Spectral-density estimation with the Gaussian integral transform

A. Roggero*
Institute for Nuclear Theory, University of Washington, Seattle, Washington 98195, USA

(Received 7 May 2020; accepted 24 July 2020; published 14 August 2020; corrected 9 August 2021)

The spectral-density operator $\hat{\rho}(\omega) = \delta(\omega - \hat{H})$ plays a central role in linear response theory as its expectation value, the dynamical response function, can be used to compute scattering cross sections. In this work, we describe a near optimal quantum algorithm providing an approximation to the spectral density with energy resolution Δ and error ϵ using $O(\sqrt{\log_2(1/\epsilon)[\log_2(1/\Delta) + \log_2(1/\epsilon)]}/\Delta)$ operations. This is achieved without using expensive approximations to the time-evolution operator, but instead exploiting qubitization to implement an approximate Gaussian integral transform of the spectral density. We also describe appropriate error metrics to assess the quality of the spectral function approximations more generally.

DOI: 10.1103/PhysRevA.102.022409

I. INTRODUCTION

Since the first seminal works of Feynman [1] and Lloyd [2], quantum computing has been recognized as a possible avenue to explore quantum dynamics of strongly correlated many-body systems beyond what is possible with classical computational tools. Recent progress in Hamiltonian simulation algorithms [3–6] has allowed a dramatic reduction of the computational cost for applications as diverse as computing out-of-equilibrium dynamics [7], exclusive scattering cross sections [8,9], and ground-state energy estimation [10]. Most of the proposed algorithms still require a number of gates too large for possible applications on noisy intermediate-scale quantum (NISQ) devices [11] and more work is required to bring these costs down (see, e.g., [9] for a recent analysis of the requirements for neutrino-nucleus scattering).

In the same spirit of the recent work by Somma [12], we propose in this work a quantum algorithm with near optimal computational cost (in terms of oracle calls) to study the problem of spectral-density estimation. In particular, given a Hermitian operator $\hat{O}$, the goal of this work is to obtain an efficient algorithm to approximate the spectral density operator, $\hat{\rho}(\omega) = \delta(\omega - \hat{O})$, with δ the Dirac delta function. Using the eigenstates $|k\rangle$ of the operator $\hat{O}$, we have the following spectral representation:

$$\hat{\rho}(\omega) = \sum_k^{\Gamma} \delta(\omega - O_k)|k\rangle\langle k|, \tag{1}$$

with O_k the eigenvalue for $|k\rangle$, and Γ the total number of eigenvalues. Without loss of generality, we will consider normalized operators $\hat{O}$ with $\|\hat{O}\| \leqslant 1$ so that the spectrum is contained in the interval $[-1, 1]$.

One of the most popular applications of the spectral-density operator is in the theory of linear response, where it is directly connected with the *dynamical response function* $S(\omega)$. More precisely, given a state vector $|\Psi\rangle$, we can define the following response function:

$$S(\omega) = \langle\Psi|\hat{\rho}(\omega)|\Psi\rangle = \sum_k^{\Gamma} |\langle\Psi|k\rangle|^2 \delta(\omega - O_k). \tag{2}$$

The response function can be used to compute, among other things, the energy-resolved inclusive cross section for a scattering process that maps an initial state $|\Phi_0\rangle$ to the final state $|\Psi\rangle = \hat{Q}|\Phi_0\rangle$ through the action of the (possibly nonunitary) vertex operator $\hat{Q}$. In this case, the relevant operator $\hat{O}$ coincides with the Hamiltonian of the physical system, and for this reason we will often call its eigenvalues "frequencies." The technique we describe here is, however, applicable to any Hermitian operator.

The approach we follow in this work is to consider approximations to the response function obtained through an integral transform of the type

$$\Phi_K(\nu) = \int d\omega K(\nu, \omega) S(\omega) = \sum_k^{\Gamma} |\langle\Psi|k\rangle|^2 K(\nu, O_k). \tag{3}$$

The integral kernel $K(\nu, \omega)$ that defines the transform can also be used directly as an approximation to the spectral-density operator: $\hat{\rho}_K(\nu) = K(\nu, \hat{O}) \approx \hat{\rho}(\omega = \nu)$. For this to be a good approximation, the kernel function should be chosen as a finite-width representation of the Dirac delta function.

We note that the approach of computing response functions by a direct inversion of integral transforms such as Eq. (3) is a common strategy in many-body physics. In quantum Monte Carlo calculations, for instance, it is common to consider the Laplace kernel $K(\omega, \nu) = \exp(-\nu\omega)$ due to its connection with Euclidean path integrals (see, e.g., [13,14]), but other alternatives such as the Sumudu [15] and Lorentz [16–18] transforms have also been considered in the past. The main difficulty encountered by these methods is the problem that for any compact kernel function, the inversion of the integral transform is a numerically ill-posed problem: any errors in the estimate of Φ_K will get exponentially amplified by the

*roggero@uw.edu

inversion procedure (see, e.g., [19,20]). In this work, we avoid the problem by directly using the integral transform $\Phi_K(\omega)$ as the approximate reconstruction of the original signal $S(\omega)$. We note at this point that the idea of directly using the integral transform to extract physical information has been explored in the past with great success. For example, in [21], the dipole polarizability α_D of ^{22}O was computed using the coupled cluster method and using a direct mapping between α_D and a Lorentz integral transform of the response function, and, in [22], the contribution of impurity scattering in the thermal conductivity in the outer crust of neutron stars was successfully extracted by mapping it into features of the Laplace transform of the response.

A possible future extension of our work would be to consider approximate inversion schemes such as the maximum entropy method [23] to try to reduce the computational cost of the quantum algorithm at the possible expense of introducing an uncontrollable error.

The paper is organized as follows. In Sec. II, we first provide a detailed description of the error metrics that we use to judge the quality of the approximation in Eq. (3), and then we summarize the main results of the paper and compare them to the recent work of Ref. [12], which can be understood as a particular instance of the method that we propose. We also provide an argument for the near optimality of both techniques. We then present two integral kernels: the Fejer kernel naturally generated using quantum phase estimation [8,9] (Sec. III) and the Gaussian kernel which allows one to achieve near optimal scaling of the computational cost (Sec. IV). We conclude in Sec. V, providing a summary of our findings and proposing possible avenues for future improvements. We also provide a pseudocode implementation in Appendix B.

II. DEFINITIONS AND COMPARISON TO PREVIOUS WORK

In order to precisely quantify the accuracy of the approximation procedure presented in Sec. I and connect with recent work on quantum algorithms exploring similar problems [12,24], we now introduce the following definitions.

(a) We will call an integral kernel Σ-*accurate* with *resolution* Δ if the following condition holds:

$$\inf_{\omega_0\in[-1,1]} \sum\!\!\!\!\!\!\int_{\omega_0-\Delta}^{\omega_0+\Delta} d\nu K(\nu,\omega_0) \geqslant 1-\Sigma, \tag{4}$$

where the symbol $\sum\!\!\!\!\!\!\int$ indicates an integral when the transformed variable ν is defined over a continuous interval or a sum if ν is defined on a discrete set.

(b) We will call a distribution $\widetilde{\Phi}(\omega)$ a β-*approximation* to the true distribution $\Phi(\omega)$ with confidence $1-\eta_\beta$ if the total variation is bounded as

$$\delta_V(\Phi,\widetilde{\Phi}) := \sup_{\omega\in[-1,1]} |\Phi(\omega)-\widetilde{\Phi}(\omega)| \leqslant \beta, \tag{5}$$

with probability $P > 1-\eta_\beta$.

(c) If the estimator $\widetilde{\Phi}_K$ is obtained as a β-*approximation* with *confidence* $1-\eta_\beta$ of a Σ-*accurate* integral transform Φ_K of the response function $S(\omega)$ with *resolution* Δ, we will call it a $(\Sigma, \Delta, \beta, \eta_\beta)$-approximation to the response S.

These definitions are similar to those introduced in the recent work [24]. In particular, the first definition is similar in spirit to, but more stringent than, the condition of having resolution Δ and confidence $\eta = (1-\Sigma)$ (Def. 1 of [24]), while the second condition is equivalent to the β-approximation (Def. 3 of [24]).

The reason for these definitions, and the mild departure from those introduced in Ref. [24], is rooted in the fact that for a physics application, we are ultimately interested in frequency observables of the form

$$Q(S,f) = \int_{-1}^{1} d\omega S(\omega) f(\omega) \tag{6}$$

for some bounded function f. If we estimate the observable Q using a $(\Sigma, \Delta, \beta, \eta_\beta)$-approximation $\widetilde{\Phi}_K$, we have, in fact, with confidence $1-\eta_\beta$, the following bound:

$$|Q(S,f) - Q(\widetilde{\Phi}_K, f)| \leqslant f_{\max}^{\Delta} + 2 f_{\max}\Sigma + \beta f_{\text{int}}, \tag{7}$$

where we have defined the quantities

$$f_{\max} = \sup_{\omega\in[-1,1]} |f(\omega)|, \quad f_{\text{int}} = \int_{-1}^{1} d\omega |f(\omega)| \leqslant 2 f_{\max}, \tag{8}$$

and the upper bound on the maximum variation,

$$f_{\max}^{\Delta} = \sup_{\omega\in[-1,1]} \sup_{x\in[-\Delta,\Delta]} |f(\omega+x) - f(\omega)|. \tag{9}$$

A full derivation of this is provided in Appendix A.

At this point, it is important to point out another difference with Ref. [24]. In our work, the second error metric β captures both the statistical error coming from estimating the distribution $\widetilde{\Phi}_K(\omega)$ with a finite number of samples, and the possible systematic error coming from using an approximation of the quantum circuit needed to obtain the desired integral transform $\Phi_K(\omega)$. In this sense, Σ-accurate with resolution Δ is a property of the kernel function $K(\nu,\omega)$, while β-approximation with confidence $1-\eta_\beta$ is a property that characterizes the implementation of the algorithm that generates the desired integral transform.

Comparison to previous work

The approximation problem that we are trying to solve is very similar to the quantum eigenvalue estimation problem (QEEP) considered in Ref. [12]. In this section, we will anticipate the main results of our work and provide a comparison with the time-series analysis (TSA) algorithm proposed in Ref. [12]. In particular, we will compare the computational cost in terms of the number M of oracle calls to a base unitary W_Q, and the total number N_S of samples needed to generate a $(\Sigma, \Delta, \beta, \eta_\beta)$-approximation to the spectral function $S(\omega)$. In order to simplify the comparison, we will consider here the limit $\Sigma = \beta = \varepsilon$, which is sensible given the definition in Eq. (7). Detailed results for the more general case will be provided in the sections below.

The TSA approach from Ref. [12] starts by decomposing the frequency domain into N disjoint intervals of size 2Δ and then obtaining the response in each of these bins using the Fourier expansion of the bump function. In light of the definitions provided above, this can be understood as using

an integral transform with the kernel function given by the approximate frequency comb,

$$K(\nu_j,\omega)=\sum_{j=1}^{N} f_b(\nu_j,\omega), \tag{10}$$

where ν_j is the central value of the jth frequency bin, and the function f_b is obtained from bump functions and has support on $[\nu_j-\Delta,\nu_j+\Delta]$ only. Due to this property, it is straightforward to see that this kernel allows one to achieve accuracies $\Sigma=0$ in Eq. (4). Note, however, that using $\Sigma\ll\beta$ will not help reduce the final error in Eq. (7) (unless $f_{\max}\gg f_{\text{int}}$), and in fact here we only require them to both be equal to ε.

The TSA algorithm requires one to apply the (controlled) time-evolution operator $U_O(t)=\exp(-it\hat{O})$ for a maximum time $t_{\max}$ scaling as (see Appendix A of [12])

$$t_{\max}=O\left[\frac{1}{\Delta}\log_2\left(\frac{1}{\varepsilon}\right)^2\right], \tag{11}$$

together with a total number of samples scaling as

$$N_S=O\left[\frac{1}{\Delta^3\varepsilon^2}\log_2\left(\frac{1}{\varepsilon}\right)^6\log_2\left(\frac{1}{\eta_\beta}\right)\right], \tag{12}$$

in order to achieve a $(\Sigma=0,\Delta,\varepsilon,\eta_\beta)$-approximation. Note that if we require the final approximation over N frequency to have total error less than ε (as done in [12]), the ε-dependent logarithmic terms above will include an additional $1/\Delta$ like the Gaussian integral transform (GIT) (see, also, Appendix B).

In order to compare these asymptotic scalings with the bounds provided in our work, while at the same time account for the unavailable bound on the time-evolution error for the TSA method, we consider here the situation where we use the optimal time-evolution scheme of Ref. [4] (which is based on qubitization [5]) and neglect the mild overhead needed to improve the precision to the desired level. Using this implementation, the number of applications of the qubiterate unitary W_Q (see Sec. III for more details) is simply $M=O(t_{\max})$.

In this work, we consider two different integral transforms. The first is associated with the Fejer kernel that is naturally produced by using the quantum phase estimation (QPE) algorithm [25] to approximate the response as described in Refs. [8,9]. The second is a Gaussian integral transform (GIT) obtained using the connection between quantum walks and Chebyshev polynomials [26]. We will analyze these integral transforms in detail in the following sections and anticipate here the main results.

Due to the choice $\beta=\Sigma=\varepsilon$, both the standard Fejer method of Ref. [8] and the qubitization-based variant from Ref. [9] have the same asymptotic scaling. We will anticipate here results for the latter, which can produce a $(\varepsilon,\Delta,\varepsilon,\eta_\beta)$-approximation using

$$M=O\left(\frac{1}{\Delta\varepsilon}\right)\quad\text{and}\quad N_S=O\left[\frac{1}{\varepsilon^2}\log_2\left(\frac{1}{\eta_\beta}\right)\right]. \tag{13}$$

Even though the sample complexity is greatly reduced, for small target errors ε the gate count of this scheme will be larger than the estimate obtained from Eq. (11). Despite this, as described in detail in Sec. III, this scheme could still be beneficial as it avoids performing an approximation to the time-evolution operator.

As we will show in more detail in Sec. IV, using the GIT provides a considerable reduction of the quantum computational cost (i.e., the gate count) compared to both methods described above. This comes at the cost of requiring a larger number of measurements N_S than the Fejer-based methods, but still less or comparable to Eq. (12). In particular, we will find that a $(\varepsilon,\Delta,\varepsilon,\eta_\beta)$-approximation to the response function requires only

$$M=O\left[\frac{1}{\Delta}\sqrt{\log_2\left(\frac{1}{\varepsilon}\right)\log_2\left(\frac{1}{\Delta\varepsilon}\right)}\right] \tag{14}$$

calls to the qubiterate unitary W_Q, together with

$$N_S=O\left\{\frac{1}{\Delta^3\varepsilon^2}\left[\log_2\left(\frac{1}{\varepsilon}\right)\log_2\left(\frac{1}{\Delta\varepsilon}\right)\right]^{3/2}\log_2\left(\frac{1}{\eta_\beta}\right)\right\} \tag{15}$$

samples. We summarize these estimates in Table I and provide a pseudocode implementation in Appendix B.

That the quantum query complexity given by Eq. (14) is almost optimal can be seen by looking at our approach as a technique to estimate the ground-state energy of some Hamiltonian as in Ref. [27]. In particular, optimality can be shown by considering a Hamiltonian with spectral gap $(\omega_1-\omega_0)>2\Delta$, an initial state $|\Psi\rangle$ with an overlap on the ground-state state $|\langle\Psi|0\rangle|\geqslant\varepsilon$, and asking for an approximation of the ground-state energy with probability $P>1-\varepsilon$ and confidence $1-\eta_\beta$. Using the results from Ref. [27] (Lemmas 3 and 5 and Theorems 8 and 9), we know that this requires at least $M=O[1/\Delta\log_2(1/\varepsilon)]$ oracle calls to W_Q. We can also solve this problem by considering a $(\varepsilon,\Delta,\varepsilon,\eta_\beta)$-approximation $\widetilde{\Phi}_K$ to the response $S_2(\omega)=\langle\Psi|\hat{\rho}(\omega)|\Psi\rangle$. Our result is then only a factor $O[\sqrt{\log_2(1/\varepsilon)\log_2(1/\Delta)}]$ away from the optimal result and provides a quadratic speedup in the logarithmic factors compared to the TSA scheme of Ref. [12].

III. FEJER KERNEL

The standard quantum phase estimation (QPE) algorithm [25,28] uses n applications of the (controlled) time-evolution unitary $U_O(t)=\exp(-it\hat{O})$ and $N=2^n$ ancilla qubits to approximately diagonalize the "Hamiltonian" operator $\hat{O}$ [29]. As we proposed in Ref. [8], this technique can be used to perform an integral transform generated by a rescaled Fejer kernel,

$$K_F(\sigma_q,\omega,N)=\frac{1}{N^2}\frac{\sin^2[N\pi(\sigma_q-\omega)/2]}{\sin^2[\pi(\sigma_q-\omega)/2]}, \tag{16}$$

where the discrete frequencies σ_q are defined on a grid with N points: $\sigma_k=(2k/N)-1$ for $k=\{0,\ldots,N-1\}$. The integer parameter $N>1$ controls the maximum propagation time $t_{\max}$ used in QPE as $t_{\max}=\pi N$. In this case, ensuring the resulting integral transform $\Phi_F(\omega)$ is Σ-accurate with resolution Δ is equivalent to requiring the probability of measuring a phase σ_k with error more than Δ to be less than Σ. This probability

TABLE I. Comparison of the computational cost required to obtain a $(\varepsilon, \Delta, \varepsilon, \eta_\beta)$-approximation to the response function using the time-series analysis (TSA) method of Ref. [12], the Fejer-based methods from Refs. [8,9], and the GIT-based method proposed in this work. See, also, Appendix B for the asymptotic scaling in a different limit.

Method	No. calls to W_Q	Total no. samples
TSA	$O\left[\frac{1}{\Delta}\log_2\left(\frac{1}{\varepsilon}\right)^2\right]$	$O\left[\frac{1}{\Delta^3\varepsilon^2}\log_2\left(\frac{1}{\varepsilon}\right)^6\log_2\left(\frac{1}{\eta_\beta}\right)\right]$
Fejer	$O\left(\frac{1}{\Delta\varepsilon}\right)$	$O\left[\frac{1}{\varepsilon^2}\log_2\left(\frac{1}{\eta_\beta}\right)\right]$
GIT	$O\left[\frac{1}{\Delta}\sqrt{\log_2\left(\frac{1}{\varepsilon}\right)\log_2\left(\frac{1}{\Delta\varepsilon}\right)}\right]$	$O\left\{\frac{1}{\Delta^3\varepsilon^2}\left[\log_2\left(\frac{1}{\varepsilon}\right)\log_2\left(\frac{1}{\Delta\varepsilon}\right)\right]^{3/2}\log_2\left(\frac{1}{\eta_\beta}\right)\right\}$

can be bounded using standard techniques (see, e.g., 5.2.1 of [30]) as

$$P\left(\left|k-\frac{N}{2}(\omega+1)\right|>\frac{N\Delta}{2}\right)\leqslant\frac{1}{N\Delta-2}, \tag{17}$$

which then implies we can take the closest power of 2 of

$$N\geqslant\frac{1}{\Delta}\left(\frac{1}{\Sigma}+2\right), \tag{18}$$

in order to satisfy Eq. (4). The dependence on the resolution Δ is already optimal and the constant factors could be improved using optimized preparations of the ancilla register [28]. The scaling with the error Σ instead could be improved to $N=O[\log_2(1/\Sigma)1/\Delta]$ in the special situation where the signal $S(\omega)$ is composed of a single-frequency mode by using schemes such as Kitaev's original algorithm [31] or the more efficient Iterative Phase Estimation Algorithm [32]. In the general case where the number of modes in the response of Eq. (2) satisfies $\Gamma\gg 1$, this is not, in general, possible (see, e.g., [33]). We can now use $N_S=O(1/\beta^2)$ samples to produce the β-approximate estimator $\widetilde{\Phi}_F$ by collecting a histogram of the measured frequencies. More precisely, using Hoeffding's inequality [34], we find it sufficient to take

$$N_S=\frac{1}{2\beta^2}\log_2\left(\frac{2}{\eta_\beta}\right), \tag{19}$$

with η_β the confidence of the β-approximation.

In general, the time-evolution operator $U_O(t)$ needs to be approximated with additive error δ_t, using available quantum operations, and a proper consideration of this approximation error is critical for a fair assessment of the overall computational cost. As discussed in Sec. II, we will consider these errors as contributions to the total variation given by Eq. (5), which define the β-approximation. In particular, if we denote by $\Phi_F^e(\omega)$ the transform obtained by using the approximate time-evolution unitary and $\widetilde{\Phi}_F^e(\omega)$ its finite population estimator, we have

$$\delta_V\left(\Phi_F,\widetilde{\Phi}_F^e\right)\leqslant\delta_V\left(\Phi_F,\Phi_F^e\right)+\delta_V\left(\Phi_F^e,\widetilde{\Phi}_F^e\right). \tag{20}$$

The second term measures statistical fluctuations and can be dealt with using, again, the Hoeffding bound, for the first term instead in Appendix C, we show that

$$\delta_V\left(\Phi_F,\Phi_F^e\right)\leqslant\log_2(N)\delta_t, \tag{21}$$

with δ_t an upper bound to the approximation error of the time-evolution operator for times up to $t_{\max}=\pi N$. The finite-population estimator of the approximate Fejer transform is then β-accurate with confidence η_β if

$$N_S=\frac{2}{\beta^2}\log_2\left(\frac{2}{\eta_\beta}\right),\quad \delta_t\leqslant\frac{\beta}{2\log_2(N)}. \tag{22}$$

Using optimal scaling algorithms for time evolution such as quantum signal processing [4], the total gate count is

$$M=O\left[\frac{1}{\Delta\Sigma}+\log_2\left(\frac{1}{\beta}\right)\right], \tag{23}$$

in terms of oracle queries to the basic quantum subroutine: the qubiterate W_Q. This unitary is defined as

$$W_Q=\exp[i\widetilde{Y}\arccos(\hat{O})], \tag{24}$$

where $\widetilde{Y}$ is an isometry defined over a two-dimensional space for each energy eigenvalue (see [4,5] and the Appendix of Ref. [10] for a complete derivation). The most important property of W_Q for our purposes is that it can be implemented exactly and efficiently. It is important to point out that short-time approximation methods based on the Trotter-Suzuki [35] expansion are not able to achieve the optimal scaling in Eq. (23).

A slight modification to this scheme, with the same scaling but possibly greatly reduced prefactors, can be obtained by applying the QPE algorithm directly on the qubiterate W_Q (see Ref. [9]). One can easily show that this leads to a modified Fejer kernel given by

$$K_{FQ}(\sigma_q,\omega,N)=\frac{K_F(\sigma_q,\theta_\omega,N)+K_F(\sigma_q,-\theta_\omega,N)}{2}, \tag{25}$$

where we have defined $\cos(\theta_\omega)=\omega$. In order to distinguish the two peaks at $\pm\theta_\omega$, we can shift and rescale the excitation operator $\hat{O}$ so that its spectrum lies in [0,1] only. The needed resolution in this transformed space (apart from the trivial factor of 1/2 coming from the rescaling) will need to satisfy

$$|\cos(\theta_\omega\pm\Delta_\theta)-\cos(\theta_\omega)|\leqslant\frac{\Delta}{2}, \tag{26}$$

which amounts to require $\Delta_\theta\leqslant\sqrt{1+\Delta}-1$. We then find that in order to obtain a $(\Sigma,\Delta,\beta,\eta_\beta)$-approximation to the response function, the qubitization-based Fejer transform of Ref. [9] requires the closest power of 2,

$$M\geqslant\frac{2}{\Delta_\theta}\left(\frac{1}{\Sigma}+2\right)\gtrsim\frac{4}{\Delta}\left(\frac{1}{\Sigma}+2\right), \tag{27}$$

black box invocations of the qubiterate W_Q [36], together with the same number of samples reported in Eq. (19). Despite the possible slight increase in oracle calls with respect to the time-evolution-based Fejer scheme presented before, by avoiding

the overhead in approximating the time-evolution operator $U_O(t)$, we expect this strategy to require shorter circuit depths and, at the same time, less cumbersome controlled operations.

In the next section, we consider algorithms with exponentially better dependence on Σ.

IV. GAUSSIAN KERNEL

We now consider a Gaussian integral transform (GIT) defined through the following kernel function:

$$K_G(\sigma,\omega,\Lambda)=\frac{1}{\sqrt{2\pi}\,\Lambda}\exp\left[-\frac{(\sigma-\omega)^2}{2\Lambda^2}\right], \tag{28}$$

where $\Lambda>0$ controls the resolution, and the transformed frequency σ is defined over the whole real line [37]. The first step is to determine the conditions for which the approximate response obtained using the GIT is Σ-accurate with resolution Δ. Using the translational invariance of the kernel $K_G(\sigma,\omega,\Lambda)$ for $\sigma\in\mathbb{R}$, we can rewrite the condition given by Eq. (4) in terms of the error function as

$$\frac{1}{\sqrt{2\pi}\,\Lambda}\int_{-\Delta}^{\Delta}d\sigma\,\exp\left(-\frac{\sigma^2}{2\Lambda^2}\right)=\operatorname{erf}\left(\frac{\Delta}{\sqrt{2}\Lambda}\right)\geqslant 1-\Sigma. \tag{29}$$

A sufficient condition for this to hold is to choose the kernel resolution Λ according to

$$\Lambda\leqslant\frac{\Delta}{\sqrt{2\log_2(1/\Sigma)}}. \tag{30}$$

We now move on to find the condition for the GIT to be β-approximate with confidence η_β according to Eq. (5). As we mentioned in Sec. I, this property is directly connected with the specific implementation of the GIT and the way we estimate it. Here we consider an approximate implementation of the Gaussian kernel $K_G(\sigma,\omega,\Lambda)$ using an expansion in a set of orthogonal polynomials. Due to its direct connection with quantum walks [26,38] and the qubitization method [5], we consider here the basis spanned by the Chebyshev polynomials T_k. In particular, one can show that if we indicate with $|G\rangle$ the flag state in the ancilla register used for the block encoding of the excitation operator $\langle G|W_Q|G\rangle=\hat{O}$ and entering in the definition of the qubiterate W_Q, we have (see the proof of Lemma 16 of Ref. [26] and Appendix D of Ref. [38])

$$W_Q^k|G\rangle\otimes|\Psi\rangle=|G\rangle\otimes T_k(\hat{O})|\Psi\rangle+|\Phi^\perp\rangle, \tag{31}$$

with $|\Psi\rangle$ the initial state that defines the response function $S(\omega)$ in Eq. (2) and $|\Phi^\perp\rangle$ not normalized and orthogonal to the flag state $|G\rangle$. The expectation value of the kth Chebyshev polynomial can then be obtained as

$$\langle\Psi|T_k(\hat{O})|\Psi\rangle=\langle\Psi_G|W_Q^k|\Psi_G\rangle, \tag{32}$$

where, for convenience, we have defined $|\Psi_G\rangle:=|G\rangle\otimes|\Psi\rangle$. Note that this procedure is deterministic since we are computing a single polynomial at a time. An exact representation for the GIT can be obtained considering first the series expansion of the Gaussian function,

$$\exp\left(-\frac{\omega^2}{2\Lambda^2}\right)=\sum_{k=0}^{\infty}a_k(\Lambda)T_k(\omega), \tag{33}$$

and then expanding the integral kernel as

$$\begin{aligned}K_G(\sigma,\omega,\Lambda)&=\frac{1}{\sqrt{2\pi}\,\Lambda}\sum_{k=0}^{\infty}a_k\left(\frac{\Lambda}{2}\right)T_k\left(\frac{\sigma-\omega}{2}\right)\\&=\sum_{k=0}^{\infty}c_k(\Lambda,\sigma)T_k(\omega).\end{aligned} \tag{34}$$

The step leading to the second line is necessary to be able to implement the GIT using qubitization, and the new expansion coefficients c_k can be obtained from the bare a_k and polynomials in σ. Explicit expressions for these coefficients can be found in Eq. (D7) of Appendix D.

In order for this to be useful, we need to truncate the series given by Eq. (34) at some finite order L. This leads to an approximate kernel function,

$$K_{GL}(\sigma,\omega,\Lambda)=K_G(\sigma,\omega,\Lambda)-R_L(\sigma,\omega,\Lambda), \tag{35}$$

where we have defined R_L to be the approximation error. We note in passing that such truncated expansions of the kernel function are routinely used to perform reasonable inversions of the Lorentzian kernel by neglecting the error term R_L as a way of performing a regularization to the ill-posed problem [39]. The final approximate integral transform $\Phi_{GL}(\omega)$ is then obtained as

$$\Phi_{GL}(\omega)=\sum_{k=0}^{L}c_k(\Lambda,\sigma)\langle\Psi_G|W_Q^k|\Psi_G\rangle. \tag{36}$$

As described in Sec. II, the approximation error contributes to the total variation given by Eq. (5) similarly to the approximation error of the time-evolution operator for the simpler Fejer transform. As we did in Sec. III, we can decompose the total variation as

$$\begin{aligned}\delta_V(\Phi_G,\widetilde{\Phi_{GL}})&\leqslant\delta_V(\Phi_G,\Phi_{GL})+\delta_V(\Phi_{GL},\widetilde{\Phi_{GL}})\\&\leqslant R_L(\sigma,\omega,\Lambda)+\delta_V(\Phi_{GL},\widetilde{\Phi_{GL}}),\end{aligned} \tag{37}$$

where Φ_G is the exact GIT, Φ_{GL} is the approximate integral transform obtained by truncating the series in Eq. (34) at order L, and $\widetilde{\Phi_{GL}}$ is its finite-population estimator. As we did for the Fejer kernel above, we will now require that both error terms be less than $\beta/2$ with confidence η_β. In order to bound the total statistical error of the finite-population estimator $\widetilde{\Phi_{GL}}$ of the GIT in Eq. (36), and assuming for simplicity the same number of measurements for each one of the L expectation values in the expansion, we can take a number of samples given by

$$\begin{aligned}N_S&=2L\log_2\left(\frac{2}{\eta_\beta}\right)\max_{k=\{0,\ldots,L\}}\left(L\frac{|c_k|}{\beta}\right)^2\\&\lesssim 2L^3\left(1+\frac{2.2}{\beta}\right)^2\log_2\left(\frac{2}{\eta_\beta}\right),\end{aligned} \tag{38}$$

where we used the upper bound on Eq. (D13) on c_k obtained in Appendix D. Note that for technical reasons explained in Appendix D, this is valid only after rescaling the operator $\hat{O}$ by a factor of 2. Since it is possible to also find an appropriate bound in the general case, we do not correspondingly rescale the resolution Δ here.

The rest of this section will be dedicated to determine an appropriate value for L to ensure $R_L \leqslant \beta/2$.

In order to find optimal truncation schemes, it is now convenient to distinguish between two different situations depending on the desired value β as a function of Σ and the resolution Δ. More precisely, if we define two critical values β_U and β_L as follows:

$$\beta_L = \frac{1}{\Sigma}\exp\left(-\frac{1}{\Delta^2}\right), \quad \beta_U = \frac{1}{\Delta}\sqrt{\frac{\log_2(1/\Sigma)}{2}}, \tag{39}$$

we will try to optimally truncate the polynomial expansion in Eq. (34) in two regimes, i.e., the asymptotic regime $\beta \leqslant \beta_L$ and the intermediate regime where the target accuracy satisfies $\beta_L \leqslant \beta \leqslant \beta_U$. Note that the convention we chose in Sec. II is compatible with the latter.

As we show in detail in Appendix D [see Eqs. (D21) and (D30)], we can ensure a truncation error $R_L \leqslant \beta/2$ by choosing the maximum order L according to the following:

(i) In the asymptotic regime $\beta \leqslant \beta_L$, we need

$$L_{\text{asy}} = \left\lceil \frac{2e}{\Delta^2}\log_2\left(\frac{1}{\Sigma}\right) + g_a\left(\frac{\Delta^2}{e}\frac{\log_2(6.8/\beta)}{\log_2(1/\Sigma)}\right)\right\rceil - 2, \tag{40}$$

where, for convenience, we have introduced the function $g_a(x) = x/W(x)$, with W the Lambert W function [40] (see, also, Appendix D for details).

(ii) In the intermediate regime $\beta_L \leqslant \beta \leqslant \beta_U$, we need

$$L_{\text{int}} = \left\lceil \frac{\alpha_1}{\Delta}\sqrt{\log_2\left(\frac{1}{\Sigma}\right) g_i\left[\frac{\alpha_2}{\Delta\beta}\log_2\left(\frac{1}{\Sigma}\right)\right]}\right\rceil - 1, \tag{41}$$

with $\alpha_1 \lesssim 2.93$, $\alpha_2 \lesssim 4.14$, while the function g_i is

$$g_i(x) = \log_2(x) - \tfrac{1}{4}\log_2[\log_2(x^2)]. \tag{42}$$

As is apparent from the definition of the truncated GIT Φ_{GL} in Eq. (36), this is the maximum required number of invocations to the qubiterate W_Q in a single run since the L expectation values can be computed in parallel. In order to have a better understanding of these results and connect to the discussion in Sec. II, we can write these estimates in terms of asymptotic scaling as

$$L_{\text{asy}} = O\left\{\frac{1}{\Delta^2}\log_2\left(\frac{1}{\Sigma}\right) + \frac{\log_2(1/\beta)}{\log_2[\log_2(1/\beta)]}\right\}, \tag{43}$$

for the regime with $\beta \leqslant \beta_L$; in the second regime case with $\beta_L \leqslant \beta \leqslant \beta_U$, we instead find

$$L_{\text{int}} = O\left\{\frac{1}{\Delta}\sqrt{\log_2\left(\frac{1}{\Sigma}\right)\log_2\left[\frac{1}{\Delta\beta}\log_2\left(\frac{1}{\Sigma}\right)\right]}\right\}. \tag{44}$$

Note that in applications of the GIT scheme, the concrete values for the truncation order L provided above can be much more useful than the looser bounds given by Eq. (44). Finally note that as mentioned in Sec. II, the asymptotic regime $\beta \leqslant \beta_L$ is possibly not directly relevant for the approximate estimation of observables of the form given by Eq. (6), but could still be helpful in different scenarios. The same argument holds for the ability of the TSA method of Ref. [12] to achieve $\Sigma = 0$ directly.

V. SUMMARY AND CONCLUSIONS

In this work, we have studied a family of quantum algorithms for the approximate estimation of the spectral-density operator using ideas from integral transform methods and applied it to the problem of estimating, with bounded errors, the dynamical response function $S(\omega)$ from linear response theory. In particular, we find it useful to consider an integral transform defined by a Gaussian kernel, i.e., the Gaussian integral transform (GIT). This is in line with the success enjoyed by another integral transform whose kernel is a representation of the delta function, i.e., the Lorentz integral transform (LIT) [16].

Recently, Somma introduced an algorithm to evaluate multiple eigenvalues based on a time-series analysis [12]. We show that this technique can be understood in the general framework of integral transform methods introduced here. By comparing it with our GIT, we found a quadratic improvement in the regime of interest for the response function approximation problem that we are interested in. Notably, our scheme also uses potentially much simpler unitary operations as it completely avoids the need to simulate time evolution under a Hamiltonian. This will be important in applications of the GIT-based algorithm on NISQ devices. To help with implementation of the method, together with pseudocode implementation of these algorithms in Appendix B, we also provide concrete values for the constant factors of all the quantities needed in the practical design of the algorithm.

A possible extension of our algorithm for applications in future fault-tolerant devices is to reduce the sample complexity by employment of techniques, such as the method of Ref. [41] (which uses QPE and amplitude amplification), to estimate the expectation values in Eq. (32) at the expense of longer circuit depths. Another interesting possibility is to use either Quantum signal processing [5] or the Linear Combination of Unitaries method [26] to directly implement the approximate spectral density,

$$\hat{\rho}_K(\omega) = K(\omega, \hat{O}) = \sum_k^{\Gamma} K(\omega, O_k)|k\rangle\langle k|. \tag{45}$$

This would allow one, together with amplitude amplification, to selectively prepare the final states of scattering processes within a predetermined energy window allowing the application of the algorithm proposed in Ref. [8] to study rare processes. In such applications, the algorithm ceases to be deterministic and a detailed analysis of the failure probability would be needed.

The same strategy can of course be used as a near optimal state preparation scheme similar in many ways to the one recently proposed in Ref. [27]. Finally, the general framework introduced in this work, and the accuracy metrics defined in Sec. II, could prove useful to devise alternative approximation schemes based on integral transforms. The interesting question of whether the Gaussian provides the optimal integral kernel for these approximation is left for future work.

Note added. Recently, we became aware of a similar work by Rall [42] where an interesting construction for a polynomial representation of a window function was proposed. As we show in Appendix E, one can use this result to obtain an

algorithm for approximating the spectral density with a query complexity, $O[\frac{1}{\Delta}\log_2(\frac{1}{\Sigma\Delta})]$. This is an improvement over the scaling of the TSA method by Somma [12], but not quite as efficient as the GIT-based method proposed here.

ACKNOWLEDGMENTS

I want to thank N. Wiebe for useful comments. This work was supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research (ASCR) quantum algorithm teams program, under field work Proposal No. ERKJ333 and by the Institute for Nuclear Theory under U.S. Department of Energy Grant No. DE-FG02-00ER41132.

APPENDIX A: ERROR BOUND FOR FREQUENCY OBSERVABLES

As in the main text, we start with a response function

$$S(\omega)=\sum_k^\Gamma \alpha_k\delta(\omega-\omega_k),\qquad \sum_k^\Gamma \alpha_k=1, \tag{A1}$$

where we also have $\alpha_k>0$. This decomposition follows directly from the spectral representation reported in Eq. (2) of the main text. We also define an observable Q which generalizes sum rules as the integral,

$$Q(S,f)=\int_{-1}^{1} d\omega S(\omega)f(\omega). \tag{A2}$$

If we use a $(\Sigma,\Delta,\beta,\eta_\beta)$-approximation $\widetilde{\Phi}(\omega)$ to the response $S(\omega)$ obtained using a β-approximate estimator for the integral transform $\Phi(\omega)$, we want to find an upper bound for the total error,

$$\delta_Q(\widetilde{\Phi})=|Q(S,f)-Q(\widetilde{\Phi},f)|, \tag{A3}$$

where the approximate observables are expressed as

$$Q(\widetilde{\Phi},f)=\iint d\nu\widetilde{\Phi}(\nu). \tag{A4}$$

Using the triangle inequality, we find

$$\begin{aligned}\delta_Q(\widetilde{\Phi})&=\left|\int_{-1}^{1} d\omega S(\omega)f(\omega)-\iint d\nu\widetilde{\Phi}(\nu)f(\nu)\right|\\&\leqslant\left|\int_{-1}^{1} d\omega S(\omega)f(\omega)-\iint d\nu\Phi(\nu)f(\nu)\right|\\&\quad+\left|\iint d\nu[\Phi(\nu)-\widetilde{\Phi}(\nu)]f(\nu)\right|\\&=\delta_Q(\Phi)+\left|\iint d\nu[\Phi(\nu)-\widetilde{\Phi}(\nu)]f(\nu)\right|.\end{aligned} \tag{A5}$$

In order to find a bound for the first term, note that thanks to the spectral representation given by Eq. (A1), we can decompose the total integral transform in the sequence

$$\Phi(\omega)=\sum_k^\Gamma \alpha_k\Phi_k(\omega), \tag{A6}$$

with Φ_K the transform of a single peaked response, $S_k(\omega)=\delta(\omega-\omega_k)$. We can therefore write

$$Q(S,f)=\sum_k^\Gamma \alpha_k\int_{-1}^{1} d\omega S_k(\omega)f(\omega)=\sum_k^\Gamma \alpha_k f(\omega_k), \tag{A7}$$

while, for the integral transform approximator,

$$\begin{aligned}Q(\Phi,f)&=\sum_k^\Gamma \alpha_k\iint d\nu\Phi_k(\nu)f(\nu)\\&=\sum_k^\Gamma \alpha_k\iint d\nu\int d\omega K(\nu,\omega)S(\omega)f(\nu)\\&=\sum_k^\Gamma \alpha_k\iint d\nu K(\nu,\omega_k)f(\nu),\end{aligned} \tag{A8}$$

where, in the last line, we performed the frequency integral using the decomposition given by Eq. (A1). Using the definition (4) of a Σ-accurate kernel with resolution Δ, we can find a bound for the first term in Eq. (A5) as follows:

$$\begin{aligned}\delta_Q(\Phi)&=\left|\sum_k^\Gamma \alpha_k\left[f(\omega_k)-\iint d\nu K(\nu,\omega_k)f(\nu)\right]\right|\\&\leqslant\left|\sum_k^\Gamma \alpha_k\left[f(\omega_k)-\iint_{\omega_k-\Delta}^{\omega_k+\Delta} d\nu K(\nu,\omega_k)f(\nu)\right]\right|\\&\quad+\left|\sum_k^\Gamma \alpha_k\left[\iint_{\omega_k+\Delta} d\nu K(\nu,\omega_k)f(\nu)\right.\right.\\&\quad\left.\left.-\iint^{\omega_k-\Delta} d\nu K(\nu,\omega_k)f(\nu)\right]\right|\\&\leqslant\left|\sum_k^\Gamma \alpha_k\left[f(\omega_k)-\iint_{\omega_k-\Delta}^{\omega_k+\Delta} d\nu K(\nu,\omega_k)f(\nu)\right]\right|\\&\quad+f_{\max}\Sigma\\&\leqslant\left|\sum_k^\Gamma \alpha_k f(\omega_k)\left[1-\iint_{\omega_k-\Delta}^{\omega_k+\Delta} d\nu K(\nu,\omega_k)\right]\right|\\&\quad+\left|\sum_k^\Gamma \alpha_k\iint_{\omega_k-\Delta}^{\omega_k+\Delta} d\nu K(\nu,\omega_k)[f(\nu)-f(\omega_k)]\right|\\&\quad+f_{\max}\Sigma\\&\leqslant\left|\sum_k^\Gamma \alpha_k\iint_{\omega_k-\Delta}^{\omega_k+\Delta} d\nu K(\nu,\omega_k)[f(\nu)-f(\omega_k)]\right|\\&\quad+2f_{\max}\Sigma\\&\leqslant f_{\max}^\Delta\left[\sup_{\omega\in[-1,1]}\left|\iint_{\omega-\Delta}^{\omega+\Delta} d\nu K(\nu,\omega)\right|\right]+2f_{\max}\Sigma,\end{aligned}$$

with $f_{\max}\geqslant|f(\omega)|$ for all $\omega\in[-1,1]$ and

$$f_{\max}^\Delta=\sup_{\omega\in[-1,1]}\sup_{x\in[-\Delta,\Delta]}|f(\omega+x)-f(\omega)|. \tag{A9}$$

Algorithm 1. Fejer-based approximator.

1: Given integers $M = 2^m$ and N_S,
2: **for** $i = 1$ to N_S, **do:**
3: prepare target state $|\Psi\rangle$,
4: apply QPE with unitary V and maximum order $V^{M/2}$,
5: measure m qubits in ancilla register in frequency ν_i,
6: add result to frequency histogram
7: Return frequency histogram

Finally, the second term in Eq. (A5) is bounded as

$$\left|\sum\!\!\!\!\!\!\int d\nu[\Phi(\nu) - \widetilde{\Phi}(\nu)]f(\nu)\right| \leqslant \beta \int_{-1}^{1} d\omega |f(\omega)|. \tag{A10}$$

Bringing it all together, we can finally prove the upper bound,

$$\delta_Q \leqslant f_{\max}^{\Delta} + 2f_{\max}\Sigma + \beta \int_{-1}^{1} d\omega |f(\omega)|. \tag{A11}$$

APPENDIX B: PSEUDOCODE IMPLEMENTATION

We present here a pseudocode implementation for the spectral-density estimation algorithms that we discuss in the main text. The goal of our algorithm is to return a $(\Sigma, \Delta, \beta, \eta_\beta)$-approximation $\widetilde{\Phi}_k(\nu)$ to the response function $S(\omega) = \langle\Psi|\hat{\rho}(\omega)|\Psi\rangle$ at a single frequency point ν. This is reasonable since we might not want to estimate $\widetilde{\Phi}_k(\nu)$ on a whole grid composed of the maximal number $O(1/\Delta)$ of frequency points (as done instead in [12]).

This is, however, not possible with the Fejer-based method of Sec. III since the transformed frequencies η are *sampled* from the distribution $\widetilde{\Phi}_k(\nu)$ instead. For this reason, we provide two independent implementations.

For the Fejer-based strategies, we use Algorithm 1 with

$$M = O\left(\frac{1}{\Delta\Sigma}\right), \quad N_S = O\left[\frac{1}{\beta^2}\log_2\left(\frac{1}{\eta_\beta}\right)\right], \tag{B1}$$

while $V \equiv \exp(-i2\pi\hat{O})$ for the time-dependent method and $V = W_Q$ for the qubitization-based method.

In the case of either the TSA algorithm or the GIT-based method, we can use Algorithm 2 instead, with a maximum order M given by

$$M_{\rm TSA} = O\left[\frac{1}{\Delta}\log_2\left(\frac{1}{\beta}\right)^2\right], \tag{B2}$$

Algorithm 2. Orthogonal polynomial-based approximator.

1: Given integers M and $N_S = M \times N$,
2: **for** $k = 1$ to M, **do**
3: **for** $i = 1$ to N, **do**
4: prepare target state $|\Psi\rangle$,
5: measure expectation value $v_k = \langle\Psi|V^k|\Psi\rangle$,
6: store estimator of v_k with error $O(1/\sqrt{N})$
7: Compute expansion coefficients $\vec{c}(\nu)$ corresponding to the integral transform being evaluated at target frequency ν
8: Return $\widetilde{\Phi}_k(\nu) = \vec{c}(\nu)\cdot\vec{v}$

for the TSA algorithm of [12], while for the GIT,

$$M_{\rm GIT} = O\left\{\frac{1}{\Delta}\sqrt{\log_2\left(\frac{1}{\Sigma}\right)\log_2\left[\frac{1}{\Delta\beta}\log_2\left(\frac{1}{\Sigma}\right)\right]}\right\}, \tag{B3}$$

together with a number of samples per order scaling as

$$N = O\left[\frac{M^2}{\beta^2}\log_2\left(\frac{1}{\eta_\beta}\right)\right]. \tag{B4}$$

Finally, if we want the transform $\widetilde{\Phi}_k(\nu)$ at all the $O(1/\Delta)$ frequency points while keeping the *total* error β (as considered in [12]), we will need instead

$$M_{\rm TSA} = O\left[\frac{1}{\Delta}\log_2\left(\frac{1}{\Delta\beta}\right)^2\right], \tag{B5}$$

and

$$M_{\rm GIT} = O\left\{\frac{1}{\Delta}\sqrt{\log_2\left(\frac{1}{\Sigma}\right)\log_2\left[\frac{1}{\Delta^2\beta}\log_2\left(\frac{1}{\Sigma}\right)\right]}\right\}, \tag{B6}$$

respectively. In this case, we see that the logarithmic term for TSA also contains the resolution scale Δ.

APPENDIX C: ERROR ANALYSIS FOR FAULTY IMPLEMENTATION OF FEJER

Assume we have an approximation to the phase kickback (PKB) part of the QPE algorithm (the application of the controlled-U operations), which satisfies

$$\|\widetilde{V}_{\rm PKB} - V_{\rm PKB}\| \leqslant \delta. \tag{C1}$$

Define $|\Phi_A\rangle = V_{\rm PKB}|\Psi\rangle$; without loss of generality, we have

$$|\Phi_B\rangle = \widetilde{V}_{\rm PKB}|\Psi\rangle = \cos(\alpha)|\Phi_A\rangle + \sin(\alpha)|\xi\rangle, \tag{C2}$$

with $\langle\Phi_A|\xi\rangle = 0$. If we introduce the density matrices $\rho = |\Phi_A\rangle\langle\Phi_A|$ and $\sigma = |\Phi_B\rangle\langle\Phi_B|$, we can now write

$$\sqrt{2[1-\cos(\alpha)]}\||\Phi_A\rangle - |\Phi_B\rangle\|_2 \leqslant \|\widetilde{V}_{\rm PKB} - V_{\rm PKB}\| \leqslant \delta.$$

Then, $\cos(\alpha) \geqslant 1 - \delta^2/2$, but also

$$D(\rho,\sigma) = \frac{1}{2}{\rm Tr}[|\rho-\sigma|] = |\sin(\alpha)| \leqslant \delta\sqrt{1-\frac{\delta^2}{4}}. \tag{C3}$$

We can write the transform at location $\sigma_q = 2q/N - 1$ as

$$\Phi_F(\sigma_q,\Delta,N) = \Phi_F(\sigma_q) = {\rm Tr}[\Pi_q U_{\rm QFT}^\dagger \rho U_{\rm QFT}], \tag{C4}$$

where $\Pi_q = |q\rangle\langle q|$ and $U_{\rm QFT}$ is the unitary implementing the quantum Fourier transform on the ancilla register. A similar expression holds for the faulty density matrix σ. We now have, for any $0 \leqslant q < N$, that

$$\begin{aligned}\delta_V(\Phi_F,\widetilde{\Phi}_F) &\leqslant \sup_{\|\widetilde{V}_{\rm PKB}-V_{\rm PKB}\|\leqslant\delta} \left|\Phi_F(\sigma_q) - \widetilde{\Phi}_F(\sigma_q)\right| \\ &= |{\rm Tr}[\Pi_q U_{\rm QFT}^\dagger(\rho-\sigma)U_{\rm QFT}]| \\ &\leqslant \frac{1}{2}{\rm Tr}[|\rho-\sigma|] \leqslant \delta\sqrt{1-\frac{\delta^2}{4}} \leqslant \delta.\end{aligned} \tag{C5}$$

Furthermore, since V_{PKB} is a product of $n = \log_2(N)$ controlled time-evolution unitaries,

$$V_{\mathrm{PKB}} = \prod_{k=0}^{n-1} U(t = 2\pi 2^k), \tag{C6}$$

we have, by the union bound, that

$$\begin{aligned}\|\widetilde{V}_{\mathrm{PKB}} - V_{\mathrm{PKB}}\| &\leqslant \sum_{k=0}^{n-1} \|U(2\pi 2^k) - \widetilde{U}(2\pi 2^k)\| \\ &\leqslant n \max_{0\leqslant<k<n} \delta_t(2\pi 2^k).\end{aligned} \tag{C7}$$

In the last equation, $\delta_t(\tau)$ is the approximation error of the time-evolution unitary for total time $t = \tau$. If we choose all approximation errors to be the same, $\delta_t(\tau) = \delta_t$, then we find

$$\delta_V(\Phi_F, \widetilde{\Phi_F}) \leqslant \log_2(N)\delta_t. \tag{C8}$$

APPENDIX D: CHEBYSHEV EXPANSION OF THE GAUSSIAN KERNEL

Using an expansion in Chebyshev polynomials, we can express the Gaussian function as

$$\exp\left(-\frac{x^2}{2\Lambda^2}\right) = \sum_{n=0}^{L} a_n(\Lambda)T_n(x) + r_L(x, \Lambda), \tag{D1}$$

where $r_L(\Lambda)$ indicates the truncation error and the coefficients are given by (see Eq. (4) of Ref. [43])

$$a_n = \begin{cases} \frac{\gamma_n}{\sqrt{2\pi}\Lambda} i^{\frac{n}{2}} \exp\left(-\frac{1}{4\Delta^2}\right) J_{n/2}\left(\frac{i}{4\Lambda^2}\right) & \text{for even } n \\ 0 & \text{for odd } n, \end{cases} \tag{D2}$$

with $\gamma_0 = 1$ and $\gamma_{n>0} = 2$, and J_n the Bessel function of order n. Before discussing bounds on the magnitude of the truncation error, we want to first discuss how the kernel function $K_G(\sigma, \omega, \Lambda)$ can be generated using the expansion above. First note that we can write the truncated kernel function as in Eq. (34) of the main text,

$$K_{GL}(\sigma, \omega, \Lambda) = \frac{1}{\sqrt{2\pi}\Lambda} \sum_{k=0}^{L} a_k\left(\frac{\Lambda}{2}\right) T_k\left(\frac{\omega - \sigma}{2}\right). \tag{D3}$$

Since $T_k(\frac{\omega-\sigma}{2})$ is a polynomial of degree $k \leqslant L$, we have

$$T_k\left(\frac{\omega - \sigma}{2}\right) = \sum_{j=0}^{L} b_{jk}(\sigma)T_j(\omega), \tag{D4}$$

where the expansion coefficients are given by

$$\begin{aligned} b_{jk}(\sigma) &= \frac{\gamma_j}{\pi} \int_{-1}^{1} \frac{dx}{\sqrt{1-x^2}} T_k\left(\frac{x-\sigma}{2}\right) T_j(x) \\ &= \frac{\gamma_j}{L} \sum_{m=0}^{L-1} T_k\left(\frac{x_m - \sigma}{2}\right) T_j(x_m). \end{aligned} \tag{D5}$$

In the second line, we used the Gauss-Chebyshev quadrature and $x_m = \cos(\pi \frac{2m-1}{2L})$, the Chebyshev nodes (this is similar to the strategy used in Ref. [43]). Using this representation, we can rewrite the kernel function as

$$K_{GL}(\sigma, \omega, \Lambda) = \sum_{j=0}^{L} c_j(\Lambda, \sigma)T_j(\omega), \tag{D6}$$

where the new expansion coefficients are given by

$$c_j = \frac{\gamma_j}{\sqrt{2\pi}\Lambda L} \sum_{m=0}^{L-1} \sum_{k=0}^{L} a_k\left(\frac{\Lambda}{2}\right) T_k\left(\frac{x_m - \sigma}{2}\right) T_j(x_m). \tag{D7}$$

1. Bound of expansion coefficients

We can bound the magnitude of c_j as follows:

$$\begin{aligned} |c_j| &= \left| \frac{\gamma_j}{\sqrt{2\pi}\Lambda L} \sum_{m=0}^{L-1} \left[e^{-\frac{(x_m-\sigma)^2}{2\Lambda^2}} - r_L\left(x_m, \frac{\Lambda}{2}\right)\right] T_j(x_m) \right| \\ &\leqslant \frac{\gamma_j}{\sqrt{2\pi}\Lambda} r_L\left(\frac{\Lambda}{2}\right) + \frac{\gamma_j}{\sqrt{2\pi}\Lambda L} \sum_{m=0}^{L-1} e^{-\frac{(x_m-\sigma)^2}{2\Lambda^2}} \\ &:= \gamma_j R_L(\Lambda) + \Omega_j, \end{aligned} \tag{D8}$$

with $R_L(\Lambda)$ the truncation error of the kernel function [cf. Eq. (35)]. For the second term, we can use

$$\begin{aligned} \Omega_j &\leqslant \frac{\gamma_j}{\sqrt{2\pi}\Lambda L} \int_0^L dx \exp\left\{-\frac{\left[\cos\left(\pi\frac{2x-1}{2L}\right) - \sigma\right]^2}{2\Lambda^2}\right\} \\ &= \frac{\gamma_j}{\sqrt{2\pi^3}\Lambda} \int_0^\pi dy \exp\left\{-\frac{\left[\cos\left(y - \frac{\pi}{2L}\right) - \sigma\right]^2}{2\Lambda^2}\right\}. \end{aligned} \tag{D9}$$

The integral approximately measures the number of of Chebyshev nodes within the envelope of the Gaussian kernel centered at σ. Since these nodes cluster near the edges of the interval $[-1, 1]$, we can obtain coefficients with a smaller maximum magnitude by rescaling the energy spectrum into a smaller interval and considering transformed variables σ in the same restricted interval. As we mention in the main text, we work here with the assumption that $\omega \in [-1/2, 1/2]$ and the same for σ.

Now we use the following bound for the cosine term:

$$\left[\cos\left(y - \frac{\pi}{2L}\right) - \sigma\right]^2 \geqslant [\cos(y) - \sigma]^2 - \left(\frac{\pi}{L}\right)^2, \tag{D10}$$

to simplify the integrand above and obtain

$$\begin{aligned} \Omega_j &\leqslant \frac{\gamma_j}{\sqrt{2\pi^3}\Lambda} e^{\frac{\pi^2}{2L^2\Lambda^2}} \int_0^\pi dy \exp\left\{-\frac{[\cos(y) - \sigma]^2}{2\Lambda^2}\right\} \\ &= \frac{\gamma_j}{\sqrt{2\pi^3}\Lambda} e^{\frac{\pi^2}{2L^2\Lambda^2}} \int_{-1}^{1} \frac{dx}{\sqrt{1-x^2}} \exp\left[-\frac{(x-\sigma)^2}{2\Lambda^2}\right]. \end{aligned} \tag{D11}$$

Finally, using the fact that we rescaled the energies so that $\sigma \in [-1/2, 1/2]$, we can bound the integral by

$$\begin{aligned} \Omega_j &\leqslant \frac{\gamma_j}{\sqrt{2\pi^3}\Lambda} e^{\frac{\pi^2}{2L^2\Lambda^2}} \int_{-1}^{1} \frac{dx}{\sqrt{1-x^2}} \exp\left[-\frac{(x-1/2)^2}{2\Lambda^2}\right] \\ &\leqslant \frac{\gamma_j}{\sqrt{2\pi^3}\Lambda} \exp\left(\frac{\pi^2}{2L^2\Lambda^2}\right)(2.5\Lambda), \end{aligned} \tag{D12}$$

where the constant factor in the second line was obtained numerically. In summary, we found the following bound:

$$|c_j(\Lambda,\sigma)| \leqslant \gamma_j\left[R_L(\Lambda) + 0.32\exp\left(\frac{\pi^2}{2L^2\Lambda^2}\right)\right] \leqslant 2[R_L(\Lambda)+1.1], \tag{D13}$$

where we anticipated the result $L\Lambda > 2$ that will be proven in the next two sections.

2. Bound on truncation error

We turn now to providing upper bounds for the error terms $r_L(\Lambda)$ and $R_L(\Lambda)$. Using the result from Tausch and Weckiewicz [43], we can bound the magnitude of the expansion coefficients as

$$|a_n(\Lambda)| \leqslant 2\Lambda\sqrt{\pi}\exp\{-(n+1)\kappa[(n+1)2\Lambda^2]\}, \tag{D14}$$

where the auxiliary function $\kappa(x)$ is given by

$$\kappa(x) = \frac{\log_2(x+\sqrt{1+x^2})}{2} - \frac{1}{4x}\frac{(x-1+\sqrt{1+x^2})^2}{x+\sqrt{1+x^2}}. \tag{D15}$$

The total error $r_L(\Lambda)$ can then be bounded, summing a geometric series [44], and the result is

$$|r_L(\Lambda)| = \left|\sum_{n=L+1}^{\infty} a_n(\Lambda)T_n(x)\right| \leqslant \sum_{n=L+1}^{\infty}|a_n(\Lambda)| \leqslant 2\Lambda\sqrt{\pi}\frac{\exp[-L'\kappa(2L'\Lambda^2)]}{1-\exp[-\kappa(2L'\Lambda^2)]}, \tag{D16}$$

with $L' = L+2$ for L even, and $L' = L+3$ for L odd. We can obtain a simpler upper bound by first using the fact that for $x > 1$, we can bound $\kappa(x)$ with

$$\kappa(x) \geqslant \tfrac{1}{2}[\log_2(2x) - 1], \tag{D17}$$

and then using the monotonicity of the denominator in Eq. (D16) to find, for $2L'\Lambda^2 \geqslant 1$, the bound

$$|r_L(\Lambda)| \leqslant \frac{2\Lambda\sqrt{\pi}}{1-\exp[-\kappa(1)]}\left(\frac{e}{4L'\Lambda^2}\right)^{\frac{L'}{2}}. \tag{D18}$$

Using this result, we find the total error $R_N(\sigma,\Lambda)$ in the Gaussian transform given by Eq. (35) to be bounded as

$$|R_L(\sigma,\Lambda)| = \frac{1}{\sqrt{2\pi}\Lambda}\left|r_L\left(\frac{\Lambda}{2}\right)\right| \leqslant \frac{1}{\sqrt{2}}\frac{1}{1-\exp[-\kappa(1)]}\left(\frac{e}{L'\Lambda^2}\right)^{\frac{L'}{2}} \lesssim 3.4\left(\frac{e}{L'\Lambda^2}\right)^{\frac{L'}{2}}, \tag{D19}$$

valid in the asymptotic regime $L' \geqslant 2/\Lambda^2$. In order to guarantee a truncation error of, at most, ϵ_R, we now need

$$\epsilon_R \geqslant 3.4\left[\frac{e}{(L+2)\Lambda^2}\right]^{\frac{(L+2)}{2}} \tag{D20}$$

for the number of repetitions L. Note that in this last expression, we used the conservative value $L' = L+2$. The inequality in Eq. (D20) can be solved as

$$L \geqslant \frac{e}{\Lambda^2} - 2 + \frac{2\Lambda^2}{e}\frac{\log_2(3.4/\epsilon_R)}{W\left[\frac{2\Lambda^2}{e}\log_2(3.4/\epsilon_R)\right]}, \tag{D21}$$

where W is the Lambert W function [40]. In order to understand the scaling of this expression, we can use the less tight bound,

$$L \geqslant \frac{e}{\Lambda^2} + \frac{\log_2(3.4/\epsilon_R)}{\log_2[\log_2(3.4/\epsilon_R)]} - 2, \tag{D22}$$

which is usually employed in the literature (see, e.g., [4]).

3. Intermediate regime

We will now provide bounds in the second regime considered in the main text, where the upper bound on the order L is the lower limit of validity for Eq. (D19), namely, $0 < L' \leqslant \frac{2}{\Lambda^2}$. In this case, there is a minimum error which we can guarantee, the value of which we will determine in this section [see Eq. (D34)]. We can start by first noticing that for $0 < x \leqslant 1$, we have

$$x\kappa(1) \leqslant \kappa(x) \leqslant \frac{x}{4}, \tag{D23}$$

so that we can bound the total error r_N using

$$\begin{aligned}\sum_{n=L+1}^{\infty}|a_n(\Lambda)| &\leqslant 2\Lambda\sqrt{\pi}\sum_{n=L+1}^{\infty}e^{-(n+1)\kappa\left[(n+1)2\Lambda^2\right]} \\ &\leqslant 2\Lambda\sqrt{\pi}\sum_{n=L+1}^{\infty}e^{-(x+1)^2 2\kappa(1)\Lambda^2} \\ &\leqslant 2\Lambda\sqrt{\pi}\int_L^{\infty}dx e^{-(x+1)^2 2\kappa(1)\Lambda^2} \\ &= \frac{\pi}{\sqrt{2\kappa(1)}}\operatorname{erfc}[(L+1)\Lambda\sqrt{2\kappa(1)}].\end{aligned} \tag{D24}$$

This, in turn, implies the following upper bound for the error in the transform,

$$|R_L(\sigma,\Lambda)| \leqslant \frac{1}{2\Lambda}\sqrt{\frac{\pi}{\kappa(1)}}\operatorname{erfc}\left[(L+1)\Lambda\sqrt{\frac{\kappa(1)}{2}}\right]. \tag{D25}$$

This error con be bounded from above using

$$|R_L(\sigma,\Lambda)| \leqslant \frac{1}{\sqrt{2}\Lambda_\kappa^2}\frac{1}{L+1}\exp\left[-(L+1)^2\frac{\Lambda_\kappa^2}{2}\right], \tag{D26}$$

where we defined $\Lambda_\kappa = \Lambda\sqrt{\kappa(1)}$, and is valid for

$$L \geqslant \sqrt{\frac{2}{\pi}}\frac{1}{\Lambda_k} - 1. \tag{D27}$$

As we did in the general case above, if we want a truncation error of, at most, ϵ_R, we need

$$\epsilon_R \geqslant \frac{1}{\sqrt{2}\Lambda_\kappa^2}\frac{1}{L+1}\exp\left[-(L+1)^2\frac{\Lambda_\kappa^2}{2}\right]. \tag{D28}$$

The solution can again be conveniently expressed in terms of the Lambert W function as

$$L+1 \geqslant \frac{1}{\Lambda_k}\sqrt{\frac{1}{2}W\left(\frac{1}{2\Lambda_\kappa^2\epsilon_R^2}\right)}. \tag{D29}$$

We can now use another result from [45], Theorem 2.1, to find the sufficient condition,

$$L = \left\lceil \frac{1}{\Lambda}\sqrt{\frac{1}{\kappa(1)}g\left(\frac{1}{\sqrt{2\kappa(1)}\Lambda\epsilon_R}\right)} \right\rceil - 1, \tag{D30}$$

where, for convenience, we defined the auxiliary function

$$g(x) = \log_2(x) - \tfrac{1}{4}\log_2[\log_2(x^2)]. \tag{D31}$$

These estimates hold for sufficiently small target errors,

$$\Lambda\epsilon_R \leqslant \frac{1}{\sqrt{2\kappa(1)e}} \approx 0.54, \tag{D32}$$

a condition that ensures that Eq. (D27) is also satisfied. We finally note that it is also possible to find a bound on L valid for any value of the target error,

$$L = \left\lceil \frac{1}{\Lambda}\sqrt{\frac{2}{\kappa(1)}\log_2\left(\sqrt{\frac{\pi}{\kappa(1)}}\frac{1}{2\Lambda\epsilon_R}\right)} \right\rceil. \tag{D33}$$

We now need to find the minimum error that can be guaranteed in this intermediate regime. Using the upper bound from Eq. (D26), we find

$$\begin{aligned}\epsilon_R^{\min} &\leqslant \frac{1}{\sqrt{2}\kappa(1)}\frac{1}{2+\Lambda^2}\exp\left[-\frac{\kappa(1)}{2}\frac{(2+\Lambda)^2}{\Lambda^2}\right] \\ &\leqslant \frac{e^{-2\kappa(1)}}{\sqrt{8}\kappa(1)}\exp\left[-\frac{2\kappa(1)}{\Lambda^2}\right] \lesssim \exp\left(-\frac{1}{2\Lambda^2}\right),\end{aligned} \tag{D34}$$

and this is valid for any reasonable value Λ [to satisfy Eq. (D27), the condition is $\Lambda \leqslant 5$].

APPENDIX E: KERNEL BASED ON JACKSON'S THEOREM

The recent work by Rall [42] introduced an integral transform kernel based on Jackson's theorem from approximation theory. In this Appendix, we use the relevant results from Ref. [42] to construct an approximate integral transform and compare it with the GIT- and TSA-based methods described in the main text. The approximate window function introduced in Ref. [42] can be used to construct a (normalized) integral kernel as

$$K_J(\sigma,\omega,k,N) = \mathcal{N}_{kN}\omega_{kN}\left(\frac{\sigma-\omega}{2}\right), \tag{E1}$$

with $\mathcal{N}_{kN}$ a normalization factor and

$$\omega_{kN}(x) = A_k\left[\frac{4}{5}J_N(x)\right]. \tag{E2}$$

In the expression above, A_k is the amplifying polynomial from Eq. (A5) of Ref. [42], while $J_N(x)$ is the Jackson's approximation to the function $g(x)$ defined as

$$g(x) = \begin{cases} -1, & x < -\delta \\ -1 + \frac{2}{\delta}(x+\delta), & -\delta < x \leqslant 0 \\ 1 - \frac{2}{\delta}x, & 0 > x > \delta \\ -1, & x > \delta, \end{cases} \tag{E3}$$

for some fixed resolution δ. Note that in this construction, we let the approximation interval $[\bar{a},\bar{b}]$ defined in Ref. [42] shrink to zero. As shown in Ref. [42], in order to ensure J_N approximates $g(x)$ with error less than $1/4$, one can take $N = 24/\delta$. The order k of A_k controls the final approximation error by ensuring that for $x > \delta$, the final function satisfies

$$A_k\left[\frac{4}{5}J_N(x)\right] \leqslant \tau \equiv \exp(-k/6). \tag{E4}$$

The condition for the integral transform to be Σ-accurate with resolution Δ can be written as

$$\begin{aligned}\sup_{\omega_0\in[-1,1]}\Bigg[&\int_{-1}^{\omega-\Delta} d\sigma K_J(\sigma,\omega_0,k,N) \\ &+ \int_{\omega+\Delta}^{1} d\sigma K_J(\sigma,\omega_0,k,N)\Bigg] \leqslant \Sigma,\end{aligned} \tag{E5}$$

or, in the more convenient form,

$$2\mathcal{N}_{kN}\int_{\Delta/2}^{1} dx A_k\left[\frac{4}{5}J_N(x)\right] \leqslant \Sigma. \tag{E6}$$

By choosing the resolution in the g function in Eq. (E3) to be $\delta = \Delta/2$, we find that Eq. (E5) is satisfied for

$$\tau \leqslant \frac{\Sigma}{2-\Delta}\frac{1}{\mathcal{N}_{kN}}. \tag{E7}$$

The normalization constant can be bounded using

$$1 = \int_{-1}^{1} dx K_J(x,\omega,k,N) \leqslant \mathcal{N}_{kN}[2\delta + \tau(2-2\delta)], \tag{E8}$$

and this gives the following necessary condition on τ:

$$\tau \leqslant \Sigma\frac{\Delta+\tau(2-\Delta)}{2-\Delta} \Rightarrow \tau \leqslant \frac{\Sigma}{1-\Sigma}\frac{\Delta}{2-\Delta}. \tag{E9}$$

If we require the approximation to be Σ-accurate with resolution Δ, the order d of the polynomial representation of the kernel K_J needs to be larger than

$$d_{\min} = \frac{288}{\Delta}\log_2\left(\frac{1-\Sigma}{\Sigma}\frac{2-\Delta}{\Delta}\right). \tag{E10}$$

The asymptotic cost of using the Jackson kernel for the spectral-density approximation is therefore worse than the GIT-based method presented in the main text. Comparing this result with the TSA method of Ref. [12] will, however, require one to find an upper bound on the normalization constant first. We can obtain this by noticing that in the intervals $[0,\delta/2]$ and $[-\delta/2,0]$, the kernel function can be bounded from below using a linear function, while outside of this region the lower bound is zero. We can therefore write, for $\tau < 5/8$, the following:

$$\mathcal{N}_{kN} \leqslant \frac{1}{\delta}\frac{4}{5-8\tau}, \tag{E11}$$

which recovers the intuition that in general, $\mathcal{N}_{kN}$ should scale linearly with the resolution. This shows that the method presented in this Appendix also has a better complexity than the TSA algorithm.

[1] R. P. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).
[2] S. Lloyd, Science **273**, 1073 (1996).
[3] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, Phys. Rev. Lett. **114**, 090502 (2015).
[4] G. H. Low and I. L. Chuang, Phys. Rev. Lett. **118**, 010501 (2017).
[5] G. H. Low and I. L. Chuang, Quantum **3**, 163 (2019).
[6] R. Babbush, D. W. Berry, J. R. McClean, and H. Neven, npj Quantum Inf. **5**, 92 (2019).
[7] H. Lamm and S. Lawrence, Phys. Rev. Lett. **121**, 170501 (2018).
[8] A. Roggero and J. Carlson, Phys. Rev. C **100**, 034610 (2019).
[9] A. Roggero, A. C. Y. Li, J. Carlson, R. Gupta, and G. N. Perdue, Phys. Rev. D **101**, 074038 (2020).
[10] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, Proc. Natl. Acad. Sci. **115**, 9456 (2018).
[11] J. Preskill, Quantum **2**, 79 (2018).
[12] R. D. Somma, New J. Phys. **21**, 123025 (2019).
[13] J. Carlson and R. Schiavilla, Phys. Rev. Lett. **68**, 3682 (1992).
[14] D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).
[15] A. Roggero, F. Pederiva, and G. Orlandini, Phys. Rev. B **88**, 094302 (2013).
[16] V. D. Efros, W. Leidemann, and G. Orlandini, Phys. Lett. B **338**, 130 (1994).
[17] V. D. Efros, W. Leidemann, G. Orlandini, and N. Barnea, J. Phys. G: Nucl. Part. Phys. **34**, R459 (2007).
[18] S. Bacca, N. Barnea, G. Hagen, G. Orlandini, and T. Papenbrock, Phys. Rev. Lett. **111**, 122502 (2013).
[19] W. Glöckle and M. Schwamb, Few-Body Syst. **46**, 55 (2009).
[20] N. Barnea, V. D. Efros, W. Leidemann, and G. Orlandini, Few-Body Syst. **47**, 201 (2010).
[21] M. Miorelli, S. Bacca, N. Barnea, G. Hagen, G. R. Jansen, G. Orlandini, and T. Papenbrock, Phys. Rev. C **94**, 034317 (2016).
[22] A. Roggero and S. Reddy, Phys. Rev. C **94**, 015803 (2016).
[23] J. E. Gubernatis, M. Jarrell, R. N. Silver, and D. S. Sivia, Phys. Rev. B **44**, 6011 (1991).
[24] L. Novo, J. Bermejo-Vega, and R. García-Patrón, arXiv:1912.06608.
[25] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Proc. R. Soc. London A **454**, 339 (1998).
[26] A. M. Childs, R. Kothari, and R. D. Somma, SIAM J. Comput. **46**, 1920 (2017).
[27] L. Lin and Y. Tong, arXiv:2002.12508 .
[28] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. **85**, 5098 (2000).
[29] If needed, the number of ancilla qubits can be reduced to just 1 using iterative schemes.
[30] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2011).
[31] A. Y. Kitaev, Electron. Colloq. Comput. Complex. **3** (1995).
[32] N. Wiebe and C. Granade, Phys. Rev. Lett. **117**, 010503 (2016).
[33] T. E. O'Brien, B. Tarasinski, and B. M. Terhal, New J. Phys. **21**, 023022 (2019).
[34] W. Hoeffding, J. Am. Stat. Assoc. **58**, 13 (1963).
[35] M. Suzuki, J. Math. Phys. **32**, 400 (1991).
[36] Note the additional factor of two coming from the need in QPE to perform $\log_2(M)$ controlled evolutions .
[37] We can also add an additional normalization factor $\mathcal{N}(\sigma, \Lambda)$ that could be used to keep the kernel normalized (and hence maintain the validity of sum rules), while restricting the values of σ to lie in the range $[-1, 1]$ as the frequency. We did not find any significant advantage in doing this; we will take σ to be defined over the full real line.
[38] S. Subramanian, S. Brierley, and R. Jozsa, J. Phys. Commun. **3**, 065002 (2019).
[39] D. Andreasi, W. Leidemann, C. Reiß, and M. Schwamb, Eur. Phys. J. A **24**, 361 (2005).
[40] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, Adv. Comput. Math. **5**, 329 (1996).
[41] E. Knill, G. Ortiz, and R. D. Somma, Phys. Rev. A **75**, 012328 (2007).
[42] P. Rall, arXiv:2004.06832.
[43] J. Tausch and E. Weckiewicz, SIAM J. Sci. Comput, 3547 (2009).
[44] Note the missing factor of 2 from Eq. (11) of [43].
[45] A. Hoorfar and M. Hassani, J. Inequal. Pure Appl. Math. **9**, 5 (2008).

Correction: Equation (4) contained a minor error and has been fixed.