\right)^2 \zeta(1+6\zeta)^{2K}/\epsilon^2 + 3\right)^K \tag{28}$$

suffices to ensure that $\| \, |\Psi(\boldsymbol{\theta}, \boldsymbol{s})\rangle - |\omega_K\rangle \, \| \leq \epsilon$.

To prove this, we first utilize a result proven in App. C that the DB-QSP error can be bounded as

$$\| \, |\Psi(\boldsymbol{\theta}, \boldsymbol{s})\rangle - |\omega_K\rangle \, \| \leq \frac{4}{3}\zeta^{1/2}(1+6\zeta)^K/\sqrt{N} \, . \tag{29}$$

Next, we compute the circuit depth, which is defined as the total number of Hamiltonian evolution gates and the reflection gates. Given that the state is $|\omega_k\rangle = U_k |0\rangle$, we can write the recursive unitary synthesis formula as

$$U_{k+1} = U_k e^{i\theta_k |0\rangle\langle 0|} U_k^\dagger \times G^N \times U_k \tag{30}$$

with

$$G = U_k e^{is_k^{(N)}|0\rangle\langle 0|} U_k^\dagger e^{is_k^{(N)} H} \times U_k e^{-is_k^{(N)}|0\rangle\langle 0|} U_k^\dagger e^{-is_k^{(N)} H}.$$

This implies that each step involves $4N+3$ repetitions of the unitary operators U_k at the previous step. Therefore, since an additional $4N+1$ gates ($2N$ gates for Hamiltonian evolution and $2N+1$ for the reflection gates on the initial state $|0\rangle\langle 0|$) are required, the circuit depth $\mathcal{N}_{k+1}$ at step $k+1$ is given by $\mathcal{N}_{k+1} = (4N+3)\mathcal{N}_k + 4N + 1$. Thus, the total circuit depth required for a polynomial of degree K can be represented as

$$\mathcal{N}_K = \frac{(4N+1)((4N+3)^K - 1)}{4N+2} \leq (4N+3)^K \tag{31}$$

Thus, by substituting Eq. (31) into the right-hand side of Eq. (29), Eq. (28) satisfies to ensure the ϵ-precision as claimed in the theorem.

Thm. 3 indicates that the circuit depth scaling can be prohibitive for high degree polynomials. Namely, although the depth scales polynomially in the precision $1/\epsilon$, it grows super-exponentially with the degree of the polynomials K. Consequently, DB-QSP is not practically applicable to polynomials of arbitrary degrees, but should target low-degree polynomials.

3.5 Performance Analysis of Perturbations in Parameters

In this section, we analyze the effect of statistical noise. As shown in Alg. 1, DB-QSP requires the estimation of energy and variance to determine the parameters s_k and θ_k at each time step. However, due to the finite number of measurement shots in practice, precise estimation is not feasible on quantum hardware. Consequently, parameters deviate from their true values at each time step, with perturbations satisfying $|s_k - \tilde{s}_k| \leq \delta_s$ and $|\theta_k - \tilde{\theta}_k| \leq \delta_\theta$. In other words, even if the quantum hardware performs the operations perfectly, statistical errors from the measurements lead to erroneous parameters.

Under this setting, we provide an error bound for implementing a polynomial of degree K. We introduce a noisy state to handle the erroneous parameters:

$$|\tilde{\Psi}_H(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{s}})\rangle = \prod_{k=1}^{K} e^{i\tilde{\theta}_k \tilde{\Psi}_k} e^{\tilde{s}_k [\tilde{\Psi}_k, H]} \, |\Psi_0\rangle \, , \tag{32}$$

where we define $|\tilde{\Psi}_{k+1}\rangle = e^{i\tilde{\theta}_k \tilde{\Psi}_k} e^{\tilde{s}_k [\tilde{\Psi}_k, H]} \, |\tilde{\Psi}_k\rangle$, with $|\tilde{\Psi}_0\rangle = |\Psi_0\rangle$. Here, we also introduce $\tilde{\boldsymbol{\theta}} = (\tilde{\theta}_0, \ldots, \tilde{\theta}_{K-1})$ and $\tilde{\boldsymbol{s}} = (|\tilde{s}_0|, \ldots, |\tilde{s}_{K-1}|)$. Again, we assume $\|H\| \leq 1$. We can then derive the following result on the circuit error with the detailed proof provided in App. C.

Proposition 4 (Stability of Thm. 2 under erroneous estimation)**.** *Let* H *be a Hermitian matrix such that* $\|H\| \leq 1$, *and assume that the estimated parameters* $\tilde{s}_k$ *and* $\tilde{\theta}_k$ *satisfy* $|s_k - \tilde{s}_k| \leq \delta_s$ *and* $|\theta_k - \tilde{\theta}_k| \leq \delta_\theta$ *with ideal parameters* s_k *and* θ_k *for all* k. *By setting* $\zeta = \max(\boldsymbol{\theta}, \boldsymbol{s})$, *the perturbed state* $|\tilde{\Psi}_H(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{s}})\rangle$ *in Eq. (32) and the state* $|\Psi_H(\boldsymbol{\theta}, \boldsymbol{s})\rangle$ *from Thm. 2 satisfies*

$$\| \, |\Psi_H(\boldsymbol{\theta}, \boldsymbol{s})\rangle - |\tilde{\Psi}_H(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{s}})\rangle \, \| \leq \frac{1}{3\zeta}(1+6\zeta)^K \max(\delta_s, \delta_\theta) \, . \tag{33}$$

Prop. 4 indicates that the parameter deviation $\max(\delta_s, \delta_\theta)$ scales linearly with the accumulated error and hence its suppression is critical. On the other hand, since these parameters are nonlinear functions of energy and variance, analyzing the impact of statistical estimates from Prop. 4 is non-trivial. To address this, we extend our analysis to explicitly account for statistical noise in energy and variance estimation.

The recursive structure of Eq. (27) and Eq. (32) implies that, even in the limit of infinite measurement shots, the estimated energy and variance may still differ from their ideal values. This discrepancy arises because these quantities are measured on a potentially different state at each iteration. Specifically, the statistical estimate $\overline{E}_k$ ($\overline{V}_k$) converges to $\tilde{E}_k$ ($\tilde{V}_k$) obtained from the noisy state $|\tilde{\Psi}_k\rangle$, rather than the ideal values E_k and V_k. To account for this, we extend Prop. 4 and demonstrate that the statistical noise, $\delta_E = |\overline{E}_k - \tilde{E}_k|$ and $\delta_V = |\overline{V}_k - \tilde{V}_k|$, exhibits a linear dependence on the accumulated error in Eq. (33), but keeps the exponential scaling with K. That is, $\| \, |\Psi_H(\boldsymbol{\theta}, \boldsymbol{s})\rangle - |\tilde{\Psi}_H(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{s}})\rangle \, \| \leq C^K \max(\delta_E, \delta_V)$ for a constant $C \geq 1$. See App. C for further technical details.

Using the results, we also estimate the number of measurement shots needed to achieve the state ϵ-close to the ideal state at K step. Without loss of generality, we express the Hermitian matrix as a weighted sum of Pauli terms, i.e., $H = \sum_{i=1}^{J} w_i P_i$. Then, the number of samples N_E (N_V) required to estimate the energy (variance) within an error $\tilde{\epsilon}$ with probability at least $1-\delta$, $\delta \in (0,1]$, scales as $N_E \in \mathcal{O}(J\|w\|_1^2/\tilde{\epsilon}\delta)$ ($N_V \in \mathcal{O}(J^2\|w\|_1^4/\tilde{\epsilon}\delta)$), where $\|w\|_1 = \sum_{i=1}^{J} |w_i|$. See App. F for further details. Thus, the number of measurement shots needed to achieve the ϵ-precision grows exponentially with the polynomial degree K, since the estimation error must be sufficiently small to cancel a term that scales exponentially with K. However, when K scales logarithmically with $1/\epsilon$, the required measurement shots reduce to polynomial resources in $1/\epsilon$.

Lastly, in App. E and App. F, we explore two different aspects to alleviate statistical errors: (1) the potential of classical computations to reduce the impact of imprecise estimation, and (2) an analysis of the estimator for variance operators.

3.6 Application Examples: Ground-State Approximation and Matrix Inversions

A significant limitation of DB-QSP is its scaling with polynomial order K in Thm. 3. Yet, Thm. 3 implies that when the degree of polynomials K scales logarithmically in the inverse of the precision, $1/\epsilon$, then the circuit depths follow a quasi-polynomial scaling, i.e., $\mathcal{N}_K = 2^{\text{poly}\log(1/\epsilon)}$. Thus, DB-QSP is potentially applicable to polynomial functions of degree at most $K = \mathcal{O}(\log(1/\epsilon))$. Indeed, low-degree approximations, such as those using Chebyshev polynomials, allow efficient representations of certain functions using logarithmically-small degrees. For specific examples of such approximations, see App. D. In the following, we present two representative tasks that illustrate the utility and limitations of low-degree approximations: ground-state preparation and matrix inversion. We discuss DB-QSP's applicability to other tasks in App. D.

Ground-State Approximation. Given a Hermitian matrix H, the objective here is to prepare its ground state $|\lambda_0\rangle$. To achieve this, some quantum algorithms employ a variety of filtering techniques that apply suitable functions of the Hamiltonian. In what follows, we focus on two representative examples: Imaginary-Time Evolution (ITE), which employs the exponential filter, and the approach proposed in Ref. [9], which makes use of a sign-function filter.

As a first example of a method that can be implemented using DB-QSP, we consider Imaginary-Time Evolution (ITE), where the non-unitary operator $p(H) = e^{-\tau H}$ is applied to an initial state $|\Psi_0\rangle$:

$$|\Psi_\tau\rangle = \frac{e^{-\tau H}\,|\Psi_0\rangle}{\|e^{-\tau H}\,|\Psi_0\rangle\,\|} \,. \tag{34}$$

A key feature of ITE is that it guarantees convergence as long as the initial state has a nonzero overlap with the ground state. Refs. [18, 37] established that Eq. (7) serves as a first-order approximation of ITE and further extends this result by employing group commutator iterations in Eq. (19) to construct a unitary realization of ITE: See App. D for more details. In addition to the methods proposed in Refs. [18, 37], DB-QSP can be used to directly construct a polynomial approximation of the exponential function. Specifically, such an approximation requires a polynomial of degree $K = \mathcal{O}(\sqrt{2\tilde{\tau}\log(4/\epsilon)})$ with $\tilde{\tau} = \lceil \max(e^2\tau, \log(2/\epsilon)) \rceil$ [38]. This implies that a polynomial approximation of ITE can be implemented with a complexity that scales quasi-polynomially in the error precision ϵ. However, the scaling with respect to the evolution time τ remains unfavorable.

Beyond ITE, DB-QSP can also construct alternative filtering functions. Ref. [9] presents a

nearly optimal algorithm for ground-state preparation using QSVT. The core idea is to use a low-degree approximation of the sign function, whose degree scales logarithmically with $1/\epsilon$, i.e., $K = \mathcal{O}(\log(1/\epsilon)/\delta)$, for an input $x \in [-2,2]\backslash(-\delta,\delta)$ with $\delta > 0$. Hence, with the same approximation technique, similar filtering strategies could potentially be realized via DB-QSP with favorable scaling in the error precision ϵ.

We recall that the success probability of existing QSP implementations depends on the overlap between the initial state and ground state, meaning that these methods may fail entirely if the initial state is not well-prepared [9]. In contrast, our approach is applicable to any state, as long as there is a non-zero overlap. Therefore, even if DB-QSP cannot fully implement the desired polynomial functions due to resource constraints, it can still systematically improve the quality of the state.

Matrix Inversion. The goal of "matrix inversion" is to apply A^{-1} to an input state, where A is a square matrix. This is a core subroutine for solving linear systems $A|x\rangle = |b\rangle$ for $|x\rangle$ [4, 2, 3]. As shown in App. D, a polynomial of degree $K = \mathcal{O}(\kappa \log(\kappa/\epsilon))$ can approximate the inverse function $1/x$ with the precision ϵ for an input $x \in [-1,1]\backslash(-\frac{1}{\kappa},\frac{1}{\kappa})$ where $\kappa \geq 1$ is the condition number of the matrix. While our results so far apply to Hermitian matrices, we can construct a Hermitian matrix from any square matrix A by the extension:

$$H = \begin{bmatrix} 0 & A \\ A^\dagger & 0 \end{bmatrix}, \tag{35}$$

This indicates that DB-QSP has the potential to efficiently perform matrix inversion in terms of the inverse precision $1/\epsilon$. However, the circuit depth required for DB-QSP scales super exponentially with the condition number, a key factor in assessing the algorithm's efficiency. Thus, this example also highlights a fundamental challenge for DB-QSP in certain computational tasks.

3.7 Hybrid Strategy: DB-QSP with Existing Methods

The performance analysis has highlighted that, while DB-QSP holds promise for certain tasks, its circuit depth and the requirement for precise estimation of the energy and variance pose significant challenges. However, DB-QSP does not have to be used as a standalone approach. By integrating it with existing methods, these limitation can be alleviated and a hybrid approach may further enhance its feasibility.

In the following, we explore how combining DB-QSP with established techniques can improve performance. Specifically, we examine three approaches: (1) Variational Quantum Algorithms (VQA) [39] and classical computation, (2) Quantum Dynamic Programming (QDP)[40], (3) qubitization and LCU.

VQA & Classical pre-computations. A potential strategy to circumvent the challenges in DB-QSP is to employ a preconditioner that bypasses the initial steps. In this regard, classical computational methods can serve as effective preconditioners. Our target operation in Eq. (11) consists of a weighted sum of I and H with appropriate coefficients. Consequently, the feasibility of classical computation relies on the efficiency of evaluating $\langle\Psi_0|H^{2k+2}|\Psi_0\rangle$ for degree-k polynomials. We show that, if the initial state $|\Psi_0\rangle$ is sparse and a Hermitian matrix contains a limited number of Pauli terms, then classical computation is feasible. Moreover, advanced classical techniques, e.g., tensor networks, could further improve the efficiency, see Sec. E for details.

Another approach might be to first use a Variational Quantum Algorithm (VQA) where a parameterized quantum circuit U_θ is trained to approximate the target state. By leveraging a VQA, a relatively shallow-depth circuit might be found to replicate the operations of a few DB-QSP steps, allowing the trained circuit to serve as a warm start for DB-QSP; i.e. $|\Psi_k\rangle \approx U_\theta|\Psi_0\rangle$ for a small k. However, this strategy has several challenges. First, there are no theoretical guarantees of convergence for VQAs in practical regimes. Secondly, as highlighted in Thm. 3 and Prop. 4, small errors at each step can accumulate significantly as the polynomial degree increases. Consequently, errors introduced by the VQA may degrade the final result. Another issue is the barren plateau phenomenon [41, 42], where gradient magnitudes vanish exponentially with system size, making parameter training impractical. Indeed, it has been suggested that VQAs themselves may need warm starting strate-

gies [43, 44], or else they are classically simulable [45, 46, 47, 48]. In such cases, the direct use of DB-QSP may be a better option.

Quantum Dynamic Programming (QDP). We recall that the significant increase in circuit depth arises from a recursive circuit structure. Specifically, the implementation of the state-dependent reflection through $e^{is_k\Psi_k} = U_k e^{is_k|0\rangle\langle 0|}U_k^\dagger$ leads to a prohibitive number of queries to the quantum gates. Therefore, incorporating a subroutine that reduces the implementation cost would enhance the efficiency of DB-QSP. The operation $e^{is\Psi_k}$ is a special case of Density-Matrix Exponentiation (DME), for which some quantum algorithms have been proposed [28, 49, 29]. DME leverages coherent swap operations between multiple copies of $|\Psi\rangle$ to realize exponentiation. Note that, since the swap operations are independent of previous runtime, the circuit depth scales only polynomially.

Recently, Quantum Dynamic Programming (QDP) has been proposed to study the use of routines such as DME for speeding up quantum recursions. QDP is powerful in that utilizing memory leads to an exponential reduction in circuit depth [40]. This characteristic makes it a viable subroutine for DB-QSP. However, due to the no-cloning theorem, QDP has the disadvantage that one must extend the width when implementing recursion steps, meaning that multiple copies of the state must be prepared [40]. Hence, when combining DB-QSP with QDP, it becomes crucial to balance the trade-off between width and depth for practical feasibility.

Qubitization & LCU. One may envision integrating DB-QSP with QSP implementations that involve post-selection (e.g. qubitization and LCU). While these methods supplemented with amplitude amplification are highly sophisticated and can function as standalone methods, their practicality can be hindered in certain scenarios. As discussed in Sec. 2, an exponentially small success probability can limit the practicality of these methods for some tasks such as ground-state preparation. Thus, by leveraging DB-QSP as a preconditioner, we can potentially mitigate this issue and enhance the overall feasibility of these advanced algorithms.

4 Discussion

Quantum signal processing (QSP) is a fundamental framework for designing quantum algorithms. Existing implementation methods, such as qubitization and linear combinations of unitaries (LCU), are powerful but rely on post-selection of auxiliary qubits, which could limit their celebrated efficiency in certain cases. In this work, we propose a unitary synthesis formula for QSP without auxiliary qubits and post-selection. Our method, termed DB-QSP, relies on Thm. 2 together with the recently established Double-Bracket Quantum Algorithm (DBQA) framework [17, 18]. While our approach comes at the cost of circuit depth and requires precise estimation of energy and variance, it can be used to efficiently implement low-degree polynomial approximations. We further note that our method requires fewer multi-qubit controlled gates than qubitization, as no ancillary qubits are used. Namely, DB-QSP may be more advantageous under hardware constraints, which are likely to occur in near-term or upcoming quantum devices. Thus our proposal broadens the range of options for implementing the QSP framework on quantum hardware, with hybrid approaches that combines both DB-QSP and prior methods looking particularly appealing.

Further investigations of the fundamental limits of our algorithm's performance could be interesting. Thm. 2 clarifies that the time duration s and the angle θ are determined by the energy mean and fluctuation (i.e., variance). This suggests that thermodynamic quantities alone provide sufficient information to guide the implementation of the target transformation. Moreover, these quantities are key to the algorithm's efficiency, as the complexity is determined by the time duration and angles defined using them. Given that this unitary process originates from an unphysical polynomial function, the link between unphysical operations and underlying physical principles will provide an intriguing perspective.

Let us highlight a geometrical view of our algorithm. As shown in Eq. (3), the core of qubitization is that an auxiliary two-level system enables the construction of specific polynomials through a sequence of unitary operators interleaved with phase gates. Interestingly, Thm. 2

reveals that QSP implementations without auxiliary qubits exhibit a similar structure (i.e., Eq. (17)). This structural similarity suggests that both approaches can be analyzed from a geometrical perspective. More specifically, through the lens of DBQA [17, 18], it is known that the exponential of commutators, $e^{[\Psi,H]}$, approximates the steepest descent direction on the Riemannian manifold of quantum states with respect to the cost function $-\|H-\Psi\|_2^2/2$ [50, 51, 52, 53, 54, 55, 56, 57, 58]. While our formulation introduces additional state-dependent reflection gates, polynomials with real roots, such as Chebyshev polynomials, can be constructed without these additional gates. This observation suggests a promising direction for exploring a geometric interpretation of QSP, and possibly of qubitization itself, within our framework.

Finally, a natural direction for future work is to compare the runtime of the proposed approach with that of alternative methods. In particular, recent studies have introduced Lindbladian approaches requiring only a single ancilla qubit [59]. Such dissipative schemes potentially offer an advantages over DBQA and standard QSP, as the convergence to the ground state is possible even when the initial state has zero overlap with the ground state [59, 60, 61]. Interestingly, it is known that Lindbladian dynamics can be formulated as a gradient flow with respect to the quantum relative entropy [62]. This suggests that, despite being seemingly distinct approaches, DBQAs, QSP and Lindbladian simulations may have a shared foundation in gradient flow dynamics defined for appropriate cost functions and underlying Riemannian manifolds. A detailed analysis of the role of Riemannian gradients in the procedures of Refs. [59, 60, 61] is left for future investigation.

Acknowledgments. Insightful discussions with Thais Silva and Andrew Wright are acknowledged. MG, JS, BHT and NN are supported by the start-up grant of the Nanyang Assistant Professorship at the Nanyang Technological University in Singapore. ZH was supported by the Sandoz Family Foundation-Monique de Meuron program for Academic Promotion. MG was also supported by the Presidential Postdoctoral Fellowship of the Nanyang Technological University in Singapore.

References

[1] Guang Hao Low, Theodore J. Yoder, and Isaac L. Chuang. "Methodology of resonant equiangular composite quantum gates". Phys. Rev. X **6**, 041067 (2016).

[2] John M. Martyn, Zane M. Rossi, Andrew K. Tan, and Isaac L. Chuang. "Grand unification of quantum algorithms". PRX Quantum **2**, 040203 (2021).

[3] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. "Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics". In Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing. Pages 193–204. (2019).

[4] Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. "Quantum algorithm for linear systems of equations". Physical Review Letters **103**, 150502 (2009).

[5] Guang Hao Low and Isaac L. Chuang. "Hamiltonian Simulation by Qubitization". Quantum **3**, 163 (2019).

[6] Guang Hao Low and Isaac L Chuang. "Optimal Hamiltonian simulation by quantum signal processing". Physical Review Letters **118**, 010501 (2017).

[7] Guang Hao Low. "Quantum signal processing by single-qubit dynamics". PhD thesis. Massachusetts Institute of Technology. (2017). url: `http://hdl.handle.net/1721.1/115025`.

[8] Yimin Ge, Jordi Tura, and J Ignacio Cirac. "Faster ground state preparation and high-precision ground energy estimation with fewer qubits". Journal of Mathematical Physics**60** (2019).

[9] Lin Lin and Yu Tong. "Near-optimal ground state preparation". Quantum **4**, 372 (2020).

[10] Lei Zhang, Jizhe Lai, Xian Wu, and Xin Wang. "Quantum imaginary-time evolution with polynomial resources in time" (2025). arXiv:2507.00908.

[11] Long Gui-Lu and Liu Yang. "Duality computing in quantum computers". Communications in Theoretical Physics **50**, 1303 (2008).

[12] Andrew M. Childs and Nathan Wiebe. "Hamiltonian simulation using linear combinations of unitary operations". Quantum Inf. Comput. **12**, 901–924 (2012).

[13] Shantanav Chakraborty. "Implementing any linear combination of unitaries on intermediate-term quantum computers". Quantum **8**, 1496 (2024).

[14] Gilles Brassard, Peter Hø yer, Michele Mosca, and Alain Tapp. "Quantum amplitude amplification and estimation". In Quantum computation and information (Washington, DC, 2000). Volume 305 of Contemp. Math., pages 53–74. Amer. Math. Soc., Providence, RI (2002).

[15] Dominic W Berry, Andrew M Childs, Richard Cleve, Robin Kothari, and Rolando D Somma. "Exponential improvement in precision for simulating sparse Hamiltonians". In Proceedings of the forty-sixth annual ACM symposium on Theory of computing. Pages 283–292. (2014).

[16] Michelle Wynne Sze, Yao Tang, Silas Dilkes, David Muñoz Ramo, Ross Duncan, and Nathan Fitzpatrick. "Hamiltonian dynamics simulation using linear combination of unitaries on an ion trap quantum computer" (2025). arXiv:2501.18515.

[17] Marek Gluza. "Double-bracket quantum algorithms for diagonalization". Quantum **8**, 1316 (2024).

[18] Marek Gluza, Jeongrak Son, Bi Hong Tiang, Yudai Suzuki, Zoë Holmes, and Nelly H. Y. Ng. "Double-bracket quantum algorithms for quantum imaginary-time evolution" (2024). arXiv:2412.04554.

[19] Shantanav Chakraborty, András Gilyén, and Stacey Jeffery. "The Power of Block-Encoded Matrix Powers: Improved Regression Techniques via Faster Hamiltonian Simulation". In Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, and Stefano Leonardi, editors, 46th International Colloquium on Automata, Languages, and Programming (ICALP 2019). Volume 132 of Leibniz International Proceedings in Informatics (LIPIcs), pages 33:1–33:14. (2019).

[20] Abhijeet Alase. "Quantum signal processing without angle finding" (2025). arXiv:2501.07002.

[21] Yuta Kikuchi, Conor Mc Keever, Luuk Coopmans, Michael Lubasch, and Marcello Benedetti. "Realization of quantum signal processing on a noisy quantum computer". npj Quantum Information **9**, 93 (2023).

[22] Hans Hon Sang Chan, David Muñoz Ramo, and Nathan Fitzpatrick. "Simulating non-unitary dynamics using quantum signal processing with unitary block encoding" (2023). arXiv:2303.06161.

[23] Thais L. Silva, Márcio M. Taddei, Stefano Carrazza, and Leandro Aolita. "Fragmented imaginary-time evolution for early-stage quantum signal processors". Scientific Reports **13**, 18258 (2023).

[24] Matteo Robbiati, Edoardo Pedicillo, Andrea Pasquale, Xiaoyue Li, Andrew Wright, Renato Farias, Khanh Uyen Giang, Jeongrak Son, Johannes Knörzer, Siong Thye Goh, et al. "Double-bracket quantum algorithms for high-fidelity ground state preparation" (2024). arXiv:2408.03987.

[25] Li Xiaoyue, Matteo Robbiati, Andrea Pasquale, Edoardo Pedicillo, Andrew Wright, Stefano Carrazza, and Marek Gluza. "Strategies for optimizing double-bracket quantum algorithms" (2024). arXiv:2408.07431.

[26] Christopher M Dawson and Michael A Nielsen. "The Solovay-Kitaev algorithm". Quantum Information & Computation **6**, 81–95 (2006).

[27] Yu-An Chen et al. "Efficient product formulas for commutators and applications to quantum simulation". Phys. Rev. Res. **4**, 013191 (2022).

[28] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. "Quantum principal component analysis". Nature Physics **10**, 631–633 (2014).

[29] M Kjaergaard, ME Schwartz, A Greene, GO Samach, A Bengtsson, M O'Keeffe, CM McNally, J Braumüller, DK Kim, P Krantz, et al. "Demonstration of density matrix exponentiation using a supercon-

ducting quantum processor". Physical Review X **12**, 011005 (2022).

[30] Adriano Barenco, Charles H. Bennett, Richard Cleve, David P. DiVincenzo, Norman Margolus, Peter Shor, Tycho Sleator, John A. Smolin, and Harald Weinfurter. "Elementary gates for quantum computation". Physical Review A **52**, 3457 (1995).

[31] Ben Zindorf and Sougato Bose. "Efficient Implementation of Multi-Controlled Quantum Gates" (2024). arXiv:2404.02279.

[32] Ben Zindorf and Sougato Bose. "Multi-controlled quantum gates in linear nearest neighbor" (2025). arXiv:2506.00695.

[33] Andrew M Childs and Robin Kothari. "Limitations on the simulation of non-sparse Hamiltonians". Quantum Information & Computation **10**, 669–684 (2010).

[34] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma. "Simulating Hamiltonian Dynamics with a Truncated Taylor Series". Phys. Rev. Lett. **114**, 090502 (2015).

[35] Andrew M. Childs, Yuan Su, Minh C. Tran, Nathan Wiebe, and Shuchen Zhu. "Theory of Trotter Error with Commutator Scaling". Phys. Rev. X **11**, 011020 (2021).

[36] Andrew M. Childs and Yuan Su. "Nearly optimal lattice simulation by product formulas". Phys. Rev. Lett. **123**, 050503 (2019).

[37] Nathan A McMahon, Mahum Pervez, and Christian Arenz. "Equating quantum imaginary time evolution, riemannian gradient flows, and stochastic implementations" (2025). arXiv:2504.06123.

[38] Sushant Sachdeva, Nisheeth K Vishnoi, et al. "Faster algorithms via approximation theory". Foundations and Trends® in Theoretical Computer Science **9**, 125–210 (2014).

[39] Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, et al. "Variational quantum algorithms". Nature Reviews Physics **3**, 625–644 (2021).

[40] Jeongrak Son, Marek Gluza, Ryuji Takagi, and Nelly H. Y. Ng. "Quantum dynamic programming". Phys. Rev. Lett. **134**, 180602 (2025).

[41] Jarrod R McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush, and Hartmut Neven. "Barren plateaus in quantum neural network training landscapes". Nature Communications **9**, 1–6 (2018).

[42] M. Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J Coles. "Cost function dependent barren plateaus in shallow parametrized quantum circuits". Nature Communications **12**, 1–12 (2021).

[43] Ricard Puig, Marc Drudis, Supanut Thanasilp, and Zoë Holmes. "Variational quantum simulation: a case study for understanding warm starts". PRX Quantum **6**, 010317 (2025).

[44] Hela Mhiri, Ricard Puig, Sacha Lerch, Manuel S Rudolph, Thiparat Chotibut, Supanut Thanasilp, and Zoë Holmes. "A unifying account of warm start guarantees for patches of quantum landscapes" (2025). arXiv:2502.07889.

[45] Marco Cerezo, Martin Larocca, Diego García-Martín, Nelson L Diaz, Paolo Braccia, Enrico Fontana, Manuel S Rudolph, Pablo Bermejo, Aroosa Ijaz, Supanut Thanasilp, et al. "Does provable absence of barren plateaus imply classical simulability?". Nature Communications **16**, 7907 (2025).

[46] Armando Angrisani, Alexander Schmidhuber, Manuel S Rudolph, M Cerezo, Zoë Holmes, and Hsin-Yuan Huang. "Classically estimating observables of noiseless quantum circuits" (2024). arXiv:2409.01706.

[47] Pablo Bermejo, Paolo Braccia, Manuel S Rudolph, Zoë Holmes, Lukasz Cincio, and M Cerezo. "Quantum convolutional neural networks are (effectively) classically simulable" (2024). arXiv:2408.12739.

[48] Sacha Lerch, Ricard Puig, Manuel S Rudolph, Armando Angrisani, Tyson Jones, M Cerezo, Supanut Thanasilp, and Zoë Holmes. "Efficient quantum-enhanced classical simulation for patches of quantum landscapes" (2024). arXiv:2411.19896.

[49] Shelby Kimmel, Cedric Yen-Yu Lin, Guang Hao Low, Maris Ozols, and Theodore J. Yoder. "Hamiltonian simulation with optimal sample complexity". npj Quantum Inf. **3**, 13 (2017).

[50] Uwe Helmke and John B. Moore. "Optimization and dynamical systems". Springer London. (1994).

[51] JB Moore, RE Mahony, and U Helmke. "Numerical gradient algorithms for eigenvalue and singular value calculations". SIAM Journal on Matrix Analysis and Applications **15**, 881–902 (1994).

[52] Anthony M Bloch. "Steepest descent, linear programming and Hamiltonian flows". Contemp. Math. AMS **114**, 77–88 (1990).

[53] Steven Thomas Smith. "Geometric optimization methods for adaptive filtering". Harvard University. (1993).

[54] G Dirr and U Helmke. "Lie theory for quantum control". GAMM-Mitteilungen **31**, 59–93 (2008).

[55] Indra Kurniawan, Gunther Dirr, and Uwe Helmke. "Controllability aspects of quantum dynamics: a unified approach for closed and open systems". IEEE transactions on automatic control **57**, 1984–1996 (2012).

[56] T Schulte-Herbrüggen, A Spörl, N Khaneja, and SJ Glaser. "Optimal control for generating quantum gates in open dissipative systems". Journal of Physics B: Atomic, Molecular and Optical Physics **44**, 154013 (2011).

[57] Thomas Schulte-Herbrüggen, Steffen J. Glaser, Gunther Dirr, and Uwe Helmke. "Gradient flows for optimization in quantum information and quantum dynamics: Foundations and applications". Reviews in Mathematical Physics **22**, 597–667 (2010).

[58] Roeland Wiersema and Nathan Killoran. "Optimizing quantum circuits with Riemannian gradient flow". Phys. Rev. A **107**, 062421 (2023).

[59] Zhiyan Ding, Chi-Fang Chen, and Lin Lin. "Single-ancilla ground state preparation via Lindbladians". Physical Review Research **6**, 033147 (2024).

[60] Chi-Fang Chen, Michael Kastoryano, Fernando GSL Brandão, and András Gilyén. "Efficient quantum thermal simulation". Nature **646**, 561–566 (2025).

[61] Chi-Fang Chen, Hsin-Yuan Huang, John Preskill, and Leo Zhou. "Local minima in quantum systems". In Proceedings of the 56th Annual ACM Symposium on Theory of Computing. Pages 1323–1330. (2024).

[62] Markus Mittnenzweig and Alexander Mielke. "An entropic gradient structure for Lindblad equations and couplings of quantum systems to macroscopic models". Journal of Statistical Physics **167**, 205–233 (2017).

[63] Danial Motlagh and Nathan Wiebe. "Generalized quantum signal processing". PRX Quantum **5**, 020368 (2024).

[64] David Poulin and Pawel Wocjan. "Preparing ground states of quantum many-body systems on a quantum computer". Physical Review Letters **102**, 130503 (2009).

[65] Oumarou Oumarou, Pauline J Ollitrault, Cristian L Cortes, Maximilian Scheurer, Robert M Parrish, and Christian Gogolin. "Molecular Properties from Quantum Krylov Subspace Diagonalization" (2025). arXiv:2501.05286.

[66] Long Gui-Lu. "General quantum interference principle and duality computer". Communications in Theoretical Physics **45**, 825 (2006).

[67] Andrew M Childs, Robin Kothari, and Rolando D Somma. "Quantum algorithm for systems of linear equations with exponentially improved dependence on precision". SIAM Journal on Computing **46**, 1920–1950 (2017).

[68] Yulong Dong, Lin Lin, and Yu Tong. "Ground-state preparation and energy estimation on early fault-tolerant quantum computers via quantum eigenvalue transformation of unitary matrices". PRX Quantum **3**, 040305 (2022).

[69] Ruizhe Zhang, Guoming Wang, and Peter Johnson. "Computing ground state properties with early fault-tolerant quantum computers". Quantum **6**, 761 (2022).

[70] Lin Lin and Yu Tong. "Heisenberg-Limited Ground-State Energy Estimation for Early

Fault-Tolerant Quantum Computers". PRX Quantum **3**, 010318 (2022).

[71] Seth Lloyd. "Almost any quantum logic gate is universal". Physical Review Letters **75**, 346 (1995).

[72] Nicholas C Rubin, Ryan Babbush, and Jarrod McClean. "Application of fermionic marginal constraints to hybrid quantum algorithms". New Journal of Physics **20**, 053020 (2018).

Appendix

Table of Contents

A Overview of Methods for QSP Involving Post-Selection **18**
- A.1 Overview of QSP Using Qubitization . . . 18
- A.2 Overview of QSP Using Linear Combination of Unitaries (LCU) . . . 19

B Proofs of Lem. 1 and Thm. 2 **21**
- B.1 Proof of Lem. 1 . . . 21
- B.2 Useful Preliminary Results . . . 22
- B.3 Proof of Thm. 2 . . . 24

C Notions of Stability for Unitary Synthesis of Exact Formula in Thm. 2 **27**
- C.1 Convergence of DB-QSP . . . 27
- C.2 Perturbation of the Hamiltonian . . . 30
- C.3 Perturbation of Angles . . . 31
- C.4 Statistical Error Propagation . . . 33

D Applications of DB-QSP **39**
- D.1 Examples of Low-Degree Polynomial Approximations . . . 39
- D.2 Derivation of DB-QITE Using DB-QSP . . . 41
- D.3 Hamiltonian Simulation . . . 41
- D.4 Evolution under a Polynomial Function of Hamiltonian . . . 41
- D.5 Laurent Polynomials . . . 42

E Classically-Aided DB-QSP Synthesis **42**

F Unbiased Estimator of the Operator Variance for Hamiltonians **43**
- F.1 Measurement Procedure . . . 43
- F.2 Biased and Unbiased Estimator of the Operator Variance . . . 44
- F.3 Total Variance of the Unbiased Estimator of the Operator Variance . . . 46
- F.4 Alternative Unbiased Method of Estimating Operator Variance . . . 53

A Overview of Methods for QSP Involving Post-Selection

We start with a brief overview of Quantum Signal Processing (QSP) through its unitary synthesis method known as qubitization. Then, we also introduce the Linear Combination of Unitaries (LCU) as a unitary synthesis technique for implementing QSP.

A.1 Overview of QSP Using Qubitization

QSP is a framework for systematically constructing matrix-valued functions. The concept of QSP originated from a series of works which aimed at characterizing the achievable polynomial functions of a scalar value embedded in a single-qubit rotation [1]. The QSP methodology introduced in Ref. [1] was later extended to Hermitian matrices through a technique known as *qubitization* [5], using the framework of block-encodings. Subsequently, QSP was generalized to all polynomials [63] and extended to non-square matrices through Quantum Singular Value Transformation (QSVT) [3]. Notably, this has led to asymptotically optimal Hamiltonian simulation algorithms [5] and a near-optimal method for ground-state preparation [9]. Furthermore, QSP serves as a fundamental tool for constructing primitive quantum algorithms that exhibit quantum advantages [2, 3]. Therefore, its efficiency in implementing linear algebraic operations and its role as a key building block for quantum algorithms with potential advantages have made QSP a subject of significant interest.

Here, we focus on the qubitization technique [1]. Specifically, following the approach in Ref. [1], we begin with a degree-K polynomial of a scalar input $x \in [-1, 1]$. In the original work, a quantum circuit U_{YLC} was introduced with a sequential structure comprising of two types of operators: *signal operators* W and *signal processing operators* $S(\phi)$, where the phase ϕ is drawn from a set ϕ_k. The desired polynomial transformation is then obtained by performing a measurement in the so-called *signal basis*. Concretely, it was demonstrated that there exists a sequence of QSP phase $\{\phi_k\}$ such that the following circuit

$$U_{YLC} = S_z(\phi_0) \prod_{k=1}^{K} W(x) S_z(\phi_k) \tag{36}$$

with the operators

$$W(x) = e^{ixX/2} = \begin{bmatrix} x & i\sqrt{1-x^2} \\ i\sqrt{1-x^2} & x \end{bmatrix}, \qquad S_z(\phi) = e^{i\phi Z} = \begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{bmatrix},$$

followed by measurement in the basis $M = \{|+\rangle, |-\rangle\}$ can realize a degree-K real polynomial $p(x)$, provided that

1. Degree of $p(x)$ is equal to or less than K,
2. $p(x)$ has a parity K mod 2,
3. $\forall x \in [-1, 1]$, $|p(x)| \leq 1$.

This setting is referred to as the Wx convention, as the signal operator is implemented using the rotation-x gate. Alternatively, the rotation-z gate can be used, which is known as the Wz convention. For further details, see Ref. [2].

The core idea of the synthesis approach is that single-qubit rotations can implement arbitrary polynomial transformations, provided the conditions mentioned earlier are met. Similarly, this single-qubit-like structure used in Eq. (36) to synthesize polynomial functions can be extended to Hermitian matrices by employing the block-encoding technique [19], which embeds a Hermitian matrix H in the top-left block of a larger unitary matrix. More precisely, U_H is called a (α, n_a, ϵ) block-encoding of H, if it satisfies

$$\|H - \alpha(\langle 0|^{\otimes n_a} \otimes I) U_H (|0\rangle^{\otimes n_a} \otimes I)\| \leq \epsilon \tag{37}$$

with $\alpha, \epsilon \in \mathbb{R}^+$ and the number of auxiliary qubits n_a. An example of the matrix form is given by

$$W(H) = \begin{bmatrix} H & i\sqrt{I - H^2} \\ i\sqrt{I - H^2} & H \end{bmatrix}. \tag{38}$$

By substituting the signal operator $W(x)$ in the unitary U_{YLC} of Eq. (36) with the block-encoded unitary in Eq. (38), we can perform polynomial transformations of Hermitian matrices. Furthermore, QSP has been extended to non-square matrices via QSVT, which enables the manipulation of singular values for broader applications in quantum linear algebra.

We note that, given an input state $|\Psi\rangle$, the state after applying the block-encoding unitary U_H in Eq. (37) is expressed as

$$|0\rangle^{\otimes n_a} \otimes \frac{H}{\alpha} |\Omega\rangle + |\text{garbage}^{\perp}\rangle . \tag{39}$$

Here, $|\text{garbage}^{\perp}\rangle$ is a state orthogonal to $|0\rangle^{\otimes n_a} \otimes H/\alpha\,|\Omega\rangle$. Due to the normalization, the probability of getting $|0\rangle^{\otimes n_a}$ is given by $p_{\text{succ}} = \|H\,|\Omega\rangle\,\|^2/\alpha^2$. By extending Eq. (36) to controlled-unitary operations, we obtain the state

$$|0\rangle^{\otimes n_a} \otimes p(H/\alpha)\,|\Omega\rangle + |\text{garbage}^{\perp}\rangle , \tag{40}$$

which succeeds with probability

$$p_{\text{succ}} = \|p(H/\alpha)\,|\Omega\rangle\,\|^2. \tag{41}$$

As shown in the main text, a core insight is that this probability can be exponentially small in case of Imaginary-Time Evolution (ITE) where $p(H) \approx e^{-\tau H}$. In such cases, Eq. (41) is inversely proportional to the fidelity of the initial state with a thermal state, which can decay exponentially [22, 16]. More generally, this fidelity dependence holds across different scenarios. For instance, the block-encoding query complexity for nearly-optimal ground-state preparation algorithm in Ref. [9] scales as $O(\alpha/\gamma)$, where $\gamma = |\langle\lambda_0|\Psi\rangle|^2$ is the fidelity of the input state $|\Psi\rangle$ with the ground state $|\lambda_0\rangle$ of H. The query scaling $O(\alpha/\gamma)$ corresponds to the inverse success probability and thus requires repeated trials for obtaining a successful outcome. Other probabilistic methods exhibit similar sensitivity [9, 8, 64]; see, for example, Ref. [65] for a discussion focused on computing expectation values rather than preparing quantum states and Ref. [10] for modifications of the filter functions which aim to alleviate this problem. This indicates that the success of the block-encoding depends on the input state. Additionally, since the number of queries to the block-encoding unitary scales with the degree of polynomials as shown in Eq. (36), the degree K needs to be sufficiently low to ensure successful post-selection each time.

A.2 Overview of QSP Using Linear Combination of Unitaries (LCU)

Another straightforward approach to implementing QSP is the Linear Combination of Unitaries (LCU) technique [66, 12, 12, 34, 13]. More broadly, LCU is a fundamental method for realizing general matrix functions using unitary operations. The key idea is that, a given matrix $H = \sum_{j=1}^{J} w_j U_j$, which can be expressed as a weighted sum of unitary operators $\{U_j\}$, can be efficiently implemented with additional auxiliary qubits whose number grows logarithmically with the number of decomposed terms J in the matrix. The desired transformation is then realized by measuring the auxiliary qubits, which corresponds to successfully projecting the system onto a subspace where the target operation is encoded. In this sense, LCU serves as one way to implement the block-encoding framework in Eq. (38); that is, LCU can be used as a subroutine of qubitization. However, in this section, we focus on LCU as a standalone approach for realizing QSP.

We begin by outlining the LCU technique in detail. The framework is built upon two essential subroutines: PREP and SEL. The PREP encodes the J coefficients $\{w_j\}$ of the target matrix H on

auxiliary register states $|0\rangle_a = |0\rangle^{\otimes n_a}$ as follows:

$$\text{PREP}: \quad \text{PREP}\ |0\rangle_a = \sum_{j=1}^{J} \sqrt{\frac{w_j}{\|w\|_1}}\,|j\rangle\,, \tag{42}$$

where $\|w\|_1 = \sum_{j=1}^{J} |w_j|$ is the 1-norm of the matrix H. The SEL subroutine applies the unitary U_j to an input state $|\Omega\rangle$, conditioned on the control register being in state j. Combining these operations, we construct the unitary $U_{\text{LCU}} = \text{PREP}^\dagger \cdot \text{SEL} \cdot \text{PREP}$, which gives

$$U_{\text{LCU}}\,|0\rangle_a \otimes |\Omega\rangle = \frac{1}{\|w\|_1}\,|0\rangle_a \otimes H\,|\Omega\rangle + |\text{garbage}^{\perp}\rangle\,. \tag{43}$$

If a measurement of the auxiliary register yields $|0\rangle_a$, the remaining quantum state is the normalized state given by $H\,|\Omega\rangle\,/\|H\,|\Omega\rangle\,\|$. The probability of this successful projection is given by

$$p_{\text{succ}} = \|H\,|\Omega\rangle\,\|^2/\|w\|_1^2\,. \tag{44}$$

This procedure extends naturally to QSP. We exploit the fundamental theorem of algebra, which states that any univariate polynomial with complex coefficients $p(z) = \sum_{k=0}^{K} a_k z^k$ can be factorized in terms of its roots z_k to take the form $p(z) = a_K \prod_{k=0}^{K}(z - z_k)$. This directly generalizes to matrix functions and we get

$$p(H) = a_K \prod_{k=0}^{K} (H - z_k I)\,. \tag{45}$$

Setting H as a Hermitian matrix, we proceed inductively by applying a sequence of the operators $F_k = H - z_k I$ using LCU. This results in the transformation

$$|\Psi_{k+1}\rangle = \frac{F_k\,|\Psi_k\rangle}{\|F_k\,|\Psi_k\rangle\,\|}\,. \tag{46}$$

Since the leading coefficient a_K of the polynomial $p(H) = a_K \prod_{k=1}^{K} F_k$ cancels out by normalization, we obtain

$$|\Psi_K\rangle = \frac{p(H)\,|\Psi_0\rangle}{\|p(H)\,|\Psi_0\rangle\,\|}. \tag{47}$$

Let us next discuss the success probability of this procedure assuming that the Hamiltonian is decomposed into Pauli operators as $H = \sum_{i=1}^{J} w_i P_i$. From Eq. (44), the success probability of post-selection for k step is equal to the conditional probability given that the state $|\Psi_{k-1}\rangle$ at $(k-1)$ step is successfully generated: that is, we have

$$\Pr(k\text{-th step success}\ |\Psi_k) = \frac{\|F_{k-1}\,|\Psi_{k-1}\rangle\,\|^2}{(|z_{k-1}| + \|w\|_1)^2}\,. \tag{48}$$

Thus, the success probability at K step is given by

$$\Pr(\text{QSP success}) = \prod_{k=1}^{K} \Pr(k\text{-th step success}\mid \Psi_{k-1}) = \frac{\|\prod_{k=1}^{K} F_{k-1}\,|\Psi_0\rangle\,\|^2}{\prod_{k=1}^{K}(|z_{k-1}| + \|w\|_1)^{2K}}. \tag{49}$$

Suppose that the probability in Eq. (48) can be bounded by $1 - q$ with $q \in (0, 1]$, then we have $\Pr(\text{QSP success}) \leq (1-q)^{2K}$, indicating an exponential hardness of successful post-selection.

To address these limitations, we turn to our proposal, DB-QSP. Unlike LCU, DB-QSP constructs deterministic unitary operations that implement the desired state transformations without requiring post-selection and auxiliary qubits. This approach could improve the preprocessing initialization for QSP, reducing the overall hardware runtime by eliminating the need for post-selection.

B Proofs of Lem. 1 and Thm. 2

B.1 Proof of Lem. 1

For completeness, we restate the statement from the main text.

Lemma B.1 (Unitary synthesis for linear polynomials without post-selection)**.** *Suppose $p(H) = H - \alpha I$ is any linear polynomial of a Hermitian matrix H with $\alpha \in \mathbb{R}$. Given a state vector $|\Psi\rangle$ with energy mean $E_\Psi = \langle\Psi| H |\Psi\rangle$ and variance $V_\Psi = \langle\Psi| H^2 |\Psi\rangle - E_\Psi^2$, the unitary synthesis for the action of $p(H)$ can be achieved by*

$$U_\Psi = e^{s_\Psi [\Psi, H]}, \tag{50}$$

with

$$s_\Psi = \frac{-1}{\sqrt{V_\Psi}} \arccos\left(\frac{E_\Psi - \alpha}{\sqrt{V_\Psi + (E_\Psi - \alpha)^2}}\right). \tag{51}$$

Proof of Lem. B.1. Firstly, let us verify that U_Ψ is indeed a unitary operator. For a matrix of the form e^W to be unitary, W must be anti-Hermitian, i.e., $W = -W^\dagger$. Since $[\Psi, H] = -([\Psi, H])^\dagger$, the operator in Eq. (50) is therefore unitary.

Next, we demonstrate that the unitary operator e^{W_H} with $W_H = [\Psi, H]$ can be exactly represented by a linear polynomial when applied to the input state $|\Psi\rangle$. By definition, the unitary operator can be expressed as

$$e^{sW_H} = \sum_{k=0}^{\infty} \frac{s^k}{k!} W_H^k \ . \tag{52}$$

Now, we observe that

$$W_H |\Psi\rangle = E_\Psi |\Psi\rangle - H |\Psi\rangle , \tag{53}$$

and

$$W_H^2 |\Psi\rangle = E_\Psi W_H |\Psi\rangle - W_H H |\Psi\rangle = E_\Psi^2 |\Psi\rangle - E_\Psi H |\Psi\rangle - \langle\Psi| H^2 |\Psi\rangle |\Psi\rangle + E_\Psi H |\Psi\rangle = -V_\Psi |\Psi\rangle .$$

This indicates that any even power of the commutator W_H acting on the state $|\Psi\rangle$ gives

$$W_H^{2k} |\Psi\rangle = (-V_\Psi)^k |\Psi\rangle \ . \tag{54}$$

Similarly, we have $W_H^{2k+1} |\Psi\rangle = (-V_\Psi)^k W_H |\Psi\rangle$ for the cases of odd powers. Thus, by separating the odd and even terms, we obtain a weighted sum of $|\Psi\rangle$ and $W_H |\Psi\rangle$ with coefficients expressed by sine and cosine functions as

$$e^{sW_H} |\Psi\rangle = \cos\left(s\sqrt{V_\Psi}\right) |\Psi\rangle + \frac{\sin\left(s\sqrt{V_\Psi}\right)}{\sqrt{V_\Psi}} W_H |\Psi\rangle \ . \tag{55}$$

Using Eq. (53), this simplifies to $e^{sW_H} |\Psi\rangle = (a(s)I + b(s)H) |\Psi\rangle$, where $a(s), b(s)$ are real-valued coefficients for any $s \in \mathbb{R}$:

$$a(s) = \frac{E_\Psi}{\sqrt{V_\Psi}} \sin\left(s\sqrt{V_\Psi}\right) + \cos\left(s\sqrt{V_\Psi}\right), \qquad b(s) = -\frac{1}{\sqrt{V_\Psi}} \sin\left(s\sqrt{V_\Psi}\right) \ . \tag{56}$$

Finally, an explicit calculation reveals that the ansatz for the duration Eq. (51) solves the equations $a(s_\Psi) = -\alpha / \| p(H) |\Psi\rangle \|$ and $b(s_\Psi) = 1 / \| p(H) |\Psi\rangle \|$ where we utilize the equality

$$\| p(H) |\Psi\rangle \| = \sqrt{V_\Psi + (E_\Psi - \alpha)^2} \ . \tag{57}$$

The proof is concluded by noting that this means that

$$e^{s_\Psi W_H} |\Psi\rangle = \frac{(H - \alpha I)) |\Psi\rangle}{\| (H - \alpha I)) |\Psi\rangle \|} \ . \tag{58}$$

□

B.2 Useful Preliminary Results

In this section, we derive an exact formula for implementing an exponential of commutators, $e^{s[\Psi,H]}$, without any approximation or truncation error.

B.2.1 Effective Idempotence of Exponentials of $[\Omega, H]$

We derive an equivalent expression of the unitaries $e^{s[\Omega,H]}$ found in Eq. (50), involving pure states Ψ and the problem Hamiltonian H. We start with the general Taylor series of the exponential of an operator

$$e^{s[\Psi,H]} = \sum_{k=0}^{\infty} \frac{s^k}{k!} \left([\Psi, H]\right)^k \ , \tag{59}$$

where all k-th powers of $s\,[\Psi, H]$ contribute to the unitary. In general, one may approximate this infinite series expansion by truncating it to a degree-K polynomial,

$$e^{s[\Psi,H]} \approx \sum_{k=0}^{K} \frac{s^k}{k!} \left([\Psi, H]\right)^k \ . \tag{60}$$

However, the error $O(s^{K+1})$ requires additional care and investment of resources to control. Interestingly, we prove in Prop. B.2 that when Ψ is a pure state, an *exact* polynomial representation can be obtained with $K = 2$, rather than an approximation.

Proposition B.2 (Effective idempotence)**.** *Let $\Psi = |\Psi\rangle\langle\Psi|$ be a pure density matrix associated to state vector $|\Psi\rangle$ with energy fluctuation $V_\Psi = \langle\Psi| H^2 |\Psi\rangle - \langle\Psi| H |\Psi\rangle^2$. Then for any duration $s \in \mathbb{R}$ we have*

$$e^{s[\Psi,H]} = I + A(s)\,[\Psi, H] + B(s)\left([\Psi, H]\right)^2 \tag{61}$$

where

$$A(s) = \frac{\sin\left(s\sqrt{V_\Psi}\right)}{\sqrt{V_\Psi}}, \qquad B(s) = \frac{1 - \cos\left(s\sqrt{V_\Psi}\right)}{V_\Psi} \ . \tag{62}$$

Proof. We will make a technical calculation showing that third power of the commutator is, in fact, directly proportional to the first power of the commutator, with a scaling factor that depends on energy fluctuation:

$$\left([\Psi, H]\right)^3 = -V_\Psi\,[\Psi, H] \ . \tag{63}$$

We call this effective idempotence. Indeed, in general an operator W is indempotent if $W^2 = 1$ which implies that $e^{sW} = \cos(s)I + \sin(s)W$. Here we have the form $A^3 = \alpha A$ with $\alpha \in \mathbb{R}$, similar to idempotence. Effective idempotence has analogous consequences for the solution to the exponential series. It implies that the $(2k+1)$-th power and the $2k$-th power of the commutator can be written as

$$\left([\Psi, H]\right)^{2k+1} = (-V_\Psi)^k\,[\Psi, H] \ , \qquad \left([\Psi, H]\right)^{2k} = (-V_\Psi)^{k-1}\left([\Psi, H]\right)^2 . \tag{64}$$

Thus we find for the series of representation of the unitary

$$\begin{aligned} e^{s[\Psi,H]} &= I + \left(\sum_{k=0} (-1)^k \frac{s^{2k+1}\left(V_\Psi\,[H]\right)^k}{(2k+1)!}\right)[\Psi, H] + \left(\sum_{k=1} (-1)^{k-1} \frac{s^{2k}\left(V_\Psi\,[H]\right)^{k-1}}{(2k)!}\right)\left([\Psi, H]\right)^2 \\ &= I + A(s)\,[\Psi, H] + B(s)\left([\Psi, H]\right)^2 , \end{aligned} \tag{65}$$

where $A(s), B(s)$ defined in Eq. (62), and we have utilized the Taylor series for sine and cosine in the last equality. We complete the proof by deriving the effective idempotence namely

$$\left([\Psi, H]\right)^3 = (\Psi H - H\Psi)^3 = \left(\Psi H\Psi H - \Psi H^2\Psi - H\Psi H + H\Psi H\Psi\right)(\Psi H - H\Psi) \tag{66}$$

where we have used the assumption that the quantum state is pure, i.e., $\Psi^2 = \Psi$. Next, we switch from density matrix representation to state vector representation, i.e. we substitute back $\Psi = |\Psi\rangle\langle\Psi|$. Thus, it becomes

$$\begin{aligned}([\Psi, H])^3 &= \left(\langle H\rangle (\Psi H + H\Psi) - H\Psi H - \langle H^2\rangle \Psi\right)(\Psi H - H\Psi) \\ &= \langle H\rangle (\Psi H\Psi H + H\Psi H) - H\Psi H\Psi H - \langle H^2\rangle \Psi H \\ &\quad - \langle H\rangle \left(\Psi H^2\Psi + H\Psi H\Psi\right) + H\Psi H^2\Psi + \langle H^2\rangle \Psi H\Psi \ ,\end{aligned}$$

where we introduce the notation $\langle H\rangle = \langle\Psi| H |\Psi\rangle$ and $\langle H^2\rangle = \langle\Psi| H^2 |\Psi\rangle$ in the first line, and the second equality is a direct expansion. Finally, we repeat the same procedure and the result is given by

$$\begin{aligned}([\Psi, H])^3 &= \langle H\rangle (\langle H\rangle \Psi H + H\Psi H) - \langle H\rangle H\Psi H - \langle H^2\rangle \Psi H \\ &\quad - \langle H\rangle \left(\langle H^2\rangle \Psi + \langle H\rangle H\Psi\right) + \langle H^2\rangle H\Psi + \langle H^2\rangle \langle H\rangle \Psi \qquad (67) \\ &= \langle H\rangle^2 \Psi H - \langle H^2\rangle \Psi H - \langle H\rangle^2 H\Psi + \langle H^2\rangle H\Psi \qquad (68) \\ &= -\left(\langle H^2\rangle - \langle H\rangle^2\right)[\Psi, H] = -\left(V_\Psi[H]\right)[\Psi, H] \,, \qquad (69)\end{aligned}$$

where we again use the pure state assumption in the first equality and the definition of V_Ψ in the last equality. □

B.2.2 Exponentials of $[\Psi, H]$ Can Express the Normalized Action of Any Real-Valued Linear Polynomial in H

We next extend Lem. 1 to operators of the form $xI + yH$, where $x, y \in \mathbb{R}$ are not both zero.

Lemma B.3. *Let $x, y \in \mathbb{R}$ and $(x, y) \neq (0, 0)$. Define the parameter*

$$s_\Psi = -\frac{\mathrm{sgn}(y)}{\sqrt{V_\Psi}} \arccos\left(\frac{x + yE_\Psi}{\|(xI + yH)|\Psi\rangle\|}\right). \qquad (70)$$

Then,

$$\frac{(xI + yH)|\Psi\rangle}{\|(xI + yH)|\Psi\rangle\|} = (a(s_\Psi)I + b(s_\Psi)H)|\Psi\rangle \ , \qquad (71)$$

where $a(s_\Psi), b(s_\Psi)$ are real-valued coefficients given by

$$a(s_\Psi) = \frac{E_\Psi}{\sqrt{V_\Psi}} \sin\left(s_\Psi\sqrt{V_\Psi}\right) + \cos\left(s_\Psi\sqrt{V_\Psi}\right), \qquad (72)$$

$$b(s_\Psi) = -\frac{1}{\sqrt{V_\Psi}} \sin\left(s_\Psi\sqrt{V_\Psi}\right) \ . \qquad (73)$$

Proof. We here consider to match the weights of I and H between the polynomial operation and exponentials of $[\Psi, H]$. Specifically, we solve the following two equations;

$$\frac{x}{\|(xI + yH)|\Psi\rangle\|} = \frac{E_\Psi}{\sqrt{V_\Psi}} \sin\left(s\sqrt{V_\Psi}\right) + \cos\left(s\sqrt{V_\Psi}\right), \qquad (74)$$

$$\frac{y}{\|(xI + yH)|\Psi\rangle\|} = -\frac{1}{\sqrt{V_\Psi}} \sin\left(s\sqrt{V_\Psi}\right) \ . \qquad (75)$$

By computing Eq. (74)+E_Ψ×Eq. (75), we get

$$\frac{x + yE_\Psi}{\|(xI + yH)|\Psi\rangle\|} = \cos\left(s_\Psi\sqrt{V_\Psi}\right) \ . \qquad (76)$$

Thus, computing the inverse of the cosine function yields the desired solution. However, since both s_Ψ and $-s_\Psi$ satisfy the equation, the sign must be determined explicitly. Eq. (75) indicates that the sign of $\sin(s_\Psi\sqrt{V_\Psi})$ is given by $-\mathrm{sgn}(y)$. This leads to the expression for the duration s_Ψ shown in Eq. (70).

We lastly verify that this solution is consistent. By substituting Eq. (70) into Eq. (73), we find

$$b(s_\Psi) = \frac{\mathrm{sgn}(y)}{\sqrt{V_\Psi}} \sin\left(\arccos\left(\frac{x + yE_\Psi}{\|(xI + yH)\,|\Psi\rangle\,\|}\right)\right) . \tag{77}$$

Applying $\sin(x) = \sqrt{1 - \cos^2(x)}$ further reveals that

$$b(s_\Psi) = \frac{\mathrm{sgn(y)}}{\sqrt{V_\Psi}} \sqrt{1 - \frac{(x + yE_\Psi)^2}{\|(xI + yH)\,|\Psi\rangle\,\|^2}} . \tag{78}$$

With the identity $\|(xI + yH)\,|\Psi\rangle\,\|^2 = (x + yE_\Psi)^2 + y^2V_\Psi$, this simplifies to

$$b(s_\Psi) = \frac{\mathrm{sgn}(y)}{\sqrt{V_\Psi}} \frac{|y|\sqrt{V_\Psi}}{\|(xI + yH)\,|\Psi\rangle\,\|} \tag{79}$$

$$= \frac{y}{\|(xI + yH)\,|\Psi\rangle\,\|} . \tag{80}$$

Similarly, we use the relation $a(s_\Psi) = -E_\Psi b(s_\Psi) + \cos\left(s_\Psi\sqrt{V_\Psi}\right)$, which gives

$$a(s_\Psi) = \frac{-yE_\Psi}{\|(xI + yH)\,|\Psi\rangle\,\|} + \frac{x + yE_\Psi}{\|(xI + yH)\,|\Psi\rangle\,\|} \tag{81}$$

$$= \frac{x}{\|(xI + yH)\,|\Psi\rangle\,\|} . \tag{82}$$

□

B.3 Proof of Thm. 2

We again restate Thm. 2 in the main text.

Theorem B.4 (Unitary synthesis for QSP without post-selection)**.** *Suppose an input state* $|\Psi_0\rangle$ *and any polynomial* $p(H)$ *of degree* K *for a given Hermitian matrix* H *in the form of Eq.* (15)*. Given energy mean* $E_k = \langle\Psi_k|\,H\,|\Psi_k\rangle$ *and variance* $V_k = \langle\Psi_k|\,H^2\,|\Psi_k\rangle - E_k^2$*, the unitary synthesis for* $p(H)$ *can be achieved by*

$$\frac{p(H)\,|\Psi_0\rangle}{\|p(H)\,|\Psi_0\rangle\,\|} = \prod_{k=0}^{K-1} e^{i\theta_k\Psi_k} e^{s_k[\Psi_k,H]}\,|\Psi_0\rangle\,, \tag{83}$$

with

$$s_k = \frac{-1}{\sqrt{V_k}} \arccos\left(\frac{|E_k - z_k|}{\sqrt{V_k + |E_k - z_k|^2}}\right), \qquad \textit{and} \qquad \theta_k = \arg\left(\frac{E_k - z_k}{|E_k - z_k|}\right) . \tag{84}$$

Here, we recursively define the state $|\Psi_k\rangle$ *as* $|\Psi_{k+1}\rangle = e^{i\theta_k\Psi_k} e^{s_k[\Psi_k,H]}\,|\Psi_k\rangle$*.*

Proof. Let z_k be the roots of $p(H)$ as in Eq. (15). We iterate over the roots and at each step k, we will find θ_k and s_k such that the unitary $U_k = e^{i\theta_k\Psi_k} e^{s_k[\Psi_k,H]}$ will implement the state

$$|\Psi_{k+1}\rangle = \frac{(H - z_kI)\,|\Psi_k\rangle}{\|(H - z_kI)\,|\Psi_k\rangle\,\|} \tag{85}$$

as $|\Psi_{k+1}\rangle = U_k\,|\Psi_k\rangle$. Let us comment that, if we apply the k-th filter fragment $F_k = H - z_kI$, the normalization is given by

$$\|(H - z_kI)\,|\Psi\rangle\,\| = V_\Psi + |E_\Psi - z_k|^2 . \tag{86}$$

We cannot use Lem. 1 directly because in general $z_k \in \mathbb{C}$, while polynomials with only real roots such as Chebyshev polynomials can be realized by directly applying Lem. 1. Instead, in general cases, we associate to z_k the real number

$$u_k = E_k - |E_k - z_k| \tag{87}$$

which is real and using Lem. 1 we set s_k such that

$$e^{s_k[\Psi_k, H]} |\Psi_k\rangle = \frac{(H - u_k I) |\Psi_k\rangle}{\|(H - u_k I) |\Psi_k\rangle\|} . \tag{88}$$

We define θ_k to be within $[0, 2\pi)$ and satisfy $e^{i\theta_k} = \frac{E_k - z_k}{|E_k - z_k|}$. We next observe that using that Ψ_k is pure we have the form $e^{i\theta_k \Psi_k} = I + (e^{i\theta_k} - 1)\Psi_k$, we get the following expression

$$|\Psi_{k+1}\rangle = \frac{(I + (e^{i\theta_k} - 1)\Psi_k)(H - u_k I) |\Psi_k\rangle}{\|(H - u_k I) |\Psi_k\rangle\|} = \frac{(H + e^{i\theta_k}(E_k - u_k)I - E_k I) |\Psi_k\rangle}{\|(H - u_k I) |\Psi_k\rangle\|} . \tag{89}$$

The definitions above were such that $e^{i\theta_k}(E_k - w_k) = E_k - z_k$ which leads to a cancellation and

$$|\Psi_{k+1}\rangle = \frac{(H - z_k I) |\Psi_k\rangle}{\|(H - u_k I) |\Psi_k\rangle\|} . \tag{90}$$

Here the numerator involves z_k as desired but the norm is an expression involving w_k. We have

$$\|(H - u_k I) |\Psi_k\rangle\| = V_k + E_k^2 - 2u_k E_k + u_k^2 = V_k + (E_k - u_k)^2 , \tag{91}$$

which means that, using Eq. (87), we arrive at the form in Eq. (86)

$$\|(H - u_k I) |\Psi_k\rangle\| = V_k + |E_k - z_k|^2 \ = \|(H - z_k I) |\Psi_k\rangle\| . \tag{92}$$

Thus the norms match and we conclude that the unitaries implement the desired action of $F_k = H - z_k I$. □

We conclude this section by discussing the range of s_k, which is relevant for analyzing implementation costs and circuit depth. As shown in Eq. (84), the duration s_k is given by

$$s_k = \frac{-1}{\sqrt{V_k}} \arccos\left(\frac{|E_k - z_k|}{\sqrt{V_k + |E_k - z_k|^2}}\right) . \tag{93}$$

First, to see if $|s_k|$ is a decreasing function with respect to V_k we differentiate the two components $g(V_k) = \sqrt{V_k}$ and $f(V_k) = \arccos\left(\frac{|E_k - z_k|}{\sqrt{V_k + |E_k - z_k|^2}}\right)$;

$$g'(V_k) = \frac{1}{2\sqrt{V_k}} \tag{94}$$

$$f'(V_k) = -\frac{1}{\sqrt{1 - \left(\frac{|E_k - z_k|}{\sqrt{V_k + |E_k - z_k|^2}}\right)^2}} \cdot \left(\frac{|E_k - z_k|}{\sqrt{V_k + |E_k - z_k|^2}}\right)' = \frac{|E_k - z_k|}{2\sqrt{V_k}(V_k + |E_k - z_k|^2)} \tag{95}$$

where we used $(\arccos(y))' = -1/\sqrt{1 - y^2}$. Thus, using quotient rule, we have

$$(|s_k|)' = \frac{f'(V_k)g(V_k) - g'(V_k)f(V_k)}{g^2(V_k)} = \frac{1}{V_k}\left(\frac{1}{2}\frac{|E_k - z_k|}{V_k + |E_k - z_k|^2} - \frac{1}{2}\frac{1}{\sqrt{V_k}} \arccos\left(\frac{|E_k - z_k|}{\sqrt{V_k + |E_k - z_k|^2}}\right)\right) . \tag{96}$$

Next we define $x = \dfrac{|E_k - z_k|}{\sqrt{V_k + |E_k - z_k|^2}} = \cos(\alpha)$ which implies that $\dfrac{1}{\sqrt{V_k}} = \dfrac{x}{|E_k - z_k|}\dfrac{1}{\sqrt{1-x^2}}$ and so we find

$$(|s_k|)' = \frac{x}{2V_k|E_k - z_k|}\left(x - \frac{1}{\sqrt{1-x^2}}\arccos(x)\right) \tag{97}$$

$$= \frac{x}{2V_k|E_k - z_k|}\left(\cos(\alpha) - \frac{\alpha}{\sin(\alpha)}\right) \tag{98}$$

$$= \frac{x}{2V_k|E_k - z_k|}\frac{\frac{1}{2}\sin(2\alpha) - \alpha}{\sin(\alpha)} \tag{99}$$

$$= \frac{x}{4V_k|E_k - z_k|}\frac{\sin(2\alpha) - 2\alpha}{\sin(\alpha)} \leq 0\ , \tag{100}$$

where we use the fact $\sin x \leq x$ in the last line. Then, the maximum value of s_k arises when $V_k = 0$. However, since V_k appears in the denominator, we cannot simply compute the value of s_k at $V_k = 0$. Thus, we apply the L'Hôpital's rule, we get

$$\lim_{V_k \to 0} |s_k| = \lim_{V_k \to 0} \frac{f'(V_k)}{g'(V_k)} = \frac{|E_k - z_k|}{V_k + |E_k - z_k|^2}\bigg|_{V_k = 0} = \frac{1}{|E_k - z_k|}. \tag{101}$$

Thus, the duration $|s_k|$ is upper-bounded by $1/|E_k - z_k|$.

C Notions of Stability for Unitary Synthesis of Exact Formula in Thm. 2

In this section we explore the unitary synthesis of Thm. 2 in more detail. in Sec. C.1 we prove that discretizations using group commutator approximation can converge to the desired QSP.

We then analyze the sensitivity of the exact formula in Thm. 2 to perturbations in the input parameters. We begin by studying a question similar to an existing stability result for QSP using block-encodings. Concretely, the output of QSP synthesis using qubitization will depend on any errors in the block-encoding of the input operator H. This enjoys a certain degree of stability; namely, given block-encodings of H and H' for $\|H\| \leq 1$, $\|H'\| \leq 1$, their transformed block-encodings are also close, $\|p(H) - p(H')\| \leq 4K\|H - H'\|$ [3]. In Sec. C.2, we derive a bound in this scenario.

We then focus on the impact of imperfect parameters θ and s on the performance. In Sec. C.3, we first analyze the impact of deviations in the parameters from their ideal values. However, this analysis alone is insufficient for practical scenarios, as statistical errors arise when estimating energy and variance from a noisy state. To address limitation, we extend the results to the situation where the estimated energy and variance may still differ from their ideal values even in the limit of finite measurement shots. Sec. C.4 explores this extension, beginning with the single-step case before generalizing to arbitrary steps. These results provide insight into the statistical estimation requirements necessary for achieving a converging QSP synthesis.

Hence, we further extend the result to the case where the statistical noise happens when the energy and variance is different from the ideal situation even if we have the infinite number of measurement shots. To address this, Sec. C.4 starts with a single-step case, followed by the arbitrary steps. This result sheds light on the demands of statistical estimation required to obtain a converging QSP synthesis.

C.1 Convergence of DB-QSP

Proposition C.1 (DB-QSP convergence)**.** *Suppose H is a Hermitian matrix whose spectral radius does not exceed unity, i.e., $\|H\| \leq 1$. Let $\zeta = \max(\boldsymbol{\theta}, \boldsymbol{s})$ be the maximum value across all elements in $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_{K-1})$ and $\boldsymbol{s} = (|s_0|, \ldots, |s_{K-1}|)$. For the analysis, we define the state constructed by DB-QSP with $s_k^{(N)} = \sqrt{|s_k|/N}$*

$$|\omega_K\rangle = \prod_{k=0}^{K-1} e^{i\theta_k\omega_k} \left(e^{is_k^{(N)}\omega_k} e^{is_k^{(N)}H} e^{-is_k^{(N)}\omega_k} e^{-is_k^{(N)}H}\right)^N |\omega_0\rangle \tag{102}$$

We also define the exact QSP state derived from Thm. 2

$$|\Psi(\boldsymbol{\theta}, \boldsymbol{s})\rangle = \prod_{k=0}^{K-1} e^{i\theta_k\Psi_k} e^{s_k[\Psi_k, H]} |\Psi_0\rangle \ . \tag{103}$$

Then we have

$$\| \, |\Psi(\boldsymbol{\theta}, \boldsymbol{s})\rangle - |\omega_K\rangle \, \| \leq \frac{4}{3}\sqrt{\frac{\zeta}{N}}(1 + 6\xi)^K \ . \tag{104}$$

Proof. Let us define the intermediate QSP states as

$$|\Psi_k\rangle = \prod_{k'=0}^{k-1} e^{i\theta_{k'}\Psi_{k'}} e^{s_{k'}[\Psi_{k'}, H]} |\Psi_0\rangle \tag{105}$$

and the intermediate DB-QSP states as

$$|\omega_{k+1}\rangle = e^{i\theta_k\omega_k} \left(e^{is_k^{(N)}\omega_k} e^{is_k^{(N)}H} e^{-is_k^{(N)}\omega_k} e^{-is_k^{(N)}H}\right)^N |\omega_k\rangle \ . \tag{106}$$

Section	Main Focus	Proof	Final Results
Sec. C.1	Difference between the exact formula in Thm. 2 and DB-QSP	Prop. C.1	$\| \|\Psi(\boldsymbol{\theta},\boldsymbol{s})\rangle - |\omega_K\rangle \| \leq \frac{4}{3}\sqrt{\frac{\zeta}{N}}(1+6\xi)^K$.
Sec. C.2	Stability with respect to the difference in Hamiltonians	Prop. C.2	$\left\||\Psi_{\theta,s}(H)\rangle - |\Psi_{\theta,s}(\tilde{H})\rangle\right\| \leq \frac{1}{3}(1+6\zeta)^K\|H-\tilde{H}\|$.
Sec. C.3	Sensitivity to changes from the exact angles θ and s	Prop. C.3	$\| \|\Psi_H(\boldsymbol{\theta},\boldsymbol{s})\rangle - |\tilde{\Psi}_H(\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{s}})\rangle \| \leq \frac{\max(\delta_s,\delta_\theta)}{3\zeta}(1+6\zeta)^K$.
Sec. C.4	Error in a single step caused by erroneous estimation of energy and variance	Prop. C.4	$\left\|e^{i\theta_\Psi \Psi} e^{s_\Psi[\Psi,H]}|\Psi\rangle - e^{i\bar{\theta}\Psi}e^{\bar{s}_\Psi[\Psi,H]}|\Psi\rangle\right\| \leq 20\eta^4\max(\delta_{V'},\delta_{E'})$
Sec. C.4	Error in K steps using the estimated QSP parametrization $(\overline{\theta},\overline{s})$	Prop. C.5	$\| \|\Psi_H(\boldsymbol{\theta},\boldsymbol{s})\rangle - |\overline{\Psi}_H(\overline{\boldsymbol{\theta}},\overline{\boldsymbol{s}})\rangle \| \leq (14+120\eta^4)^K\max(\delta_V,\delta_E)$.

Table 1: **Summary of the results explored in this section.** Props. C.1 and C.5 are the key results, but the other derivations should be helpful in understanding their proof. For notation, please refer to the corresponding sections. In addition, we introduce the shorthand $\delta_{E'} = |E_\Psi - E'|$ and $\delta_{V'} = |V_\Psi - V'|$ in this table.

First, we decompose the difference between the updated QSP states and the DB-QSP states as follows:

$$\| |\Psi_{k+1}\rangle - |\omega_{k+1}\rangle \| = \| e^{i\theta_k\Psi_k} e^{s_k[\Psi_k,H]}|\Psi_k\rangle - e^{i\theta_k\omega_k}\left(e^{i\tilde{s}_k\omega_k}e^{i\tilde{s}_kH}e^{-i\tilde{s}_k\omega_k}e^{-i\tilde{s}_kH}\right)^N|\omega_k\rangle \| \tag{107}$$

$$\begin{aligned}
=\Big\| &\left(e^{i\theta_k\Psi_k}\, e^{s_k[\Psi_k,H]}|\Psi_k\rangle - e^{i\theta_k\omega_k}\, e^{s_k[\Psi_k,H]}|\Psi_k\rangle\right)\\
&+\left(e^{i\theta_k\omega_k}\, e^{s_k[\Psi_k,H]}|\Psi_k\rangle - e^{i\theta_k\omega_k}\, e^{s_k[\omega_k,H]}|\Psi_k\rangle\right)\\
&+\left(e^{i\theta_k\omega_k}\, e^{s_k[\omega_k,H]}|\Psi_k\rangle - e^{i\theta_k\omega_k}\, e^{s_k[\omega_k,H]}|\omega_k\rangle\right)\\
&+\left(e^{i\theta_k\omega_k}\, e^{s_k[\omega_k,H]}|\omega_k\rangle - e^{i\theta_k\omega_k}\left(e^{is_k^{(N)}\omega_k}\, e^{is_k^{(N)}H}\, e^{-is_k^{(N)}\omega_k}\, e^{-is_k^{(N)}H}\right)^N|\omega_k\rangle\right)\Big\|
\end{aligned} \tag{108}$$

$$\begin{aligned}
\leq &\| \left(e^{i\theta_k\Psi_k} - e^{i\theta_k\omega_k}\right) e^{s_k[\Psi_k,H]}|\Psi_k\rangle \|\\
&+\| e^{i\theta_k\omega_k}\left(e^{s_k[\Psi_k,H]} - e^{s_k[\omega_k,H]}\right)|\Psi_k\rangle \|\\
&+\| e^{i\theta_k\omega_k}\, e^{s_k[\omega_k,H]}\left(|\Psi_k\rangle - |\omega_k\rangle\right) \|\\
&+\| e^{i\theta_k\omega_k}\left(e^{s_k[\omega_k,H]} - \left(e^{is_k^{(N)}\omega_k}\, e^{is_k^{(N)}H}\, e^{-is_k^{(N)}\omega_k}\, e^{-is_k^{(N)}H}\right)^N\right)|\omega_k\rangle \|
\end{aligned} \tag{109}$$

where we use triangle inequality to obtain the last inequality. Next, we evaluate these terms separately.

1. Using the definition of the operator norm

$$\| \left(e^{i\theta_k \Psi_k} - e^{i\theta_k \omega_k}\right) e^{s_k[\Psi_k, H]} |\Psi_k\rangle \| \leq \|e^{i\theta_k \omega_k} - e^{i\theta_k \Psi_k}\| \leq |\theta_k| \, \|\Psi_k - \omega_k\| \tag{110}$$

where we utilize the inequality $\|e^A - e^B\| \leq \|A - B\|$ for unitary operator and the fact $\|AB\| \leq \|A\|\|B\|$. Moreover, note that $\|\Psi_k - \omega_k\| \leq 2 \, \||\Psi_k\rangle - |\omega_k\rangle\|$ and thus we obtain

$$\| \left(e^{i\theta_k \Psi_k} - e^{i\theta_k \omega_k}\right) e^{s_k[\Psi_k, H]} |\Psi_k\rangle \| \leq 2|\theta_k| \, \| \, |\Psi_k\rangle - |\omega_k\rangle \, \| \tag{111}$$

2. For the second term, we have

$$\begin{aligned} \|e^{i\theta_k \omega_k} \left(e^{s_k[\Psi_k, H]} - e^{i\theta_k \omega_k}\right) e^{s_k[\omega_k, H]} |\Psi_k\rangle \| &\leq \|e^{i\theta_k \omega_k}\| \cdot \| \left(e^{s_k[\Psi_k, H]} - e^{s_k[\omega_k, H]}\right) |\Psi_k\rangle \| && (112) \\ &\leq \|e^{s_k[\Psi_k, H]} - e^{s_k[\omega_k, H]}\| && (113) \\ &\leq |s_k| \, \| \, [\Psi_k - \omega_k, H] \, \| && (114) \end{aligned}$$

where we again employ the unitary invariance and normalised state assumption in the second line; and the property $\|e^A - e^B\| \leq \|A - B\|$ in the last line. Next, using the bound $\|[A, B]\| \leq 2\|A\|\|B\|$, it can be further simplified to

$$\|e^{i\theta_k \omega_k} \left(e^{s_k[\Psi_k, H]} - e^{i\theta_k \omega_k}\right) e^{s_k[\omega_k, H]} |\Psi_k\rangle \| \leq 2|s_k| \, \|\Psi_k - \omega_k\| \cdot \|H\| \tag{115}$$

Similar to the first term, employing the bound $\|\Psi_k - \omega_k\| \leq 2 \, \||\Psi_k\rangle - |\omega_k\rangle\|$ and the assumption that $\|H\| \leq 1$, we get

$$\|e^{i\theta_k \omega_k} \left(e^{s_k[\Psi_k, H]} - e^{i\theta_k \omega_k}\right) e^{s_k[\omega_k, H]} |\Psi_k\rangle \| \leq 4|s_k| \, \| \, |\Psi_k\rangle - |\omega_k\rangle \, \| \tag{116}$$

3. For the third term, since $e^{i\theta_k \omega_k} \, e^{s_k[\omega_k, H]}$ is unitary operator, the third term can be simplified to

$$\|e^{i\theta_k \omega_k} \, e^{s_k[\omega_k, H]} \left(|\Psi_k\rangle - |\omega_k\rangle\right) \| \leq \| \, |\Psi_k\rangle - |\omega_k\rangle \, \| \tag{117}$$

where we use the unitary invariance property of norm.

4. Finally, for the fourth term, it becomes

$$\begin{aligned} &\|e^{i\theta_k \omega_k} \left(e^{s_k[\omega_k, H]} - \left(e^{is_k^{(N)} \omega_k} \, e^{is_k^{(N)} H} \, e^{-is_k^{(N)} \omega_k} \, e^{-is_k^{(N)} H}\right)^N\right) |\omega_k\rangle \| \\ &\leq \|e^{s_k[\omega_k, H]} - \left(e^{i\tilde{s}_k \omega_k} e^{i\tilde{s}_k H} e^{-i\tilde{s}_k \omega_k} e^{-i\tilde{s}_k H}\right)^N \| \end{aligned} \tag{118}$$

Using upper bound in Lemma. (9) from [17] by replacing $s_k \to s_k^{(N)}$, we have

$$\left\| e^{is_k^{(N)} \omega_k} e^{is_k^{(N)} H} e^{-is_k^{(N)} \omega_k} e^{-is_k^{(N)} H} - e^{s_k^{(N)}[\omega_k, H]} \right\| \leq |s_k|^{3/2} N^{-3/2} \left(\|[H, [H, \omega_k]]\| + \|[\omega_k, [\omega_k, H]]\| \right) , \tag{119}$$

By the definition of $s_k^{(N)}$ and telescoping, we have

$$\begin{aligned} &\|e^{s_k[\omega_k, H]} - \left(e^{is_k^{(N)} \omega_k} e^{is_k^{(N)} H} e^{-is_k^{(N)} \omega_k} e^{-is_k^{(N)} H}\right)^N \| \\ &\leq |s_k|^{3/2} / \sqrt{N} \times (\|[H, [H, \omega_k]]\| + \|[\omega_k, [\omega_k, H]]\|) && (120) \\ &\leq 2|s_k|^{3/2} / \sqrt{N} \times (\|[H, \omega_k]\| \times \|H\| + \|[\omega_k, H]\| \times \|\omega_k\|) && (121) \\ &\leq 4|s_k|^{3/2} / \sqrt{N} \times \left(\|\omega_k\| \times \|H\|^2 + \|H\| \times \|\omega_k\|^2\right) && (122) \end{aligned}$$

where we recall the bound $\|[A, B]\| \leq 2\|A\|\|B\|$ in the second and third line. Since we assume that $\|H\| \leq 1$ and $\|\omega_k = 1\|$, we achieve

$$\|e^{s_k[\omega_k, H]} - \left(e^{is_k^{(N)} \omega_k} e^{is_k^{(N)} H} e^{-is_k^{(N)} \omega_k} e^{-is_k^{(N)} H}\right)^N \| \leq 8|s_k|^{3/2} / \sqrt{N} \, . \tag{123}$$

Collecting all terms, Eq. (109) becomes

$$\| |\Psi_{k+1}\rangle - |\omega_{k+1}\rangle \| \leq (1+2|\theta_k|+4|s_k|)\| |\Psi_k\rangle - |\omega_k\rangle \| + 8|s_k|^{3/2}/\sqrt{N} \tag{124}$$

$$\leq (1+6\zeta)\| |\Psi_k\rangle - |\omega_k\rangle \| + 8\zeta^{3/2}/\sqrt{N} \ . \tag{125}$$

where we use the definition $\zeta = \max_{k=1,\ldots,K}(\theta_k, s_k)$ to obtain last line. Iterating this recursive bound, we get

$$\| |\Psi_{k+1}\rangle - |\omega_{k+1}\rangle \| \leq \frac{8\zeta^{3/2}}{\sqrt{N}} \sum_{i=0}^{k} (1+6\zeta)^i = \frac{8\zeta^{3/2}}{\sqrt{N}} \times \frac{(1+6\xi)^{k+1}-1}{(1+6\xi)-1} \tag{126}$$

$$\leq \frac{4}{3}\sqrt{\frac{\xi}{N}}(1+6\xi)^{k+1} \ . \tag{127}$$

Setting $K = k+1$, the proposition statement is justified.

□

C.2 Perturbation of the Hamiltonian

Using Thm. 2 we define $|\Psi_{\boldsymbol{\theta},\boldsymbol{s}}(H)\rangle = \prod_{k=1}^{K} e^{i\theta\Psi_k} e^{s_k[\Psi_k,H]} |\Psi_0\rangle$. This definition indicates that we will hold the angles θ_k and s_k fixed but consider what happens if the Hamiltonian is perturbed.

Proposition C.2 (QSP task stability)**.** *Suppose H is a Hermitian matrix whose spectral radius does not exceed unity, i.e., $\|H\| \leq 1$. Let $\zeta = \max(\boldsymbol{\theta}, \boldsymbol{s})$ be the maximum value across all elements in $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_{K-1})$ and $\boldsymbol{s} = (|s_0|, \ldots, |s_{K-1}|)$. Then, we have*

$$\left\| |\Psi_{\theta,s}(H)\rangle - |\Psi_{\theta,s}(\tilde{H})\rangle \right\| \leq \frac{1}{3}(1+6\zeta)^K \|H - \tilde{H}\| \ . \tag{128}$$

Proof. Let us define the intermediate QSP states

$$|\Psi_k\rangle = \prod_{k'=0}^{k-1} e^{i\theta_{k'}\Psi_{k'}} e^{s_{k'}[\Psi_{k'},H]} |\Psi_0\rangle \tag{129}$$

and analogously $|\tilde{\Psi}_k\rangle$ are the intermediate states of QSP with $\tilde{H}$. Thus, the difference between $|\Psi_{k+1}\rangle$ and $|\tilde{\Psi}_{k+1}\rangle$ is given by

$$\| |\Psi_{k+1}\rangle - |\tilde{\Psi}_{k+1}\rangle \| = \| e^{i\theta_k\Psi_k} e^{s_k[\Psi_k,H]} |\Psi_k\rangle - e^{i\theta_k\tilde{\Psi}_k} e^{s_k[\tilde{\Psi}_k,\tilde{H}]} |\tilde{\Psi}_k\rangle \| \ . \tag{130}$$

Next, following the same procedure in Eq. (109) from Subsec. C.1, we add and subtract the term $\{e^{i\theta_k\Psi_k} e^{s_k[\Psi_k,H]} |\tilde{\Psi}_k\rangle \ , \ e^{i\theta_k\tilde{\Psi}_k} e^{s_k[\Psi_k,H]} |\tilde{\Psi}_k\rangle\}$ to split them into multiple norm calculations via triangle inequality.

Consequently, the result is

$$\| |\Psi_{k+1}\rangle - |\tilde{\Psi}_{k+1}\rangle \| \leq \| |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \| + \|e^{i\theta_k\Psi_k} - e^{i\theta_k\tilde{\Psi}_k}\| + \|e^{s_k[\Psi_k,H]} - e^{s_k[\tilde{\Psi}_k,\tilde{H}]}\| \tag{131}$$

$$\leq \| |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \| + |\theta_k| \cdot \|\Psi_k - \tilde{\Psi}_k\| + |s_k| \cdot \|[\Psi_k, H] - [\tilde{\Psi}_k, \tilde{H}]\| \ , \tag{132}$$

where we recall the unitary invariance property of norm in the first inequality and we utilize the formula $\|e^A - e^B\| \leq \|A - B\|$ in the second inequality. We then simplify these three terms separately.

1. For the first term $\| |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \|$, it remains unchanged.

2. For the second term, it becomes

$$|\theta_k| \cdot \|\Psi_k - \tilde{\Psi}_k\| \leq 2|\theta_k| \cdot \| |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \| \ , \tag{133}$$

where we use the relation $\|\Psi_k - \tilde{\Psi}_k\| \leq 2\| |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \|$.

3. For the third term, we rewrite it as

$$|s_k| \cdot \|[\Psi_k, H] - [\tilde{\Psi}_k, \tilde{H}]\| = |s_k| \cdot \|\Psi_k H - H\Psi_k - \left(\tilde{\Psi}_k \tilde{H} - \tilde{H}\tilde{\Psi}_k\right)\| \tag{134}$$

$$= |s_k| \cdot \|\Psi_k H - \Psi_k \tilde{H} + \Psi_k \tilde{H} - \tilde{\Psi}_k \tilde{H} - H\Psi_k + H\tilde{\Psi}_k - H\tilde{\Psi}_k + \tilde{H}\tilde{\Psi}_k\| \tag{135}$$

$$= |s_k| \cdot \|\Psi_k \left(H - \tilde{H}\right) + \left(\Psi_k - \tilde{\Psi}_k\right)\tilde{H} - H\left(\Psi_k - \tilde{\Psi}_k\right) - \left(H - \tilde{H}\right)\tilde{\Psi}_k\| \ . \tag{136}$$

By triangle inequality and operator norm's definition, we obtain

$$|s_k| \cdot \|[\Psi_k, H] - [\tilde{\Psi}_k, \tilde{H}]\| \leq 2|s_k| \cdot \|\Psi_k - \tilde{\Psi}_k\| \cdot \|H\| + 2|s_k| \cdot \|H - \tilde{H}\| \ . \tag{137}$$

Similarly, using $\|\Psi_k - \tilde{\Psi}_k\| \leq 2\| \, |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \, \|$, it becomes

$$\|[\Psi_k, H] - [\tilde{\Psi}_k, \tilde{H}]\| \leq 4|s_k| \cdot \| \, |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \, \| + 2|s_k| \cdot \|H - \tilde{H}\| \ . \tag{138}$$

Collecting all the terms, Eq. (132) reduces to

$$\| \, |\Psi_{k+1}\rangle - |\tilde{\Psi}_{k+1}\rangle \, \| \leq (1 + 2|\theta_k| + 4|s_k|)\| \, |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \, \| + 2|s_k| \cdot \|H - \tilde{H}\| \tag{139}$$

$$\leq (1 + 6\zeta)\| \, |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \, \| + 2|s_k| \cdot \|H - \tilde{H}\| \ , \tag{140}$$

where we use the definition $\zeta = \max(\boldsymbol{\theta}, \boldsymbol{s})$ in the last line. Finally, iterating this recursive bound and it yields

$$\| \, |\Psi_{k+1}\rangle - |\tilde{\Psi}_{k+1}\rangle \, \| \leq \frac{|s_k| \cdot \|H - \tilde{H}\|}{3\zeta}(1 + 6\zeta)^{k+1} \leq \frac{1}{3}\|H - \tilde{H}\|(1 + 6\zeta)^{k+1} \ , \tag{141}$$

where we again used the definition $\zeta = \max(\boldsymbol{\theta}, \boldsymbol{s})$, i.e. $\dfrac{|s_k|}{\zeta} \leq 1$. Setting $K = k + 1$, the proposition's statement is justified. □

C.3 Perturbation of Angles

In order to study sensitivity of the parametrization in Thm. 2 we define

$$|\Psi_H(\boldsymbol{\theta}, \boldsymbol{s})\rangle = \prod_{k=0}^{K-1} e^{i\theta\Psi_k} e^{s_k[\Psi_k, H]} \, |\Psi_0\rangle \ . \tag{142}$$

In practice, we first measure the energy and variance, then compute and then compute s_k and θ_k to implement the operation. From this perspective, the time duration s_k and phase θ_k vary at each step; that is, the perturbations satisfy $|s_k - \tilde{s}_k| \leq \delta_s$ and $|\theta_k - \tilde{\theta}_k| \leq \delta_\theta$. In other words, even if the unitary implementation is perfect, the determined values for time duration and phase can cause errors.

Under this setting, we establish an error bound for implementing a non-unitary polynomial of degree K. For simplicity, we assume the errors in time duration and phase remain constant across all steps. In what follows, we denote the ideal state and operations as $|\Psi_{k+1}\rangle$ and $e^{i\theta_k\Psi_k}e^{s_k[\Psi_k,H]}$, whereas erroneous counterparts are given by $e^{i\tilde{\theta}_k\tilde{\Psi}_k}e^{\tilde{s}_k[\tilde{\Psi}_k,H]}$. Finally, we note that no group commutator approximation is performed in this analysis.

Proposition C.3 (QSP parametrization stability)**.** *Let H be a Hermitian matrix such that $\|H\| \leq 1$, and assume that the estimated parameters $\tilde{s}_k$ and $\tilde{\theta}_k$ satisfy $|s_k - \tilde{s}_k| \leq \delta_s$ and $|\theta_k - \tilde{\theta}_k| \leq \delta_\theta$ with ideal parameters s_k and θ_k for all k. By setting $\zeta = \max(\boldsymbol{\theta}, \boldsymbol{s})$, the perturbed state $|\tilde{\Psi}_H(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{s}})\rangle$ in Eq. (32) and the state $|\Psi_H(\boldsymbol{\theta}, \boldsymbol{s})\rangle$ from Thm. 2 satisfies*

$$\| \, |\Psi_H(\boldsymbol{\theta}, \boldsymbol{s})\rangle - |\tilde{\Psi}_H(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{s}})\rangle \, \| \leq \frac{1}{3\zeta}(1 + 6\zeta)^K \max(\delta_s, \delta_\theta) \ . \tag{143}$$

Proof. Let us define the intermediate QSP states

$$|\Psi_k\rangle = \prod_{k'=0}^{k-1} e^{i\theta_{k'}\Psi_{k'}} e^{s_{k'}[\Psi_{k'},H]} |\Psi_0\rangle \tag{144}$$

and analogously $|\tilde{\Psi}_k\rangle$ are the intermediate states of QSP with $\tilde{\theta}_k$ and $\tilde{s}_k$. The difference between $|\Psi_{k+1}\rangle$ and $|\tilde{\Psi}_{k+1}\rangle$ is given by

$$\| \, |\Psi_{k+1}\rangle - |\tilde{\Psi}_{k+1}\rangle \, \| = \| e^{i\theta_k\Psi_k} e^{s_k[\Psi_k,H]} |\Psi_k\rangle - e^{i\tilde{\theta}_k\tilde{\Psi}_k} e^{\tilde{s}_k[\tilde{\Psi}_k,H]} |\tilde{\Psi}_k\rangle \, \| \, . \tag{145}$$

Again, following the same procedure in Eq. (109) from Subsec. C.1, we add and subtract the term $\{e^{i\theta_k\Psi_k} e^{s_k[\Psi_k,H]} |\tilde{\Psi}_k\rangle \, , \;\; e^{i\theta_k\tilde{\Psi}_k} e^{s_k[\Psi_k,H]} |\tilde{\Psi}_k\rangle \, , \;\; e^{i\theta_k\tilde{\Psi}_k} e^{s_k[\tilde{\Psi}_k,H]} |\tilde{\Psi}_k\rangle\}, \;\; e^{i\tilde{\theta}_k\tilde{\Psi}_k} e^{s_k[\tilde{\Psi}_k,H]} |\tilde{\Psi}_k\rangle\}$ to split them into multiple norm calculation via triangle inequality. Therefore, the result is

$$\begin{aligned}
\| \, |\Psi_{k+1}\rangle - |\tilde{\Psi}_{k+1}\rangle \, \| &\leq \| \, |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \, \| + \| e^{i\theta_k\Psi_k} - e^{i\theta_k\tilde{\Psi}_k} \| + \| e^{s_k[\Psi_k,H]} - e^{s_k[\tilde{\Psi}_k,H]} \| \\
&\quad + \| e^{i\theta_k\tilde{\Psi}_k} - e^{i\tilde{\theta}_k\tilde{\Psi}_k} \| + \| e^{s_k[\tilde{\Psi}_k,H]} |\tilde{\Psi}_k\rangle - e^{\tilde{s}_k[\tilde{\Psi}_k,H]} |\tilde{\Psi}_k\rangle \, \| \qquad (146) \\
&\leq \| \, |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \, \| \\
&\quad + |\theta_k| \cdot \| \Psi_k - \tilde{\Psi}_k \| \\
&\quad + |s_k| \cdot \| [\Psi_k - \tilde{\Psi}_k, H] \| \\
&\quad + |\theta_k - \tilde{\theta}_k| \\
&\quad + \| \left(e^{s_k[\tilde{\Psi},H]} - e^{\tilde{s}_k[\tilde{\Psi},H]} \right) |\tilde{\Psi}_k\rangle \, \| \, , \qquad (147)
\end{aligned}$$

where we use the formula $\|e^A - e^B\| \leq \|A - B\|$ in the second inequality. Next, we proceed to evaluate these terms separately.

1. For the first term $\| \, |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \, \|$, it remains unchanged.

2. For the second term, it becomes

$$|\theta_k| \cdot \| \Psi_k - \tilde{\Psi}_k \| \leq 2|\theta_k| \cdot \| \, |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \, \| \, , \tag{148}$$

   where we use the relation $\|\Psi_k - \tilde{\Psi}_k\| \leq 2\| \, |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \, \|$.

3. For the third term, it is

$$\begin{aligned}
|s_k| \cdot \| [\Psi_k - \tilde{\Psi}_k, H] \| &\leq 2|s_k| \cdot \| \Psi_k - \tilde{\Psi}_k \| \cdot \| H \| \qquad (149) \\
&\leq 4|s_k| \cdot \| \, |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \, \| \, , \qquad (150)
\end{aligned}$$

   where we use the bound $\|[A,B]\| \leq 2\|A\|\|B\|$ in the first line and the relation $\|\Psi_k - \tilde{\Psi}_k\| \leq 2\| \, |\Psi_k\rangle - |\tilde{\Psi}_k\rangle \, \|$. Note that we also exploited the assumption that $\|H\| \leq 1$ in the second line.

4. For the fourth term, we recall the definition of δ_θ, i.e. $|\theta_k - \tilde{\theta}_k| \leq \delta_\theta$.

5. For the fifth term, we observe that

$$\begin{aligned}
e^{\tilde{s}[\Psi,H]} |\tilde{\Psi}_k\rangle &= \left(\left(\frac{E_\Omega}{\sqrt{V_\Omega}} \sin(\tilde{s}\sqrt{V_\Omega}) + \cos(\tilde{s}\sqrt{V_\Omega}) \right) I - \frac{1}{\sqrt{V_\Omega}} \sin(\tilde{s}\sqrt{V_\Omega}) H \right) |\tilde{\Psi}_k\rangle \\
&= \cos(\delta_s\sqrt{V_\Omega}) \left(e^{s[\Psi,H]} |\tilde{\Psi}_k\rangle \right) + \sin(\delta_s\sqrt{V_\Omega}) \left(e^{(s+\pi/2\sqrt{V_\Omega})[\Psi,H]} |\tilde{\Psi}_k\rangle \right) .
\end{aligned} \tag{151}$$

   Using this expression, the fifth term can be simplified to

$$\begin{aligned}
\| e^{\tilde{s}[\Psi,H]} |\Psi\rangle - e^{s[\Psi,H]} |\Psi\rangle \, \| &= \sqrt{2 - 2|\, \langle\Psi| e^{\tilde{s}[\Psi,H]} e^{-s[\Psi,H]} |\Psi\rangle \,|} \\
&= \sqrt{2 - 2|\cos(\delta_s\sqrt{V_\Omega})|} \qquad (152) \\
&= \sqrt{4 - 4|\cos^2(\delta_s\sqrt{V_\Omega/2})|} = 2|\sin(\delta_s\sqrt{V_\Omega}/2)| \leq \delta_s\sqrt{V_\Omega},
\end{aligned}$$

Using the fact that $\sqrt{V_{\Psi_k}} \leq \|H\|$ and the assumption that $\|H\| \leq 1$, we have

$$\|e^{\tilde{s}[\Psi,H]}\,|\Psi\rangle - e^{s[\Psi,H]}\,|\Psi\rangle\,\| \leq \delta_s\ . \tag{153}$$

Collecting all the terms, Eq. (147) reduces to

$$\begin{aligned}\|\,|\Psi_{k+1}\rangle - |\tilde{\Psi}_{k+1}\rangle\,\| &\leq (1+2|\theta_k|+4|s_k|)\|\,|\Psi_k\rangle - |\tilde{\Psi}_k\rangle\,\| + \delta_\theta + \delta_s &(154)\\ &\leq (1+6\zeta)\|\,|\Psi_k\rangle - |\tilde{\Psi}_k\rangle\,\| + 2\gamma. &(155)\end{aligned}$$

where we utilize the definition of γ in the last line, i.e. $\gamma = \max(\delta_s, \delta_\theta)$. Now, solving the iterative sequence, we get

$$\|\,|\Psi_{k+1}\rangle - |\tilde{\Psi}_{k+1}\rangle\,\| \leq 2\gamma\sum_{i=0}^{k}(1+6\zeta)^i = 2\gamma\frac{1-(1+6\zeta)^{k+1}}{1-(1+6\zeta)} \leq \frac{\gamma}{3\zeta}(1+6\zeta)^{k+1}\ . \tag{156}$$

Setting $K = k+1$, the proposition statement is justified.

□

C.4 Statistical Error Propagation

In this section, we study sensitivity of the parametrization in Thm. 2 to estimation errors of the energy and variance. More precisely, for $k = 0, \ldots K-1$, we define the energy $\overline{E}_k$ and the variance $\overline{V}_k$ for states $|\overline{\Psi}_k\rangle$, which is recursively determined by

$$|\overline{\Psi}_{k+1}\rangle = e^{i\overline{\theta}_k\overline{\Psi}_k}e^{\overline{s}_k[\overline{\Psi}_k,H]}\,|\overline{\Psi}_k\rangle\ , \tag{157}$$

with

$$\overline{s}_k = \frac{-1}{\sqrt{\overline{V}_k}}\arccos\left(\frac{\overline{E}_k - u}{\sqrt{\overline{V}_k + (\overline{E}_k - u)^2}}\right) \tag{158}$$

and

$$\overline{\theta}_k = \arg\left(\frac{\overline{E}_k - z_k}{|\overline{E}_k - z_k|}\right). \tag{159}$$

With this, the final state reads

$$|\overline{\Psi}_H(\overline{\boldsymbol{\theta}}, \overline{\boldsymbol{s}})\rangle = \prod_{k=0}^{K-1} e^{i\overline{\theta}_k\overline{\Psi}_k}e^{\overline{s}_k[\overline{\Psi}_k,H]}\,|\overline{\Psi}_0\rangle\ . \tag{160}$$

Note that, while Prop. C.3 characterizes the sensitivity to differences in parameters s_k and θ_k, its direct application to analyzing the impact of statistical estimates is non-trivial. To address this, we establish a lemma that circumvents this challenge by directly considering the relevant quantum states. In the analysis, we define

$$s(E,V) = \frac{-1}{\sqrt{V}}\arccos\left(\frac{|E-z|}{\sqrt{V+|E-z|^2}}\right) \tag{161}$$

and

$$\theta(E,V) = \arg\left(\frac{E-z}{|E-z|}\right)\ . \tag{162}$$

for any $E \in \mathbb{R}$ and $V \geq 0$.

Proposition C.4 (Statistical error propagation)**.** *Suppose H is a Hermitian matrix whose spectral radius does not exceed unity, i.e., $\|H\| \leq 1$. Consider the linear polynomial $p(H) = H - zI$, which we implement using Thm. 2 for some $z \in \mathbb{C}$. Let $|\Psi\rangle$ be a state with energy E_Ψ and variance V_Ψ. Then, for any $E' \in \mathbb{R}$ and $V' \geq 0$, we have*

$$\| \left(e^{i\theta(E_\Psi,V_\Psi)\Psi} e^{s(E_\Psi,V_\Psi)[\Psi,H]} - e^{i\theta(E',V')\Psi} e^{s(E',V')[\Psi,H]} \right) |\Psi\rangle \| \leq 20\eta^4 \max(|E_\Psi - E'|, |V_\Psi - V'|) \ , \tag{163}$$

where $\eta = \max(\frac{1}{\sqrt{V_\Psi}}, \frac{1}{\sqrt{V'}}, \frac{1}{|E_\Psi - z|}, \frac{1}{|E' - z|}, 1 + |z|)$ is the maximal characteristic instability scale.

Proof. For brevity, we define $\theta = \theta(E_\Psi, V_\Psi)$, and $\overline{\theta} = \theta(E', V')$, as well as $s = s(E, V)$ and $\overline{s} = s(E', V')$.

First we reduce the problem into two separate bounds

$$\begin{aligned} \left\| e^{i\theta\Psi} e^{s[\Psi,H]} |\Psi\rangle - e^{i\overline{\theta}\Psi} e^{\overline{s}[\Psi,H]} |\Psi\rangle \right\| &\leq \left\| e^{i\theta\Psi} e^{s[\Psi,H]} |\Psi\rangle - e^{i\theta\Psi} e^{\overline{s}[\Psi,H]} |\Psi\rangle \right\| \\ &\quad + \left\| e^{i\theta\Psi} e^{\overline{s}[\Psi,H]} |\Psi\rangle - e^{i\overline{\theta}\Psi} e^{\overline{s}[\Psi,H]} |\Psi\rangle \right\| \qquad (164) \\ &= \left\| e^{i\theta\Psi} - e^{i\overline{\theta}\Psi} \right\| + \left\| e^{s[\Psi,H]} |\Psi\rangle - e^{\overline{s}[\Psi,H]} |\Psi\rangle \right\| \ , \qquad (165) \end{aligned}$$

where we use the unitary invariance in the second inequality. Next, we evaluate these two terms individually.

1. First, by utilizing the fact that Ψ is a pure state, we have

$$\begin{aligned} \left\| e^{i\overline{\theta}\Psi} - e^{i\theta\Psi} \right\| &= \| I + (e^{i\overline{\theta}} - 1)\Psi - \left(I + (e^{i\theta} - 1)\Psi \right) \| \qquad (166) \\ &= |e^{i\overline{\theta}} - e^{i\theta}| \qquad (167) \\ &= \left| \frac{E' - z}{|E' - z|} - \frac{E_\Psi - z}{|E_\Psi - z|} \right| \qquad (168) \\ &\leq |E' - z| \left| \frac{1}{|E' - z|} - \frac{1}{|E_\Psi - z|} \right| + \frac{|E' - E_\Psi|}{|E_\Psi - z|} \qquad (169) \\ &\leq \eta \Big| |E_\Psi - z| - |E' - z| \Big| + \eta |E' - E_\Psi| \qquad (170) \\ &\leq 2\eta |E' - E_\Psi|, \qquad (171) \end{aligned}$$

   where we utilize the triangle inequality for the forth line, while we use reverse triangle inequality in the last line.

2. The implementation with $\overline{s}$ results in $e^{\overline{s}[\Psi,H]} |\Psi\rangle = (a(\overline{s})I + b(\overline{s})H) |\Psi\rangle$ with

$$a(\overline{s}) = \frac{E_\Psi}{\sqrt{V_\Psi}} \sin\left(\overline{s}\sqrt{V_\Psi}\right) + \cos\left(\overline{s}\sqrt{V_\Psi}\right), \tag{172}$$

$$b(\overline{s}) = -\frac{1}{\sqrt{V_\Psi}} \sin\left(\overline{s}\sqrt{V_\Psi}\right) \ . \tag{173}$$

   Recall that the equalities are derived in Lem. 1; see Sec. B for more details. We stress that E_Ψ and V_Ψ could be different from the estimated ones used for determining $\overline{s}$. This stands in contrast to implementing the polynomial which we wanted $e^{s[\Psi,H]} |\Psi\rangle = (a(s)I + b(s)H) |\Psi\rangle$ with

$$a(s) = \frac{E_\Psi}{\sqrt{V_\Psi}} \sin\left(s\sqrt{V_\Psi}\right) + \cos\left(s\sqrt{V_\Psi}\right), \tag{174}$$

$$b(s) = -\frac{1}{\sqrt{V_\Psi}} \sin\left(s\sqrt{V_\Psi}\right) \ . \tag{175}$$

With these expressions, we have

$$\left\| e^{\overline{s}[\Psi,H]} |\Psi\rangle - e^{s[\Psi,H]} |\Psi\rangle \right\| = \|(a(\overline{s})I + b(\overline{s})H) |\Psi\rangle - (a(s)I + b(s)H) |\Psi\rangle\| \tag{176}$$

$$\leq |a(s) - a(\overline{s})| + |b(s) - b(\overline{s})| \, \|H\| \tag{177}$$

$$\leq |a(s) - a(\overline{s})| + |b(s) - b(\overline{s})| \ . \tag{178}$$

In the last line, we used the spectral assumption $\|H\| \leq 1$.

(a) We begin by bounding $|b(s) - b(\overline{s})|$ because this will help with the bound for $a(s)$. First, for ease of notation, we introduce $\alpha = \sqrt{V_\Psi} s = \arccos\left(\frac{|E_\Psi - z|}{\sqrt{V_\Psi + |E_\Psi - z|^2}}\right)$ to denote $b(s) = -\sin(\alpha)/\sqrt{V_\Psi}$. Similarly, $b(\overline{s})$ with the estimated $\overline{s}$ is expressed using $\overline{\alpha} = \sqrt{V'}\overline{s}$ as

$$b(\overline{s}) = -\frac{1}{\sqrt{V_\Psi}} \sin\left(\sqrt{V_\Psi / V'}\overline{\alpha}\right) \ . \tag{179}$$

Thus, we have

$$|b(s) - b(\overline{s})| = \frac{1}{\sqrt{V_\Psi}} |\sin(\sqrt{V_\Psi / V'}\overline{\alpha}) - \sin(\alpha)| \tag{180}$$

$$\leq \frac{1}{\sqrt{V_\Psi}} |\sin(\sqrt{V_\Psi / V'}\overline{\alpha}) - \sin(\overline{\alpha})| + \frac{1}{\sqrt{V_\Psi}} |\sin(\overline{\alpha}) - \sin(\alpha)| \ . \tag{181}$$

i. For the first part, using $\sin(a) - \sin(b) = 2\sin(\frac{a-b}{2})\cos(\frac{a+b}{2})$, we get

$$\frac{1}{\sqrt{V_\Psi}} |\sin(\sqrt{V_\Psi / V'}\overline{\alpha}) - \sin(\overline{\alpha})|$$

$$\leq \frac{1}{\sqrt{V_\Psi}} \left| 2\sin\left((\sqrt{V_\Psi / V'} - 1)\overline{\alpha}/2\right) \cos\left((\sqrt{V_\Psi / V'} + 1)\overline{\alpha}/2\right) \right| \tag{182}$$

$$\leq \frac{\pi}{2\sqrt{V_\Psi}} \left| \sqrt{V_\Psi / V'} - 1 \right| \tag{183}$$

$$\leq \frac{\pi}{2} \left| \frac{1}{\sqrt{V_\Psi}} - \frac{1}{\sqrt{V'}} \right| \tag{184}$$

$$\leq \frac{\pi}{2} \frac{|V_\Psi - V'|}{\sqrt{V_\Psi}\sqrt{V'}(\sqrt{V_\Psi} + \sqrt{V'})} \tag{185}$$

$$\leq \frac{\pi}{2} \frac{|V_\Psi - V'|}{V_\Psi \sqrt{V'}} \tag{186}$$

$$\leq 2\eta^3 |V_\Psi - V'|, \tag{187}$$

where we use $\cos(x) \leq 1$, $\sin(x) \leq x$ and $\alpha \leq \pi/2$ in the second inequality. As for the third equality, we utilize

$$\frac{1}{\sqrt{A}} - \frac{1}{\sqrt{B}} = \frac{B - A}{\sqrt{AB}(\sqrt{A} + \sqrt{B})},$$

while we use $1/(x + y) \leq 1/x$ for $x > 0$ and $y > 0$ in the fourth inequality.

ii. For the second part, we notice that

$$\sin(\alpha) = \sqrt{1 - |E_\Psi - z|^2/(V_\Psi + |E_\Psi - z|^2)} = \sqrt{V_\Psi/(V_\Psi + |E_\Psi - z|^2)}. \tag{188}$$

with $\cos\alpha = \frac{|E_\Psi - z|}{\sqrt{V_\Psi + |E_\Psi - z|^2}}$. Hence, we have

$$\begin{aligned}
&\frac{1}{\sqrt{V_\Psi}}|\sin(\overline{\alpha}) - \sin(\alpha)| \\
&\leq \frac{1}{\sqrt{V_\Psi}}\left|\sqrt{\frac{V_\Psi}{V_\Psi + |E_\Psi - z|^2}} - \sqrt{\frac{V'}{V' + |E' - z|^2}}\right| && (189)\\
&\leq \left|\frac{1}{\sqrt{V_\Psi + |E - z|^2}} - \frac{1}{\sqrt{V' + |E' - z|^2}}\right| + \frac{1}{\sqrt{V_\Psi}\sqrt{V' + |E' - z|^2}}\left|\sqrt{V_\Psi} - \sqrt{V'}\right| \ . && (190)
\end{aligned}$$

The first component is further upper bounded by

$$\begin{aligned}
&\left|\frac{1}{\sqrt{V_\Psi + |E_\Psi - z|^2}} - \frac{1}{\sqrt{V' + |E' - z|^2}}\right| \\
&\leq \frac{|V_\Psi + |E - z|^2 - (V' + |E' - z|^2)|}{\sqrt{V_\Psi + |E - z|^2}\sqrt{V' + |E' - z|^2}(\sqrt{V_\Psi + |E - z|^2} + \sqrt{V' + |E' - z|^2})} && (191)\\
&\leq \frac{|V_\Psi - V'| + ||E_\Psi - z|^2 - |E' - z|^2|}{|E_\Psi - z|^2|E' - z|} && (192)\\
&\leq \frac{|V_\Psi - V'|}{|E_\Psi - z|^2|E' - z|} + \frac{2(1 + |z|)|E_\Psi - E'|}{|E_\Psi - z|^2|E' - z|} && (193)\\
&\leq \eta^3|V_\Psi - V'| + 2\eta^4|E_\Psi - E'| \ , && (194)
\end{aligned}$$

where we exploit $E_\Psi + E' \leq 2$, and $|E_\Psi - z|^2 - |E' - z|^2 \leq 2(1+|z|)|E_\Psi - E'| \leq 2\eta|E_\Psi - E'|$ because of the assumption $\|H\| \leq 1$.
Also, the second component is given by

$$\begin{aligned}
\frac{1}{\sqrt{V_\Psi}\sqrt{V' + |E' - z|^2}}\left|\sqrt{V_\Psi} - \sqrt{V'}\right| &\leq \frac{|V_\Psi - V'|}{\sqrt{V_\Psi}\sqrt{V' + |E' - z|^2}(\sqrt{V'} + \sqrt{V_\Psi})} && (195)\\
&\leq \frac{|V_\Psi - V'|}{V_\Psi|E' - z|} && (196)\\
&\leq \eta^3|V_\Psi - V'|. && (197)
\end{aligned}$$

Consequently, we have

$$\begin{aligned}
\frac{1}{\sqrt{V_\Psi}}|\sin(\overline{\alpha}) - \sin(\alpha)| &\leq \eta^3|V_\Psi - V'| + 2\eta^4|E_\Psi - E'| + \eta^3|V_\Psi - V'| && (198)\\
&\leq 2\eta^3|V_\Psi - V'| + 2\eta^4|E_\Psi - E'|. && (199)
\end{aligned}$$

Therefore, the upper bound of $|b(s) - b(\overline{s})|$ is expressed as

$$\begin{aligned}
|b(s) - b(\overline{s})| &\leq 2\eta^3|V_\Psi - V'| + 2\eta^4|E_\Psi - E'| + 2\eta^3|V_\Psi - V'| && (200)\\
&= 4\eta^3|V_\Psi - V'| + 2\eta^4|E_\Psi - E'|. && (201)
\end{aligned}$$

(b) Using similar procedure, we arrive at

$$\begin{aligned}
|a(s) - a(\overline{s})| &\leq E_\Psi|b(s) - b(\overline{s})| + |\cos(\sqrt{V_\Psi/V'}\alpha) - \cos(\alpha)| + |\cos(\overline{\alpha}) - \cos(\alpha)| && (202)\\
&\leq |b(s) - b(\overline{s})| + |2\sin((\sqrt{V_\Psi/V'} - 1)\alpha/2)\sin((\sqrt{V_\Psi/V'} + 1)\alpha/2)| \\
&\quad + |\cos(\overline{\alpha}) - \cos(\alpha)| && (203)\\
&\leq |b(s) - b(\overline{s})| + \frac{\pi}{2}\left|\frac{1}{\sqrt{V_\Psi}} - \frac{1}{\sqrt{V'}}\right||\cos(\overline{\alpha}) - \cos(\alpha)| \ . && (204)
\end{aligned}$$

For the last component, recalling $\cos\alpha = \dfrac{|E_\Psi - z|}{\sqrt{V_\Psi + |E_\Psi - z|^2}}$, we have

$$
\begin{aligned}
&|\cos(\overline{\alpha}) - \cos(\alpha)| \\
&\leq \left| \frac{|E_\Psi - z|}{\sqrt{V_\Psi + |E_\Psi - z|^2}} - \frac{|E' - z|}{\sqrt{V' + |E' - z|^2}} \right| && (205)\\
&\leq |E_\Psi - z| \left| \frac{1}{\sqrt{V_\Psi + |E_\Psi - z|^2}} - \frac{1}{\sqrt{V' + |E' - z|^2}} \right| + \frac{||E_\Psi - z| - |E' - z||}{\sqrt{V' + |E' - z|^2}} && (206)\\
&\leq |E_\Psi - z| \left| \frac{|V_\Psi - V'|}{|E_\Psi - z|^2 |E' - z|} + \frac{2(1+|z|)|E_\Psi - E'|}{|E_\Psi - z|^2 |E' - z|} \right| + \frac{|E_\Psi - E'|}{\sqrt{V'}} && (207)\\
&\leq \eta^2 |V_\Psi - V'| + 2\eta^3 |E_\Psi - E'| + \eta |E_\Psi - E'| && (208)\\
&= \eta^2 |V_\Psi - V'| + (2\eta^3 + \eta) |E_\Psi - E'| \,, && (209)
\end{aligned}
$$

where we use the result Eq. (193) in the third inequality. Hence, we obtain

$$
\begin{aligned}
|a(s) - a(\overline{s})| &\leq 4\eta^3 |V_\Psi - V'| + 2\eta^4 |E_\Psi - E'| + \frac{\pi}{2}\eta^3 |V_\Psi - V'| + \eta^2 |V_\Psi - V'| \\
&\quad + (2\eta^3 + \eta) |E_\Psi - E'| && (210)\\
&\leq (6\eta^3 + \eta^2) |V_\Psi - V'| + (2\eta^4 + 2\eta^3 + \eta) |E_\Psi - E'|. && (211)
\end{aligned}
$$

By substituting Eq. (201) and Eq. (211) into Eq. (178), we have

$$
\begin{aligned}
\left\| e^{\overline{s}[\Psi,H]} |\Psi\rangle - e^{s[\Psi,H]} |\Psi\rangle \right\| &\leq (6\eta^3 + \eta^2) |V_\Psi - V'| + (2\eta^4 + 2\eta^3 + \eta) |E_\Psi - E'| + 4\eta^3 |V_\Psi - V'| \\
&\quad + 2\eta^4 |E_\Psi - E'| && (212)\\
&= (10\eta^3 + \eta^2) |V_\Psi - V'| + (4\eta^4 + 2\eta^3 + \eta) |E_\Psi - E'| && (213)\\
&\leq 18\eta^4 \max(|V_\Psi - V'|, |E_\Psi - E'|) \,. && (214)
\end{aligned}
$$

where we use $10\eta^3 + \eta^2 \leq 11\eta^4$ and $2\eta^3 + \eta \leq 3\eta^4$ in the last line since $\eta \geq 1$ by definition. Lastly, combining Eq. (171) and Eq. (214) leads to the constants claimed above.

□

Finally, leveraging the techniques developed thus far, we establish a bound on the deviation between the ideal state and the state affected by noisy estimations of energy and variance in terms of their statistical errors.

Proposition C.5 (QSP estimation stability)**.** *Suppose H is a Hermitian matrix whose spectral radius does not exceed unity, i.e., $\|H\| \leq 1$ and consider the polynomial $p(H)$ of degree K with roots z_k satisfying $|z_k| \leq |z|$. We denote the ideal parameter sequences as $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_{K-1})$ and $\boldsymbol{s} = (|s_0|, \ldots, |s_{K-1}|)$, which yield the exact state $|\Psi_H(\boldsymbol{\theta}, \boldsymbol{s})\rangle$. Similarly, let $\overline{\boldsymbol{\theta}} = (\overline{\theta}_0, \ldots, \overline{\theta}_{K-1})$ and $\overline{\boldsymbol{s}} = (|\overline{s}_0|, \ldots, |\overline{s}_{K-1}|)$ be the parameters obtained from statistical estimates of the energy mean and variance, as used in the states defined in Eq. (157). Also, we define the maximal instability scale across all states as $\eta = \max(\frac{1}{\sqrt{V_\Psi}}, \frac{1}{\sqrt{V'}}, \frac{1}{|E_\Psi - z|}, \frac{1}{|E' - z|}, 1 + |z|)$. Then, we have*

$$
\| \, |\Psi_H(\boldsymbol{\theta}, \boldsymbol{s})\rangle - |\overline{\Psi}_H(\overline{\boldsymbol{\theta}}, \overline{\boldsymbol{s}})\rangle \, \| \leq (14 + 120\eta^4)^K \cdot \max(\delta_V, \delta_E) = O(\max(\delta_V, \delta_E)) \,. \tag{215}
$$

where $\delta_E \geq |E_k - \overline{E}_k|$ and $\delta_V \geq |V_k - \overline{V}_k|$ is the statistical errors of energy and variance.

Proof. Let us define the intermediate QSP states

$$
|\Psi_k\rangle = \prod_{k'=0}^{k-1} e^{i\theta_{k'}\Psi_{k'}} e^{s_{k'}[\Psi_{k'},H]} |\Psi_0\rangle \tag{216}
$$

and analogously $|\overline{\Psi}_k\rangle$ will be the intermediate states of QSP with $\overline{\theta}_k$ and $\overline{s}_k$. Thus, the difference between $|\Psi_{k+1}\rangle$ and $|\overline{\Psi}_{k+1}\rangle$ is given by

$$\| \, |\Psi_{k+1}\rangle - |\overline{\Psi}_{k+1}\rangle \, \| = \| e^{i\theta_k \Psi_k} e^{s_k[\Psi_k, H]} \, |\Psi_k\rangle - e^{i\overline{\theta}_k \overline{\Psi}_k} e^{\overline{s}_k[\overline{\Psi}_k, H]} \, |\overline{\Psi}_k\rangle \, \| \tag{217}$$

Again, following the same procedure in Eq. (109) from Sec. C.1, we add and subtract the term $\{e^{i\theta_k \Psi_k} e^{s_k[\Psi_k,H]} \, |\overline{\Psi}_k\rangle \, , \; e^{i\theta_k \overline{\Psi}_k} e^{s_k[\Psi_k,H]} \, |\overline{\Psi}_k\rangle \, , \; e^{i\theta_k \overline{\Psi}_k} e^{s_k[\overline{\Psi}_k,H]} \, |\overline{\Psi}_k\rangle\}, \; e^{i\overline{\theta}_k \overline{\Psi}_k} e^{s_k[\overline{\Psi}_k,H]} \, |\overline{\Psi}_k\rangle\}$ to split them into multiple norm calculation via triangle inequality. Therefore, the result is

$$\begin{aligned} \| \, |\Psi_{k+1}\rangle - |\overline{\Psi}_{k+1}\rangle \, \| &\leq \| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \| + \| e^{i\theta_k \Psi_k} - e^{i\theta_k \overline{\Psi}_k} \| + \| e^{s_k[\Psi_k,H]} - e^{s_k[\overline{\Psi}_k,H]} \| \\ &\quad + \| e^{i\theta_k \overline{\Psi}_k} - e^{i\overline{\theta}_k \overline{\Psi}_k} \| + \| e^{s_k[\overline{\Psi}_k,H]} \, |\overline{\Psi}_k\rangle - e^{\overline{s}_k[\overline{\Psi}_k,H]} \, |\overline{\Psi}_k\rangle \, \| \qquad (218) \\ &\leq \| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \| \\ &\quad + |\theta_k| \cdot \| \Psi_k - \overline{\Psi}_k \| \\ &\quad + |s_k| \cdot \| [\Psi_k - \overline{\Psi}_k, H] \| \\ &\quad + \| e^{i\theta_k \overline{\Psi}_k} - e^{i\overline{\theta}_k \overline{\Psi}_k} \| \\ &\quad + \| \left(e^{s_k[\overline{\Psi},H]} - e^{\overline{s}_k[\overline{\Psi},H]} \right) |\overline{\Psi}_k\rangle \, \| \; . \qquad (219) \end{aligned}$$

1. For the second term, we recall that $\| \Psi_k - \overline{\Psi}_k \| \leq 2 \| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \|$. Thus, we get

$$|\theta_k| \cdot \| \Psi_k - \overline{\Psi}_k \| \leq 2\pi \cdot 2 \| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \| \leq 13 \| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \|. \tag{220}$$

2. For the third term, since $|s_k| \leq \pi/2\sqrt{V_k} \leq 2\eta$, we have

$$|s_k| \cdot \| [\Psi_k - \overline{\Psi}_k, H] \| \leq 2\eta \cdot 2 \| \Psi_k - \overline{\Psi}_k \| \cdot \| H \| \leq 8\eta \, \| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \|. \tag{221}$$

3. By a similar consideration as in Eq. (171), we arrive at

$$\begin{aligned} \| e^{i\theta_k \overline{\Psi}_k} - e^{i\overline{\theta}_k \overline{\Psi}_k} \| &\leq \| e^{i\theta_k \overline{\Psi}_k} - e^{i\tilde{\theta}_k \overline{\Psi}_k} \| + \| e^{i\tilde{\theta}_k \overline{\Psi}_k} - e^{i\overline{\theta}_k \overline{\Psi}_k} \| \qquad (222) \\ &\leq 4\eta \| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \| + 2\eta\delta_E \; , \qquad (223) \end{aligned}$$

   where we used $|E_k - \tilde{E}_k| \leq 2 \| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \|$ to bound $|\theta_k - \tilde{\theta}_k|$.

4. Here, most steps proceeded as in the Prop. C.3, but the last remaining term needs a separate treatment as we do not a-priori have a bound on $|s_k - \overline{s}_k|$. To proceed with the bound, we use the exact expectation values $\tilde{E}_k$ and $\tilde{V}_k$ of the states $|\overline{\Psi}_k\rangle$ in Eq. (157) to introduce

$$\tilde{s}_k = s(\tilde{E}_k, \tilde{V}_k) = \frac{1}{\sqrt{\tilde{V}_k}} \arccos\left(\frac{\tilde{E}_k - u}{\sqrt{\tilde{V}_k + (\tilde{E}_k - u)^2}} \right) \; . \tag{224}$$

   Next, Prop. C.4 is used twice. Indeed, we use s_k, $\tilde{s}_k$ and $\overline{s}_k$ to denote different durations:

   - s_k: a time duration computed using the exact expectation values of the exact states $|\Psi_k\rangle$
   - $\tilde{s}_k$: a time duration computed using exact expectation values for states $|\overline{\Psi}_k\rangle$
   - $\overline{s}_k$: a time duration computed using the estimated expectation values of the states $|\overline{\Psi}_k\rangle$.

   Then, we have

$$\left\| \left(e^{s_k[\overline{\Psi},H]} - e^{\overline{s}_k[\overline{\Psi},H]} \right) |\overline{\Psi}_k\rangle \right\| \leq \left\| \left(e^{s_k[\overline{\Psi},H]} - e^{\tilde{s}_k[\overline{\Psi},H]} \right) |\overline{\Psi}_k\rangle \right\| + \left\| \left(e^{\tilde{s}_k[\overline{\Psi},H]} - e^{\overline{s}_k[\overline{\Psi},H]} \right) |\overline{\Psi}_k\rangle \right\| \; . \tag{225}$$

   We remark that both terms are already computed in Eq. (214) for the proof of Prop. C.4. Hence, we can compute them as follows.

(a) The first term is equal to Eq. (214) when V' (E') is replaced with $\tilde{V}_k$ ($\tilde{E}_k$), i.e.,

$$\left\| \left(e^{s_k[\overline{\Psi},H]} - e^{\tilde{s}_k[\overline{\Psi},H]} \right) |\overline{\Psi}_k\rangle \right\| \leq 18\eta^4 \cdot \left(\max(|V_k - \tilde{V}_k|, |E_k - \tilde{E}_k|) \right) . \tag{226}$$

Thus, the remaining task is to account the differences of the exact expectation values and variances, which are given by

$$|E_k - \tilde{E}_k| \leq 2\| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \|, \tag{227}$$

$$|V_k - \tilde{V}_k| \leq |\, \langle \Psi_k| \, H^2 \, |\Psi_k\rangle - \langle \overline{\Psi}_k| \, H^2 \, |\overline{\Psi}_k\rangle \,| + |E_k^2 - \tilde{E}_k^2| \leq 6\| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \|. \tag{228}$$

Using this observation, we have

$$\left\| \left(e^{s_k[\overline{\Psi},H]} - e^{\tilde{s}_k[\overline{\Psi},H]} \right) |\overline{\Psi}_k\rangle \right\| \leq 18\eta^4 \cdot \left(6\| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \| \right) = 108\eta^4 \| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \| \, . \tag{229}$$

(b) For the second term, the gaps in the energy and variance arise from the inaccurate estimation caused by statistical noise. Thus we have

$$\left\| \left(e^{\tilde{s}_k[\overline{\Psi},H]} - e^{\tilde{s}_k[\overline{\Psi},H]} \right) |\overline{\Psi}_k\rangle \right\| \leq 18\eta^4 \max(\delta_V, \delta_E) \, . \tag{230}$$

Overall, by substituting the above calculations into Eq. (219), we have

$$\| \, |\Psi_{k+1}\rangle - |\overline{\Psi}_{k+1}\rangle \, \| \leq (14 + 120\eta^4) \| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \| + 20\eta^4 \max(\delta_V, \delta_E) \, . \tag{231}$$

Thus, using these bounds inductively, we get

$$\| \, |\Psi_k\rangle - |\overline{\Psi}_k\rangle \, \| \leq 20\eta^4 \max(\delta_V, \delta_E) \sum_{i=0}^{k-1} (14 + 120\eta^4)^i \tag{232}$$

$$= 20\eta^4 \max(\delta_V, \delta_E) \frac{1 - (14 + 120\eta^4)^k}{1 - (14 + 120\eta^4)} \tag{233}$$

$$\leq (14 + 120\eta^4)^k \max(\delta_V, \delta_E) \, . \tag{234}$$

By setting $k = K$, we get the claimed scaling. □

D Applications of DB-QSP

In this section, we explore the potential applications of DB-QSP. As noted in the main text, a limitation of DB-QSP is that only low-degree polynomials is feasible. Therefore, we first provide useful approximation techniques that can expand its applicability. We then discuss applicability of DB-QSP for other tasks.

D.1 Examples of Low-Degree Polynomial Approximations

To assess the effectiveness of DB-QSP, it is crucial to understand what kind of functions can be realized using low-degree polynomials. In QSP with post-selection, this issue is directly linked to the success probability of post-selection, as it depends on the degree of polynomials as well as an input state $|\Psi_0\rangle$. Importantly, several polynomial approximation techniques have been explored in the literature [2, 3]. To illustrate that DB-QSP can achieve ϵ-precision while maintaining a logarithmically small polynomial degree, we present three representative examples below.

We first show an approximation of the sign function, which can be used for the ground-state preparation task [9].

Example D.1 (Approximation of the sign function $\mathrm{sgn}(x)$ [9, 2])**.** *Suppose $\delta > 0$, $x \in \mathbb{R}$ and $\epsilon \in (0, 1/2)$. Given a degree $K = \mathcal{O}(\log(1/\epsilon)/\delta)$, there exists an odd polynomial $p(x)$, such that*

- *for all* $x \in [-2,2]$*:* $|p(x)| \leq 1$ *and*
- *for all* $x \in [-2,2]\backslash(-\delta,\delta)$*:* $|p(x) - sgn(x)| \leq \epsilon$

where

$$sgn(x) = \begin{cases} 1 & \text{if } x > 0, \\ -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0. \end{cases} \tag{235}$$

As another polynomial function for filtering in the ground-state preparation task, we also demonstrate the approximation of trigonometric functions.

Example D.2 (Polynomial approximation of trigonometric functions by Jacobi-Anger expansion [3, 2])**.** *Suppose* $s \in \mathbb{R}$ *and* $\epsilon \in (0, \frac{1}{e})$*. Given a degree* $K = \lfloor \frac{1}{2} r\left(\frac{e|s|}{2}, \frac{5}{4}\epsilon\right) \rfloor$*, trigonometric functions can be approximated as follows;*

$$\| \cos(sx) - J_0(s) + 2\sum_{l=1}^{K} (-1)^l J_{2l}(t) T_{2l}(x) \|_{[-1,1]} \leq \epsilon, \tag{236}$$

$$\| \sin(sx) - 2\sum_{l=1}^{K} (-1)^l J_{2l+1}(t) T_{2l+1}(x) \|_{[-1,1]} \leq \epsilon, \tag{237}$$

where $J_m(s)$ *is the Bessel functions of the first kind and* $T_m(x)$ *is the Chebyshev polynomials of the first kind. Also,* $r(t,\epsilon)$ *is a function that asymptotically scales as*

$$r(t,\epsilon) = \Theta\left(|t| + \frac{\log(1/\epsilon)}{\log(e + \frac{\log(1/\epsilon)}{|t|})} \right). \tag{238}$$

This indicates that trigonometric functions can be approximated using a polynomials of degree $d = \lfloor \frac{1}{2} r\left(\frac{e|s|}{2}, \frac{5}{4}\epsilon\right) \rfloor$ to achieve ϵ-precision.

Next, we move to an polynomial approximation for matrix inversion used for, e.g., solving linear system of equations.

Example D.3 (Polynomial approximation of the inverse [67, 3])**.** *Suppose* $\kappa > 1$*,* $\epsilon \in (0, \frac{1}{2})$ *and* $x \in [-1,1]\backslash(-\frac{1}{\kappa}), \frac{1}{\kappa}$*. Then*

$$f(x) = \frac{1 - (1 - x^2)^a}{x} \tag{239}$$

is an odd function with $a = \lceil \kappa^2 \log(\kappa/\epsilon) \rceil$ *is* ϵ*-close to the inverse* $1/x$*. Given a degree* $K = \lceil \sqrt{a \log(4a/\epsilon)} \rceil = \mathcal{O}(\kappa \log(\kappa/\epsilon))$*, the odd real function*

$$g(x) = 4\sum_{l=1}^{K} (-1)^l \left[\frac{\sum_{j=l+1}^{a} \binom{2a}{a+j}}{2^{2a}} \right] T_{2l+1}(x) \tag{240}$$

is ϵ*-close to* $f(x)$ *on the domain* $[-1,1]$*.*

This indicates that DB-QSP has the potential to efficiently perform matrix inversion in terms of the inverse precision $1/\epsilon$. However, the circuit depth required for DB-QSP scales super exponentially with the condition number, a key factor in assessing the algorithm's efficiency. Thus, this example also highlights a fundamental challenge for DB-QSP in certain computational tasks.

With these approximations, our approach could circumvent the exponential costs for some cases. Furthermore, approximations for other functions have been provided, e.g., in Ref. [7]. This suggests that certain tasks benefiting from these polynomials can also be performed using DB-QSP. Thus, DB-QSP remains practically viable in some cases.

D.2 Derivation of DB-QITE Using DB-QSP

As discussed in the main text, DB-QSP can be utilized to implement Imaginary-Time Evolution (ITE), which is a key technique for ground-state preparation. The non-unitary operator in imaginary time evolution, $e^{-\tau H}$, can be approximated up to the first order as $p(H) = I - \tau H$. Thus, Lem. 1 immediately suggests that DB-QSP can be used to approximate the imaginary time evolution. Then, by using the group commutator

$$e^{s_0[\Psi,H]} = e^{i\sqrt{s_0}H} e^{i\sqrt{s_0}\Psi} e^{-i\sqrt{s_0}H} e^{-i\sqrt{s_0}\Psi} + O(s_0^{3/2}) \tag{241}$$

and noticing that the last unitary has a trivial action on $|\Psi\rangle$, we arrive at the proposal in Ref. [18]

$$|\omega_{k+1}\rangle = e^{i\sqrt{s_0}H} e^{i\sqrt{s_0}\omega_k} e^{-i\sqrt{s_0}H} |\omega_k\rangle \ . \tag{242}$$

This is essentially the same as choosing $N = 1$ in DB-QSP with $\theta_k = 0$.

One subtlety is that DB-QSP suggests to use state-dependent scheduling s_k. Ref. [18] proved that using a sufficiently small constant s_0 allows to converge to the ground state and in every step have a cooling rate matching imaginary time evolution. The iterative use of Thm. 1 shows that this quantum algorithm can be devised based on QSP as the design approach.

D.3 Hamiltonian Simulation

Next, we move onto the Hamiltonian simulation task, where the goal is to implement the real-time evolution e^{-itH}. To perform this task, a common assumption in QSP implementation is direct access to a subroutine which applies the input Hermitian matrix H to an input state; see App. A. One approach to Hamiltonian simulation is to approximate the evolution operator using a Taylor series expansion

$$p_{\mathrm{HS}}(H) = \sum_{k=0}^{K} \frac{(-it)^n}{n!} H^n \ , \tag{243}$$

which allows the Hamiltonian evolution to be approximated via polynomial transformations.

At first glance, this polynomial decomposition suggests that DB-QSP might also be applicable to Hamiltonian simulation. However, DB-QSP is not designed for this task. Since Alg. 1 assumes direct access to e^{itH}, using DB-QSP for Hamiltonian simulation would be vacuous. This query model is also known as the Hamiltonian evolution model and has been widely used in tasks such as ground-state energy estimation with early fault-tolerant quantum computers [68, 69, 70]. From this perspective, the fact that DB-QSP does not target Hamiltonian simulation is not a limitation but rather an inherent feature of the query model.

D.4 Evolution under a Polynomial Function of Hamiltonian

Interestingly, though, DB-QSP can be used to effectively transform Hamiltonians, while working in the Hamiltonian evolution model. In its simplest variant, we aim to implement the Hamiltonian simulation of H^2, the second power of the input matrix. Strikingly, DB-QSP can be applied to this scenario by interpreting $p_{\mathrm{HS}}(H^2)$ as a polynomial of doubled degree in the variable H, which leads to the factorization

$$p_{\mathrm{HS}}(H^2) = a_{2K} \prod_{k=1}^{2K} (H - \sqrt{z_k} I)(H + \sqrt{z_k} I) \tag{244}$$

in contrast to the alternative formulation

$$p_{\mathrm{HS}}(H^2) = a_K \prod_{k=1}^{K} (H^2 - z_k I), \tag{245}$$

which treats H^2 as the primary variable. More generally, for evolution under $e^{itg(H)}$ where g is a polynomial, we observe that if $h = p \circ g$ is also a polynomial, then $p_{\mathrm{HS}}(g(H))$ can be factorized accordingly, allowing us to proceed in an analogous manner.

The possibility to systematically use the Hamiltonian simulation e^{-itH} to simulate e^{-itH^2} is implied by classic results in Lie group theory [71], but an explicit construction of the type provided by DB-QSP is new to our knowledge. In particular, DB-QSP could provide a convergence rate and a circuit lower bound to a large class of instances of this classic question. Finally, we remark that Thm. 2 is required in this case, because $p_{\mathrm{HS}}(H) = I - iH - H^2/2$ has complex roots $z_\pm = \pm 1 - i$ for $K = 2$ and $t = 1$.

D.5 Laurent Polynomials

Another application is the Laurent polynomials, which include terms with negative powers, i.e.,

$$p_{\mathrm{L}}(H) = \sum_{k=-K}^{K} a_k H^k \ . \tag{246}$$

While it is useful for QSP to consider "polynomials" involving inverse powers, Thm. 2 does not directly provide unitary synthesis for this type. Yet, assuming the matrix H satisfies $\|I - H\| < 1$, we can consider $H^{-K} \approx p_{\mathrm{INV}}(H)^K$ and thus write

$$p_{\mathrm{DBL}}(H) = p_{\mathrm{INV}}(H)^K \sum_{k=0}^{K} a_{k-K} H^k \ . \tag{247}$$

This gives the approximation $p_{\mathrm{DBL}}(H) \approx p_{\mathrm{L}}(H)$ and provides an example how one can implement Laurent polynomials using DB-QITE. The general case for Laurent polynomials with arbitrary Hermitian matrices is left for future work.

E Classically-Aided DB-QSP Synthesis

Statistical error is unavoidable when estimating energy mean and variance on quantum hardware, because of the finite number of measurement shots. Error analysis in Eq. (33) further suggests that this issue becomes more pronounced as the polynomial degree increases. This suggests the need for an approach to circumvent this challenge. One potential solution is to leverage classical computation in initial steps. Motivated by this, we explore conditions under which energy and variance can be efficiently computed using classical resources.

Assume that the initial state $|\Psi_0\rangle$ is expressed in a basis where only m of its components are nonzero. For instance, if the initial state is a tensor-product of zero states, i.e., $|0\rangle^{\otimes L}$ with L-qubits, then $m = 1$. Additionally, suppose the target Hermitian matrix is given by $H = \sum_{i=1}^{J} w_i P_i$ with Pauli operators $\{P_i\}$. Using Eq. (11) together with the effect of state-dependent phase gate, the resultant state can always be written as

$$|\Psi_k\rangle = \left(\prod_{j=1}^{k} (a'(s_j) I + b(s_j) H) \right) |\Psi_0\rangle \tag{248}$$

with $a'(s_j) \in \mathbb{C}$ and $b(s_j) \in \mathbb{R}$. Note that the coefficients $a'(s_j)$ and $b(s_j)$ are determined by the energy and variance of the state at $(k-1)$-th step. Our goal is to estimate the energy $O = H$ and the variance $O = (H - \langle H \rangle)^2$. By substituting Eq. (248) into the expectation value $\langle \Psi_k | O | \Psi_k \rangle$, we obtain

$$\langle \Psi_k | H | \Psi_k \rangle = \sum_{l=1}^{2k+1} \xi_l \langle \Psi_0 | H^l | \Psi_0 \rangle \, , \tag{249}$$

$$\langle \Psi_k | (H - \langle H \rangle)^2 | \Psi_k \rangle = \sum_{l=0}^{2k+2} \xi'_l \langle \Psi_0 | H^l | \Psi_0 \rangle \, . \tag{250}$$

where coefficients $\{\xi_l\}$ and $\{\xi'_l\}$ are determined by computing Eq. (248). Implication of Eqs. (249), (250) is that, if we can compute $\langle\Psi_0|H^l|\Psi_0\rangle$ up to $l = 2k+2$ classically, the energy and variance at k step is tractable using classical computers.

With these criteria in mind, we analyze the conditions under which the classical computation of energy and variance is feasible. First, the number of Pauli operators in H^{2k+2} is at most $J^{2k+2}+1$. Furthermore, since each Pauli operator has exactly one nonzero entry per row and column, the total number of nonzero elements that need to be stored scales as $\mathcal{O}(mJ^{2k+2})$. To ensure classical tractability in terms of memory and computational cost, we require this scaling to remain within $\mathcal{O}(\mathrm{poly}(n))$. Thus, the condition on k for energy and variance to be classically computable is given by

$$\begin{aligned} & m^2 J^{2k+2} \leq \mathrm{poly}(n) \\ \Leftrightarrow & k \leq \frac{\log\left(\mathrm{poly}(n)\right)/2 - \log(m)}{\log\left(J\right)} - 1 \\ \Leftrightarrow & k = \mathcal{O}\left(\frac{\log\left(\mathrm{poly}(n)\right)/2 - \log(m)}{\log\left(J\right)}\right). \end{aligned}$$

This indicates that, if $m, J = \mathcal{O}(1)$, we can classically compute up to $k = \mathcal{O}(\log(n))$. On the other hand, computing only constant step k is possible if $m, J = \mathcal{O}(\mathrm{poly}(n))$ This clearly captures the classical difficulty: if the initial state contains many non-zero elements and the number of Pauli terms becomes prohibitively large, it becomes infeasible to compute the energy even for a single step. However, for instance, the Pauli terms for Ising models scale linearly in the number of qubits. Furthermore, some situations involve easy-to-prepare initial states like the tensor product of zero states, where $m = 1$. Thus, this result suggests that a few steps of classical computation may be feasible in some cases. We also note that this estimation is straightforward, and advanced classical techniques could further improve the efficiency, which we will leave for future work.

F Unbiased Estimator of the Operator Variance for Hamiltonians

In this section, we derive an unbiased estimator for the variance of an observable expressed as a weighted sum of Pauli operators. We first describe the measurement procedure used to estimate the expectation values of individual Pauli operators and their products. Next, we construct a straightforward variance estimator and demonstrate its bias arising from finite-sample effects. To address this, we derive a corrected formula that provides an unbiased estimate of the operator variance.

F.1 Measurement Procedure

Here, we focus on the observables $\hat{O}$ that can be decomposed as the weighted sum of Pauli operators, i.e., the observable can be expressed in the form

$$\hat{O} = \sum_{i=1}^{L} w_i P_i \ , \tag{251}$$

where $P_i \in \{I, X, Y, Z\}^{\otimes n}$ denotes the Pauli operators for n qubits and w_i represents its corresponding weights.

Then, the variance of the observable $\hat{O}$ is defined as $V = \langle\hat{O}^2\rangle - \langle\hat{O}\rangle^2$, where $\langle\cdot\rangle = \langle\psi|\cdot|\psi\rangle$ for a pure quantum state $|\psi\rangle$. Thus, the square of the observable is given by

$$\hat{O}^2 = \left(\sum_{i=1}^{L} w_i P_i\right)\left(\sum_{j=1}^{L} w_j P_j\right) = \sum_{i=1}^{L} w_i^2 I + \sum_{\substack{i,j=1 \\ i\neq j}}^{L} w_i w_j P_{ij} \ , \tag{252}$$

where we use the Pauli operators's identity $P_i^2 = I$ and we introduce the notation $P_{ij} = P_i P_j$. Using the expression of $\hat{O}$ and $\hat{O}^2$, the variance of the observable is given by

$$V = \langle \hat{O}^2 \rangle - \langle \hat{O} \rangle^2 = \sum_{i=1}^{L} w_i^2 \langle I \rangle + \sum_{\substack{i,j=1 \\ i \neq j}}^{L} w_i w_j \langle P_{ij} \rangle - \left(\sum_{i=1}^{L} w_i \langle P_i \rangle \right)^2 . \tag{253}$$

To estimate the variance, we measure each Pauli component multiple times. Using the measurement outcomes, we can then estimate the expectation value of each term in Eq. (253). The procedure is:

1. $\overline{\langle I \rangle} = 1$ by the assumption of normalised state.

2. Suppose we measure P_i a total of N_i times, yielding outcomes (a set of measured bit-strings)

$$\{b_1^{(P_i)}, b_2^{(P_i)}, \ldots, b_{N_i}^{(P_i)}\} \quad \text{with each} \quad b_k \in \{-1, +1\} \quad \text{for } 1 \leq k \leq N_i \ . \tag{254}$$

   Then the estimation of a single Pauli operator P_i is $\overline{\langle P_i \rangle} = \dfrac{1}{N_i} \displaystyle\sum_{k=1}^{N_i} b_k^{(P_i)}$.

3. Similarly, if we measure the product operator $P_{ij} = P_i P_j$ (for $i \neq j$) N_{ij} times, we obtain the estimator

$$\overline{\langle P_i P_j \rangle} = \frac{1}{N_{ij}} \sum_{k=1}^{N_{ij}} b_k^{(P_{ij})}. \tag{255}$$

F.2 Biased and Unbiased Estimator of the Operator Variance

F.2.1 Biased and Unbiased Estimator

First, we mention the definition of biased and unbiased estimator.

Definition F.1 (Estimator)**.** *Let $\overline{A}$ be an estimator of the parameter A. The estimator is said to be:*

- ***unbiased*** *if* $\mathbb{E}[\overline{A}] = A$,
- ***biased*** *if* $\mathbb{E}[\overline{A}] \neq A$,

where we use $\mathbb{E}$ to denote the expected value over the sampling process in this section.

Using the statistics, a natural choice of the estimator of Eq. (253) would be

$$\tilde{V} = \sum_{i=1}^{L} w_i^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{L} w_i w_j \, \overline{\langle P_i P_j \rangle} - \left(\sum_{i=1}^{L} w_i \, \overline{\langle P_i \rangle} \right)^2 \tag{256}$$

$$= \sum_{i=1}^{L} w_i^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{L} w_i w_j \left(\frac{1}{N_{ij}} \sum_{k=1}^{N_{ij}} b_k^{(P_{ij})} \right) - \left[\sum_{i=1}^{L} w_i \left(\frac{1}{N_i} \sum_{k=1}^{N_i} b_k^{(P_i)} \right) \right]^2 . \tag{257}$$

However, we demonstrate a simple example where the last term of this estimator introduces bias into the estimator.

Example F.1. *Let $P \in \{I, X, Y, Z\}^{\otimes n}$ be a Pauli operator and suppose that we estimate its expectation value by performing N independent and identically distributed (i.i.d.) measurements, yielding*

outcomes $\{b_i\}_{i=1}^N$ *(with* $b_i \in \{-1,+1\}$*). Using our construction, the natural estimator for* $\langle P\rangle^2$ *is* $\langle \tilde{P}\rangle^2 := \left(\frac{1}{N}\sum_{i=1}^N b_i\right)^2$*, and its expectation value yields*

$$\mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^N b_i\right)^2\right] = \frac{1}{N^2}\left(\sum_{i=1}^N \mathbb{E}[b_i^2] + \sum_{\substack{i,j=1\\ i\neq j}}^N \mathbb{E}[b_i b_j]\right) = \frac{1}{N^2}\left(\sum_{i=1}^N 1 + \sum_{\substack{i,j=1\\ i\neq j}}^N \langle P_i\rangle^2\right) \tag{258}$$

$$= \frac{1}{N^2}\left(N + N(N-1)\langle P\rangle^2\right) = \langle P\rangle^2 + \frac{1-\langle P\rangle^2}{N}\ . \tag{259}$$

where we use the i.i.d. assumption ($\mathbb{E}[b_i b_j] = \mathbb{E}[b_i]\,\mathbb{E}[b_j]$ *for* $i \neq j$*) and the relation* $b_i^2 = 1$ *in the second line. Clearly,* $\mathbb{E}\left[\widetilde{\langle P\rangle^2}\right] \neq \langle P\rangle^2$ *for any finite sample size* N*, and hence it is a biased estimator by the definition. To remove this bias, we introduce a correction factor and define the unbiased estimator as*

$$\overline{\langle P\rangle^2} := \frac{N}{N-1}\left[\left(\frac{1}{N}\sum_{i=1}^N b_i\right)^2 - \frac{1}{N}\right]. \tag{260}$$

Directly evaluating the expectation value of new estimator of $\overline{\langle P\rangle^2}$ *yields*

$$\mathbb{E}\left[\frac{N}{N-1}\left[\left(\frac{1}{N}\sum_{i=1}^N b_i\right)^2 - \frac{1}{N}\right]\right] = \frac{N}{N-1}\left[\langle P\rangle^2 + \frac{1-\langle P\rangle^2}{N} - \frac{1}{N}\right] = \langle P\rangle^2\ . \tag{261}$$

Thus, this is indeed an unbiased estimator for $\langle P\rangle^2$ *as the expectation value of the estimator is consistent with the true value.*

Proposition F.2 (Unbiased estimator for the variance of an observable)**.** *Consider an observable* $\hat{O}$ *which can be written as* $\hat{O} = \sum_{i=1}^L w_i P_i$*, where* $P_i \in \{I,X,Y,Z\}^{\otimes n}$ *denotes the Pauli operators for* n *qubits and* w_i *represents its corresponding weights. The unbiased estimator for the variance of this observable is then given by*

$$\overline{V} = \sum_{i=1}^L w_i^2 + \sum_{\substack{i,j=1\\ i\neq j}}^L w_i w_j \left(\frac{1}{N_{ij}}\sum_{k=1}^{N_{ij}} b_k^{(P_{ij})}\right)$$

$$- \sum_{i=1}^L \frac{w_i^2 N_i}{N_i - 1}\left[\left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)^2 - \frac{1}{N_i}\right] - \sum_{\substack{i,j=1\\ i\neq j}}^L w_i w_j \left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)\left(\frac{1}{N_j}\sum_{k=1}^{N_j} b_k^{(P_j)}\right)\ , \tag{262}$$

where we measure the operator P_i *a total of* N_i *times and the product operator* $P_{ij} = P_i P_j$ *(for* $i \neq j$*) a total of* N_{ij} *times.*

Proof. The expected value of $\overline{V}$ is

$$\mathbb{E}\left[\overline{V}\right] = \sum_{i=1}^L w_i^2 + \sum_{\substack{i,j=1\\ i\neq j}}^L w_i w_j\ \mathbb{E}\left[\left(\frac{1}{N_{ij}}\sum_{k=1}^{N_{ij}} b_k^{(P_{ij})}\right)\right]$$

$$- \sum_{i=1}^L \frac{w_i^2 N_i}{N_i - 1}\mathbb{E}\left[\left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)^2 - \frac{1}{N_i}\right] - \sum_{\substack{i,j=1\\ i\neq j}}^L w_i w_j\ \mathbb{E}\left[\left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)\left(\frac{1}{N_j}\sum_{k=1}^{N_j} b_k^{(P_j)}\right)\right]\ , \tag{263}$$

where we use the fact that the expected value of a constant is the same constant for the first term. Next, we address the remaining terms separately.

1. Since the term $\frac{1}{N_{ij}}\sum_{k=1}^{N_{ij}} b_k^{(P_{ij})}$ is unbiased estimator for $\langle P_{ij}\rangle$, we have

$$\sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j \,\mathbb{E}\left[\left(\frac{1}{N_{ij}}\sum_{k=1}^{N_{ij}} b_k^{(P_{ij})}\right)\right] = \sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j \,\langle P_{ij}\rangle \ . \tag{264}$$

2. For the third term, we have

$$\sum_{i=1}^{L} \frac{w_i^2 N_i}{N_i-1}\mathbb{E}\left[\left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)^2 - \frac{1}{N_i}\right] = \sum_{i=1}^{L} \frac{w_i^2 N_i}{N_i-1}\left\{\mathbb{E}\left[\left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)^2\right] - \frac{1}{N_i}\right\} \tag{265}$$

$$= \sum_{i=1}^{L} \frac{w_i^2 N_i}{N_i-1}\left\{\langle P_i\rangle^2 + \frac{1-\langle P_i\rangle^2}{N_i} - \frac{1}{N_i}\right\} = \sum_{i=1}^{L} w_i^2 \langle P_i\rangle^2 \ , \tag{266}$$

   where we use Eq. (279) of Lem. F.4 in the second line.

3. For the last term, since the samples for different indices ($i \neq j$) are i.i.d., we have

$$\sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j \,\mathbb{E}\left[\left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)\left(\frac{1}{N_j}\sum_{k=1}^{N_j} b_k^{(P_j)}\right)\right] = \sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j \,\mathbb{E}\left[\left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)\right]\mathbb{E}\left[\left(\frac{1}{N_j}\sum_{k=1}^{N_j} b_k^{(P_j)}\right)\right] \tag{267}$$

$$= \sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j \,\langle P_i\rangle\,\langle P_j\rangle \ . \tag{268}$$

Collecting all the terms, the expected value of $\overline{V}$ becomes

$$\mathbb{E}\left[\overline{V}\right] = \sum_{i=1}^{L} w_i^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j \,\langle P_i P_j\rangle - \sum_{i=1}^{L} w_i^2 \,\langle P_i\rangle^2 - \sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j \,\langle P_i\rangle\,\langle P_j\rangle \tag{269}$$

$$= \sum_{i=1}^{L} w_i^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j \,\langle P_{ij}\rangle - \left(\sum_{i=1}^{L} w_i \,\langle P_i\rangle\right)^2 \tag{270}$$

Since $\mathbb{E}\left[\overline{V}\right] = V$ by using Eq. (253), $\overline{V}$ is indeed an unbiased estimator. □

F.3 Total Variance of the Unbiased Estimator of the Operator Variance

Next, with the motivation to assess the query complexity in DB-QSP, we compute the uncertainty of the unbiased estimator of the operator variance. Here, we consider variance as the uncertainty metric, which is given by:

$$\mathrm{Var}[\overline{V}] = \mathrm{Var}\left[\overline{\langle\hat{O}^2\rangle} - \overline{\langle\hat{O}\rangle}^2\right] = \mathrm{Var}\left[\overline{\langle\hat{O}^2\rangle}\right] + \mathrm{Var}\left[\overline{\langle\hat{O}\rangle}^2\right] - 2\,\mathrm{Cov}\left[\overline{\langle\hat{O}^2\rangle}, \overline{\langle\hat{O}\rangle}^2\right], \tag{271}$$

where we use the identity $\mathrm{Var}[A+B] = \mathrm{Var}[A]+\mathrm{Var}[B]+2\mathrm{Cov}(A,B)$. The estimator for $\overline{\langle\hat{O}^2\rangle}$ and $\overline{\langle\hat{O}\rangle}^2$ are determined by the measurement on the Pauli operators P_{ij} and P_k respectively. Assuming that the measurements on P_{ij} and P_k are independent, their covariance is zero, i.e., $\mathrm{Cov}\left[\overline{\langle\hat{O}^2\rangle}, \overline{\langle\hat{O}\rangle}^2\right] = 0$.

Therefore, the remaining task is to evaluate $\mathrm{Var}\left[\overline{\langle\hat{O}^2\rangle}\right]$ and $\mathrm{Var}\left[\overline{\langle\hat{O}\rangle^2}\right]$. We address these two terms in Lem. F.3 and Lem. F.5.

Lemma F.3. *Suppose we have an observable $\hat{O}$ which is of the form $\hat{O} = \sum_{i=1}^{L} w_i P_i$, where $P_i \in \{I, X, Y, Z\}^{\otimes n}$ denotes the Pauli operators for n qubits and w_i represents its corresponding weights. Assuming that measurements performed for all operators are i.i.d., then the uncertainty (variance) of the estimation of $\hat{O}^2$ can be expressed as*

$$Var\left[\overline{\langle\hat{O}^2\rangle}\right] = \sum_{\substack{i,j=1\\ i\neq j}}^{L} \frac{w_i^2 w_j^2}{N_{ij}} \left(1 - \langle P_{ij}\rangle^2\right) \ , \tag{272}$$

where we define $\langle P_{ij}\rangle = \langle P_i\rangle \langle P_j\rangle$ and N_{ij} is the number of sample used to estimate $\langle P_{ij}\rangle$.

Proof. We start with the expression

$$\mathrm{Var}\left[\overline{\langle\hat{O}^2\rangle}\right] = \mathrm{Var}\left[\sum_{i=1}^{L} w_i^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j \left(\frac{1}{N_{ij}} \sum_{k=1}^{N_{ij}} b_k^{(P_{ij})}\right)\right] . \tag{273}$$

Since the first term $\sum_{i=1}^{L} w_i^2$ is a constant, its variance is zero. Therefore, we have

$$\mathrm{Var}\left[\overline{\langle\hat{O}^2\rangle}\right] = \mathrm{Var}\left[\sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j \left(\frac{1}{N_{ij}} \sum_{k=1}^{N_{ij}} b_k^{(P_{ij})}\right)\right] . \tag{274}$$

Assuming that the contributions from different pairs of Pauli operators (i, j) are independent, it reduces to

$$\mathrm{Var}\left[\overline{\langle\hat{O}^2\rangle}\right] = \sum_{\substack{i,j=1\\ i\neq j}}^{L} \mathrm{Var}\left[\frac{w_i w_j}{N_{ij}} \sum_{k=1}^{N_{ij}} b^{(P_{ij})}\right] = \sum_{\substack{i,j=1\\ i\neq j}}^{L} \frac{w_i^2 w_j^2}{N_{ij}^2} \mathrm{Var}\left[\sum_{i=1}^{N_{ij}} b_i^{(P_{ij})}\right] \ , \tag{275}$$

where we use the property $\mathrm{Var}[aX] = a^2\,\mathrm{Var}[X]$ (for any nonnegative constant a) in the last line. Furthermore, assuming that the $b_i^{(P_{ij})}$ are i.i.d., it can be further simplified to

$$\mathrm{Var}\left[\overline{\langle\hat{O}^2\rangle}\right] = \sum_{\substack{i,j=1\\ i\neq j}}^{L} \frac{w_i^2 w_j^2}{N_{ij}^2} \left(N_{ij}\,\mathrm{Var}\left[b_i^{(P_{ij})}\right]\right) \ . \tag{276}$$

Next, recall that each b_i satisfies $b_i^2 = 1$, and hence we obtain the following relation

$$\mathrm{Var}\left[b_i^{(P_{ij})}\right] = \langle b_i^{(P_{ij})} \times b_i^{(P_{ij})}\rangle - \langle b_i^{(P_{ij})}\rangle^2 = 1 - \langle b_i^{(P_{ij})}\rangle^2 = 1 - \langle P_{ij}\rangle^2 \ . \tag{277}$$

Combining Eq. (276) and Eq. (277) yields Eq. (272). □

Before proceeding to Lem. F.5, let us first show a technical Lem. F.4, which will be useful in the proof of Lem. F.5.

Lemma F.4. *Given a single Pauli operator of n qubits $P_i \in \{I, X, Y, Z\}^{\otimes n}$, we estimate the expectation value of P_i as*

$$X_i = \overline{\langle P_i\rangle} = \frac{1}{N_i} \sum_{k=1}^{N_i} b_k^{(P_i)} \ . \tag{278}$$

where $b_k^{(P_i)}$ denotes the outcome of k-th measurement for the Pauli operator P_i. Assuming the measurements are i.i.d., we obtain

1. *the first moment of* X_i *as* $\mathbb{E}[X_i] = \overline{\langle P_i \rangle}$.

2. *the second moment of* X_i *as*

$$\mathbb{E}\left[X_i^2\right] = \langle P_i \rangle^2 + \frac{1 - \langle P_i \rangle^2}{N_i} \ . \tag{279}$$

3. *the third moment of* X_i *as*

$$\mathbb{E}\left[X_i^3\right] = \langle P_i \rangle^3 \left(1 - \frac{3}{N_i} + \frac{2}{N_i^2}\right) + \langle P_i \rangle \left(\frac{3}{N_i} - \frac{2}{N_i^2}\right) \ . \tag{280}$$

4. *the fourth moment of* X_i *as*

$$\mathbb{E}\left[X_i^4\right] = \langle P_i \rangle^4 + \frac{6\langle P_i \rangle^2 \left(1 - \langle P_i \rangle^2\right)}{N_i} + \frac{\left(11\langle P_i \rangle^2 - 3\right)\left(\langle P_i \rangle^2 - 1\right)}{N_i^2} + \frac{2(3\langle P_i \rangle^2 - 1)(1 - \langle P_i \rangle^2)}{N_i^3} \ . \tag{281}$$

where $\langle P_i \rangle$ *denotes the true expectation value of* P_i.

Proof. To simplify the notation, we use b_i to represent $b_i^{(P_i)}$ throughout this proof.

1. ***First moment of*** X_i***:***

   Since the expectation value of a scalar is just the scalar itself, we obtain $\mathbb{E}[X_i] = \mathbb{E}\left[\overline{\langle P_i \rangle}\right] = \overline{\langle P_i \rangle}$.

2. ***Second moment of*** X_i***:***

   Taking the expectation value of it yields

$$\mathbb{E}\Big[X_i^2\Big] = \mathbb{E}\left[\left(\frac{1}{N_i}\sum_{i=1}^{N_i} b_i\right)^2\right] = \frac{1}{N_i^2}\left(\sum_{i=1}^{N_i}\mathbb{E}[b_i^2] + \sum_{\substack{i,j=1 \\ i\neq j}}^{N_i}\mathbb{E}[b_i b_j]\right) \ . \tag{282}$$

   Since each b_i satisfies $b_i^2 = 1$, we have $\mathbb{E}[b_i^2] = 1$ for all i. Furthermore, assuming the samples are i.i.d., we obtain $\mathbb{E}[b_i b_j] = \mathbb{E}[b_i]\,\mathbb{E}[b_j]$ for $i \neq j$. Thus, it becomes

$$\mathbb{E}\Big[X_i^2\Big] = \frac{1}{N_i^2}\left(\sum_{i=1}^{N_i} 1 + \sum_{\substack{i,j=1 \\ i\neq j}}^{N_i}\mathbb{E}[b_i]^2\right) = \frac{1}{N_i^2}\left(\sum_{i=1}^{N_i} 1 + \sum_{\substack{i,j=1 \\ i\neq j}}^{N_i}\langle P_i \rangle^2\right) = \frac{1}{N_i^2}\left(N_i + N_i(N_i-1)\langle P_i \rangle^2\right) \tag{283}$$

$$= \langle P_i \rangle^2 + \frac{1 - \langle P_i \rangle^2}{N_i} \ . \tag{284}$$

   where we recall the definition of the true expectation value $\mathbb{E}[b_i] = \langle P_i \rangle$ in the second equality.

3. ***Third moment of*** X_i***:***

   Similarly, for the third moment, we split the summation into multiple parts, i.e. we classify the

3–tuple (i,j,k) according to the "equivalence class" of the three indices. Thus, we have

$$\mathbb{E}\left[X_i^3\right] = \frac{1}{N_i^3}\left(\sum_{i=1}^{N_i}\mathbb{E}[b_i^3] + \sum_{\substack{i,j=1\\ i\neq j}}^{N_i}\mathbb{E}[b_i^2 b_j] + \sum_{\substack{i,j,k=1\\ i\neq j\neq k}}^{N_i}\mathbb{E}[b_i b_j b_k]\right) \tag{285}$$

$$= \frac{1}{N_i^3}\left(\sum_{i=1}^{N_i}\mathbb{E}[b_i] + \sum_{\substack{i,j=1\\ i\neq j}}^{N_i}\mathbb{E}[b_j] + \sum_{\substack{i,j,k=1\\ i\neq j\neq k}}^{N_i}\mathbb{E}[b_i]\mathbb{E}[b_j]\mathbb{E}[b_k]\right) \tag{286}$$

$$= \frac{1}{N_i^3}\left(\sum_{i=1}^{N_i}\langle P_i\rangle + \sum_{\substack{i,j=1\\ i\neq j}}^{N_i}\langle P_i\rangle + \sum_{\substack{i,j,k=1\\ i\neq j\neq k}}^{N_i}\langle P_i\rangle\langle P_i\rangle\langle P_i\rangle\right), \tag{287}$$

where we again use the identity $b_i^2 = 1$ and the i.i.d. assumption in the second line. By counting the possible configurations of each summation, we arrive at

$$\mathbb{E}\left[X_i^3\right] = \frac{1}{N_i^3}\left(N_i\langle P_i\rangle + 3N_i(N_i-1)\langle P_i\rangle + N_i(N_i-1)(N_i-2)\langle P_i\rangle^3\right) \tag{288}$$

$$= \langle P_i\rangle^3\left(1 - \frac{3}{N_i} + \frac{2}{N_i^2}\right) + \langle P_i\rangle\left(\frac{3}{N_i} - \frac{2}{N_i^2}\right). \tag{289}$$

4. ***Fourth moment of X_i:***

Lastly, for the fourth moment $\mathbb{E}\left[X_i^4\right]$, we again split the summation into multiple parts in the last line, i.e., we classify the 4–tuple (i,j,k,l) according to the "equivalence class" of the four indices.

$$\mathbb{E}\left[X_i^4\right] = \frac{1}{N_i^4}\left(\sum_{i=1}^{N_i}\mathbb{E}[b_i^4] + \sum_{\substack{i,j=1\\ i\neq j}}^{N_i}\mathbb{E}[b_i^3 b_j] + \sum_{\substack{i,j=1\\ i\neq j}}^{N_i}\mathbb{E}[b_i^2 b_j^2] + \sum_{\substack{i,j,k=1\\ i\neq j\neq k}}^{N_i}\mathbb{E}[b_i^2 b_j b_k] + \sum_{\substack{i,j,k,l=1\\ i\neq j\neq k\neq l}}^{N_i}\mathbb{E}\left[b_i b_j b_k b_l\right]\right) \tag{290}$$

$$= \frac{1}{N_i^4}\left(\sum_{i=1}^{N_i}1 + \sum_{\substack{i,j=1\\ i\neq j}}^{N_i}\mathbb{E}[b_i]\mathbb{E}[b_j] + \sum_{\substack{i,j=1\\ i\neq j}}^{N_i}1 + \sum_{\substack{i,j,k=1\\ i\neq j\neq k}}^{N_i}\mathbb{E}[b_j]\mathbb{E}[b_k] + \sum_{\substack{i,j,k,l=1\\ i\neq j\neq k\neq l}}^{N_i}\mathbb{E}\left[b_i\right]\mathbb{E}[b_j]\mathbb{E}[b_k]\mathbb{E}[b_l]\right) \tag{291}$$

$$= \frac{1}{N_i^4}\left(\sum_{i=1}^{N_i}1 + \sum_{\substack{i,j=1\\ i\neq j}}^{N_i}\langle P_i\rangle^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{N_i}1 + \sum_{\substack{i,j,k=1\\ i\neq j\neq k}}^{N_i}\langle P_i\rangle^2 + \sum_{\substack{i,j,k,l=1\\ i\neq j\neq k\neq l}}^{N_i}\langle P_i\rangle^4\right). \tag{292}$$

where we also employ the identity $b_i^2 = 1$ and the i.i.d. assumption in the second line. By accounting for all possible arrangements in each summation, we derive

$$\mathbb{E}\left[X_i^4\right] = \frac{1}{N_i^4}(N_i + 4N_i(N_i-1)\langle P_i\rangle^2 + 3N_i(N_i-1) + 6N_i(N_i-1)(N_i-2)\langle P_i\rangle^2 + N_i(N_i-1)(N_i-2)(N_i-3)\langle P_i\rangle^4) \tag{293}$$

$$= \langle P_i\rangle^4 + \frac{6\langle P_i\rangle^2\left(1-\langle P_i\rangle^2\right)}{N_i} + \frac{\left(11\langle P_i\rangle^2 - 3\right)\left(\langle P_i\rangle^2 - 1\right)}{N_i^2} + \frac{2(3\langle P_i\rangle^2 - 1)(1-\langle P_i\rangle^2)}{N_i^3}. \tag{294}$$

□

Now, we are ready to present Lem. F.5, which is the second term of Eq. (271).

Lemma F.5. *Suppose we have an observable $\hat{O}$ which can be decomposed to $\hat{O} = \sum_{i=1}^{L} w_i P_i$, where $P_i \in \{I, X, Y, Z\}^{\otimes n}$ denotes the Pauli operators for n qubits and w_i represents its corresponding weights. Assuming that measurements performed for all operators are i.i.d., then the uncertainty (variance) of the square of the estimation of $\hat{O}$ can be expressed as*

$$\begin{aligned} Var\left[\overline{\langle\hat{O}\rangle}^2\right] = &\sum_{i=1}^{L}\left[\frac{w_i^4\left(1-\langle P_i\rangle^2\right)}{(N_i-1)^2}\times\left(4\langle P_i\rangle^2 N_i + 2(1-\langle P_i\rangle^2)\right)\right] \\ &+\sum_{\substack{i,j=1\\ i\neq j}}^{L}\left[w_i^2 w_j^2\left(\frac{\langle P_i\rangle^2\left(1-\langle P_j\rangle^2\right)}{N_j}+\frac{\langle P_j\rangle^2\left(1-\langle P_i\rangle^2\right)}{N_i}+\frac{\left(1-\langle P_i\rangle^2\right)\left(1-\langle P_j\rangle^2\right)}{N_i N_j}\right)\right] \\ &+4\sum_{\substack{i,j=1\\ i\neq j}}^{L}\left(\frac{w_i^3 w_j}{N_i}\langle P_i\rangle\langle P_j\rangle\left(1-\langle P_i\rangle^2\right)\right), \end{aligned} \tag{295}$$

where N_i is the number of sample used to estimate $\langle P_i\rangle$ for each $1 \le i \le L$.

Proof. We start with the expression

$$\begin{aligned} &\mathrm{Var}\left[\overline{\langle\hat{O}\rangle}^2\right] \\ &= \mathrm{Var}\left[\sum_{i=1}^{L}\frac{w_i^2 N_i}{N_i-1}\left[\left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)^2-\frac{1}{N_i}\right]+\sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j\left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)\left(\frac{1}{N_j}\sum_{k=1}^{N_j} b_k^{(P_j)}\right)\right] \qquad (296) \\ &= \mathrm{Var}\left[\sum_{i=1}^{L}\frac{w_i^2 N_i}{N_i-1}\left[\left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)^2-\frac{1}{N_i}\right]\right]+\mathrm{Var}\left[\sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j\left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)\left(\frac{1}{N_j}\sum_{k=1}^{N_j} b_k^{(P_j)}\right)\right] \\ &\quad+2\,\mathrm{Cov}\left[\sum_{i=1}^{L}\frac{w_i^2 N_i}{N_i-1}\left[\left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)^2-\frac{1}{N_i}\right],\sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j\left(\frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}\right)\left(\frac{1}{N_j}\sum_{k=1}^{N_j} b_k^{(P_j)}\right)\right], \qquad (297) \end{aligned}$$

where we use the identity $\mathrm{Var}[A+B] = \mathrm{Var}[A]+\mathrm{Var}[B]+2\,\mathrm{Cov}(A,B)$ (for any variable A, B). Before we proceed to evaluate these three terms, let us define the shorthand notation $X_i = \frac{1}{N_i}\sum_{k=1}^{N_i} b_k^{(P_i)}$. The final expression of these three terms are:

1. For the first term, we have

$$\mathrm{Var}\left[\sum_{i=1}^{L}\frac{w_i^2 N_i}{N_i-1}\left(X_i^2-\frac{1}{N_i}\right)\right] = \sum_{i=1}^{L}\mathrm{Var}\left[\frac{w_i^2 N_i}{N_i-1}\left(X_i^2-\frac{1}{N_i}\right)\right] = \sum_{i=1}^{L}\mathrm{Var}\left[\frac{w_i^2 N_i}{N_i-1}X_i^2\right] \tag{298}$$

$$= \sum_{i=1}^{L}\frac{w_i^4 N_i^2}{(N_i-1)^2}\mathrm{Var}\left[X_i^2\right], \tag{299}$$

where we assume that P_i and P_j are independent measurement for $i \neq j$ in the first equality and the property $\mathrm{Var}[X+c] = \mathrm{Var}[X]$ (for arbitrary constant c) in the second equality. Next, by

definition we have $\mathrm{Var}\left[X_i^2\right] = \mathbb{E}\left[X_i^4\right] - \mathbb{E}\left[X_i^2\right]^2$, where using Lem. F.4 further gives

$$\mathrm{Var}\left[X_i^2\right] = \langle P_i\rangle^4 + \frac{6\langle P_i\rangle^2\left(1-\langle P_i\rangle^2\right)}{N_i} + \frac{\left(11\langle P_i\rangle^2-3\right)\left(\langle P_i\rangle^2-1\right)}{N_i^2} + \frac{2(3\langle P_i\rangle^2-1)(1-\langle P_i\rangle^2)}{N_i^3} \tag{300}$$

$$- \left(\langle P_i\rangle^2 + \frac{1-\langle P_i\rangle^2}{N_i}\right)^2 \tag{301}$$

$$= \frac{4\langle P_i\rangle^2\left(1-\langle P_i\rangle^2\right)}{N_i} + \frac{2\left(5\langle P_i\rangle^2-1\right)\left(\langle P_i\rangle^2-1\right)}{N_i^2} + \frac{2(3\langle P_i\rangle^2-1)(1-\langle P_i\rangle^2)}{N_i^3}\ . \tag{302}$$

Using this result, Eq. (299) simplifies to

$$\mathrm{Var}\left[\sum_{i=1}^{L}\frac{w_i^2 N_i}{N_i-1}\left(X_i^2-\frac{1}{N_i}\right)\right]$$

$$= \sum_{i=1}^{L}\frac{w_i^4 N_i^2}{(N_i-1)^2}\times$$

$$\left(\frac{4\langle P_i\rangle^2\left(1-\langle P_i\rangle^2\right)}{N_i} + \frac{2\left(5\langle P_i\rangle^2-1\right)\left(\langle P_i\rangle^2-1\right)}{N_i^2} + \frac{2(3\langle P_i\rangle^2-1)(1-\langle P_i\rangle^2)}{N_i^3}\right) \tag{303}$$

$$= \sum_{i=1}^{L}\frac{w_i^4(1-\langle P_i\rangle^2)}{(N_i-1)^2}\left[4N_i\langle P_i\rangle^2 - 2\left(5\langle P_i\rangle^2-1\right) + \frac{2(3\langle P_i\rangle^2-1)}{N_i}\right] \tag{304}$$

$$= \sum_{i=1}^{L}\frac{w_i^4(1-\langle P_i\rangle^2)}{N_i(N_i-1)^2}\left[4N_i^2\langle P_i\rangle^2 - 10N_i\langle P_i\rangle^2 + 2N_i + 6\langle P_i\rangle^2 - 2\right] \tag{305}$$

$$= \sum_{i=1}^{L}\frac{w_i^4(1-\langle P_i\rangle^2)}{N_i(N_i-1)^2}\left[2\langle P_i\rangle^2\left(2N_i-3\right)\left(N_i-1\right) + 2(N_i-1)\right] \tag{306}$$

$$= \sum_{i=1}^{L}\left[\frac{2w_i^4}{N_i(N_i-1)}\left(1+2(N_i-2)\langle P_i\rangle^2 - (2N_i-3)\langle P_i\rangle^4\right)\right]\ . \tag{307}$$

2. For the second term, we have

$$\mathrm{Var}\left[\sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j X_i X_j\right] = \mathrm{Var}\left[\sum_{i<j}^{L} 2w_i w_j X_i X_j\right] = \sum_{i<j}^{L} 4w_i^2 w_j^2 \mathrm{Var}\left[X_i X_j\right]\ , \tag{308}$$

where we use the property $\mathrm{Var}[aX] = a^2\,\mathrm{Var}[X]$. Next, by definition of variance, we obtain

$$\mathrm{Var}\left[X_i X_j\right] = \mathbb{E}\left[X_i^2 X_j^2\right] - \mathbb{E}\left[X_i X_j\right]^2 = \mathbb{E}\left[X_i^2\right]\mathbb{E}\left[X_j^2\right] - \mathbb{E}\left[X_i\right]^2\mathbb{E}\left[X_j\right]^2\ . \tag{309}$$

Using Lem. F.4, we have

$$\mathrm{Var}\left[X_i X_j\right] = \left(\langle P_i\rangle^2 + \frac{1-\langle P_i\rangle^2}{N_i}\right)\left(\langle P_j\rangle^2 + \frac{1-\langle P_j\rangle^2}{N_i}\right) - \langle P_i\rangle^2\langle P_j\rangle^2 \tag{310}$$

$$= \frac{\langle P_i\rangle^2\left(1-\langle P_j\rangle^2\right)}{N_j} + \frac{\langle P_j\rangle^2\left(1-\langle P_i\rangle^2\right)}{N_i} + \frac{\left(1-\langle P_i\rangle^2\right)\left(1-\langle P_j\rangle^2\right)}{N_i N_j}\ . \tag{311}$$

Consequently, Eq. (308) is

$$\begin{aligned}&\mathrm{Var}\left[\sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j X_i X_j\right]\\ &= 4\sum_{i<j}^{L}\left[\frac{w_i^2 w_j^2}{N_i N_j}\left(\left(1-\langle P_i\rangle^2\right)\left(1-\langle P_j\rangle^2\right)+N_i\langle P_i\rangle^2\left(1-\langle P_j\rangle^2\right)+N_j\langle P_j\rangle^2\left(1-\langle P_i\rangle^2\right)\right)\right] .\end{aligned} \tag{312}$$

3. For the third term, it is

$$\begin{aligned}&\mathrm{Cov}\left[\sum_{i=1}^{L}\frac{w_i^2 N_i}{N_i-1}\left[X_i^2-\frac{1}{N_i}\right], \sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j X_i X_j\right]\\ &= \sum_{\substack{i,j=1\\ i\neq j}}^{L}\left(\mathrm{Cov}\left[\frac{w_i^2 N_i}{N_i-1}\left[X_i^2-\frac{1}{N_i}\right], w_i w_j X_i X_j\right]+\mathrm{Cov}\left[\frac{w_i^2 N_i}{N_i-1}\left[X_i^2-\frac{1}{N_i}\right], w_i w_j X_j X_i\right]\right) ,\end{aligned} \tag{313}$$

where we use the bilinear property of the covariance. By symmetry, the two covariance contributions are equal and hence we have

$$\mathrm{Cov}\left[\sum_{i=1}^{L}\frac{w_i^2 N_i}{N_i-1}\left[X_i^2-\frac{1}{N_i}\right], \sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j X_i X_j\right] = 2\sum_{\substack{i,j=1\\ i\neq j}}^{L}\left(\mathrm{Cov}\left[\frac{w_i^2 N_i}{N_i-1}\left[X_i^2-\frac{1}{N_i}\right], w_i w_j X_i X_j\right]\right) , \tag{314}$$

Since $\mathrm{Cov}[A,B]=\mathbb{E}[AB]-\mathbb{E}[A]\,\mathbb{E}[B]$, we have

$$\begin{aligned}&\mathrm{Cov}\left[\frac{w_i^2 N_i}{N_i-1}\left[X_i^2-\frac{1}{N_i}\right], w_i w_j X_i X_j\right]\\ &= \mathbb{E}\left[\left(\frac{w_i^2 N_i}{N_i-1}\left[X_i^2-\frac{1}{N_i}\right]\right)(w_i w_j X_i X_j)\right]-\mathbb{E}\left[\frac{w_i^2 N_i}{N_i-1}\left[X_i^2-\frac{1}{N_i}\right]\right]\mathbb{E}\left[w_i w_j X_i X_j\right] &(315)\\ &= \frac{w_i^3 w_j N_i}{N_i-1}\,\mathbb{E}[X_j]\times\left\{\mathbb{E}[X_i^3]-\mathbb{E}[X_i^2]\,\mathbb{E}[X_i]\right\} , &(316)\end{aligned}$$

Using Lem. F.4, Eq. (316) reduces to

$$\begin{aligned}&\mathrm{Cov}\left[\frac{w_i^2 N_i}{N_i-1}\left[X_i^2-\frac{1}{N_i}\right], w_i w_j X_i X_j\right]\\ &= \frac{w_i^3 w_j N_i}{N_i-1}\langle P_j\rangle\times\left\{\langle P_i\rangle^3\left(1-\frac{3}{N_i}+\frac{2}{N_i^2}\right)+\langle P_i\rangle\left(\frac{3}{N_i}-\frac{2}{N_i^2}\right)-\left(\langle P_i\rangle^2+\frac{1-\langle P_i\rangle^2}{N_i}\right)\langle P_i\rangle\right\} &(317)\\ &= \frac{w_i^3 w_j N_i}{N_i-1}\langle P_j\rangle\times\left\{\frac{\langle P_i\rangle^3(2-2N_i)+2\langle P_i\rangle(N_i-1)}{N_i^2}\right\}=\frac{2w_i^3 w_j}{N_i}\langle P_j\rangle\langle P_i\rangle\left(1-\langle P_i\rangle^2\right) . &(318)\end{aligned}$$

Therefore, Eq. (314) is given by

$$\mathrm{Cov}\left[\sum_{i=1}^{L}\frac{w_i^2 N_i}{N_i-1}\left[X_i^2-\frac{1}{N_i}\right], \sum_{\substack{i,j=1\\ i\neq j}}^{L} w_i w_j X_i X_j\right] = 4\sum_{\substack{i,j=1\\ i\neq j}}^{L}\left(\frac{w_i^3 w_j}{N_i}\langle P_j\rangle\langle P_i\rangle\left(1-\langle P_i\rangle^2\right)\right) . \tag{319}$$

Collecting Eq. (307), Eq. (312) and Eq. (319), we arrive at the final expression:

$$\begin{aligned}\mathrm{Var}\left[\overline{\langle\hat{O}\rangle^2}\right] &= 2\sum_{i=1}^{L}\left[\frac{w_i^4}{N_i(N_i-1)}\left(1+2(N_i-2)\langle P_i\rangle^2-(2N_i-3)\langle P_i\rangle^4\right)\right]\\ &+4\sum_{i<j}^{L}\left[\frac{w_i^2w_j^2}{N_iN_j}\left(\left(1-\langle P_i\rangle^2\right)\left(1-\langle P_j\rangle^2\right)+N_i\langle P_i\rangle^2\left(1-\langle P_j\rangle^2\right)+N_j\langle P_j\rangle^2\left(1-\langle P_i\rangle^2\right)\right)\right]\\ &+4\sum_{i<j}^{L}\left(\frac{w_i^3w_j}{N_i}\langle P_i\rangle\langle P_j\rangle\left(1-\langle P_i\rangle^2\right)\right)\ . \end{aligned} \tag{320}$$

□

Theorem F.6 (Uncertainty of the estimated variance of an observable)**.** *Suppose we have an observable $\hat{O}$ which is of the form $\hat{O}=\sum_{i=1}^{L}w_iP_i$, where $P_i\in\{I,X,Y,Z\}^{\otimes n}$ denotes the Pauli operators for n qubits and w_i represents its corresponding weights. Assuming that measurements performed for all operators including $P_{ij}=P_iP_j$ are i.i.d., then the uncertainty (variance) of the estimated variance of $\hat{O}$ can be expressed as*

$$\begin{aligned}Var[\overline{V}] &= \sum_{\substack{i,j=1\\ i\neq j}}^{L}\frac{w_i^2w_j^2}{N_{ij}}\left(1-\langle P_{ij}\rangle^2\right)+2\sum_{i=1}^{L}\left[\frac{w_i^4}{N_i(N_i-1)}\left(1+2(N_i-2)\langle P_i\rangle^2-(2N_i-3)\langle P_i\rangle^4\right)\right]\\ &+4\sum_{i<j}^{L}\left[\frac{w_i^2w_j^2}{N_iN_j}\left(\left(1-\langle P_i\rangle^2\right)\left(1-\langle P_j\rangle^2\right)+N_i\langle P_i\rangle^2\left(1-\langle P_j\rangle^2\right)+N_j\langle P_j\rangle^2\left(1-\langle P_i\rangle^2\right)\right)\right]\\ &+4\sum_{i<j}^{L}\left(\frac{w_i^3w_j}{N_i}\langle P_i\rangle\langle P_j\rangle\left(1-\langle P_i\rangle^2\right)\right)\ . \end{aligned} \tag{321}$$

Proof. Since the set of Pauli operators $\{P_i\}$ and $P_{ij}=P_iP_j$ are i.i.d. and mutually independent, we obtain $\mathrm{Var}[\overline{V}]=\mathrm{Var}\left[\overline{\langle\hat{O}^2\rangle}\right]-\mathrm{Var}\left[\overline{\langle\hat{O}\rangle^2}\right]$, where the first and second terms are given by Lem. F.3 and Lem. F.5 respectively:

$$\begin{aligned}\mathrm{Var}[\overline{V}] &= \sum_{\substack{i,j=1\\ i\neq j}}^{L}\frac{w_i^2w_j^2}{N_{ij}}\left(1-\langle P_{ij}\rangle^2\right)+2\sum_{i=1}^{L}\left[\frac{w_i^4}{N_i(N_i-1)}\left(1+2(N_i-2)\langle P_i\rangle^2-(2N_i-3)\langle P_i\rangle^4\right)\right]\\ &+4\sum_{i<j}^{L}\left[\frac{w_i^2w_j^2}{N_iN_j}\left(\left(1-\langle P_i\rangle^2\right)\left(1-\langle P_j\rangle^2\right)+N_i\langle P_i\rangle^2\left(1-\langle P_j\rangle^2\right)+N_j\langle P_j\rangle^2\left(1-\langle P_i\rangle^2\right)\right)\right]\\ &+4\sum_{i<j}^{L}\left(\frac{w_i^3w_j}{N_i}\langle P_i\rangle\langle P_j\rangle\left(1-\langle P_i\rangle^2\right)\right)\ . \end{aligned} \tag{322}$$

□

F.4 Alternative Unbiased Method of Estimating Operator Variance

Here, we provide another way of computing an unbiased estimator of the variance operator.

Lemma F.7. *Suppose we have an observable $\hat{O}$ of the form $\hat{O}=\sum_{i=1}^{L}w_iP_i$, where $P_i\in\{I,X,Y,Z\}^{\otimes n}$ denotes the Pauli operators for n qubits and w_i represents its corresponding weights. Assuming measurements performed for all operators including $P_{ij}=P_iP_j$ are i.i.d., then the unbiased estimator of the variance operator can be alternatively written as*

$$\overline{V}=\sum_{i=1}^{L}w_i^2+\sum_{\substack{i,j=1\\ i\neq j}}^{L}w_iw_j\left(\frac{1}{N_{ij}}\sum_{k=1}^{N_{ij}}b_k^{(P_{ij})}\right)-\sum_{i,j=1}^{L}w_iw_j\left(\frac{1}{N_{(i\otimes j)}}\sum_{i=1}^{N_{(i\otimes j)}}b_i^{(P_{i\otimes j})}\right)\ , \tag{323}$$

where we perform N_{ij} times measurement on the operator $P_{ij} = P_i P_j$ and $N_{(i\otimes j)}$ times joint measurement on $P_i \otimes P_j$.

Proof. Recall that the variance for an observable $\hat{O}$ is given by

$$V = \langle \hat{O}^2 \rangle - \langle \hat{O} \rangle^2 = \sum_{i=1}^{L} w_i^2 \langle I \rangle + \sum_{\substack{i,j=1 \\ i \neq j}}^{L} w_i w_j \langle P_{ij} \rangle - \left(\sum_{i=1}^{L} w_i \langle P_i \rangle \right)^2 . \tag{324}$$

To estimate the second term $\langle \hat{O} \rangle^2$, we now perform joint measurements on two copies of quantum states. For each independent measurement $\{P_i \otimes P_j\}$, we collects the results of measured bit string $\{b_i^{(P_{i\otimes j})}\}$. Thus, the unbiased estimator of the product $\langle P_i \rangle \langle P_j \rangle$ is given by

$$\overline{\langle P_i \rangle \langle P_j \rangle} = \frac{1}{N_{(i\otimes j)}} \sum_{i=1}^{N_{(i\otimes j)}} b_i^{(P_{i\otimes j})} , \tag{325}$$

where $N_{(i\otimes j)}$ denotes the number of samples for the measurement $(P_i \otimes P_j)$. Hence, the term $\langle \hat{O} \rangle^2$ can be estimated as

$$\overline{\langle \hat{O} \rangle^2} = \sum_{i,j=1}^{L} w_i w_j \langle P_i \rangle \langle P_j \rangle = \sum_{i,j=1}^{L} w_i w_j \left(\frac{1}{N_{(i\otimes j)}} \sum_{i=1}^{N_{(i\otimes j)}} b_i^{(P_{i\otimes j})} \right) . \tag{326}$$

Thus, the unbiased estimator of the variance is

$$\overline{V} = \sum_{i=1}^{L} w_i^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{L} w_i w_j \left(\frac{1}{N_{ij}} \sum_{k=1}^{N_{ij}} b_k^{(P_{ij})} \right) - \sum_{i,j=1}^{L} w_i w_j \left(\frac{1}{N_{(i\otimes j)}} \sum_{i=1}^{N_{(i\otimes j)}} b_i^{(P_{i\otimes j})} \right) . \tag{327}$$

□

Next, we derive the uncertainty of this estimated variance based on this alternative measurement protocol.

Lemma F.8. *Let $\hat{O}$ be an observable of the form $\hat{O} = \sum_{i=1}^{L} w_i P_i$, where $P_i \in \{I, X, Y, Z\}^{\otimes n}$ denotes the Pauli operators for n qubits and w_i represents its corresponding weights. Assuming the measurements $\{P_{i\otimes j}\}$ and $\{P_{ij} = P_i P_j\}$ are i.i.d. and they are mutually independent to each other, then the uncertainty (variance) of the estimated variance $\overline{V}$ of the observable is*

$$Var[\overline{V}] = \sum_{\substack{i,j=1 \\ i \neq j}}^{L} \frac{w_i^2 w_j^2 \left(1 - \langle P_{ij} \rangle^2\right)}{N_{ij}} + \sum_{i,j=1}^{L} \frac{w_i^2 w_j^2 \left(1 - \langle P_i \rangle^2 \langle P_j \rangle^2\right)}{N_{(i\otimes j)}} , \tag{328}$$

where we perform N_{ij} times measurement on the operator $P_{ij} = P_i P_j$ and $N_{(i\otimes j)}$ times joint measurement on $P_i \otimes P_j$.

Proof. First, since the measurements $\{P_{i\otimes j}\}$ and $\{P_{ij}\}$ are i.i.d. and mutually independent, the variance of the operator is

$$\overline{V} = \sum_{i=1}^{L} w_i^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{L} w_i w_j \left(\frac{1}{N_{ij}} \sum_{k=1}^{N_{ij}} b_k^{(P_{ij})} \right) - \sum_{i,j=1}^{L} w_i w_j \left(\frac{1}{N_{(i\otimes j)}} \sum_{i=1}^{N_{(i\otimes j)}} b_i^{(P_{i\otimes j})} \right) . \tag{329}$$

where we use Lem. F.7. Expanding the uncertainty (variance) of this estimation gives

$$\mathrm{Var}[\overline{V}] = \mathrm{Var}\left[\sum_{i=1}^{L} w_i^2 + \sum_{\substack{i,j=1 \\ i\neq j}}^{L} w_i w_j \left(\frac{1}{N_{ij}}\sum_{k=1}^{N_{ij}} b_k^{(P_{ij})}\right) - \sum_{i,j=1}^{L} w_i w_j \left(\frac{1}{N_{(i\otimes j)}}\sum_{i=1}^{N_{(i\otimes j)}} b_i^{(P_{i\otimes j})}\right)\right] \tag{330}$$

$$= \mathrm{Var}\left[\sum_{\substack{i,j=1 \\ i\neq j}}^{L} w_i w_j \left(\frac{1}{N_{ij}}\sum_{k=1}^{N_{ij}} b_k^{(P_{ij})}\right)\right] + \mathrm{Var}\left[\sum_{i,j=1}^{L} w_i w_j \left(\frac{1}{N_{(i\otimes j)}}\sum_{i=1}^{N_{(i\otimes j)}} b_i^{(P_{i\otimes j})}\right)\right], \tag{331}$$

where we use the fact that the first term $\sum_{i=1}^{L} w_i^2$ has zero variance in the last line. Next, as $\mathrm{Var}[aX] = a^2\,\mathrm{Var}[X]$ for any scalar factor a, it can be simplified to

$$\mathrm{Var}[\overline{V}] = \sum_{\substack{i,j=1 \\ i\neq j}}^{L} \frac{w_i^2 w_j^2}{N_{ij}^2}\mathrm{Var}\left[\sum_{k=1}^{N_{ij}} b_k^{(P_{ij})}\right] + \sum_{i,j=1}^{L} \frac{w_i^2 w_j^2}{N_{(i\otimes j)}^2}\mathrm{Var}\left[\sum_{i=1}^{N_{(i\otimes j)}} b_i^{(P_{i\otimes j})}\right]. \tag{332}$$

Since the bit string b_i satisfies $b_i^2 = 1$, we have the following identity:

$$\mathrm{Var}\left[\sum_{k=1}^{N_{ij}} b_k^{(P_{ij})}\right] = \sum_{k=1}^{N_{ij}} \mathrm{Var}\left[b_k^{(P_{ij})}\right] = \sum_{k=1}^{N_{ij}} \left(1 - \langle P_{ij}\rangle^2\right) = N_{ij}\left(1 - \langle P_{ij}\rangle^2\right). \tag{333}$$

Similarly, for the joint measurements, it becomes

$$\mathrm{Var}\left[\sum_{i=1}^{N_{(i\otimes j)}} b_i^{(P_{i\otimes j})}\right] = \sum_{i=1}^{N_{(i\otimes j)}} \mathrm{Var}\left[b_i^{(P_{i\otimes j})}\right] = \sum_{i=1}^{N_{(i\otimes j)}} \left(1 - \langle P_i\rangle^2 \langle P_j\rangle^2\right) = N_{(i\otimes j)}\left(1 - \langle P_i\rangle^2 \langle P_j\rangle^2\right). \tag{334}$$

Therefore, Eq. (332) becomes

$$\mathrm{Var}[\overline{V}] = \sum_{\substack{i,j=1 \\ i\neq j}}^{L} \frac{w_i^2 w_j^2\left(1 - \langle P_{ij}\rangle^2\right)}{N_{ij}} + \sum_{i,j=1}^{L} \frac{w_i^2 w_j^2\left(1 - \langle P_i\rangle^2 \langle P_j\rangle^2\right)}{N_{(i\otimes j)}}. \tag{335}$$

□

Using this result, we now present a proposition that tells us how many samples we need to achieve precision ϵ when estimating the variance of the observable $\hat{O}$.

Proposition F.9. *Suppose we have an observable $\hat{O}$ of the form $\hat{O} = \sum_{i=1}^{L} w_i P_i$, where $P_i \in \{I, X, Y, Z\}^{\otimes n}$ denotes the Pauli operators for n qubits and w_i represents its corresponding weights. Let us denote $\overline{V}$ as the estimated variance of the observable. Assume that the measurements $\{P_{i\otimes j}\}$ and $\{P_{ij}\}$ are i.i.d. and mutually independent, the number of samples required to achieve a target precision ϵ for $\overline{V}$ scales as the fourth power of the Hamiltonian's L1 norm and quadratically in the inverse of ϵ.*

Proof. First, note that the total number of the measurements are given by

$$N = \sum_{\substack{i,j=1 \\ i\neq j}}^{L} N_{ij} + \sum_{i,j=1}^{L} N_{(i\otimes j)}. \tag{336}$$

In this setting, the optimal allocation of measurement shots can be determined via Lagrange multipliers. Our objective is to minimize the total number of shots while ensuring that the uncertainty of $\overline{V}$

remains below the desired precision ϵ^2. We follow the approach outlined in Ref. [72], which provides the optimal allocation of measurement shots for estimating each term of the Hamiltonian. Note that Ref. [72] demonstrates that the number of measurements required to achieve ϵ-precision is given by $\mathcal{O}(|w|^2/\epsilon^2)$ with $|w| = \sum_i^L |w_i|$. First, the corresponding Lagrangian $\mathcal{L}$ can be expressed as

$$\mathcal{L} = \sum_{\substack{i,j=1 \\ i\neq j}}^{L} N_{ij} + \sum_{i,j=1}^{L} N_{(i\otimes j)} + \lambda \left(\mathrm{Var}\left[\overline{V_N}\right] - \epsilon^2 \right). \tag{337}$$

According to Lem. F.8, we have

$$\mathrm{Var}[\overline{V}] = \sum_{\substack{i,j=1 \\ i\neq j}}^{L} \frac{w_i^2 w_j^2 \left(1 - \langle P_{ij}\rangle^2\right)}{N_{ij}} + \sum_{i,j=1}^{L} \frac{w_i^2 w_j^2 \left(1 - \langle P_i\rangle^2 \langle P_j\rangle^2\right)}{N_{(i\otimes j)}} . \tag{338}$$

Using this result, we can proceed to evaluate $\mathcal{L}$. By taking the derivative of $\mathcal{L}$, Eq. (337) yields

$$\frac{\partial \mathcal{L}}{\partial N_{ij}} = 1 - \lambda \frac{w_i^2 w_j^2 \left(1 - \langle P_{(i,j)}\rangle^2\right)}{N_{ij}^2} , \qquad \frac{\partial \mathcal{L}}{\partial N_{(i\otimes j)}} = 1 - \lambda \frac{w_i^2 w_j^2 \left(1 - \langle P_i\rangle^2 \langle P_j\rangle^2\right)}{N_{(i\otimes j)}^2} . \tag{339}$$

To obtain zero derivatives, we require

$$N_{ij} = |w_i|\,|w_j| \sqrt{\lambda \left(1 - \langle P_{(i,j)}\rangle^2\right)} , \qquad N_{(i\otimes j)} = |w_i|\,|w_j| \sqrt{\lambda \left(1 - \langle P_i\rangle^2 \langle P_j\rangle^2\right)} . \tag{340}$$

Recall that we set the target precision to be ϵ, i.e. we would like to achieve $\mathrm{Var}\left[\overline{V_N}\right] = \epsilon^2$ and hence Eq. (328) yields

$$\epsilon^2 = \sum_{\substack{i,j=1 \\ i\neq j}}^{L} \frac{w_i^2 w_j^2 \left(1 - \langle P_{ij}\rangle^2\right)}{N_{ij}} + \sum_{i,j=1}^{L} \frac{w_i^2 w_j^2 \left(1 - \langle P_i\rangle^2 \langle P_j\rangle^2\right)}{N_{(i\otimes j)}} \tag{341}$$

$$= \frac{1}{\sqrt{\lambda}} \left(\sum_{\substack{i,j=1 \\ i\neq j}}^{L} |w_i|\,|w_j| \sqrt{\left(1 - \langle P_{ij}\rangle^2\right)} + \sum_{i,j=1}^{L} |w_i|\,|w_j| \sqrt{\left(1 - \langle P_i\rangle^2 \langle P_j\rangle^2\right)} \right) . \tag{342}$$

where we substitute back Eq. (340) to obtain last line. Thus, we have

$$\sqrt{\lambda} = \frac{1}{\epsilon^2} \left(\sum_{\substack{i,j=1 \\ i\neq j}}^{L} |w_i|\,|w_j| \sqrt{\left(1 - \langle P_{(i,j)}\rangle^2\right)} + \sum_{i,j=1}^{L} |w_i|\,|w_j| \sqrt{\left(1 - \langle P_i\rangle^2 \langle P_j\rangle^2\right)} \right) . \tag{343}$$

Finally, the optimal number of total measurements are

$$N = \sum_{\substack{i,j=1 \\ i\neq j}}^{L} N_{ij} + \sum_{i,j=1}^{L} N_{(i\otimes j)} = \sqrt{\lambda} \left(\sum_{\substack{i,j=1 \\ i\neq j}}^{L} |w_i|\,|w_j| \sqrt{\left(1 - \langle P_{(i,j)}\rangle^2\right)} + \sum_{i,j=1}^{L} |w_i|\,|w_j| \sqrt{\left(1 - \langle P_i\rangle^2 \langle P_j\rangle^2\right)} \right)$$

$$= \frac{1}{\epsilon^2} \left(\sum_{\substack{i,j=1 \\ i\neq j}}^{L} |w_i|\,|w_j| \sqrt{\left(1 - \langle P_{(i,j)}\rangle^2\right)} + \sum_{i,j=1}^{L} |w_i|\,|w_j| \sqrt{\left(1 - \langle P_i\rangle^2 \langle P_j\rangle^2\right)} \right)^2 . \tag{344}$$

Since $b_i \in \{\pm 1\}$, $\langle P_i \rangle \leq 1$ for all i. The total number of measurements is then upper bounded by

$$N \leq \frac{1}{\epsilon^2} \left(\sum_{\substack{i,j=1 \\ i \neq j}}^{L} |w_i|\,|w_j| + \sum_{i,j=1}^{L} |w_i|\,|w_j| \right)^2 \leq \frac{1}{\epsilon^2} \left(2 \sum_{i,j=1}^{L} |w_i|\,|w_j| \right)^2 \tag{345}$$

$$= \frac{4}{\epsilon^2} \left(\sum_{i=1}^{L} |w_i| \right)^2 \left(\sum_{j=1}^{L} |w_j| \right)^2 = \frac{4}{\epsilon^2} \left(\sum_{i=1}^{L} |w_i| \right)^4 . \tag{346}$$

So, the proposition's statement is justified. □