Spectral density estimation with the Gaussian Integral Transform

A. Roggero
Institute for Nuclear Theory, University of Washington, Seattle, WA 98195, USA
(Dated: April 13, 2020)

The spectral density operator $\hat{\rho}(\omega) = \delta(\omega - \hat{H})$ plays a central role in linear response theory as it’s 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 $\Delta$ and error $\epsilon$ using $O\left(\log(1/\epsilon) (\log(1/\Delta) + \log(1/\epsilon))/\Delta\right)$ operations. This is achieved without using expensive approximations to the time-evolution operator but exploiting instead qubitization to implement an approximate Gaussian Integral Transform (GIT) of the spectral density. We also describe appropriate error metrics to assess the quality of spectral function approximations more generally.

Since the 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 [3], 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 NISQ devices [11] and more work is required to bring these costs down (see eg. [11] 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 new quantum algorithm with near optimal computational cost (in terms of oracle calls) to study the problem of spectral density estimation. In particular, given an 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 $\delta$ the Dirac delta function. Using the eigenstates $|k\rangle$ of the operator $\hat{O}$ we have the following representation

$$\hat{\rho}(\omega) = \sum_{k=0}^{\Gamma} \delta(\omega - O_k)|k\rangle\langle k|$$

with $O_k$ the eigenvalue for $|k\rangle$, and $\Gamma$ the total number of eigenvalues. Without loss of generality, we will consider normalized operators $\hat{O}$ with $||\hat{O}|| \leq 1$ 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

$$S(\omega) = \langle \Psi | \hat{\rho}(\omega) | \Psi \rangle = \sum_{k=0}^{\Gamma} |\langle \Psi | k \rangle|^2 \delta(\omega - O_k)$$

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 non unitary) vertex operator $\hat{Q}$. In this case, the relevant operator $\hat{Q}$ coincides with the Hamiltonian of the physical system, and for this reason we will often call it’s 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 response function obtained through an integral transform of the type

$$\Phi_K(\nu) = \int d\omega K(\nu, \omega) S(\omega) = \sum_{k=0}^{\Gamma} |\langle \Psi | k \rangle|^2 K(\nu, O_k) \quad (3)$$

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 like Eq. (3) is a common strategy in many-body physics. In Quantum Monte Carlo calculations, for instance, is common to consider the Laplace kernel $K(\omega, \nu) = \exp(-\nu \omega)$ due to it’s connection with euclidean path integrals (see eg. [13, 14]), but other alternatives such as the Sumudu [15] and Lorentz [16, 17] 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 $\Phi_K$ will get exponentially amplified by the inversion procedure (see eg. [13, 20]). In this work we avoid the problem by using directly the integral transform $\Phi_K(\omega)$ as the approximate reconstruction of the original signal $S(\omega)$. A possible future extension of our work would be to consider approximate inversion schemes like the Maximum Entropy Method [21] to try and 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 we use to judge the quality of the approximation in Eq. 3 and in Sec. III we summarize the main results of the paper and compare them to the recent work from Ref. [12] which can be understood as a particular instance of the method 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] in Sec. II and the Gaussian kernel which allows to achieve near optimal scaling of the computational cost in Sec. III. We also provide a pseudocode implementation in Appendix B. We conclude in Sec. IV providing a summary of our findings and proposing possible avenues for future improvements.

I. DEFINITIONS AND COMPARISON TO PREVIOUS WORK

In order to precisely quantify the accuracy of the approximation procedure presented in the introduction above, and connect with recent work on quantum algorithms exploring similar problems [12, 22], we now introduce the following definitions:

- we will call an integral kernel Σ-accurate with resolution ∆ if the following condition holds
  \[
  \sup_{\omega_0 \in [-1,1]} \int_{\omega_0 - \Delta}^{\omega_0 + \Delta} d\nu K(\nu, \omega_0) \geq 1 - \Sigma ,
  \]
  where the symbol \( \int \) indicates: an integral when the transformed variable \( \nu \) is defined over a continuous interval or a sum if \( \nu \) is defined on a discrete set.

- we will call a distribution \( \Phi(\omega) \) a \( \beta \)-approximation to the true distribution \( \Phi(\omega) \) with confidence \( 1 - \eta_\beta \) if the total variation is bounded as
  \[
  \delta_\nu(T, \tilde{T}) := \sup_{\omega \in [-1,1]} |T(\omega) - \tilde{T}(\omega)| \leq \beta .
  \]
  with probability \( P > 1 - \eta_\beta \).

- if the estimator \( \Phi_K \) is obtained as a \( \beta \)-approximation with confidence \( 1 - \eta_\beta \) of a Σ-accurate integral transform \( \Phi_K \) of the response function \( S(\omega) \) with resolution \( \Delta \) 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 Ref. [22]. In particular, the first definition is similar in spirit to, but more stringent than, the condition of having resolution \( \Delta \) and confidence \( \eta = (1 - \Sigma) \) (Definition 1 of [22]) while the second condition is equivalent to the \( \beta \)-approximation (Definition 3 of [22]).

The reason for these definitions, and the mild departure from those introduced in Ref. [22], is rooted in the fact that for physics application we are ultimately interested in frequency observables of the form

\[
Q(S, f) = \int_1^1 d\omega S(\omega) f(\omega)
\]

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

\[
|Q(S, f) - Q(\Phi_K, f)| \leq f_{max} + 2f_{max}\Sigma + \beta f_{int} ,
\]

where we have defined the quantities

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

and the upperbound on the maximum variation

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

A full derivation of this is provided in Appendix A.

At this point it is important to point out another difference with Ref. [22]. In our work, the second error metric \( \beta \) captures both the statistical error coming from estimating the distribution \( \Phi_K(\omega) \) with a finite number of samples, but also 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 being Σ-accurate with resolution \( \Delta \) is a property of the kernel function \( K(\nu, \omega) \), while being a \( \beta \)-approximation with confidence \( 1 - \eta_\beta \) is a property that characterizes the implementation of the algorithm that generates the desired integral transform.

A. Comparison to previous work

The approximation problem 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 algorithm (TSA) 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 \( \Sigma \to \beta \to \varepsilon \) 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 \( 2\Delta \) 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 kernel function given by the approximate frequency comb

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

where \( \nu_j \) is the central value of the \( j \)-th 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 to achieve accuracies \( \Sigma = 0 \) in Eq. (11). Note however that using \( \Sigma \ll \beta \) will not help reduce the final error in Eq. (12) (unless \( f_{\text{max}} \gg f_{\text{int}} \)), and in fact here we only require them to be both equal to \( \epsilon \).

The TSA algorithm requires to apply the (controlled) time evolution operator \( U_0(t) = \exp(-it\hat{O}) \) for a maximum time \( t_{\text{max}} \) scaling as (see Appendix A of [12])

\[ t_{\text{max}} = \mathcal{O}\left(\frac{1}{\Delta} \log\left(\frac{1}{\epsilon^2}\right)\right), \tag{11} \]

together with a total number of samples scaling as

\[ N_S = \mathcal{O}\left(\frac{1}{\Delta^3 \epsilon^2} \log\left(\frac{1}{\epsilon}\right) \log\left(\frac{1}{\eta\beta}\right)\right), \tag{12} \]

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

In order to compare these asymptotic scaling 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. [12] (which is based on qubitization [3]) 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 \( M = \mathcal{O}(t_{\text{max}}) \).

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 [22] 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 [24].

We will analyze these integral transform in detail in the next sections and anticipate here the main results.

Due to the choice \( \beta = \Sigma = \epsilon \), both the standard Fejer method of Ref. [8] and the qubitization-based variant from Ref. [3] have the same asymptotic scaling. We will anticipate here results for the latter, which can produce a \((\epsilon, \Delta, \epsilon, \eta\beta)\)-approximation using

\[ M = \mathcal{O}\left(\frac{1}{\Delta^3 \epsilon^2} \log\left(\frac{1}{\epsilon}\right) \log\left(\frac{1}{\Delta \epsilon}\right)\right) \quad \text{and} \quad N_S = \mathcal{O}\left(\frac{1}{\Delta^3 \epsilon^6} \log\left(\frac{1}{\epsilon}\right) \log\left(\frac{1}{\Delta \epsilon}\right)\right), \tag{13} \]

even though the sample complexity is greatly reduced, for small target errors \( \epsilon \) 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 using the GIT provides a considerable reduction of the quantum computational cost (ie. 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 \((\epsilon, \Delta, \epsilon, \eta\beta)\)-approximation to the response function requires only

\[ M = \mathcal{O}\left(\frac{1}{\Delta^3 \epsilon^2} \log\left(\frac{1}{\epsilon}\right) \log\left(\frac{1}{\Delta \epsilon}\right)\right) \tag{14} \]
calls to the qubiterate unitary \( W_Q \), together with

\[ N_S = \mathcal{O}\left(\frac{1}{\Delta^3 \epsilon^6} \log\left(\frac{1}{\epsilon}\right) \log\left(\frac{1}{\Delta \epsilon}\right)\right)^{3/2} \log\left(\frac{1}{\eta\beta}\right), \tag{15} \]

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

The quantum query complexity 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. [23]. In particular, optimality can be shown by considering: an hamiltonian with spectral gap \((\omega_1 - \omega_0) > 2\Delta\), an initial state \(|\Psi\rangle\) with an overlap on the ground state state \(|\Psi(0)\rangle \geq \epsilon \) and ask for an approximation of the ground state energy with probability \( P > 1 - \epsilon \) and confidence \( 1 - \eta\beta \). Using the results from Ref. [23] (Lemmas 3 and 5 and Theorems 8 and 9) we know that this requires at least \( M = \mathcal{O}(1/\Delta \log(1/\epsilon)) \) oracle calls to \( W_Q \). We can also solve this problem by

| Method                          | Number of calls to \( W_Q \) | Total number of samples |
|--------------------------------|-------------------------------|-------------------------|
| Time Series Analysis (TSA) [12] | \( \mathcal{O}\left(\frac{1}{\Delta} \log\left(\frac{1}{\epsilon^2}\right)\right) \) | \( \mathcal{O}\left(\frac{1}{\Delta^3 \epsilon^2} \log\left(\frac{1}{\epsilon}\right) \log\left(\frac{1}{\Delta \epsilon}\right)\right) \) |
| Fejer Integral Transform [8, 9] | \( \mathcal{O}\left(\frac{1}{\Delta} \log\left(\frac{1}{\epsilon}\right)\right) \) | \( \mathcal{O}\left(\frac{1}{\Delta^3 \epsilon^2} \log\left(\frac{1}{\epsilon}\right) \log\left(\frac{1}{\Delta \epsilon}\right)\right) \) |
| Gaussian Integral Transform (GIT) | \( \mathcal{O}\left(\frac{1}{\Delta} \log\left(\frac{1}{\epsilon}\right) \log\left(\frac{1}{\Delta \epsilon}\right)\right) \) | \( \mathcal{O}\left(\frac{1}{\Delta^3 \epsilon^6} \log\left(\frac{1}{\epsilon}\right) \log\left(\frac{1}{\Delta \epsilon}\right)\right)^{3/2} \log\left(\frac{1}{\eta\beta}\right) \) |
considering a $\langle \varepsilon, \Delta, \varepsilon, \eta_3 \rangle$-approximation $\Phi_K$ to the response $S_\Omega(\omega) = \langle \Psi | \hat{\rho}(\omega) | \Psi \rangle$. Our result is then only a factor $O(\sqrt{\log(1/\varepsilon) \log(1/\Delta)})$ away from the optimal result and provides a quadratic speedup in the logarithmic factors compared to the TSA scheme of Ref. 12.

II. FEJER KERNEL

The standard Quantum Phase Estimation (QPE) algorithm 23 26 uses $n$ applications of the (controlled) time evolution unitary $U_Q(t) = \exp(-it\hat{O})$ and $N = 2^n$ ancilla qubits to approximately diagonalize the "Hamiltonian" operator $\hat{O}$ 27. As we proposed in Ref. 3, this technique can be used to perform an integral transform generated by a rescaled Fejer kernel

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

(16)

where the discrete frequencies $\sigma_q$ 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 $t_{\text{max}}$ used in QPE as $t_{\text{max}} = \pi N$. In this case, ensuring the resulting integral transform $\Phi_F(\omega)$ is $\Sigma$-accurate with resolution $\Delta$ is equivalent to requiring the probability of measuring a phase $\sigma_k$ with error more than $\Delta$ to be less than $\Sigma$. This probability can be bounded using standard techniques (see eg. 5.2.1 of 28) as

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

(17)

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

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

(18)

in order to satisfy Eq. (14). The dependence on the resolution $\Delta$ is already optimal and the constant factors could be improved using optimized preparations of the ancilla register 20. The scaling with the error $\Sigma$ instead could be improved to $N = O(\log(1/\Sigma) / \Delta)$ in the special situation where the signal $S(\omega)$ is composed of a single frequency mode by using schemes like Kitaev’s original algorithm 29 or the more efficient IPEA 31. 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 eg. 31). We can now use $N_S = O(1/\beta^2)$ samples to produce the $\beta$-approximate estimator $\Phi_F$ by collecting an histogram of the measured frequencies. More precisely, using Hoeffding’s inequality 32 we find it sufficient to take

$$N_S = \frac{1}{2\beta^2} \log \left( \frac{2}{\eta_3} \right),$$

(19)

with $\eta_3$ the confidence of the $\beta$-approximation.

In general, the time-evolution operator $U_Q(t)$ needs to be approximated with additive error $\delta_t$, 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. 11 we will consider these errors as contributions to the total variation Eq. (3) which define the $\beta$-approximation. In particular, if we denote by $\Phi^e_F(\omega)$ the transform obtained by using the approximate time-evolution unitary and $\Phi_F^e(\omega)$ it’s finite population estimator, we have

$$\delta_V(\Phi_F, \Phi_F^e) \leq \delta_V(\Phi_F, \Phi^e_F) + \delta_V(\Phi^e_F, \Phi_F^e).$$

(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

$$\delta_V(\Phi_F, \Phi_F^e) \leq \log_2(N) \delta_t,$$

(21)

with $\delta_t$ an upperbound to the approximation error of the time-evolution operator for times up to $t_{\text{max}} = \pi N$. The finite population estimator of the approximate Fejer transform is then $\beta$-accurate with confidence $\eta_3$ if

$$N_S = \frac{2}{\beta^2} \log \left( \frac{2}{\eta_3} \right),$$

(22)

Using optimal scaling algorithms for time evolution like Quantum Signal Processing 4, the total gate count is

$$M = O\left( \frac{1}{\Delta \Sigma} + \log \left( \frac{1}{\beta} \right) \right),$$

(23)

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

$$W_Q = \exp \left( i \bar{Y} \text{arccos}(\hat{O}) \right),$$

(24)

where $\bar{Y}$ 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 33 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 qubitocate $W_Q$ (see Ref. 3). One can easily show that this leads to a modified Fejer kernel given by

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

(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

$$|\cos(\theta_\omega + \Delta_M) - \cos(\theta_\omega)| \leq \frac{\Delta}{2},$$

(26)
which amounts to require $\Delta_\vartheta \leq \sqrt{1 + \Delta} - 1$. 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. [34] requires the closest power of 2

$$M \geq \frac{2}{\Delta_\vartheta} \left( \frac{1}{\Sigma} + 2 \right) \geq \frac{4}{\Delta} \left( \frac{1}{\Sigma} + 2 \right),$$

(27)

black box invocations of the qubitizer $W_Q$ [34] 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 $U_G(t)$ 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 $\Sigma$.

III. GAUSSIAN KERNEL

We consider now a Gaussian Integral Transform (GIT) defined through the following kernel function

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

(28)

where $\Lambda > 0$ controls the resolution, and the transformed frequency $\sigma$ is defined over the whole real line $\mathbb{R}$. The first step is to determine the conditions for which the approximate response obtained using the GIT is $\Sigma$-accurate with resolution $\Delta$. Using the translational invariance of the kernel $K_G(\sigma, \omega, \Lambda)$ for $\sigma \in \mathbb{R}$, we can rewrite the condition Eq. (11) in terms of the error function as

$$\frac{1}{\sqrt{2\pi}\Lambda} \int_{-\Delta}^{\Delta} d\sigma \exp\left(-\frac{\sigma^2}{2\Lambda^2}\right) = \text{erf}\left(\frac{\Delta}{\sqrt{2}\Lambda}\right) \geq 1 - \Sigma.$$

(29)

A sufficient condition for this to hold is to choose the kernel resolution $\Lambda$ according to

$$\Lambda \leq \frac{\Delta}{\sqrt{2 \log(1/\Sigma)}}.$$

(30)

We now move on to find the condition for the GIT to be $\beta$-approximate with confidence $\eta_\beta$ according to Eq. [35]. As we mentioned in the introduction, 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 it’s direct connection with quantum walks [24, 30] and the qubitization method [3], 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 $(G|W_Q|G) = \hat{O}$ and entering in the definition of the qubitizer $W_Q$, we have (see the proof of Lemma 16 of Ref. [21] and Appendix D of Ref. [30])

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

(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 $k$-the Chebyshev polynomial can then be obtained as

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

(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

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

(33)

and then expanding the integral kernel as

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

(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 $\sigma$. Explicit expressions for these coefficients can be found in Eq. (17) of Appendix D.

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

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

(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 reasonably inversions of the Lorentzian kernel by neglecting the error term $R_L$ as a way of performing a regularization to the ill-posed problem [37]. The final approximate integral transform $\Phi_{GL}(\omega)$ is then obtained as

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

(36)

As described in Sec. III the approximation error contributes to the total variation 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

$$\delta_V(\Phi_G, \Phi_{GL}) \leq \delta_V(\Phi_G, \Phi_{G}) + \delta_V(\Phi_{G}, \Phi_{GL}) \leq R_L(\sigma, \omega, \Lambda) + \delta_V(\Phi_{G}, \Phi_{GL}).$$

(37)
where $\Phi_G$ is the exact GIT, $\Phi_{GL}$ the approximate integral transform obtained by truncating the series in Eq. (34) at order $L$ and $\Phi_{GL}^\alpha$ its finite population estimator. As we did for the Fejer kernel above, we will now require that both error terms to be less than $\beta/2$ with confidence $\eta_\beta$.

In order to bound the total statistical error of the finite population estimator $\Phi_{GL}^\alpha$ 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

$$N_S = 2L \log \left( \frac{2}{\eta_\beta} \right) \max_{k=0,\ldots,L} \left( \frac{L|c_k|}{\beta} \right)^2,$$

(38)

where we used the upperbound on $c_k$ obtained in Appendix D.

Note that, for technical reasons explained in Appendix D this is valid only after rescaling the operator $O$ by a factor of 2. Since it is possible to find an appropriate bound also in the general case, we do not correspondingly rescale the resolution $\Delta$ here.

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

In order to find optimal truncation schemes it is now convenient to distinguish between two different situation depending on the desired value $\beta$ as a function of $\Sigma$ and the resolution $\Delta$. More precisely, we define two critical values $\beta_U$ and $\beta_L$ as follows

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

(39)

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

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

- in the asymptotic regime $\beta \leq \beta_L$ we need

$$L_{asy} = \left[ \frac{2e^2 \log \left( \frac{1}{\Sigma} \right) + g_\delta \left( \frac{\Delta^2}{e} \log \left( \frac{6.8/\beta}{\Delta^2} \right) \right)}{2} \right] - 2,$$

where for convenience we have introduced the function $g_\delta(x) = x/W(x)$ with $W$ is the Lambert W-function [38] (see also Appendix D for details).

- in the intermediate regime $\beta_L \leq \beta \leq \beta_U$ we need

$$L_{int} = \left[ \frac{\alpha_1}{\Delta} \sqrt{\log \left( \frac{1}{\Sigma} \right) g_i \left( \frac{\alpha_2}{\Delta \beta} \log \left( \frac{1}{\Sigma} \right) \right)} \right] - 1,$$

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

$$g_i(x) = \log \left( x \right) - \frac{1}{4} \log \left( \log \left( x^2 \right) \right).$$

(42)

As apparent from the definition of the truncated GIT $\Phi_{GL}^\alpha$ 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. IIA, we can write these estimates in terms of asymptotic scaling as

$$L_{asy} = \mathcal{O} \left( \frac{1}{\Delta^2} \log \left( \frac{1}{\Sigma} \right) + \frac{\log(1/\beta)}{\log(1/\beta)} \right),$$

(43)

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

$$L_{int} = \mathcal{O} \left( \frac{1}{\Delta} \sqrt{\log \left( \frac{1}{\Sigma} \right) \log \left( \frac{1}{\Delta \beta} \log \left( \frac{1}{\Sigma} \right) \right)} \right).$$

(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 Eq. (41). Finally note that, as mentioned in Sec. IIA the asymptotic regime $\beta \leq \beta_L$ is possibly not directly relevant for the approximate estimation of observables of the form 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 accuracies $\Sigma = 0$.

### IV. 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, 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: the Lorentz Integral Transform (LIT) [10].

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 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 implementations of the method, together with a pseudocode implementation of these algorithms in Appendix D 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 reducing the sample
complexity by employing techniques like the method of Ref. [39] (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 LCU method [22] to implement directly the approximate spectral density

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

This would allow, together with Amplitude Amplification, to selectively prepare final states of scattering processes within a pre-determined 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. [23]. Finally, the general framework introduced in this work, and the accuracy metrics defined in Sec. I 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 approximations is left for future work.

ACKNOWLEDGMENTS

I want to thank M. Savage for his continued support during the preparation of this manuscript, and 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 number ERKJ333 and by the Institute for Nuclear Theory under U.S. Department of Energy grant No. DE-FG02-00ER41132.

[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 Information 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, “Quantum computing for neutrino-nucleus scattering,” (2019), arXiv:1911.06368 [quant-ph].
[10] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, Proc. Natl. Acad. Sci. U. S. A. 115, 9456 (2018).
[11] J. Preskill, Quantum 2, 79 (2018).
[12] R. D. Somma, New Journal of Physics 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. Letters B 338, 130 (1994).
[17] V. D. Efros, W. Leidemann, G. Orlandini, and N. Barnea, J Phys. G Nucl. Partic. 34, R459 (2007).
[18] S. Bacca, N. Barnea, G. Hagen, G. Orlandini, and T. Papenbrock, Phys. Rev. Lett. 111, 122502 (2013).
[19] W. Glickle and M. Schwamb, Few-Body Systems 46, 55 (2009).
[20] N. Barnea, V. D. Efros, W. Leidemann, and G. Orlandini, Few-Body Systems 47, 201 (2010).
[21] J. E. Gubernatis, M. Jarrell, R. N. Silver, and D. S. Sivia, Phys. Rev. B 44, 6011 (1991).
[22] L. Novo, J. Bermejo-Vega, and R. Garca-Patr, “Quantum advantage from energy measurements of many-body quantum systems,” (2019), arXiv:1912.06608 [quant-ph].
[23] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 454, 339 (1998).
[24] A. M. Childs, R. Kothari, and R. D. Somma, SIAM Journal on Computing 46, 1920 (2017).
[25] L. Lin and Y. Tong, arXiv e-prints, arXiv:2002.12508 (2020), arXiv:2002.12508.
[26] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. 85, 5098 (2000).
[27] If needed, the number of ancilla qubits can be reduced to just 1 using iterative schemes.
[28] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, 2011).
[29] A. Y. Kitaev, Electronic Colloquium on Computational Complexity (ECCC) 3 (1995).
[30] N. Wiebe and C. Granade, Phys. Rev. Lett. 117, 010503 (2016).
[31] T. E. O’Brien, B. Tarasinski, and B. M. Terhal, New Journal of Physics 21, 023022 (2019).
[32] W. Hoefding, J. Am. Stat. Assoc. 58, 13 (1963).
[33] M. Suzuki, Journal of Mathematical Physics 32, 400 (1991).
[34] Note the additional factor of two coming from the need in QPE to perform $\log_2(M)$ controlled evolutions.
[35] 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 $\sigma$ to lie in the range $[-1,1]$ as the frequency. We didn’t find any significant advantage in doing this we will take $\sigma$ to be defined over the full...
Appendix A: Error bound for frequency observables

As in the main text we start with a response function

\[ S(\omega) = \sum_k \alpha_k \delta(\omega - \omega_k) \sum_k \alpha_k = 1 , \] (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

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

If we use a \((\Sigma, \Delta, \beta, \eta, \delta)\)-approximation \( \tilde{\Phi}(\omega) \) to the response \( S(\omega) \) obtained using a \( \beta \)-approximate estimator for the integral transform \( \Phi(\omega) \), we want to find an upperbound for the total error

\[ \delta_Q(\tilde{\Phi}) = |Q(S, f) - Q(\tilde{\Phi}, f)| , \] (A3)

where the approximate observables is expressed as

\[ Q(\tilde{\Phi}, f) = \int f d\nu \tilde{\Phi}(\nu) . \] (A4)

Using the triangle inequality we find

\[ \delta_Q(\tilde{\Phi}) = \left| \int_{-1}^{1} d\omega S(\omega) f(\omega) - \int d\nu \tilde{\Phi}(\nu) f(\nu) \right| \]
\[ \leq \left| \int_{-1}^{1} d\omega S(\omega) f(\omega) - \int d\nu \Phi(\nu) f(\nu) \right| + \int d\nu (\Phi(\nu) - \tilde{\Phi}(\nu)) f(\nu) \]
\[ = \delta_Q(\Phi) + \left| \int d\nu (\Phi(\nu) - \tilde{\Phi}(\nu)) f(\nu) \right| . \] (A5)

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

\[ \Phi(\omega) = \sum_k \alpha_k \Phi_k(\omega) , \] (A6)

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

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

while for the integral transform approximator

\[ Q(\Phi, f) = \sum_k \alpha_k \int d\nu \Phi_k(\nu) f(\nu) \]
\[ = \sum_k \int d\nu K(\nu, \omega_k) S(\omega) f(\nu) \]
\[ = \sum_k \alpha_k \int d\nu K(\nu, \omega_k) f(\nu) , \] (A8)

where in the last line we performed the frequency integral using the decomposition Eq. (A1). Using the definition Eq. (A4) of a \( \Sigma \)-accurate kernel with resolution \( \Delta \) we can find a bound for the first term in Eq. (A5) as follows

\[ \delta_Q(\Phi) = \left| \sum_k \alpha_k \left( f(\omega_k) - \int d\nu K(\nu, \omega_k) f(\nu) \right) \right| \]
\[ \leq \sum_k \alpha_k \left( f(\omega_k) - \int_{-\Delta}^{\Delta} d\nu K(\nu, \omega_k) f(\nu) \right) \]
\[ + \sum_k \alpha_k \left( \int_{-\Delta}^{\Delta} d\nu K(\nu, \omega_k) f(\nu) - \int_{-\Delta}^{\Delta} d\nu K(\nu, \omega_k) f(\nu) \right) \]
\[ \leq \sum_k \alpha_k \left( f(\omega_k) - \int_{-\Delta}^{\Delta} d\nu K(\nu, \omega_k) f(\nu) \right) + f_{\max} \Sigma \]
\[ \leq \sum_k \alpha_k f(\omega_k) \left( 1 - \int_{-\Delta}^{\Delta} d\nu K(\nu, \omega_k) \right) \]
\[ + \sum_k \alpha_k \int_{-\Delta}^{\Delta} d\nu K(\nu, \omega_k) f(\nu) - f(\omega_k) \]
\[ \leq \sum_k \alpha_k \int_{-\Delta}^{\Delta} d\nu K(\nu, \omega_k) f(\nu) - f(\omega_k) \]
\[ \leq f_{\max} \Sigma \]

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

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

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

\[ \left| \int d\nu (\Phi(\nu) - \tilde{\Phi}(\nu)) f(\nu) \right| \leq \beta \int_{-1}^{1} d\omega |f(\omega)| \] (A10)

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

\[ \delta_Q \leq f_{\max}^{\Delta} + 2f_{\max} \Sigma + \beta \int_{-1}^{1} d\omega |f(\omega)| \] (A11)
Appendix B: Pseudocode implementation

We present here a pseudocode implementation for the spectral density estimation algorithms we discuss in the main text. The goal of our algorithm is to return a \((\Sigma, \Delta, \delta, \beta, \eta_\beta)\)-approximation \(\Phi_k(\nu)\) to the response function \(S(\omega) = \langle \Psi | \rho(\omega) | \Psi \rangle\) at a single frequency point \(\nu\). This is reasonable since we might not want to estimate \(\Phi_k(\nu)\) on a whole grid composed by 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 Section 5 since the transformed frequencies \(\eta\) are sampled from the distribution \(\Phi_k(\nu)\) instead. For this reason we provide two independent implementations.

For the Fejer-based strategies we use Algorithm 1 with

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

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

**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 \(\nu_i\)
6: add result to frequency histogram
7: return frequency histogram

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

\[ M_{TSA} = O\left(\frac{1}{\Delta} \log \left(\frac{1}{\beta^2}\right)\right) \quad (B2) \]

for the TSA algorithm of \([12]\), while for the GIT

\[ M_{GIT} = O\left(\frac{1}{\Delta} \sqrt{\log \left(\frac{1}{\Sigma}\right) \log \left(\frac{1}{\Delta \beta} \log \left(\frac{1}{\eta_\beta}\right)\right)}\right), \quad (B3) \]

together with a number of samples per order scaling as

\[ N = O\left(\frac{M^2}{\beta^2} \log \left(\frac{1}{\eta_\beta}\right)\right), \quad (B4) \]

**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 \(\tilde{c}(\nu)\) corresponding to the integral transform being evaluated at target frequency \(\nu\)
8: return \(\Phi_k(\nu) = \tilde{c}(\nu) \cdot \tilde{v}\)

Finally, if we want the transform \(\Phi_k(\nu)\) at all the \(O(1/\Delta)\) frequency points while keeping the total error \(\beta\) (as considered in \([12]\)) we will need instead

\[ M_{TSA} = O\left(\frac{1}{\Delta} \log \left(\frac{1}{\Delta \beta}\right)^2\right) \quad (B5) \]

and

\[ M_{GIT} = O\left(\frac{1}{\Delta} \sqrt{\log \left(\frac{1}{\Sigma}\right) \log \left(\frac{1}{\Delta \beta} \log \left(\frac{1}{\eta_\beta}\right)\right)}\right), \quad (B6) \]

respectively. In this case we see that the logarithmic term for TSA also contain the resolution scale \(\Delta\).

**Appendix C: Error analysis for faulty implementation of Fejer**

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

\[ \| \hat{V}_{PKB} - V_{PKB} \| \leq \delta. \quad (C1) \]

Define \(|\Phi_A\rangle = V_{PKB} |\Psi\rangle\), without loss of generality we have

\[ |\Phi_B\rangle = \hat{V}_{PKB} |\Psi\rangle = \cos(\alpha) |\Phi_A\rangle + \sin(\alpha) |\xi\rangle \quad (C2) \]

with \(|\Phi_A\rangle |\xi\rangle = 0\). 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

\[ \sqrt{2(1 - \cos(\alpha))} \| |\Phi_A\rangle - |\Phi_B\rangle \|_2 \leq \| \hat{V}_{PKB} - V_{PKB} \| \leq \delta \]

then \(\cos(\alpha) \geq 1 - \delta^2/2\), but also

\[ D(\rho, \sigma) = \frac{1}{2} Tr[|\rho - \sigma|] = |\sin(\alpha)| \leq \sqrt{1 - \delta^2/4}. \quad (C3) \]

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

\[ \Phi_F(\sigma_q, \Delta, N) = \Phi_F(\sigma_q) = Tr\left[\Pi_q U_{QFT}^\dagger \rho U_{QFT}^\dagger\right], \quad (C4) \]

where \(\Pi_q = |q\rangle \langle q|\) and \(U_{QFT}\) the unitary implementing the Quantum Fourier Transform on the ancilla register. A similar expression holds for the faulty density matrix \(\sigma\). We now have for any \(0 \leq q < N\) that

\[ \delta_V(\Phi_F, \hat{\Phi}_F) \leq \sup_{\| \hat{V}_{PKB} - V_{PKB} \| \leq \delta} |\Phi_F(\sigma_q) - \hat{\Phi}_F(\sigma_q)| \]

\[ = \left| Tr\left[\Pi_q U_{QFT}^\dagger (\rho - \sigma) U_{QFT}^\dagger\right]\right| \quad (C5) \]

\[ \leq \frac{1}{2} Tr[|\rho - \sigma|] \leq \delta \sqrt{1 - \frac{\delta^2}{4}} \leq \delta. \]

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

\[ V_{PKB} = \prod_{k=0}^{n-1} U(t = 2\pi 2^k), \quad (C6) \]
we have by the union bound that

$$\| \tilde{V}_{\text{KHB}} - V_{\text{KHB}} \| \leq \sum_{k=0}^{n-1} \| U(2\pi 2^k) - \tilde{U}(2\pi 2^k) \|$$

$$\leq n \max_{0 \leq k < n} \delta_t(2\pi 2^k) \tag{C7}$$

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

$$\delta_v(\Phi F, \tilde{\Phi} F) \leq \log_2(N) \delta_t . \tag{C8}$$

### Appendix D: Chebyshev expansion of the gaussian kernel

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

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

where $r_L(\Lambda)$ indicates the truncation error and the coefficients are given by (see Eq.(4) of Ref. [40])

$$a_n = \left\{ \begin{array}{ll}
\frac{2n+1}{2\sqrt{\pi} \Lambda} \exp\left(-\frac{1}{4\Lambda^2}\right) J_{n/2}\left(\frac{1}{2\Lambda}\right) & \text{for even } n \\
0 & \text{for odd } n
\end{array} \right. \tag{D2}$$

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

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

Since $T_k\left(\frac{\omega - \sigma}{2}\right)$ is a polynomial of degree $k \leq L$ 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

$$b_{jk}(\sigma) = \frac{\gamma_j}{\pi} \int_{-1}^{1} \frac{dx}{\sqrt{1-x^2}} T_k\left(\frac{x-\sigma}{2}\right) T_j(x) \tag{D5}$$

$$= \frac{\gamma_j}{\pi} \sum_{m=0}^{L} T_k\left(\frac{x_m - \sigma}{2}\right) T_j(x_m).$$

In the second line we used Gauss-Chebyshev quadrature and $x_m = \cos\left(\frac{2m\pi}{2L}\right)$ the Chebyshev nodes (this is similar to the strategy used in Ref. [40]). Using this representation we can rewrite the kernel function as

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

where the new expansion coefficients are given by

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

1. **Bound of expansion coefficients**

We can bound the magnitude of $c_j$ as follows

$$|c_j| = \left| \frac{\gamma_j}{\sqrt{2\pi} \Lambda} \sum_{n=0}^{L-1} \exp\left(\frac{\gamma_j}{2\pi} d_{x_m} \right) r_L\left(\frac{\Lambda}{2}\right) T_j(x_m) \right| \leq \frac{\gamma_j}{\sqrt{2\pi} \Lambda} r_L\left(\frac{\Lambda}{2}\right) \tag{D8}$$

$$+ \frac{\gamma_j}{\sqrt{2\pi} \Lambda} \sum_{n=0}^{L-1} \exp\left(\frac{\gamma_j}{2\pi} d_{x_m} \right)$$

$$\approx \gamma_j r_L(\Lambda) + \Omega_j , \tag{D9}$$

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

$$\Omega_j \leq \frac{\gamma_j}{\sqrt{2\pi} \Lambda} \int_{0}^{\pi} dx \exp\left(\frac{\gamma_j}{2\pi} d_{x_m} \right) \frac{(\cos(\frac{\pi y}{2L}) - \sigma)^2}{2\Lambda^2} \tag{D10}$$

The integral approximately measures the number of Chebyshev nodes within the envelope of the gaussian kernel centered at $\sigma$. 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 $\sigma$ 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 $\sigma$.

Now we use the following bound for the cosine term

$$\left(\cos\left(\frac{\pi y}{2L}\right) - \sigma\right)^2 \geq (\cos(y) - \sigma)^2 - \frac{\pi^2}{L^2} \tag{D11}$$

to simplify the integrand above and obtain then

$$\Omega_j \leq \frac{\gamma_j}{\sqrt{2\pi} \Lambda} \exp\left(\frac{\gamma_j}{2\pi} d_{x_m} \right) \int_{0}^{\pi} dy \exp\left(\frac{\gamma_j}{2\pi} d_{x_m} \right) \frac{(\cos(y) - \sigma)^2}{2\Lambda^2} \tag{D12}$$

$$= \frac{\gamma_j}{\sqrt{2\pi} \Lambda} \exp\left(\frac{\gamma_j}{2\pi} d_{x_m} \right) \frac{\pi^2 L^2}{2L^2\Lambda^2} \frac{2.5\Lambda}{2}\tag{D13}$$

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

$$\Omega_j \leq \frac{\gamma_j}{\sqrt{2\pi} \Lambda} \exp\left(\frac{\gamma_j}{2\pi} d_{x_m} \right) \int_{-1/2}^{1/2} dx \exp\left(\frac{(x - 1/2)^2}{2\Lambda^2}\right) \tag{D14}$$

$$\leq \frac{\gamma_j}{\sqrt{2\pi} \Lambda} \exp\left(\frac{\pi^2 L^2}{2L^2\Lambda^2}\right) \left(2.5\Lambda\right) , \tag{D15}$$

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

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

where we anticipated the result $LA > 2$ that will be proved in the next two sections.
2. Bound on truncation error

We turn now to providing upperbounds for the error terms \( r_L(A) \) and \( r_N(A) \). Using the result from Tausch and Weckiewicz [40] we can bound the magnitude of the expansion coefficients as

\[
|a_n(A)| \leq 2\Lambda \sqrt{\pi} \exp\left(-(n+1)\kappa \left((n+1)2\Lambda^2\right)\right),
\]

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

\[
\kappa(x) = \frac{\log(x + \sqrt{1 + x^2}) - 1}{4x} \left(1 - \frac{1}{x - \sqrt{1 + x^2}}\right)^2.
\]

The total error \( r_L(A) \) can then be bounded summing a geometric series [41], the result is

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

where \( L' = L + 2 \) for \( L \) even and \( L' = L + 3 \) for \( L \) odd.

Using this result we find the total error \( r_N(x) \) in the gaussian transform Eq. (35) to be bounded as

\[
|R_L(\sigma, A)| = \frac{1}{\sqrt{2\pi\Lambda}} \left| r_L\left(\frac{\Lambda}{2}\right)\right| \leq 1 - 2\Lambda \sqrt{\pi} \exp\left(-\kappa(1)\right)\left(\frac{e}{4L'A^2}\right)^{L'}.
\]

valid in the asymptotic regime \( L' \geq 2/\Lambda^2 \). In order to guarantee a truncation error of at most \( \epsilon_R \) we now need

\[
\epsilon_R \geq 3.4 \left(\frac{e}{(L + 2)\Lambda^2}\right)^{(L+2)/2}.
\]

for the number of repetitions \( L \). Note that in this last expression we used the conservative value \( L' = L + 2 \). The inequality in Eq. (20) can be solved as

\[
L \geq \frac{2\Lambda^2}{\epsilon_R} - 2 + \frac{2\Lambda^2}{e} \log\left(\frac{3.4/\epsilon_R}{\sqrt{2\Lambda^2}}\right)
\]

where \( W \) is the Lambert W-function [39]. In order to understand the scaling of this expression we can use the less tight bound

\[
L \geq \frac{e}{\Lambda^2} + \frac{\log(3.4/\epsilon_R)}{\log(3.4/\epsilon_R)} - 2.
\]

which is usually employed in the literature (see eg. [4]).

3. Intermediate regime

We will now provide bounds in the second regime considered in the main text where the upperbound on the order \( L \) is the lower limit of validity for Eq. (D19), namely \( 0 < L' \leq \frac{2}{\Lambda^2} \). 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 \( 0 < x \leq 1 \) we have

\[
x\kappa(1) \leq \kappa(x) \leq \frac{x}{4}
\]

so that we can bound the total error \( r_N \) using

\[
\sum_{n=L+1}^{\infty} |a_n(A)| \leq 2\Lambda \sqrt{\pi} \sum_{n=L+1}^{\infty} e^{-(n+1)\kappa((n+1)2\Lambda^2)} \leq 2\Lambda \sqrt{\pi} \sum_{n=L+1}^{\infty} e^{-(n+1)^2\kappa(1)\Lambda^2} \leq 2\Lambda \sqrt{\pi} \int_{L}^{\infty} dx e^{-(x+1)^2\kappa(1)\Lambda^2} = \frac{\pi}{\sqrt{2\kappa(1)}} \text{erfc}\left((L + 1)\Lambda\sqrt{2\kappa(1)}\right).
\]

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

\[
|R_L(\sigma, A)| \leq \frac{1}{2\Lambda} \sqrt{\pi} \kappa(1) \text{erfc}\left((L + 1)\Lambda\sqrt{\kappa(1)/2}\right).
\]

This error can be bounded from above using

\[
|R_L(\sigma, A)| \leq \frac{1}{\sqrt{2\Lambda^2}} \frac{1}{L + 1} \exp\left(-(L + 1)^2\frac{\Lambda^2}{2}\right).
\]

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

\[
L \geq \sqrt{\frac{2}{\pi}} \frac{1}{\Lambda_n} - 1.
\]

As we did in the general case above, if we want a truncation error of at most \( \epsilon_R \) we need

\[
\epsilon_R \geq \frac{1}{\sqrt{2\Lambda^2}} \frac{1}{L + 1} \exp\left(-(L + 1)^2\frac{\Lambda^2}{2}\right).
\]

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

\[
L + 1 \geq \frac{1}{\Lambda_n} \left\lfloor \frac{1}{2} \frac{W\left(\frac{1}{2\Lambda_n^2\epsilon_R^2}\right)}{\epsilon_R} \right\rfloor.
\]
We can now use another result from [42], Theorem 2.1, to find the sufficient condition
\[
L = \left[ \frac{1}{\Lambda} \sqrt{\frac{1}{\kappa(1)} g \left( \frac{1}{\sqrt{2\kappa(1)\Lambda e_R}} \right)} \right] - 1, \quad (D30)
\]
where for convenience we defined the auxiliary function
\[
g(x) = \log(x) - \frac{1}{4} \log\left( \log(x^2) \right). \quad (D31)
\]
These estimates hold for sufficiently small target errors
\[
\Lambda e_R \leq \frac{1}{\sqrt{2\kappa(1)e}} \approx 0.54, \quad (D32)
\]
a conditions that ensures that also Eq. (D27) is satisfied. We finally note that it is also possible to find a bound on \(L\) valid for any value of the target error
\[
L = \left[ \frac{1}{\Lambda} \sqrt{\frac{2}{\kappa(1)} \log\left( \sqrt{\frac{\pi}{\kappa(1)} \frac{1}{2\Lambda e_R}} \right)} \right]. \quad (D33)
\]
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
\[
\epsilon_R^{\text{min}} \leq \frac{1}{\sqrt{2\kappa(1)}} \frac{1}{2 + \Lambda^2} \exp\left( -\frac{\kappa(1) (2 + \Lambda)^2}{2} \right) \leq \frac{e^{-2\kappa(1)}}{\sqrt{8\kappa(1)}} \exp\left( -\frac{2\kappa(1)}{\Lambda^2} \right) \leq \exp\left( -\frac{1}{2\Lambda^2} \right). \quad (D34)
\]
and this is valid for any reasonable value \(\Lambda\) (to satisfy Eq. (D27) the condition is \(\Lambda \leq 5\)).