APPROXIMATING SPECTRAL DENSITIES OF LARGE MATRICES

LIN LIN *, YOUSEF SAAD †, AND CHAO YANG ‡

Abstract. In physics, it is sometimes desirable to compute the so-called *Density Of States* (DOS), also known as the *spectral density*, of a Hermitian (or symmetric) matrix A. The spectral density can be viewed as a probability density distribution that measures the likelihood of finding eigenvalues near some point on the real line. The most straightforward way to obtain this density is to compute all eigenvalues of A. But this approach is generally costly and wasteful, especially for matrices of large dimension. There exist alternative methods that allow us to estimate the spectral density function at much lower cost. The major computational cost of these methods is in multiplying A with a number of vectors, which makes them appealing for large-scale problems where products of the matrix A with arbitrary vectors are inexpensive. This paper defines the problem of estimating the spectral density carefully. It then surveys a few known methods for estimating the spectral density, and proposes some new variations of existing methods. All methods are discussed from a numerical linear algebra point of view.

Key words. spectral density, density of states, large scale sparse matrix, approximation of distribution, quantum mechanics

AMS subject classifications. 15A18, 65F15

1. Introduction.

Given an $n \times n$ Hermitian (sparse real symmetric in what follows) matrix A, scientists in various disciplines often want to compute its *Density Of States* (DOS), or *spectral density*. Formally, the DOS is defined as

$$\phi(t) = \frac{1}{n} \sum_{j=1}^{n} \delta(t - \lambda_j), \tag{1.1}$$

where δ is the Dirac δ-function or Dirac distribution, and the λ_j's are the eigenvalues of A, assumed here to be labeled increasingly. We also denote by u_j the eigenvector corresponding to λ_j, so A has the eigen-decomposition

$$A = \sum_{j=1}^{n} \lambda_j u_j u_j^T. \tag{1.2}$$

The number of eigenvalues in an interval $[a, b]$ can be expressed as

$$\eta_{[a,b]} = \int_a^b \sum_j \delta(t - \lambda_j)\ dt \equiv \int_a^b n\phi(t) dt\ . \tag{1.3}$$

Therefore, one can view $\phi(t)$ as a probability distribution function which gives the probability of finding eigenvalues of A in a given infinitesimal interval near t. Ideally, if one has access to all the eigenvalues of A, computing the DOS would become a trivial task. However, in many scientific applications the matrix size is too large for eigenvalue solvers to be efficiently applied, especially when a large number of

*Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 linlin@lbl.gov

†Department of Computer Science and Engineering, University of Minnesota, Twin Cities, Minneapolis, MN 55455 saad@cs.umn.edu

‡Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 cyang@lbl.gov

eigenvalues are needed. Alternative algorithms that do not require all the eigenvalues of A are therefore of great interest in practice.

In solid state physics and electronic structure theory [3, 40], eigenvalues of a Hamiltonian matrix usually represent single-particle energy levels. The corresponding DOS gives the number of single-particle energy levels per unit energy. The DOS can also be used to characterize the so-called valence band (energy levels below the chemical potential) and the conduction band (energy levels beyond the chemical potential) along the energy spectrum. The gap between the valence band and the conduction band is called the band-gap, and the DOS inside the band-gap is zero.

In the study of excited state properties of electrons, the joint DOS associated with both the conduction and valence bands allows one to examine the interband transitions owing to the absorption and emission of light [40, 35].

At the molecular level, the DOS of the phonon spectrum or the normal modes of molecules, which are related to the eigenvalues of the Hessian of the potential function with respect to atomic coordinates, characterizes thermodynamic properties of solids or molecules. The DOS can also be used as a spectral measure for integrating certain quantities to yield the heat capacity and the entropy of a molecular system. For example, the heat capacity C_v of a molecular system has the expression [58]:

$$C_v = k_B \int_0^\infty \frac{(\hbar\omega c/k_B T)^2 e^{-\hbar\omega c/k_B T}\phi(\omega)}{(1-e^{-\hbar\omega c/k_B T})^2} d\omega, \tag{1.4}$$

where k_B is the Boltzmann constant, c is the speed of light, $\hbar$ is Planck's constant, T is the temperature and $\omega = \sqrt{\lambda}$ is a vibration frequency.

Efficient algorithms for computing the DOS are also developed and employed for optical absorption spectra calculations [55], the transition state theory for chemical reactions [28], quantum transport simulations [17], dynamical correlation function calculations [10], and many other recent applications [18, 27, 2, 1, 7, 47, 9, 4, 48].

Recall that a Dirac δ-function is not a proper function, it is a member of the adjoint space of a Hilbert space. It is formally defined through applications to a test function g:

$$\langle \delta(\cdot - \lambda), g\rangle = \int_{-\infty}^{\infty} \delta(t-\lambda) g(t) dt \equiv g(\lambda),$$

where g is a rapidly decreasing $\mathcal{C}^\infty$ function with bounded support. More details on distributions can be found, e.g., in [44]. Because the spectral density is formally defined as the sum of Dirac distributions, it may be thought at first that it cannot be approximated by a smooth function. Indeed, when one views this as a function, it appears to be highly discontinuous, and so it may seem unwise to try to approximate it by smooth functions, e.g., polynomials. However, a distribution can also be viewed as a discretized version of some smooth function after it is sampled at a finite number of points, and our goal is to approximate that original smooth function. Often in quantum mechanical applications, the spectrum of the Hamiltonian defined in the continuous space is continuous, and the function (1.1) shown above can then be viewed as just a discrete version of a continuous function.

A rigorous approach to mitigate the issue, one that has both theoretical and practical importance, is to consider a continuous version of the distribution ϕ obtained, in essence, by "blurring" it. For example, we can consider a function of the form:

$$\phi_\sigma(t) = \frac{1}{n} \sum_{j=1}^{n} h_\sigma(t-\lambda_j), \tag{1.5}$$

in which $h_\sigma(t)$ can be any infinitely differentiable function whose integral in $(-\infty, \infty)$ is one, and which is zero or very small outside of a narrow interval $[-C\sigma, C\sigma]$ where $C > 0$ is a small constant independent of σ. One example is the Gaussian function:

$$h_\sigma(t) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-\frac{t^2}{2\sigma^2}}. \tag{1.6}$$

Such a setup may be appealing and can help establish some theory as well as develop algorithms as will be seen later.

A related question one might have is: How can we compare the original density ϕ with an approximate one $\widetilde{\phi}$? The distribution ϕ is obtained as the limit of a sequence of functions $\phi_\sigma \in \mathcal{C}^\infty$ with compact support, in the sense that

$$\lim_{\sigma \to 0} \langle \phi_\sigma, g \rangle = \langle \phi, g \rangle,$$

for every $\mathcal{C}^\infty$ and compactly supported function g. Since

$$\begin{aligned}
|\langle \phi_\sigma - \widetilde{\phi}, g \rangle| &\le \|\phi_\sigma - \widetilde{\phi}\|_1 \|g\|_\infty, \\
|\langle \phi_\sigma - \widetilde{\phi}, g \rangle| &\le \|\phi_\sigma - \widetilde{\phi}\|_2 \|g\|_2, \\
|\langle \phi_\sigma - \widetilde{\phi}, g \rangle| &\le \|\phi_\sigma - \widetilde{\phi}\|_\infty \|g\|_1,
\end{aligned}$$

where $\|\cdot\|_p$ stands for the L^p norm, we will only need to first approximate the "true" DOS ϕ by a "surrogate" DOS ϕ_σ, and then measure the distance between ϕ_σ and $\widetilde{\phi}$ either in the L^1, L^2 or L^∞ sense to quantify the error.

The goal then is to find a smooth approximation to the spectral density function ϕ. Because calculating the spectral density is such an important problem in quantum mechanics, there is an abundant literature devoted to this problem and research in this areas was extremely active in the 1970s and 1980s. Clever and powerful methods have been developed by physicists and chemists [15, 54, 13, 57] for this purpose and this paper will review a few of them from a linear algebra viewpoint. What is striking about research in this area is the wide variety of methods used as well as the mathematical depth of the various techniques.

The Kernel Polynomial Method [50, 56] expands the spectral density in a basis of orthogonal polynomials by resorting to the reproducing kernel property of orthogonal polynomials. The method, which is widely used in a variety of calculations that require the DOS [16], has continued to receive a tremendous amount of interest in the last few years [16, 8, 32, 49]. Haydock's method [26] exploits the Lanczos algorithm [36] in an effort to evaluate functions of the form $f(t) = v^T(A - (t + i\eta)I)^{-1}v$ whose peaks reveal the presence of nearby poles, i.e. approximate eigenvalues of A. Its use for DOS calculation has recently been discussed in [5, 39, 6]. The Lanczos algorithm itself can be used to estimate the spectrum and the spectral density associated with each estimated eigenvalue. Lanczos also proposed a method [37], termed a "spectroscopic" approach, which turns out to be better suited for calculating spectral densities than actual spectra of Hermitian matrices. Lanczos's technique also relies on detecting peaks of a certain function, that are indicative of the presence of eigenvalues.

In this paper we wish to take a careful look at the problem of computing the DOS and to describe a few common techniques used in the literature. One of our goals is to expose these techniques from a numerical linear algebra viewpoint. Not all methods described are new, but we will see a new way of developing the Kernel Polynomial Method, which may have a few advantages. In order to directly compute the surrogate DOS ϕ_σ, we will also describe a new algorithm based on orthogonal expansions of Gaussian functions, as well a method based on using the standard Lanczos algorithm. The next section will discuss spectral density functions in some detail.

2. Spectral densities as regular functions. One of the first questions one might ask when studying spectral densities is: how can we evaluate accuracy? This was briefly discussed in the introduction. A practical viewpoint is to evaluate $\|\phi_\sigma - \widetilde{\phi}\|_p$ with the norm p being $1, 2$ or ∞ where ϕ_σ is a certain smoothed version of ϕ. The function ϕ_σ can be viewed as a surrogate of ϕ which can be used for practical purposes. This viewpoint becomes useful for example for the purpose of plotting a spectral distribution for visually comparing it with some approximation.

Let us consider this issue in some more detail. When drawing spectral density curves, we will be interested in some approximations to the density functions shown above. The first approximation that comes to mind is to exploit the "probabilistic" interpretation of ϕ: we can just subdivide the interval $[a, b]$ into non-overlapping subintervals $I_j = (t_j - h/2, t_j + h/2)$ and define the value of ϕ in I_j to be the constant n_i/n where n_i is the number of eigenvalues in I_j. This "histogram" approach will give a surrogate which is a piecewise constant function $\phi_{P,h}$ that depends on the interval width h. A slight variation of this scheme is to replace each of the Dirac functions $\delta(t - \lambda_i)$ in (1.3) by a piecewise constant function which has a value of $1/h$ in an interval of width h centered at λ_i and zero elsewhere. We would have:

$$\phi(t) \approx \phi_{P,h}(t) = \frac{1}{nh} \sum_{j=1}^{n} \chi \left(\frac{t - \lambda_j}{h/2} \right)$$

where $\chi(t)$ represents the indicator function for the interval $[-1, 1]$, which equals to one if t is in the interval and zero otherwise. The integral of this function on the whole real line is one as required. We note that piecewise constant surrogates will tend to give us rough functions with big variations if the selected width h is small.

A better alternative is to use Gaussians, replacing again each $\delta(t - \lambda_i)$ by the function $h_\sigma(t - \lambda_i)$ where h_σ is defined by (1.6). The parameter σ plays the role of h in the above case of piecewise constant surrogates. Larger values of σ will lead to smooth curves at the expense of accuracy. Too small values of σ will again lead to rough curves that have peaks at the eigenvalues and zeros elsewhere. This is illustrated in Figure 2.1 where σ takes 4 different values. We can see that as σ increases, ϕ_σ becomes smoother. When $\sigma = 0.96$, which corresponds to a very smooth spectral density, we can still see the global profile of the eigenvalue distribution, although local variation of the spectral density is mostly averaged out.

One of the methods to be described later will consist of seeking an approximation of a smoothed version of ϕ obtained by discretizing the interval $[a,\ b]$ uniformly and placing Gaussians at each knot. In this way the function that is approximated is a regular function.

A heuristic criterion we adopt for choosing σ in the case where uniform intervals are used is based on starting with some width h that represents the discretization width of the interval. To obtain a compromise between smoothness and accuracy we select σ so that the value of each Gaussian at its peak, is κ times its value at the end point of each sub-interval, where κ is a ratio larger than one. This means that

$$1 = \kappa e^{-(h/2)^2/(2\sigma^2)} \to \log(\kappa) = \frac{h^2}{8\sigma^2} \to \sigma = \frac{h}{2\sqrt{2\log(\kappa)}}. \qquad (2.1)$$

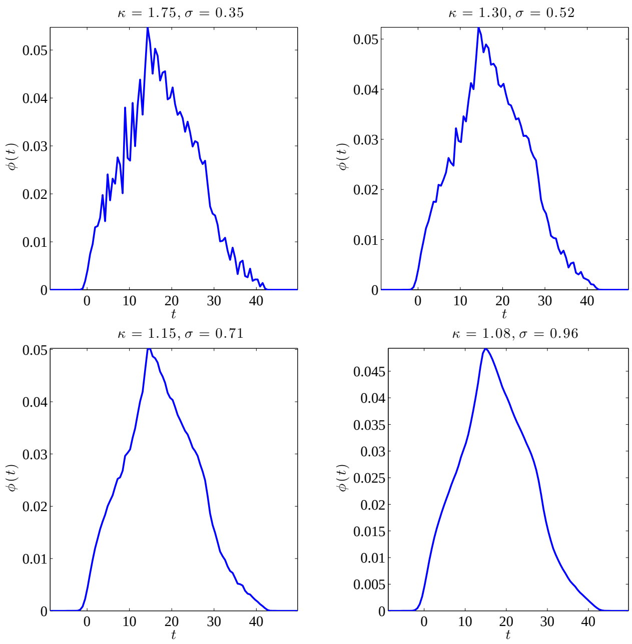

Fig. 2.1: Various surrogate DOS ϕ_σ obtained by blurring the exact DOS (sum of δ-functions positioned at eigenvalues) of a Hamiltonian matrix associated with a quantum mechanical system.

If κ is close to one, then the overall curve, which is the sum of all the Gaussians will be very smooth. This is illustrated in Figure 2.2.

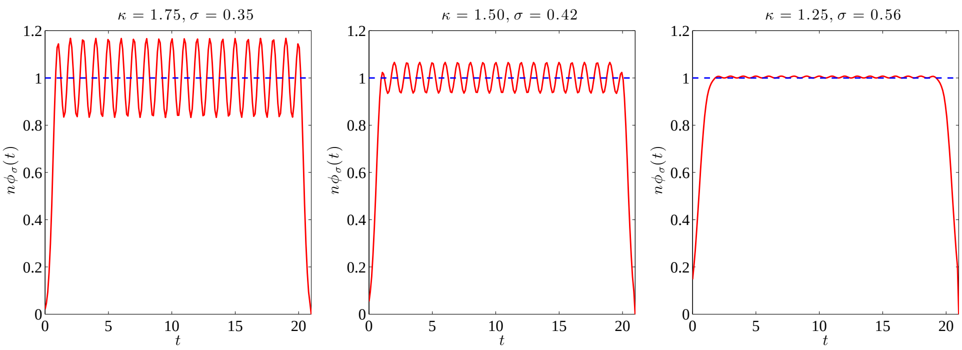

Fig. 2.2: The sum of Gaussians of the form $n\phi_\sigma(t)$ when h_σ is given by (1.6), and σ is selected according to (2.1) for $h = 0.75$, $\kappa = 1.75, 1.50, 1.25$. The eigenvalues are located at $1, 2, \cdots, 20$.

The criterion given by (2.1) is motivated by the case when the distribution is uniform, as Figure 2.2 illustrates. In this situation, we would like to have a function that has values not too far from one in the middle of the interval and at the interval intersections. One can decide on the level of smoothness desired in the final approximation and select h and κ accordingly. One can choose $h = \Delta E/N$, where ΔE is the width of the spectrum, and N is the intended resolution of the DOS. Values of N between 30 and 100 and κ between 1.25 and 2.0 seem to serve as reasonable "default" selections.

When calculating these functions for the different λ_i we only need to evaluate the Gaussians at points nearby a given site λ_i because of the rapid decay of Gaussians. A calculation similar to the one given above will tell us which intervals are to be kept if we need the accuracy to be within a certain tolerance.

We remark that the optimal choice of σ, and therefore the smoothness of the approximate DOS, is application dependent. Ideally, σ should be chosen to be as large as possible so that the surrogate DOS ϕ_σ is easy to approximate numerically. However, increasing σ could cause an undesirable loss of detail and yield an erroneous result, e.g, when computing the heat capacity discussed in Section 1. It is up to the user to select a value of σ that balances accuracy and efficiency.

3. The Kernel Polynomial Method. The Kernel Polynomial Method (KPM) was proposed by Silver and Röder [50] and Wang [56] in the mid-1990s to calculate the DOS. See also [51, 52, 14, 41] among others where similar approaches were also used.

The KPM method essentially determines the exact DOS of a matrix as a sum of Dirac δ-functions by its expansion in Chebyshev polynomials. For simplicity we assume that the eigenvalues are in the interval $[-1, 1]$. As is the case for all methods which rely on Chebyshev expansions, a change of variables is first performed to map the interval $[\lambda_{\min}, \lambda_{\max}]$ into $[-1, 1]$. The goal is to estimate the spectral density function (1.1). For this, we will approximate $\phi(t)$ by a finite expansion in a basis of orthogonal polynomials, in this case, Chebyshev polynomials of the first kind. Following the Silver-Röder paper [50], we include, for convenience, the inverse of the weight function into the spectral density function, so we expand instead the distribution:

$$\hat{\phi}(t) = \sqrt{1-t^2}\phi(t) = \sqrt{1-t^2} \times \frac{1}{n}\sum_{j=1}^{n} \delta(t - \lambda_j). \tag{3.1}$$

Then, we have the (full) expansion

$$\hat{\phi}(t) = \sum_{k=0}^{\infty} \mu_k T_k(t). \tag{3.2}$$

where the expansion coefficients μ_k are formally defined by

$$\begin{aligned}
\mu_k &= \frac{2-\delta_{k0}}{\pi} \int_{-1}^{1} \frac{1}{\sqrt{1-t^2}} T_k(t)\hat{\phi}(t)dt \\
&= \frac{2-\delta_{k0}}{\pi} \int_{-1}^{1} \frac{1}{\sqrt{1-t^2}} T_k(t)\sqrt{1-t^2}\phi(t)dt \\
&= \frac{2-\delta_{k0}}{n\pi} \sum_{j=1}^{n} T_k(\lambda_j). \qquad (3.3)
\end{aligned}$$

Here δ_{ij} is the Kronecker δ symbol so that $2 - \delta_{k0}$ is equal to 1 when $k = 0$ and to 2 otherwise.

Thus, apart from the scaling factor $(2 - \delta_{k0})/(n\pi)$, μ_k is the trace of $T_k(A)$ and this can be estimated by various methods including, but not limited to, stochastic approaches. There are variations on this idea starting from the use of different orthogonal polynomials, to alternative ways in which the traces can be estimated.

A common solution to the problem of estimating $\text{Trace}(T_k(A))$ is to use a stochastic argument as suggested in [50] and, e.g., also in [29, 53]. This entails generating a large number of random vectors $v_0^{(1)}, v_0^{(2)}, \cdots, v_0^{(n_{\text{vec}})}$ with each component obtained from a normal distribution with zero mean and unit standard deviation, and each vector is normalized such that $\|v_0^{(l)}\| = 1, l = 1, \ldots, n_{\text{vec}}$. The subscript 0 is added to indicate that the vector has not been multiplied by the matrix A. Then we can estimate the trace of $T_k(A)$ as follows:

$$\text{Trace}(T_k(A)) \approx \frac{1}{n_{\text{vec}}} \sum_{l=1}^{n_{\text{vec}}} \left(v_0^{(l)}\right)^T T_k(A) v_0^{(l)}. \tag{3.4}$$

Then this will lead to the desired estimate:

$$\mu_k \approx \frac{2 - \delta_{k0}}{n\pi n_{\text{vec}}} \sum_{l=1}^{n_{\text{vec}}} \left(v_0^{(l)}\right)^T T_k(A) v_0^{(l)}. \tag{3.5}$$

Now we consider the computation of each term $\left(v_0^{(l)}\right)^T T_k(A) v_0^{(l)}$. For simplicity we drop the superscript l and denote by $v_0 \equiv v_0^{(l)}$. The 3-term recurrence of the Chebyshev polynomial is exploited to compute $T_k(A)v_0$:

$$T_{k+1}(A)v_0 = 2AT_k(A)v_0 - T_{k-1}(A)v_0 \tag{3.6}$$

so if we let $v_k \equiv T_k(A)v_0$, we have

$$v_{k+1} = 2Av_k - v_{k-1}.$$

Once the scalars $\{\mu_k\}$ are determined, we would in theory get the expansion for $\phi(t) = \frac{1}{\sqrt{1-t^2}}\hat{\phi}(t)$. Practically however, the approximate density of states will be limited to Chebyshev polynomials of degree M, so ϕ is approximated by:

$$\widetilde{\phi}_M(t) = \frac{1}{\sqrt{1-t^2}} \sum_{k=0}^{M} \mu_k T_k(t). \tag{3.7}$$

For a general matrix A whose eigenvalues are not necessarily in the interval $[-1, 1]$, a linear transformation is first applied to A to bring its eigenvalues to the desired interval. Specifically, we will apply the method to the matrix

$$B = \frac{A - cI}{d},$$

where

$$c = \frac{\lambda_{\min} + \lambda_{\max}}{2}, \quad d = \frac{\lambda_{\max} - \lambda_{\min}}{2}. \tag{3.8}$$

It is important to ensure that the eigenvalues of B are within the interval $[-1, 1]$. In an application requiring a similar approach [59], we obtain the upper and lower bounds of the spectrum from Ritz values provided by a standard Lanczos iteration.

We ran a few Lanczos steps but extended the interval $[\lambda_{\min}, \lambda_{\max}]$ by using the bounds obtained from the Lanczos algorithm.

Specifically, the upper bound for $\lambda_{\max}$, for example, is set to $[\widetilde{\lambda}_n, \widetilde{\lambda}_n + \eta]$ where $\eta = \|(A - \widetilde{\lambda}_n I)\widetilde{u}_n\|$, and $(\widetilde{\lambda}_n, \widetilde{u}_n)$ is the (algebraically) largest Ritz pair of A. This is plausible because we know that there is an eigenvalue in the interval $[\widetilde{\lambda}_n, \widetilde{\lambda}_n + \eta]$ (see e.g., [46].) Even though there is no theoretical guarantee that $[\widetilde{\lambda}_n, \widetilde{\lambda}_n + \eta]$ is indeed an upper bound for $\lambda_{\max}$, in practice it often is. Refined bounds which invoke the square of η ([46]) can also be used. Note that η can be computed without explicitly computing the eigenvector $\widetilde{u}_n$. To summarize, we outline the major steps of the KPM for approximating the spectral density of a sparse matrix in Algorithm 1.

Algorithm 1: The Kernel Polynomial Method.

Input: Real symmetric matrix A. A set of points $\{t_i\}$ at which DOS is to be evaluated, the degree M of the expansion polynomial.
Output: Approximate DOS $\{\widetilde{\phi}_M(t_i)\}$.

Set $\mu_k = 0$ for $k = 0, \cdots, M$;
for $l = 1 : n_{\text{vec}}$ **do**
  Select a new random vector $v_0^{(l)}$;
  for $k = 0 : M$ **do**
    Compute $T_k(A)v_0^{(l)}$ using 3-term recurrence (3.6);
    Update μ_k using (3.5);
  end for
end for
Evaluate the average value of $\{\widetilde{\phi}_M(t_i)\}$ at the given set of points $\{t_i\}$ using (3.7);

Because the KPM method amounts to approximating a discontinuous function, it may lead to some numerical difficulties. The Dirac δ-function can be viewed as the derivative of the Heaviside step function $H(t)$ such that

$$H(t) = \begin{cases} 1, & t \geq 0, \\ 0, & t < 0. \end{cases} \tag{3.9}$$

The discontinuity at 0 causes the polynomial approximation of $H(t)$ and $\delta(t)$ to exhibit rapid oscillations around $t = 0$. These so-called Gibbs oscillations can lead to loss of accuracy in the approximate DOS generated by KPM. A common approach used to damp these oscillation, is to use the Chebyshev-Jackson approximation [30, 45, 31], which modulates the coefficients μ_k with a damping factor g_k^M defined by

$$g_k^M = \frac{\left(1 - \frac{k}{M+1}\right)\sin(\alpha_M)\cos(k\alpha_M) + \frac{1}{M+1}\cos(\alpha_M)\sin(k\alpha_M)}{\sin(\alpha_M)}, \tag{3.10}$$

where $\alpha_M = \frac{\pi}{M+1}$. Consequently, the damped Chebyshev expansion has the form

$$\widetilde{\phi}_M(t) = \sum_{k=0}^{M} \mu_k g_k^M T_k(t).$$

Approximations to the δ-function $\delta(t)$ by Chebyshev polynomial expansions with and without the Jackson damping are compared in the left subfigure in Figure 3.1. The

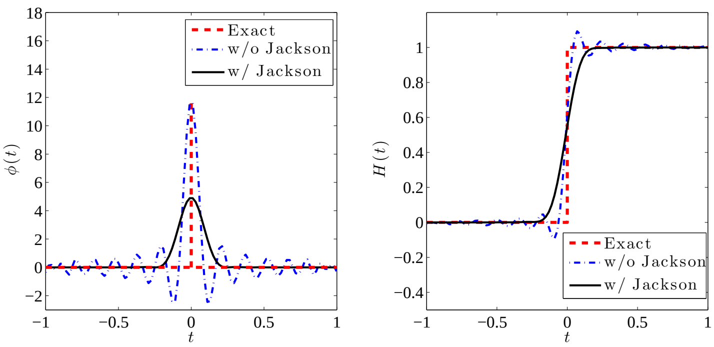

Fig. 3.1: Chebyshev expansion with and without the Jackson damping for Dirac δ-function $\delta(t)$ (left) and Heaviside function $H(t)$ (right). The Chebyshev polynomial degree is set to 40.

degrees of both polynomial approximations are set to 40. The right subfigure in Figure 3.1 shows similar approximations for the Heaviside function $H(t)$. Observe how the approximate function obtained with Jackson damping is smooth and free of the rapid oscillations seen in the Chebyshev expansion.

To illustrate how well this algorithm works, we now give an example on using KPM to approximate the spectral density of a relatively small matrix. The matrix we use for this example is a modified two-dimensional (2D) Laplacian operator with zero Dirichlet boundary condition discretized on a domain of size $[0, 30] \times [0, 30]$ using a five-point finite difference stencil. The modification involves adding a diagonal matrix as a discretized potential function. The diagonal matrix is generated by adding two Gaussians, one centered at the point (4,5) of the domain and the other at the point (25,15). In Figure 3.2, we plot both the approximate spectral density constructed by KPM with a 80-degree Chebyshev polynomial and 10 random vectors for estimating the trace in (3.4), and the smoothed "exact" spectral density constructed by Gaussian smoothing with the smoothing parameter $\sigma = 0.56$ (obtained by setting $\kappa = 1.25, h = 0.75$).

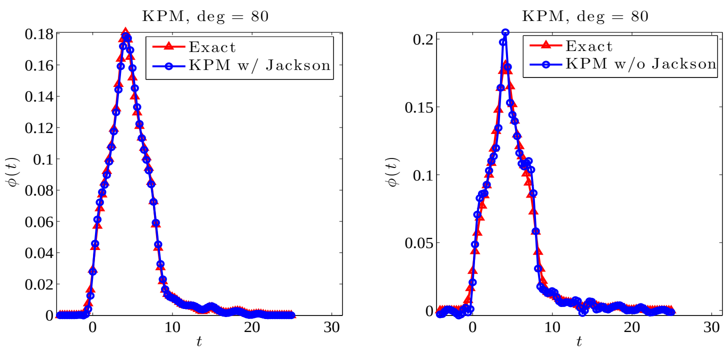

Fig. 3.2: The approximate spectral density constructed by KPM with a 80-degree Chebyshev polynomial for the modified 2D Laplacian example with Jackson damping (left) and without Jackson damping (right).

From a computational point of view, some savings in time can be achieved if we are willing to store more vectors. This is due to the formula:

$$T_p(t)T_q(t) = \frac{1}{2}\left[T_{p+q}(t) - T_{|p-q|}(t)\right],$$

from which we obtain

$$T_{p+q}(t) = 2\ T_p(t)T_q(t) + T_{|p-q|}(t).$$

For a given k we can use the above formula with $p = \lceil k/2 \rceil$ and $q = k - p$. This requires that we compute and store $v_r = T_r(A)v_0$ for $r \leq p$. Then the moments $v_0^T T_r(A)v_0$ for $r \leq p$ can be computed in the usual way, and for $r = p + q > p$ we can use the formula:

$$v_0^T T_{p+q}(A)v_0 = 2\ v_p^T v_q + v_0^T v_{|p-q|}.$$

This saves 1/2 of the matrix-vector products at the expense of storing all the previous $\{v_r\}$. It is not practical for high degree polynomials.

Chebyshev polynomials are not the only types of orthogonal polynomials that can be used in the expansion. We can use any other type of orthogonal polynomials. The only practical requirement is that we explicitly know the 3-term recurrence for the polynomials. For example, we can use the Legendre polynomials $L_k(t)$ which obey the following 3-term recursion

$$L_0(t) = 1, \quad L_1(t) = t, \quad (k+1)L_{k+1}(t) = (2k+1)tL_k(t) - kL_{k-1}(t).$$

See, for example [11], for three-term recurrences for a wide class of such polynomials, e.g., all those belonging to the Jacobi class, which include Legendre and Chebyshev polynomials as particular cases.

Figure 3.3 shows the quality of the approximation to the spectral density of the modified 2D Laplacian matrix obtained from the KPM with a Legendre polynomial expansion. The degree of the polynomial approximation is set to 80. We observe that the performance of the Legendre polynomial expansion is similar to that of a Chebyshev expansion without Jackson damping.

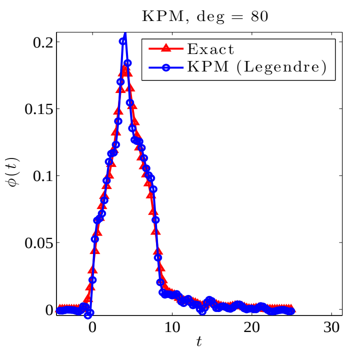

Fig. 3.3: The approximate spectral density constructed by KPM with a 80-degree Legendre polynomial for the modified 2D Laplacian matrix.

4. Spectroscopic approaches. In his 1956 book titled "Applied Analysis" [37], Lanczos described a method for computing spectra of Hermitian matrices, which he termed "spectroscopic". This approach, which also relies heavily on Chebyshev polynomials, is rather unusual in that it assimilates the spectrum of a matrix to a collection of frequencies and the goal is to detect these frequencies by Fourier analysis. Because it is not competitive with modern methods for computing eigenvalues, this technique has lost its appeal. However, it appears to be better suited for computing approximate spectral densities and so we consider here an adaptation of this approach for this purpose.

Assume that B is the same matrix as the one in Section 3 which has eigenvalues in [-1,1]. Let v_0 be any initial vector and assume that it is expanded using the eigenvectors of A as

$$v_0 = \sum_{j=1}^{n} \beta_j u_j, \tag{4.1}$$

where the coefficient $\beta_j = u_j^T v_0$. Consider the vector $v_k = T_k(A)v_0$, for $k = 0, \cdots, M$. If we set $\lambda_j = \cos\theta_j$, then

$$v_0^T v_k = \sum_{j=1}^{n} \beta_j^2 T_k(\lambda_j) = \sum_{j=1}^{n} \beta_j^2 \cos(k\theta_j). \tag{4.2}$$

The number $v_0^T v_k$ can be viewed as a discretized version of the function

$$f(t) = \sum_{j=1}^{n} \beta_j^2 \cos(t\theta_j) \tag{4.3}$$

sampled at the integer values $t = 0, \cdots, M$. The problem is then to find all values of θ_j. In order to find the θ_j's, Lanczos reasoned as follows. The function f is a periodic function and we can compute its Fourier (in this case cosine) transform:

$$F(p) = \frac{1}{2}\left(f(0) + (-1)^p f(M)\right) + \sum_{k=1}^{M-1} f(k) \cos\frac{kp\pi}{M}, \quad p = 0, \cdots, M. \tag{4.4}$$

Note that as is customary the end values are halved to account for the discontinuity of the data at the interval boundaries.

The idea then is that if f has an eigenvalue $\lambda = \cos\theta$, there should be a component $\cos(\theta t)$, which should be revealed by a peak at the point

$$p = \frac{l\theta}{\pi} \ .$$

Indeed, if we had exactly one term of the form $f(t) = \cos\left(t\frac{p\pi}{M}\right)$ where $0 \le p \le M$ is some integer, then only $F(M)$ would be nonzero. In the end, if we identify a peak value at p_j then, the corresponding eigenvalue λ_j corresponds to the angle $\theta_j = (p_j/M)\pi$, and so, $\lambda_j = \cos(\theta_j) = \cos(p_j\pi/M)$. In other words, after obtaining $F(p), p = 0, \cdots, M$, one can define an associated function

$$\hat{F}(\hat{p}) \equiv F\left(\frac{M}{\pi}\arccos\hat{p}\right), \quad \hat{p} = \cos(p\pi/M), p = 0, \cdots, M. \tag{4.5}$$

Then $\hat{F}(\hat{p})$ can be interpolated on the entire interval of $[-1,1]$, and finally the approximate DOS can be evaluated as

$$\widetilde{\phi}_M(t) = C_M \hat{F}\left(\frac{t-c}{d}\right), \tag{4.6}$$

where c, d are defined in Eq. (3.8), and C_M is a normalization constant so that $\int \widetilde{\phi}_M(t)dt = 1$.

Note that the quantity $\eta_k \equiv v_0^T v_k$ is based on a single vector v_0 and this will cause problems when one or more of the corresponding components β_j in (4.1) are small. To avoid this we will proceed in the same way as with the KPM method by taking many vectors and averaging the above quantities.

The basic steps of the spectroscopic method are outlined in Algorithm 2. Fig. 4.1 illustrates the performance of the method for the modified 2D Laplacian matrix with the polynomial degrees set to 40 and 100, respectively. The performance of the spectroscopic method is comparable to that of the KPM method without Jackson damping. However, the usage of Jackson damping is more difficult to justify.

Algorithm 2: Lanczos spectroscopic method

Input: Real symmetric matrix A. A set of points $\{t_i\}$ at which the spectral density is to be evaluated and the degree of Chebyshev polynomial approximation M.

Output: Approximate DOS $\{\widetilde{\phi}_M(t_i)\}$.

1: Set $\eta_k = 0$ for $k = 0, \cdots, M$;
2: **for** $l = 1 : n_{\text{vec}}$ **do**
3: Select a new random vector $v_0^{(l)}$;
4: **for** $k = 0 : M$ **do**
5: Compute $v_{k+1}^{(l)}$ via the three-term recurrence $v_{k+1}^{(l)} = 2Av_k^{(l)} - v_{k-1}^{(l)}$ (for $k = 0$, $v_1^{(l)} = Av_0^{(l)}$);
6: Compute $\eta_k \leftarrow \eta_k + \left(v_0^{(l)}\right)^T v_k^{(l)}$;
7: **end for**
8: **end for**
9: Set $\eta_k = \eta_k / n_{\text{vec}}$ for all $k = 0, 1, ..., M$;
10: Take the discrete cosine transform of $\{\eta_k\}$ according to Eq. (4.4) to obtain $F(p)$ for $p = 0, 1, ..., M$;
11: Evaluate $\widetilde{\phi}_M(t_i)$ according to Eq. (4.6) for each t_i.

5. The Delta-Chebyshev expansion algorithm. In this section, we will use the insight provided by the spectroscopic method seen in the previous section to derive a new algorithm which has some of the features of the spectroscopic method and which will turn out to be mathematically equivalent to KPM. The main idea relies on an expansion of the δ-function at a given set of points $\{t_i\}$.

A δ-function defined at t_i can be viewed as a spectral probe at t_i. In theory, the presence of an eigenvalue at t_i can be detected by integrating the δ-function over the entire spectrum of A with respect to a spectral point measure defined at eigenvalues only, i.e., $\int \delta(t-t_i)dt \equiv \sum_{i=1}^{n} \delta(\lambda_i - t_i)/n$. The integral returns $+\infty$ if t_i is an eigenvalue of A and 0 otherwise. However, in practice, this integration cannot be performed without knowing the eigenvalues of A in advance.

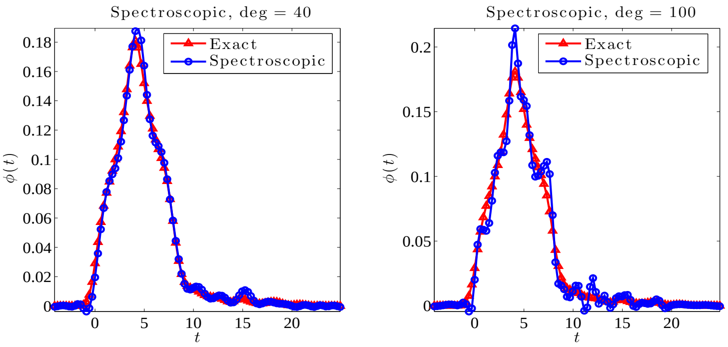

Fig. 4.1: The approximate spectral density constructed by the spectroscopic method for the modified 2D Laplacian matrix, obtained with polynomial degrees set to 40 (left) and 100 (right), respectively.

A practical probing scheme can be devised by replacing the δ-function with a polynomial approximation, and approximating the spectral measure using the same stochastic approach we introduced earlier for the KPM.

For example, if the problem has been scaled and shifted so that the spectrum of the transformed problem has been mapped to the interval $[-1, 1]$, we may approximate the δ-function centered at t_i by using an M_i-degree Chebyshev polynomial expansion of the form

$$\delta(t - t_i) \approx p_{M_i}(t) \equiv \sum_{k_i=0}^{M_i} \mu_{k_i}(t_i) T_{k_i}(t), \tag{5.1}$$

where the expansion coefficient $\mu_{k_i}(t_i)$ can be determined from

$$\mu_{k_i}(t_i) = \frac{2 - \delta_{k0}}{\pi} \int_{-1}^{1} \frac{1}{\sqrt{1-t^2}} T_{k_i}(t) \delta(t - t_i) dt = \frac{2 - \delta_{k0}}{\pi} \frac{T_{k_i}(t_i)}{\sqrt{1 - t_i^2}}. \tag{5.2}$$

If v_0 is some normalized random vector, whose expansion in the eigenbasis is given by (4.1), then $v_{M_i} \equiv p_{M_i}(A) v_0$ will have the expansion

$$v_{M_i} = \sum_{j=1}^{n} \beta_j p_{M_i}(\lambda_j) u_j. \tag{5.3}$$

Taking the Euclidean inner product between v_0 and v_{M_i} yields

$$\langle v_{M_i}, v_0 \rangle = \sum_{j=1}^{n} \beta_j^2 p_{M_i}(\lambda_j). \tag{5.4}$$

Since $\sum_{j=1}^{n} \beta_j^2 = 1$, (5.4) can be viewed as an integral of $p_{M_i}(t)$ associated with a point measure $\{\beta_j^2\}$ defined at eigenvalues of the transformed matrix A. If $\beta_j^2 = 1/n$ for all j, we can then simply rewrite the integral as $\mathrm{Trace}[p_{M_i}(A)]/n$. As we have already shown in previous sections, such a trace can be approximated by choosing multiple random vectors v_0 and averaging $\langle v_0, p_{M_i}(A) v_0 \rangle$ for all these vectors.

Formally, the ideas described above can be organized as in Algorithm 3, which we will refer to as the Delta-Chebyshev expansion algorithm. First, we select a number of nodes t_i at which the spectral density of A is to be evaluated. At each of these points, we evaluate the expansion coefficients μ_k in (5.1). We then loop through a large number of random vectors and compute the averages of the estimates (5.4) for each of the sample points t_i. These steps are outlined more precisely in Algorithm 3.

Algorithm 3: Multi-point Delta-Chebyshev expansion

Input: Real symmetric matrix A. A set of points $\{t_i\}$, and associated Chebyshev polynomial degree $\{M_i\}$, and $M_{\max}$ is the maximum degree employed for all the points.
Output: Approximate DOS $\{\widetilde{\phi}_M(t_i)\}$.

for each t_i **do**
  Compute and store the expansion coefficients μ_{k_i} for $k = 0, \cdots, M_i$;
end for
for $l = 1 : n_{\text{vec}}$ **do**
  Select a new random vector $v_0^{(l)}$;
  for $k = 0 : M_{\max}$ **do**
    Compute $v_{k+1}^{(l)}$ via the three-term recurrence $v_{k+1}^{(l)} = 2Av_k^{(l)} - v_{k-1}^{(l)}$ (for $k = 0$, $v_1^{(l)} = Av_0^{(l)}$);
    Compute $\eta_k \leftarrow \eta_k + \left(v_0^{(l)}\right)^T v_k^{(l)}$;
  end for
end for
Set $\eta_k = \eta_k / n_{\text{vec}}$ for all $k = 0, 1, ..., M_{\max}$;
Evaluate $\widetilde{\phi}_M(t_i)$ using η_k and the stored μ_{k_i}'s;

Note that this algorithm would have been an unacceptably expensive procedure if it were not for the fact that the same Chebyshev sequence $v_{k_i} = T_{k_i}(A)v_0$ can be used for all points t_i at the same time. Since the average value $\langle v_0, p_{M_i}(A)v_0 \rangle$ can be viewed as an approximation to the trace of $p_{M_i}(A)$, we can rewrite this quantity, which serves as an approximation to the spectral density at t_i, as

$$
\begin{aligned}
\phi(t_i) &= \frac{1}{n} \sum_{k_i=0}^{M_i} \mu_{k_i}(t_i) \sum_{j=1}^{n} T_{k_i}(\lambda_j) \\
&= \frac{1}{n} \sum_{k_i=0}^{M_i} \frac{2 - \delta_{k_i,0}}{\pi} \frac{T_{k_i}(t_i)}{\sqrt{1 - t_i^2}} \mathrm{Trace}(T_{k_i}(A)) \\
&= \sum_{k_i=0}^{M_i} \left[\frac{2 - \delta_{k_i,0}}{n\pi} \mathrm{Trace}(T_{k_i}(A)) \right] \frac{T_{k_i}(t_i)}{\sqrt{1 - t_i^2}}.
\end{aligned} \tag{5.5}
$$

Note that the coefficients within the square bracket in (5.5) are exactly the same coefficients that appear in the expansion (3.2) of the function $\hat{\phi}(t) = \sqrt{1 - t^2}\phi(t)$ used in the KPM. Therefore, when $M_i = M$ for all i, the Delta-Chebyshev expansion method is identical to the KPM. Hence, the cost of this approach is the same as that of KPM if polynomials of the same degree and the same number of sampling vectors are used at each t_i.

However, when M_i is allowed to vary with respect to i, there is a slight advantage of using the Delta-Chebyshev method in terms of flexibility. We can use polynomials of different degrees in different parts of the spectrum to obtain a more accurate approximation. Note that in this situation, if $M_{\max}$ is the maximum degree employed for all the points, the number of matrix-vector products employed remains the same and equal to $M_{\max}$, since we will need to compute for each random vector v_0, the vectors $T_k(A)v_0$ for $k = 0, \cdots, M_{\max}$ as these are needed by the points requiring the highest degree. However, some of the other calculations (inner products) required to obtain the spectral density can be avoided, though in most cases applying $T_k(A)$ to v_0 dominate the computational cost in the DOS calculation. The Delta-Gauss-Legendre expansion to be developed in the next section shares similar property in this aspect.

We should point out that it is also possible to apply the Delta-Chebyshev method for approximating the accumulated spectral density, which is simply the integral of the spectral density:

$$\psi(t) = \int_{-\infty}^{t} \phi(s) ds \tag{5.6}$$

Consider the step function shown on the right side of Figure 3.1. Let $H(t)$ be the Heaviside function defined in Eq. (3.9). Then the ideal step function $H(t-s)$ which changes value at certain t and we have:

$$\psi(t) = \int_{-\infty}^{t} \phi(s) ds = \int_{-\infty}^{\infty} H(t-s)\phi(s) ds$$

So, $\psi(t_0)$ can also be estimated by an expression like (5.4) where now the polynomial is an approximation of the step function $H(t)$ instead of the δ-function.

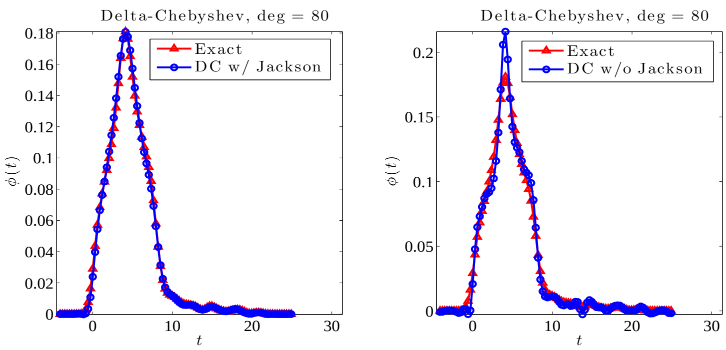

Fig. 5.1: The approximate spectral densities constructed by the Delta-Chebyshev method with Jackson damping (left) and without Jackson damping (right) for the modified 2D Laplacian matrix. The polynomial degree is 80.

6. The Delta-Gauss-Legendre expansion approach. In Section 2, we invoked the idea of comparing the approximation $\widetilde{\phi}$ of the spectral density not with the spectral density ϕ directly but with a smoothed version of ϕ which we denoted by ϕ_σ where σ is a smoothing parameter. We argued that this approach is sensible because it may be easier in practice to compare smooth functions than to compare a distribution with a function. In this section, we will take the idea a little further. Instead of approximating ϕ directly we will first select a representative ϕ_σ of ϕ for a given σ and then will approximate ϕ_σ by expanding it in orthogonal polynomials. In this way, comparisons will be relative to the same objects and the function to approximate can be as smooth as desired.

We will proceed in the same way as with the Delta-Chebyshev approach of the previous section, but instead of expanding the Dirac function at the point t we will now expand *its smoothed version* or *surrogate* defined by

$$h_\sigma(\lambda - t) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left[-\frac{(\lambda - t)^2}{2\sigma^2}\right]. \tag{6.1}$$

This is a Gaussian centered at t, and we wish to expand it in Legendre polynomials. We will drop the normalization coefficient for now and seek the expansion parameters

$$\gamma_k(t) = \int_{-1}^{1} L_k(s) e^{-\frac{1}{2}((s-t)/\sigma)^2} ds, \tag{6.2}$$

where $L_k(\lambda)$ is the Legendre polynomial of degree k. In the following we consider one fixed value t and so we will drop the variable t and use γ_k to denote $\gamma_k(t)$. We point out that Legendre polynomials do not have unit norms. Therefore, they should be properly scaled in the expansion. Specifically, since the norm of L_k is $\sqrt{2/(2k+1)}$, the expansion is actually:

$$h_\sigma(\lambda - t) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left[-\frac{(\lambda - t)^2}{2\sigma^2}\right] = \frac{1}{(2\pi\sigma^2)^{1/2}} \sum_{k=0}^{\infty} \left(k + \frac{1}{2}\right) \gamma_k L_k(\lambda) \ . \tag{6.3}$$

We now calculate the γ_k's starting with γ_0. Since $L_0(\lambda) = 1$, a change of variable $t = (s - t)/\sqrt{2\sigma^2}$ yields

$$\gamma_0 = \sigma\sqrt{\frac{\pi}{2}} \left[\operatorname{erf}\left(\frac{1-t}{\sqrt{2}\sigma}\right) - \operatorname{erf}\left(\frac{-1-t}{\sqrt{2}\sigma}\right)\right] = \sigma\sqrt{\frac{\pi}{2}} \left[\operatorname{erf}\left(\frac{1-t}{\sqrt{2}\sigma}\right) + \operatorname{erf}\left(\frac{1+t}{\sqrt{2}\sigma}\right)\right], \tag{6.4}$$

where we have used the standard error function:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \ .$$

Now consider a general coefficient γ_{k+1} with $k \geq 0$. There does not seem to exist a closed form formula for γ_k for a general k. However, these coefficients can be obtained by a recurrence relation. To this end we will need to determine concurrently the sequence:

$$\psi_k = \int_{-1}^{1} L_k'(s) e^{-\frac{1}{2}((s-t)/\sigma)^2} ds. \tag{6.5}$$

From the 3-term recurrence of the Legendre polynomials:

$$(k+1)L_{k+1}(\lambda) = (2k+1)\lambda L_k(\lambda) - kL_{k-1}(\lambda) \tag{6.6}$$

we get by integration:

$$(k+1)\gamma_{k+1} = (2k+1) \int_{-1}^{1} sL_k(s) e^{-\frac{1}{2}((s-t)/\sigma)^2} ds - k\gamma_{k-1}. \tag{6.7}$$

A useful observation is that the above formula is valid for $k = 0$ by setting $\gamma_{-1} \equiv 0$. This comes from (6.6), which is valid for $k = 0$ by setting $L_{-1}(\lambda) \equiv 0$. Next we expand the integral term in the above equality:

$$\int_{-1}^{1} s e^{-\frac{1}{2}((s-t)/\sigma)^2} L_k(s) ds = \sigma^2 \int_{-1}^{1} \frac{s-t}{\sigma^2} e^{-\frac{1}{2}((s-t)/\sigma)^2} L_k(s) ds + t\gamma_k \tag{6.8}$$

$$= \sigma^2 \int_{-1}^{1} \frac{d}{ds}[-e^{-\frac{1}{2}((s-t)/\sigma)^2}] L_k(s) ds + t\gamma_k. \tag{6.9}$$

The next step is to proceed with integration by parts for the integral in the above expression:

$$\int_{-1}^{1} \frac{d}{ds}[-e^{-\frac{1}{2}((s-t)/\sigma)^2}] L_k(s) ds = -L_k(s) e^{-\frac{1}{2}((s-t)/\sigma)^2}\Big|_{-1}^{1} + \int_{-1}^{1} e^{-\frac{1}{2}((s-t)/\sigma)^2} L_k'(s) ds. \tag{6.10}$$

Noting that $L_k(1) = 1$ and $L_k(-1) = (-1)^k$ for all k, we get

$$\int_{-1}^{1} \frac{d}{ds}[-e^{-\frac{1}{2}((s-t)/\sigma)^2}] L_k(s) ds = -e^{-\frac{1}{2}((1-t)/\sigma)^2} + (-1)^k e^{-\frac{1}{2}((1+t)/\sigma)^2} + \psi_k \tag{6.11}$$

$$= -e^{-\frac{1}{2}(1+t^2)/\sigma^2} \left[e^{t/\sigma^2} - (-1)^k e^{-t/\sigma^2} \right] + \psi_k \tag{6.12}$$

$$\equiv \psi_k - \zeta_k, \tag{6.13}$$

where we have defined

$$\zeta_k = e^{-\frac{1}{2}((1-t)/\sigma)^2} - (-1)^k e^{-\frac{1}{2}((1+t)/\sigma)^2} \tag{6.14}$$

We note in passing that according to (6.12), ζ_k can be written as

$$\zeta_k = \begin{cases} 2e^{-\frac{1}{2}(1+t^2)/\sigma^2} \mathrm{sh}(t/\sigma^2) & \text{for } k \text{ even} \\ 2e^{-\frac{1}{2}(1+t^2)/\sigma^2} \mathrm{ch}(t/\sigma^2) & \text{for } k \text{ odd.} \end{cases}$$

Substituting (6.11) into (6.9) and the result into (6.7) yields

$$(k+1)\gamma_{k+1} = (2k+1) \left[\sigma^2(\psi_k - \zeta_k) + t\gamma_k \right] - k\gamma_{k-1} \tag{6.15}$$

The only thing that is left to do is to find a recurrence for the ψ_k's. Here we use the elegant formula which can be found in, e.g., [38, p. 47]

$$L_{k+1}'(\lambda) = (2k+1) L_k(\lambda) + L_{k-1}'(\lambda). \tag{6.16}$$

Integrating in $[-1, 1]$ yields the relation:

$$\psi_{k+1} = (2k+1)\gamma_k + \psi_{k-1} \tag{6.17}$$

Note that initial values of ψ_k are $\psi_0 = 0$, $\psi_1 = \gamma_0$. In the end, we obtain the following recurrence relations:

$$\begin{cases} \gamma_{k+1} = \frac{2k+1}{k+1} \left[\sigma^2(\psi_k - \zeta_k) + t\gamma_k \right] - \frac{k}{k+1}\gamma_{k-1} \\ \psi_{k+1} = (2k+1)\gamma_k + \psi_{k-1}. \end{cases} \tag{6.18}$$

It can be noted that the above formulas work for $k = 0$ by setting $\gamma_{-1} = \psi_{-1} = 0$. The recurrence starts with $k = 0$, using the initial values γ_0 given by (6.4), $\psi_1 = \gamma_0$, and $\psi_0 = 0$. The result is now summarized in the following proposition.

PROPOSITION 6.1. *The Gaussian function (6.1) admits the following expansion in Legendre polynomials:*

$$h_\sigma(\lambda - t) = \frac{1}{(2\pi\sigma^2)^{1/2}} \sum_{k=0}^{\infty} \left(k + \frac{1}{2}\right) \gamma_k L_k(\lambda), \tag{6.19}$$

where γ_0 is given by (6.4), and the coefficients γ_k for $k \geq 1$ are defined by the recurrence (6.18), starting with $k = 0$ and using the initial values $\gamma_{-1} = 0$, $\psi_0 = \psi_{-1} = 0$.

An important remark here is that one has to be careful about the application of the recurrence (6.18). The perceptive reader may notice that such a recurrence runs the risk of being unstable. In fact we observe the following behavior. For large values of σ the function (6.1) can be very smooth and as a result a very small degree of polynomials may be needed, i.e., the value of γ_k drop to small values quite rapidly as k increases. If we ask for a high degree polynomial and continue the recurrence (6.18) beyond the point where the expansion has converged (indicated by small value of γ_k) we will essentially iterate with noise. As it turns out, this noise is amplified by the recurrence. This is because the coefficient $\psi_k - \zeta_k$ becomes just noise and this causes the recurrence to diverge. An easy remedy is to just stop iterating (6.18) as soon as two consecutive γ_k's are small. This takes care of two issues at the same time. First, it determines a sort of optimal degree to be used. Second, it avoids the unstable behavior observed by continuing the recurrence. Specifically, a test such as the following is performed:

$$|\gamma_{k-1}| + |\gamma_k| \leq k \cdot \text{tol}, \tag{6.20}$$

where tol is a small tolerance, which can be set to 10^{-6} for example. The factor k reflects the scaling $k + 1/2$ in (6.3)

With this we can now easily develop an adaptation of Algorithm 3, which we will call the Delta-Gauss-Legendre (DGL) expansion algorithm. In the DGL algorithm, we will refer to formula (5.4). But now p_M is the M-degree polynomial

$$p_M(\lambda) = \frac{1}{(2\pi\sigma^2)^{1/2}} \sum_{k=0}^{M} \left(k + \frac{1}{2}\right) \gamma_k L_k(\lambda), \tag{6.21}$$

obtained by truncating the sum (6.19) to $M + 1$ terms.

In Figure 6.1, we show the approximate spectral density for the modified 2D Laplacian matrix constructed by the DGL method. The expansion coefficient $\gamma_k(t)$ are obtained by setting $\sigma = 0.56$ (obtained by setting $\kappa = 1.25, h = 0.75$).

7. Use of the Lanczos Algorithm. Because finding a highly accurate DOS essentially amounts to computing all eigenvalues A, any method that can provide approximations to the spectrum of A can be used to construct an approximate DOS also. Since the Lanczos algorithm yields good approximations to extreme eigenvalues, it is a good candidate for computing localized spectral densities at both ends of the spectrum. In this section, we show that it is also possible to combine the Lanczos algorithm with multiple randomly generated starting vectors to construct a good approximation to the complete DOS.

Algorithm 4: Multi-point Delta-Gauss-Legendre expansion.

Input: Real symmetric matrix A. A set of points $\{t_i\}$ at which the DOS is to be evaluated, and $M_{\max}$ is the maximum degree employed for all the points.

Output: Approximate DOS $\{\widetilde{\phi}_M(t_i)\}$.

```
for each t_i do
    Compute and store the expansion coefficients γ_{k_i} for k = 0, ··· , M_i;
end for
for l = 1 : n_vec do
    Select a new random vector v_0^{(l)};
    for k = 0 : M_max do
        Compute v_{k+1}^{(l)} via the three-term recurrence v_{k+1}^{(l)} = 2Av_k^{(l)} − v_{k−1}^{(l)}   (for k = 0, v_1^{(l)} = Av_0^{(l)});
        Compute η_k ← η_k + (v_0^{(l)})^T v_k^{(l)};
    end for
end for
Set η_k = η_k/n_vec for all k = 0, 1, ..., M_max;
Evaluate φ̃_M(t_i) using η_k and the stored γ_{k_i}'s;
```

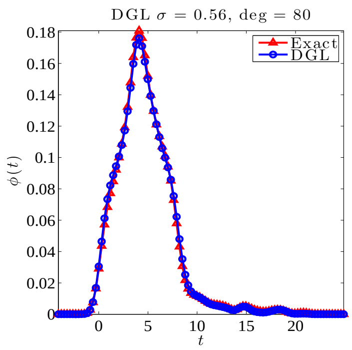

Fig. 6.1: The approximate spectral density constructed by the Delta-Gauss-Legendre method for the modified 2D Laplacian matrix example. The same value $\sigma = 0.56$ as that used in the Gaussian smearing of the exact density is used in DGL. The polynomial degree is 80.

7.1. Lanczos approximation to the spectral density. For a given starting vector v_0, an M-step Lanczos procedure for a real symmetric matrix A can be succinctly described by the following equations

$$AV_M = V_M T_M + f e_{M+1}^T, \quad V_M^T V_M = I_M, \quad V_M^T f = 0, \quad V_M e_1 = v_0, \tag{7.1}$$

here T_M is an $(M+1)\times(M+1)$ tridiagonal matrix, V_M is an $n\times(M+1)$ orthogonal matrix, and I_M is an $(M+1)\times(M+1)$ identity matrix. It is well known that the k-th column of V_M can be expressed as

$$V_M e_k = p_{k-1}(A)v_0, \quad k = 1, \cdots, M+1.$$

where $\{p_k(t)\}$, $k = 0, 1, 2, .., M$ is a set of orthogonal polynomials with respect to the weighted spectral distribution $\phi_{v_0}(t)$ taking the form

$$\phi_{v_0}(t) = \sum_{j=1}^{n} \beta_j^2 \delta(t - \lambda_j). \tag{7.2}$$

Here v_0 is assumed to be expanded in the basis of the eigenvectors of A as in Eq. (4.1). It is also well known that these orthogonal polynomials can be generated by a three-term recurrence whose coefficients are defined by the matrix elements of T_M [22].

If (θ_k, y_k), $k = 0, 1, 2, \ldots, M$ are eigenpairs of the tridiagonal matrix T_M, and τ_k is the first entry of y_k, then the distribution function defined by

$$\sum_{k=0}^{M} \tau_k^2 \delta(t - \theta_k), \tag{7.3}$$

serves as an approximation to the weighted spectral density function $\phi_{v_0}(t)$, in the sense that

$$\sum_{j=1}^{n} \beta_j^2 p_q(\lambda_j) = \sum_{k=0}^{M} \tau_k^2 p_q(\theta_k), \tag{7.4}$$

for all polynomials of degree $0 \leq q \leq 2M + 1$. The moment matching property described by (7.4) is well known [24]. It is the basis of Gaussian quadrature rules [23, 25].

Since in most cases, we are interested in the standard spectral density defined by (1.1), we would like to choose a starting vector v_0 such that $\beta_j^2 = 1/n$. However, this is generally not possible without knowing the eigenvectors $\{u_j\}$ of A in advance. To address this issue, we resort to the same stochastic approach we used in previous sections.

We repeat the Lanczos process with multiple randomly generated starting vectors $v_0^{(l)}$, $l = 1, 2, \ldots, n_{\mathrm{vec}}$. We assume the average

$$\begin{aligned} \frac{1}{n_{\mathrm{vec}} n} \sum_{l=1}^{n_{\mathrm{vec}}} \left(v_0^{(l)}\right)^T \delta(tI - A) v_0^{(l)} =& \frac{1}{n_{\mathrm{vec}} n} \sum_{l=1}^{n_{\mathrm{vec}}} \sum_{j=1}^{n} \left(v_0^{(l)}\right)^T u_j \delta(t - \lambda_j) u_j^T v_0^{(l)} \\ =& \frac{1}{n} \sum_{j=1}^{n} \left(\frac{1}{n_{\mathrm{vec}}} \sum_{l=1}^{n_{\mathrm{vec}}} \left(\beta_j^{(l)}\right)^2 \right) \delta(t - \lambda_j) \end{aligned} \tag{7.5}$$

is indeed a good approximation to the standard spectral density $\phi(t)$ in Eq. (1.1), i.e.

$$\frac{1}{n_{\mathrm{vec}}} \sum_{l=1}^{n_{\mathrm{vec}}} \left(\beta_j^{(l)}\right)^2 \approx 1/n.$$

Since each distribution (7.3) generated by the Lanczos procedure is a good approximation to (7.4), then taking the average of (7.3) over l, i.e.

$$\widetilde{\phi}(t) = \frac{1}{n_{\mathrm{vec}}} \sum_{l=1}^{n_{\mathrm{vec}}} \left(\frac{1}{n} \sum_{k=0}^{M} \left(\tau_k^{(l)}\right)^2 \delta(t - \theta_k^{(l)}) \right) \tag{7.6}$$

should yield a good approximation to the standard spectral density (1.1).

Clearly, the resolution of the approximation to (7.3) given by (7.6) becomes better as M increases. We may also replace δ-functions in (7.6) with its surrogate version using Gaussian functions, i.e.

$$\widetilde{\phi}_\sigma(t) = \frac{1}{n_{\text{vec}}} \sum_{l=1}^{n_{\text{vec}}} \left(\frac{1}{n} \sum_{k=0}^{M} \left(\tau_k^{(l)}\right)^2 h_\sigma(t - \theta_k^{(l)}) \right). \tag{7.7}$$

This will also allow us to evaluate the spectral density at some point that is different from a Ritz value θ_k. The method using Eq. (7.7) will be referred to as the Lanczos method for computing DOS in the following discussion. In Figure 7.1, we show the approximate density of states of the modified 2D Laplacian obtained from Eq. (7.7) running $M = 80$ Lanczos steps.

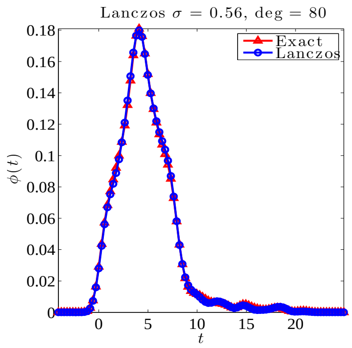

Fig. 7.1: The approximate spectral density constructed by the Lanczos method for the modified 2D Laplacian matrix. The same $\sigma = 0.56$ is used in the Gaussian smearing of both the exact DOS, and the approximation in (7.7).

An alternative way to refine the Lanczos based DOS approximation from a M-step Lanczos run is to first construct an approximate cumulative spectral density or cumulative density of states (CDOS) defined in (5.6). The approximate CDOS can be written as

$$\widetilde{\psi}(t) = \int_{-\infty}^{\infty} H(t-s)\widetilde{\phi}(s)ds = \sum_{k=0}^{M} \eta_k^2 \delta(t - \theta_k), \tag{7.8}$$

where $\eta_k^2 = \sum_{i=1}^{k} \tau_i^2$, and θ_k and τ_k are eigenvalues and the first components of the eigenvectors of the tridiagonal matrix T_M defined in (7.6). This approximation is plotted as a staircase function in Figure 7.2 for the modified 2D Laplacian. Note that both $\psi(t)$ and $\widetilde{\psi}(t)$ are monotonically non-decreasing functions. Furthermore, it can been shown [33] that $\psi(t) - \widetilde{\psi}(t)$ has precisely $2M - 1$ sign changes within the spectrum of A. A sign change occurs when $\psi(t)$ crosses either a vertical or horizontal step of $\widetilde{\psi}(t)$. These properties allow us to construct an "interpolated" CDOS that matches $\psi(t)$ and $\widetilde{\psi}(t)$ at the points where $\psi(t)$ crosses $\widetilde{\psi}(t)$.

Since we do not know $\psi(t)$, we do not know exactly where it crosses $\widetilde{\psi}(t)$ in advance. However, it can be seen from the left subfigure in Figure 7.2 that these crossing points are often near the midpoint of a vertical or horizontal step. Therefore, we may

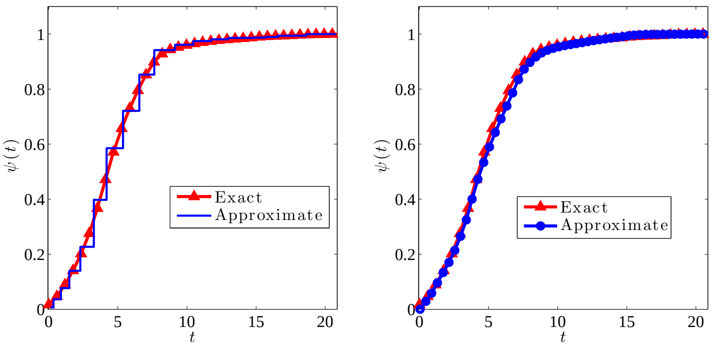

Fig. 7.2: The approximate cumulative spectral density associated with the modified 2D Laplacian constructed directly from a 20-step Lanczos run (left) and its spline-interpolated and smooth version (right).

use these midpoints to construct an interpolating function that is monotonically non-decreasing. For many practical problems, $\psi(t)$ approaches a smooth function as the dimension of the problem increases. Therefore, we can impose some constraints on the smoothness of the interpolating function in addition to monotonicity. One way to do this is to use the monotone splines introduced in [20, 21], and other choices of interpolating functions are also possible. The interpolation procedure allows us to evaluate the CDOS at any point within $[\lambda_1, \lambda_n]$. A recipe for constructing such an approximation is also given in [19]. The right subfigure in Figure 7.2 shows an interpolated approximate CDOS evaluated at 20 points evenly spaced points in $[\lambda_1, \lambda_n]$ for the modified Laplacian matrix. We observe that the interpolated CDOS is very close to the true CDOS at this resolution.

From the smoothly interpolated CDOS approximation, we can obtain an approximate DOS at any point in $[\lambda_1, \lambda_n]$ by taking the derivative of the interpolated $\widetilde{\psi}(t)$ with respect to t. We can use either the analytic derivative of the spline or simply perform a finite difference calculation. Figure 7.3 shows the approximate DOS obtained in this fashion. We observe that the approximate DOS captures the basic profile of the surrogate DOS smeared by Gaussian functions.

The quality of the approximation also improves systematically as the number of starting vectors n_{vec} increases. Figure 7.3 shows the averaged DOS approximations produced from 100 Lanczos runs looks much better than the approximation obtained from a single Lanczos run.

Furthermore, by carefully choosing v_0, we can also use this technique to obtain a localized DOS approximation. For example, if we are interested in the DOS at the low end of the spectrum, we can apply a low pass polynomial filter in A first to a randomly chosen v_l. Figure 7.4 shows that this technique indeed provide a good approximation to the local DOS.

7.2. Rational transformation. In addition to approximating the spectrum of A directly, the Lanczos algorithm can also be used to provide a rational approximation to the DOS, through the use of a well known technique for approximating a δ-function by a rational function. It follows from the Sokhotski-Plemelj formula [43]

$$\lim_{t\to 0} \frac{1}{t+i\eta} = \mathcal{P}(1/t) - i\pi\delta(t)$$

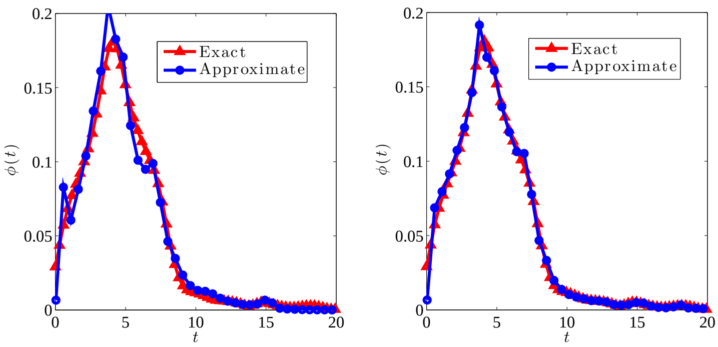

Fig. 7.3: The approximate DOS of the modified 2D Laplacian matrix obtained from the derivative of the CDOS approximation produced by a 20-step Lanczos procedure. The left subfigure is obtained from a single Lanczos run. The right subfigure is obtained from the average of 100 Lanczos runs with different random starting vectors.

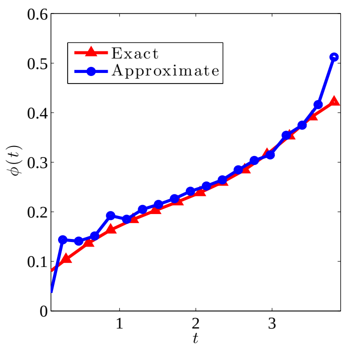

Fig. 7.4: The approximation to the local DOS at the low end of the spectrum of the modified 2D Laplacian using a filtered initial starting vector.

that a δ-function can be expressed as the imaginary part of the rational function $1/(t + i\eta)$ in the limit that $\eta \to 0$, and $\mathcal{P}$ represents the Cauchy principal value. As a result, we may express the DOS of A by

$$\begin{aligned}\phi(t) &= \lim_{\eta\to 0+} -\frac{1}{n\pi}\mathrm{Im} \sum_{j=1}^{n} \frac{1}{t + i\eta - \lambda_j} \\ &= \lim_{\eta\to 0+} -\frac{1}{n\pi}\mathrm{Im}\ \mathrm{Tr}\left[(tI - A + i\eta)^{-1}\right].\end{aligned} \tag{7.9}$$

It follows from the above discussion that estimating the spectral density of A is equivalent to computing the trace of $(tI - A + i\eta I)^{-1}$. We can use the same type of stochastic approach used previously in the KPM to compute the approximate trace by

$$\frac{1}{n_{\mathrm{vec}}} \sum_{l=1}^{n_{\mathrm{vec}}} \left(v_0^{(l)}\right)^T (tI - A + i\eta I)^{-1} v_0^{(l)}, \tag{7.10}$$

for several randomly sampled vectors v_0.

A direct calculation of (7.10) requires solving linear systems $[A-(t_i+i\eta)I]z = v_0$ repeatedly for each point t_i at which the spectral density is to be evaluated. This approach can be prohibitively expensive. We now describe an alternative approach that allows us to approximate $v_0^T(t_iI - A + i\eta I)^{-1}v_0$ for multiple t_i's at the cost of performing a single Lanczos factorization and some additional calculations that are much lower in complexity. If v_0 is used as the starting vector of the Lanczos procedure, then it follows from the shift-invariant property of the Lanczos algorithm that

$$[A - (t_i + i\eta)I]V_M = V_M[T_M - (t_i + i\eta)I] + fe_{M+1}^T, \tag{7.11}$$

where V_M and T_M are the same orthonormal and tridiagonal matrices respectively that appear in (7.1). After multiplying (7.11) from the left by $[A - (t_i + i\eta)I]^{-1}$, from the right by $[T_M - (t_i + i\eta)I]^{-1}$ and rearranging terms, we obtain

$$[A-(t_i+i\eta)I]^{-1}V_M = V_M[T_M-(t_i+i\eta)I]^{-1}-[A-(t_i+i\eta)I]^{-1}fe_{M+1}^T[T_M-(t_i+i\eta)I]^{-1}.$$

It follows that

$$\begin{aligned} v_0^T\,[A - (t_i + i\eta)I])^{-1}v_0 &= e_1^T V_M^T[A - (t_i + i\eta)I]^{-1}V_M e_1 \\ &= e_1^T[T_M - (t_i + i\eta)I]^{-1}e_1 + \xi, \end{aligned}$$

where $\xi = -\left(v_0^T[A - (t_i + i\eta)I]^{-1}f\right)\left(e_{M+1}^T[T_M - (t_i + i\eta)I]^{-1}e_1\right)$. If ξ is sufficiently small, computing $v_0^T(t_iI - A + i\eta I)^{-1}v_0$ reduces to computing the $(1,1)$-th entry of the inverse of $T_M - (t_i + i\eta)I$. Because T_M is tridiagonal with $\alpha_1, \alpha_2, ..., \alpha_{M+1}$ on the diagonal and $\beta_2, \beta_3, ..., \beta_{M+1}$ on the sub-diagonals and super-diagonals, $e_1^T(zI - T_M)^{-1}e_1$ can be computed in a recursive fashion using a continued fraction formula

$$e_1^T(zI - T_M)^{-1}e_1 = \frac{1}{z - \alpha_1 + \frac{\beta_2^2}{z-\alpha_2+\cdots}}. \tag{7.12}$$

This formula can be verified from the identity

$$(z - T_M)^{-1}_{1,1} \equiv \frac{\det(zI - T_M)}{\det(z - \hat{T}_M)}$$

where $\hat{T}_M$ is the trailing submatrix starting from the $(2,2)$ entry of T_M (7.12) and tridiagonal structure of both T_M and $\hat{T}_M$ matrices.

This approach is often referred to as Haydock's method, and was first suggested by Haydock, Heine, and Kelly [26]. It is also related to the generation of Sturm sequences which is used in bisection methods for computing eigenvalues of tridiagonal matrices [42]. The computational cost of computing $e_1^T(T_M - z_iI)^{-1}e_1$ is $\mathcal{O}(M + 1)$ for each shift $z_i = t_i + i\eta$ regardless whether this is done by applying a sparse direct method to compute $(T_M - z_iI)^{-1}e_1$ or through the use of a continued fraction. For most problems, this cost is small compared to that required to perform the Lanczos procedure to obtain T_M. Even in exact arithmetic, Haydock's method produces an exact $v_0^T(z_iI - A)^{-1}v_0$ only when the last term on the right-hand side of (7.11) is zero. This term vanishes when $f = 0$. In this case, columns of V_M form an invariant subspace of A. However, for most problems, f rarely becomes zero even for large M. But when η is relatively large, the off-diagonal elements of $(z_iI - T_M)^{-1}$ decreases sufficiently fast, away from the diagonal. Therefore, the $v_0^T(z_iI - A)^{-1}v_0$ can be evaluated accurately in this case.

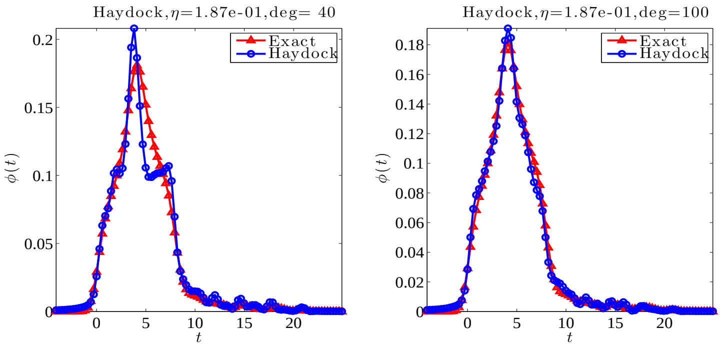

Fig. 7.5: The approximate spectral densities constructed by the Haydock method with 40-step (left) and 100-step (right) Lanczos runs for the modified 2D Laplacian matrix. The Haydock smoothing parameter η is chosen to be $\eta = 0.187$.

The performance of Haydock's approach for computing the spectral density for the modified 2D Laplacian example is shown in Fig. 7.5. The parameter η is chosen to be comparable to the resolution of the spectral density we display. Specifically, we choose

$$\eta = \frac{\varepsilon_{\max} - \varepsilon_{\min}}{2n_x}, \tag{7.13}$$

where $\varepsilon_{\max}$ and $\varepsilon_{\min}$ are the largest and smallest eigenvalues of A respectively, and n_x is the number of uniformly distributed grid points at which the spectral density is evaluated. In Fig. 7.5 we choose $n_x = 55$, which corresponds to $\varepsilon = 0.187$. This is roughly one third of the value of $\sigma = 0.56$. We observe that running 40 Lanczos steps produces a qualitatively different spectral density approximation. The quality of the approximation becomes much better when 100 Lanczos steps are taken in Haydock's approach.

8. Numerical Results. In this section, we compare all methods discussed above for approximating the spectral density of A through numerical examples. To compare these methods quantitatively, we compute the L^1, L^2 and L^∞ distances between the approximate spectral density $\widetilde{\phi}(t)$ and the surrogate spectral density $\phi_\sigma(t)$, and the latter is the true spectral density blurred by a Gaussian of the form (1.6) with a fixed σ value. Both $\phi_\sigma(t)$ and $\hat{\phi}(t)$ are normalized to have $\|\phi_\sigma(t)\|_1 = \|\hat{\phi}(t)\|_1 = 1$. It is well known that both the L^2 and L^∞ norms are sensitive to outliers in the spectral density approximation whereas the L^1 norm is relatively less sensitive to outliers.

8.1. Modified 2D Laplacian. In Table 8.1, we list the L^1, L^2 and L^∞ errors associated with different types of approximations to the spectral density of the modified Laplacian matrix we used in the previous sections to illustrate how these methods behave. The dimension of the matrix is 750, which is relatively small. We set σ to 0.56. This corresponds to setting κ to 1.25 and $h = 0.75$ in (2.1). The degree of Chebyshev or Legendre polynomials used in the KPM and DGL methods is set to 80. Similarly, the number of iterations in the Lanczos method is also set to 80. We also present the results for the spectroscopic method and Haydock's method with the number of iterations being 40 and 100, respectively. We observe that the DGL, Lanczos and KPM with Jackson damping yield the most accurate approximation among all methods in L^1, L^2 and L^∞ norms, followed by the KPM method without Jackson damping. The spectroscopic and Haydock's methods do not yield sufficiently accurate approximations if only 40 steps of Lanczos iterations are taken. In order to reach the same level of accuracy achieved by DGL and Lanczos in L^1 norm, at least 100 Lanczos iterations are required in these methods to further reduce the approximation error. In all test cases, we use 10 random vectors.

Method	L^1 error	L^2 error	L^∞ error
KPM w/ Jackson, deg=80	2.994e-02	9.2190e-03	6.036e-03
KPM w/o Jackson, deg=80	1.057e-01	3.3897e-02	2.409e-02
KPM Legendre, deg=80	1.123e-01	3.4598e-02	2.628e-02
Spectroscopic, deg=40	8.724e-02	2.6650e-02	1.566e-02
Spectroscopic, deg=100	6.818e-02	2.1735e-02	1.395e-02
DGL, deg=80	3.031e-02	9.8346e-03	5.180e-03
Lanczos, deg=80	2.047e-02	5.7810e-03	3.344e-03
Haydock, deg=40	1.583e-01	5.9017e-02	4.631e-02
Haydock, deg=100	8.074e-02	2.3214e-02	1.367e-02

Table 8.1: L^1, L^2, and L^∞ error compared with the normalized "surrogate" DOS for modified 2D Laplacian matrix, using the same set of parameters as described in this section.

8.2. The Kohn-Sham Hamiltonian associated with the Benzene molecule. To demonstrate that the methods presented in this paper are applicable to matrices other than the modified Laplacian, we now show the L^1, L^2 and L^∞ error associated with different approximations to the spectral density of a matrix, which represents a finite difference approximation of the Kohn-Sham Hamiltonian associated with a benzene molecule. This matrix is generated from the PARSEC software [34]. The dimension of this matrix is 8,219. The matrix can be obtained from the Florida Sparse Matrix repository [12]. For this problem, we also set the σ parameter for the smoothed exact spectral density to 0.56. We use 10 random Gaussian test vectors in all methods. Table 8.2 shows that the accuracy of different methods in L^1, L^2 and L^∞ norms exhibits a similar pattern to that observed for the modified Laplacian matrix. The DGL, Lanczos and the KPM with Jackson damping still yield the most accurate results. Haydock's method is less accurate compared to other methods for this example. The quality of the approximation provided by different methods can also be seen from Figures 8.1 to 8.5.

8.3. Convergence with respect to the number of random vectors. An interesting observation we made from the experimental comparison of the DGL method and the KPM is that DGL can produce reasonable results even for a very small number of random vectors. In contrast, the accuracy of KPM can vary significantly when only a few vectors are sampled. This is illustrated in Figure 8.6 which repeats the previous experiment using only one random vector. The curve on the left corresponds to KPM. It is one of many possible ones that vary quite a bit in shape. The one on the right corresponds to DGL. It remains more stable than KPM for multiple vectors used and captures the overall true spectral density quite well.

While using one random vector in KPM, DGL or Lanczos to obtain an approximate spectral density is an extreme case, we should expect the accuracy of the approx-

Item	L^1 error	L^2 error	L^∞ error
KPM w/ Jackson, deg=80	2.592e-02	5.0329e-03	2.785e-03
KPM w/o Jackson, deg=80	2.634e-02	4.4542e-03	2.002e-03
KPM Legendre, deg=80	2.504e-02	3.7881e-03	1.174e-03
Spectroscopic, deg=40	5.589e-02	8.6527e-03	2.871e-03
Spectroscopic, deg=100	4.624e-02	7.5829e-03	2.447e-03
DGL, deg=80	1.998e-02	3.3793e-03	1.149e-03
Lanczos, deg=80	2.755e-02	4.1785e-03	1.599e-03
Haydock, deg=40	6.951e-01	1.3023e-01	6.176e-02
Haydock, deg=100	2.581e-01	4.6530e-02	1.420e-02

Table 8.2: L^1, L^2, and L^∞ error compared with the normalized "surrogate" DOS for benzene matrix, using the same set of parameters as used in the pictures in this section.

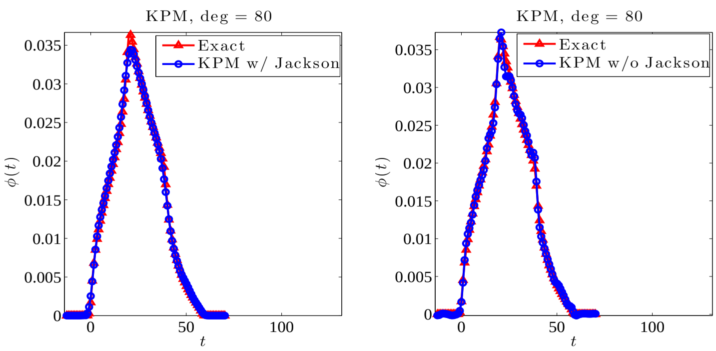

Fig. 8.1: The approximate spectral density constructed by KPM for the benzene Hamiltonian. Chebyshev polynomials of degree 80 with Jackson damping (left) and without Jackson damping (right) are used in KPM.

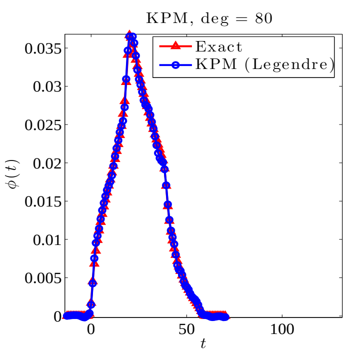

Fig. 8.2: The approximate spectral density for the benzene Hamiltonian obtained from KPM. A 80-degree Legendre polynomial is used in KPM.

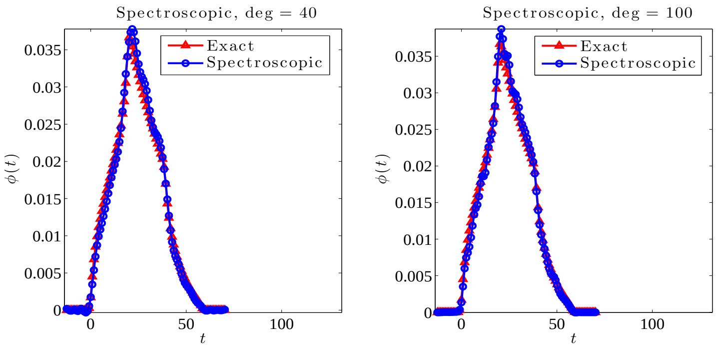

Fig. 8.3: The approximate spectral densities obtained from the Lanczos spectroscopic method in which 40 (left) and 100 (right) steps of the Lanczos algorithm are taken respectively.

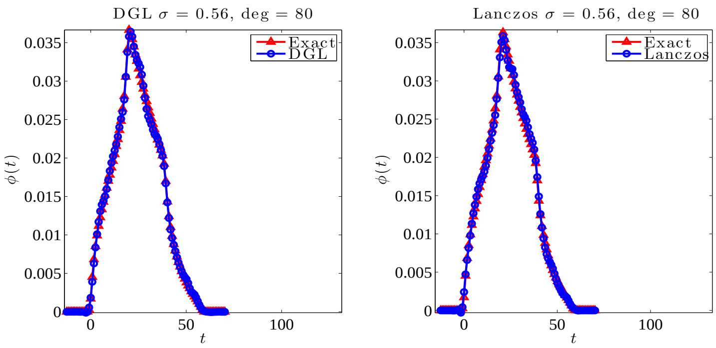

Fig. 8.4: The approximate spectral densities for the benzene Hamiltonian obtained from the Delta-Gauss-Legendre method and the Lanczos-Delta-Gauss fit. The σ value used in the Gaussian approximation of the δ-function in both methods is chosen to be the same as that used in the Gaussian smearing of the exact density.

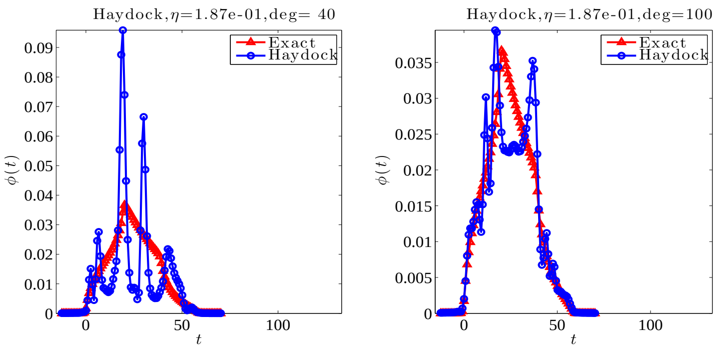

Fig. 8.5: The approximate spectral densities for the benzene Hamiltonian obtained from Haydock's method. The number of Lanczos step taken are 40 (left) and 100 (right). The smoothing parameter η is set to $\eta = 0.178$.

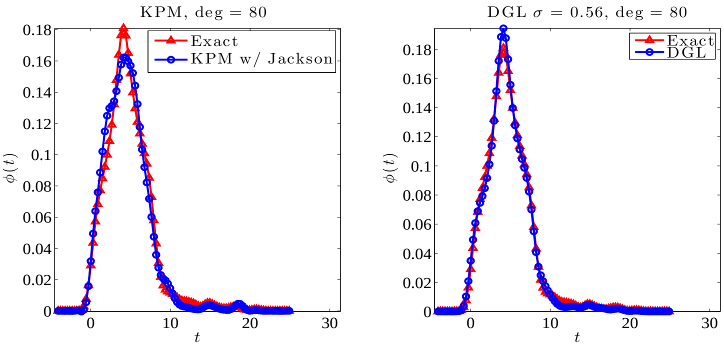

Fig. 8.6: KPM with degree 80 (left) and, Delta-Gauss-Legendre expansion with degree 80 ($\sigma = 0.56$) (right) for perturbed 2D Laplacian operator. In contrast with Figure 6.1, here only *one* random vector is used.

imation to improve systematically as we average approximations obtained from n_{vec} random vectors. This should be the case at least for the DGL and Lanczos methods, since these two methods are indeed based on the expansion of the surrogate spectral density. If the variance of the DGL or Lanczos approximation produced by sampling one random vector is denoted by var, then averaging n_{vec} approximations produced by n_{vec} independent random vectors can reduce the variance of the approximation to $\text{var}/n_{\text{vec}}$, and the standard deviation to $\sqrt{\text{var}/n_{\text{vec}}}$. In other words, the L^1, L^2 and L^∞ error should be proportional to the standard deviation and decreases as $n_{\text{vec}}^{-\frac{1}{2}}$. Fig. 8.7 shows that, for the modified 2D Laplacian matrix with $\sigma = 0.56$, the L^1, L^2 and L^∞ errors associated with DGL, Lanczos and KPM (Chebyshev expansion with Jackson damping) all decrease as the number of random vectors used (n_{vec}) increases. We observe that the convergence with respect to n_{vec} is roughly on the order of $O(n_{\text{vec}}^{-\frac{1}{2}})$. For methods (such as the KPM) that do not directly expand the surrogate DOS ϕ_σ, there is some positive distance between the result obtained from KPM and that obtained from the surrogate DOS even when n_{vec} becomes large, although the absolute value of the error can be small as $10^{-3} \sim 10^{-2}$.

8.4. DGL and Lanczos for other matrices. We now show that in addition to the modified Laplacian and the Kohn-Sham Hamiltonian for the benzene molecule, the DGL and Lanczos methods work equally well for a variety of problems listed in Table 8.3. The $Ga_{10}As_{10}H_{30}$ matrix arises from a Kohn-Sham density functional theory calculation of a Ga and As cluster. It has a relatively large dimension. The PE3K matrix arises from the molecular vibration analysis of a polyethylene molecule with 3000 atoms [58]. The CFD1 matrix arises from a computational fluid dynamics calculation. The SHWATER matrix arises from another computational fluid dynamics application involving weather shallow water equations. In order to evaluate the accuracy of the DGL and Lanczos methods, we compute all eigenvalues of these matrices in advance using the ScaLAPACK software. The algebraically smallest and largest eigenvalues are also listed in Table 8.3. Table 8.4 compares the L^1, L^2 and L^∞ errors for the DGL and Lanczos methods for these problems.

As can be seen from the table, both the DGL and the Lanczos methods yield accurate approximations to the spectral densities of these matrices. We overlay the

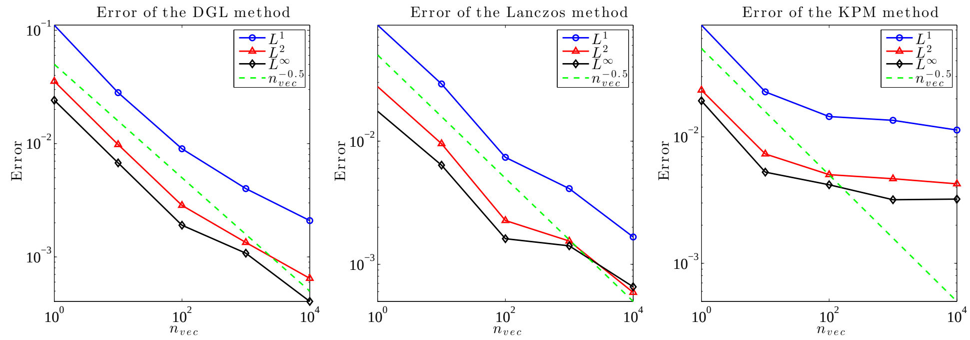

Fig. 8.7: The L^1, L^2 and L^∞ errors for the DGL , Lanczos methods and the KPM (Chebyshev expansion with Jackson damping) with respect to the number of random vectors used (n_{vec}). The matrix chosen is the modified 2D Laplacian matrix with $\sigma = 0.56$.

Matrix	n	λ_1	λ_n
$Ga_{10}As_{10}H_{30}$	113,081	-1.2	1.3×10^3
PE3K	9,000	8.1×10^{-6}	1.3×10^2
CFD1	70,656	2.0×10^{-5}	6.8
SHWATER	81,920	5.8	2.0×10^1

Table 8.3: Description of the size and the spectrum range of the test matrices.

computed spectral density (by DGL) on top of the “exact” spectral density for both the CFD1 and SHWATER matrices in Figure 8.8. As can be seen, the computed spectral densities are nearly indistinguishable from the exact ones. The same observation is made for $Ga_{10}As_{10}H_{30}$, which shows a similar spectral density profile as that associated with the benzene molecule, and the PE3K molecule, which we will show in the next section.

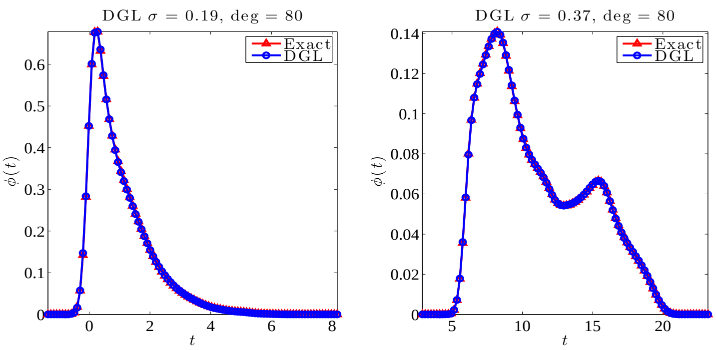

Fig. 8.8: The approximate spectral densities of the CFD1 and SHWATER matrices obtained from the DGL method match well with the corresponding smoothed “exact” spectral density.

Matrix	Method	L^1 error	L^2 error	L^∞ error
$Ga_{10}As_{10}H_{30}$	DGL	3.937e-03	3.214e-04	4.301e-05
	Lanczos	4.828e-03	3.940e-04	5.452e-05
PE3K	DGL	4.562e-03	7.368e-04	3.143e-04
	Lanczos	5.459e-03	7.372e-04	3.294e-04
CFD1	DGL	2.276e-03	1.299e-03	1.746e-03
	Lanczos	2.024e-03	1.286e-03	2.478e-03
SHWATER	DGL	3.779e-03	1.282e-03	9.328e-04
	Lanczos	3.047e-03	9.829e-04	6.100e-04

Table 8.4: L^1, L^2, and L^∞ error associated with the approximate spectral densities produced by the DGL and Lanczos methods for different test matrices in Table 8.3.

8.5. Heat capacity calculation. As we mentioned in section 1, one of the utilities of the DOS is that it provides an appropriate measure for integrating certain physical quantities of interest. One good example is the heat capacity defined in (1.4). Figure 8.9 shows that, the approximate spectral density allows us to captures the variation of the heat capacity with respect to temperature very well.

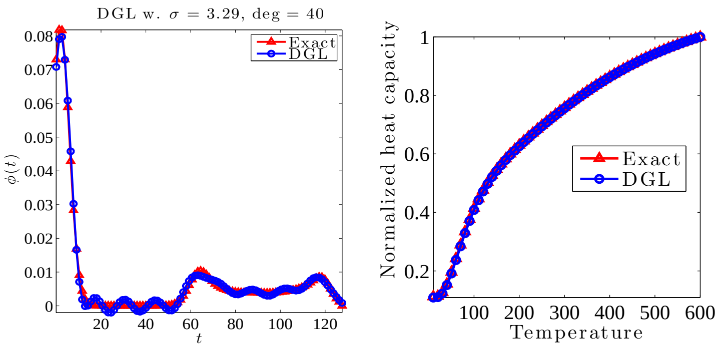

Fig. 8.9: The spectral density of the PE3K matrix and the variation of the computed heat capacity with respect to temperature change. The computed heat capacities are obtained from using the exact (red) and the DGL approximate (blue) spectral densities respectively.

9. Conclusion. We presented a number of algorithms for estimating the spectral density of a Hermitian matrix A from a numerical linear algebra perspective. Some of these algorithms (KPM, DC and DGL) are based on constructing polynomial approximations to δ-functions. In these approaches, estimating the spectral density essentially amounts to approximating the trace of A or $p(A)$ where $p(t)$ is a polynomial. An important technique for approximating the trace of a Hermitian matrix is stochastic sampling and averaging of the Rayleigh quotient $v_0^T p(A) v_0$, with $\|v_0\|_1 = 1$.

A few other methods make use of the Lanczos procedure to either directly estimate the spectrum of A or to construct a rational approximation of the δ-function. Stochastic sampling of the starting vector of the Lanczos process often improve the fidelity of the approximation. However, if the number of Lanczos steps used in these methods is sufficiently large, even one starting vector may be sufficient to produce a reliable estimation of the spectral density.

We suggest that it is more appropriate to approximate a smoothed version of the spectral density, which we sometimes call a surrogate spectral density, instead of the exact spectral density given by (1.1) directly. We can view this approach as a regularization technique to attenuate the numerical difficulties associated with approximating δ-functions directly. The smoothness of the surrogate spectral density can be controlled by one parameter σ. However, there is a trade-off between the amount of smoothness of the surrogate (hence the cost of the approximation) and the accuracy of the approximation. When σ is chosen to yield a very smooth surrogate spectral density, accurate approximations to this surrogate can often be obtained easily with a relatively low degree polynomial or a small number of Lanczos iterations. However, the surrogate itself may miss some important features of the exact spectral density required in some applications. On the other hand, if σ is chosen to minimize the difference between the surrogate and (1.1), then more work is required obtain a good approximation. In the extreme case when the exact spectral density given by (1.1) is to be obtained, then all eigenvalues of the matrix A must be computed.

Acknowledgments. This work is partially supported by the Laboratory Directed Research and Development Program of Lawrence Berkeley National Laboratory under the U.S. Department of Energy contract number DE-AC02-05CH11231 (L. L. and C. Y.), and by Scientific Discovery through Advanced Computing (SciDAC) program funded by U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research and Basic Energy Sciences DE-SC0008877 (Y. S. and C. Y.).

REFERENCES

[1] A. Alvermann, D. M. Edwards, and H. Fehske, *Boson-controlled quantum transport*, Phys. Rev. Lett., 98 (2007), p. 056602.

[2] M. Amini, S. A. Jafari, and F. Shahbazi, *Anderson transition in disordered graphene*, Europhys. Lett., 87 (2009), p. 37002.

[3] N. Ashcroft and N. Mermin, *Solid State Physics*, Thomson Learning, Toronto, 1976.

[4] D. R. Bowler and T. Miyazaki, *O(N) methods in electronic structure calculations*, Rep. Prog. Phys., 75 (2012), p. 036503.

[5] S. R. Broderick and K. Rajan, *Eigenvalue decomposition of spectral features in density of states curves*, Europhysics Letters, 95 (2011), p. 57005.

[6] R. K. Chouhan, A. Alam, S. Ghosh, and A. Mookerjee, *Ab initial study of phonon spectrum, entropy and lattice heat capacity of disordered Re-W alloys*, Journal of Physics: Condense Matter, 24 (2012), p. 375401.

[7] L. Covaci, F. M. Peeters, and M. Berciu, *Efficient numerical approach to inhomogeneous superconductivity: The Chebyshev-Bogoliubov-de Gennes method*, Phys. Rev. Lett., 105 (2010), p. 167006.

[8] L. Covaci, F. M. Peeters, and M. Berciu, *Efficient numerical approach to inhomogeneous superconductivity: The Chebyshev-Bogoliubov-de Gennes method*, Phys. Rev. Lett., 105 (2010), p. 167006.

[9] P. E. Dargel, A. Honecker, R. Peters, R. M. Noack, and T. Pruschke, *Adaptive Lanczos-vector method for dynamic properties within the density matrix renormalization group*, Phys. Rev. B, 83 (2011), p. 161104.

[10] P. E. Dargel, A. Wöllert, A. Honecker, I. P. McCulloch, U. Schollwök, and T. Pruschke, *Lanczos algorithm with matrix product states for dynamical correlation functions*, Phy. Rev. B, 85 (2012), p. 205119.

[11] P. J. Davis, *Interpolation and Approximation*, Blaisdell, Waltham, MA, 1963.

[12] T. A. Davis and Y. Hu, *The University of Florida sparse matrix collection*, ACM Trans. Math. Software, 38 (2011), p. 1.

[13] David A. Drabold and Otto F. Sankey, *Maximum entropy approach for linear scaling in the electronic structure problem*, Phys. Rev. Lett., 70 (1993), pp. 3631–3634.
[14] D. A. Drabold and O. F. Sankey, *Maximum entropy approach for linear scaling in the electronic structure problem*, Phys. Rev. Lett., 70 (1993), pp. 3631–3634.
[15] F. Ducastelle and F. Cyrot-Lackmann, *Moments developments and their application to the electronic charge distribution of d bands*, J. Phys. Chem. Solids, 31 (1970), pp. 1295–1306.
[16] A. Weisse, G. Wellein, A. Alvermann, and H. Fehske, *The kernel polynomial method*, Rev. Mod. Phys., 78 (2006), pp. 275–306.
[17] Z. Fan, A. Uppstu, T. Siro, and A. Harju, *Efficient linear-scaling quantum transport calculations on graphics processing units and applications on electron transport in graphene*, arXiv, 1307.0288v1 (2013), pp. 1–15.
[18] A. Ferreira, J. Viana-Gomes, J. Nilsson, E. R. Mucciolo, N. M. R. Peres, and A. H. C. Neto, *Unified description of the DC conductivity of monolayer and bilayer graphene at finite densities based on resonant scatterers*, Phys. Rev. B, 83 (2011), p. 165402.
[19] B. Fischer, *Polynomial Based Iteration Methods for Symmetric Linear Systems*, Wiley, New York, 1996.
[20] F. N. Fritsch and J. Butland, *A method for constructing local monotone piecewise cubic interpolations*, SIAM J. Sci. Stat. Comput., 5 (1984), pp. 300–304.
[21] F. N. Fritsch and R. E. Carlson, *Monotone piecewise cubic interpolation*, SIAM J. Numer. Anal., 17 (1980), pp. 238–246.
[22] W. Gautschi, *Computational aspects of three-term recurrence relations*, SIAM Rev., 9 (1967), pp. 24–82.
[23] ———, *Construction of Gauss-Christoffel quadrature formulas*, Math. Comp., 22 (1968), pp. 251–270.
[24] ———, *A survey of Gauss-Christoffel quadrature formulae*, in E. B. Christoffel: The influence of his work in mathematics and the physical sciences, P. Butzer and F. Feher, eds., Basel, 1981, Birkhauser, pp. 72–147.
[25] G. H. Golub and J. H. Welsch, *Calculation of Gauss quadrature rule*, Math. Comp., 23 (1969), pp. 221–230.
[26] R. Haydock, V. Heine, and M. J. Kelly, *Electronic structure based on the local atomic environment for tight-binding bands*, J. Phys. C: Solid State Phys., 5 (1972), p. 2845.
[27] A. Holzner, A. Weichselbaum, I. P. McCulloch, U. Schollwöck, and J. von Delft, *Chebyshev matrix product state approach for spectral functions*, Phys. Rev. B, 83 (2011), p. 195115.
[28] C. Huang, A. F. Voter, and D. Perez, *Scalable kernel polynomial method for calculating transition rates*, Phy. Rev. B, 87 (2013), p. 214106.
[29] M. F. Hutchinson, *A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines*, Commun. Stat. Simul. Comput., 18 (1989), pp. 1059–1076.
[30] D. Jackson, *The theory of approximation*, American Mathematical Society, New York, 1930.
[31] L. O. Jay, H. Kim, Y. Saad, and J. R. Chelikowsky, *Electronic structure calculations using plane wave codes without diagonlization*, Comput. Phys. Comm., 118 (1999), pp. 21–30.
[32] D. Jung, G. Czycholl, and S. Kettemann, *Finite size scaling of the typical density of states of disordered systems within the kernel polynomial method*, Int. J. Mod. Phys. Conf. Ser., 11 (2012), p. 108.
[33] S. Karlin and L. S. Shapley, *Geometry of Moment Spaces*, Amer. Math. Soc., 1953.
[34] L. Kronik, A. Makmal, M. L. Tiago, M. M. G. Alemany, M. Jain, X. Huang, Y. Saad, and J. R. Chelikowsky, *PARSEC-the pseudopotential algorithm for real-space electronic structure calculations: Recent advances and novel applications to nano-structures*, Phys. Status Solidi B, (2006).
[35] M. Kuno, *Introductory Nanoscience*, Garland Science, New York, NY, 2011.
[36] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Res. Nat. Bur. Stand., 45 (1950), pp. 255–282.
[37] ———, *Applied analysis*, Dover, New York, 1988.
[38] N. N. Lebedev, *Special functions and their applications*, Dover, New York, 1972.
[39] W. Li, H. Sevincili, S. Roche, and G. Cuniberti, *Efficient linear scaling method for computing the thermal conductivity of disordered materials*, Phys. Rev. B, 83 (2011), p. 155416.
[40] R. Martin, *Electronic Structure – Basic Theory and Practical Methods*, Cambridge Univ. Pr., West Nyack, NY, 2004.
[41] G. A. Parker, W. Zhu, Y. Huang, D.K. Hoffman, and D. J. Kouri, *Matrix pseudo-spectroscopy: iterative calculation of matrix eigenvalues and eigenvectors of large matrices using a polynomial expansion of the Dirac delta function*, Comput. Phys. Commun., 96 (1996), pp. 27–35.

[42] B. N. Parlett, *The symmetric eigenvalue problem*, vol. 7, SIAM, Philadelphia, 1980.

[43] J. Plemelj, *Problems in the sense of Riemann and Klein*, Interscience Publishers, New York, 1964.

[44] R. D. Richtmyer and W. Beiglböck, *Principles of advanced mathematical physics*, vol. 1, Springer-Verlag, New York, 1981.

[45] T. J. Rivlin, *An introduction to the approximation of functions*, Dover, New York, 2003.

[46] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, SIAM, Philadelpha, 2011.

[47] G. Schubert and H. Fehske, *Metal-to-insulator transition and electron-hole puddle formation in disordered graphene nanoribbons*, Phys. Rev. Lett., 108 (2012), p. 066402.

[48] Bernhard Seiser, D. G. Pettifor, and Ralf Drautz, *Analytic bond-order potential expansion of recursion-based methods*, Phys. Rev. B, 87 (2013), p. 094105.

[49] B. Seiser, D. G. Pettifor, and R. Drautz, *Analytic bond-order potential expansion of recursion-based methods*, Phys. Rev. B, 87 (2013), p. 094105.

[50] R. N. Silver and H. Röder, *Densities of states of mega-dimensional Hamiltonian matrices*, Int. J. Mod. Phys. C, 5 (1994), pp. 735–753.

[51] ———, *Calculation of densities of states and spectral functions by Chebyshev recursion and maximum entropy*, Phys. Rev. E, 56 (1997), p. 4822.

[52] R. N. Silver, H. Röder, A. F. Voter, and J. D. Kress, *Kernel polynomial approximations for densities of states and spectral functions*, J. Comput. Phys., 124 (1996), pp. 115–130.

[53] J. M. Tang and Y. Saad, *A probing method for computing the diagonal of a matrix inverse*, Numer. Lin. Alg. Appl., 19 (2012), pp. 485–501.

[54] I. Turek, *A maximum-entropy approach to the density of states within the recursion method*, J. Phys. C, 21 (1988), p. 3251.

[55] L. Wang, *Calculating the density of states and optical-absorption spectra of large quantum systems by the plane-wave moments method*, Phy. rRv. B, 49 (1994), p. 10154.

[56] L.-W. Wang, *Calculating the density of states and optical-absorption spectra of large quantum systems by the plane-wave moments method*, Phys. Rev. B, 49 (1994), p. 10154.

[57] John C. Wheeler and Carl Blumstein, *Modified moments for harmonic solids*, Phys. Rev. B, 6 (1972), pp. 4380–4382.

[58] C. Yang, D. W. Noid, B. G. Sumpter, D. C. Sorensen, and R. E. Tuzun, *An efficient algorithm for calculating the heat capacity of a large-scale molecular system*, Macromol. Theory Simul., 10 (2001), pp. 756–761.

[59] Y. Zhou, Y. Saad, M. L. Tiago, and J. R. Chelikowsky, *Parallel self-consistent-field calculations via Chebyshev-filtered subspace acceleration*, Phy. Rev. E, 74 (2006), p. 066704.