Computing Spectral Measures and Spectral Types

Matthew J. Colbrook

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, UK.
E-mail: m.colbrook@damtp.cam.ac.uk

Abstract: Spectral measures arise in numerous applications such as quantum mechanics, signal processing, resonance phenomena, and fluid stability analysis. Similarly, spectral decompositions (into pure point, absolutely continuous and singular continuous parts) often characterise relevant physical properties such as the long-time dynamics of quantum systems. Despite new results on computing spectra, there remains no general method able to compute spectral measures or spectral decompositions of infinite-dimensional normal operators. Previous efforts have focused on specific examples where analytical formulae are available (or perturbations thereof) or on classes of operators that carry a lot of structure. Hence the general computational problem is predominantly open. We solve this problem by providing the first set of general algorithms that compute spectral measures and decompositions of a wide class of operators. Given a matrix representation of a self-adjoint or unitary operator, such that each column decays at infinity at a known asymptotic rate, we show how to compute spectral measures and decompositions. We discuss how these methods allow the computation of objects such as the functional calculus, and how they generalise to a large class of partial differential operators, allowing, for example, solutions to evolution PDEs such as the linear Schrödinger equation on $L^2(\mathbb{R}^d)$. Computational spectral problems in infinite dimensions have led to the Solvability Complexity Index (SCI) hierarchy, which classifies the difficulty of computational problems. We classify the computation of measures, measure decompositions, types of spectra, functional calculus, and Radon–Nikodym derivatives in the SCI hierarchy. The new algorithms are demonstrated to be efficient on examples taken from orthogonal polynomials on the real line and the unit circle (giving, for example, computational realisations of Favard's theorem and Verblunsky's theorem, respectively), and are applied to evolution equations on a two-dimensional quasicrystal.

Contents

1. Introduction

The analysis and computation of spectral properties of operators form core parts of many branches of science and mathematics, arising in diverse fields such as differential and integral equations, orthogonal polynomials, quantum mechanics, statistical mechanics, integrable systems and optics [15,43,44,69,110,129,138]. Correspondingly, the problem of numerically computing the spectrum, $\sigma(T)$, of an operator T acting on the canonical separable Hilbert space $l^2(\mathbb{N})$ has attracted a large amount of interest over the last 60 years or so [5–8,21,22,24,26,27,34,38,40,46,49,73,79–81,94,98,99,116,120,122–124]. However, the richness, beauty and difficulties that are encountered in infinite dimensions lie not just in the set $\sigma(T) \subset \mathbb{C}$, but also in the generalisation of projections onto eigenspaces and the possibility of different spectral types. Specifically, given a normal operator T, there is an associated projection-valued measure (resolution of the identity), which we denote by E^T, whose existence is guaranteed by the spectral theorem

and whose support is $\sigma(T)$ [86,87,113]. This allows the representation of the operator T as an integral over $\sigma(T)$, analogous to the finite-dimensional case of diagonalisation:

$$Tx = \left[\int_{\sigma(T)} \lambda \, dE^T(\lambda) \right] x, \quad \forall x \in \mathcal{D}(T), \tag{1.1}$$

where $\mathcal{D}(T)$ denotes the domain of T. For example, if T is compact, then E^T corresponds to projections onto eigenspaces, familiar from the finite-dimensional setting. However, in general, the situation is much richer and more complicated, with different types of spectra (pure point, absolutely continuous and singular continuous). An excellent and readable introduction can be found in Halmos' article [77].

The computation of E^T, along with its various decompositions and their supports, is of great interest, both theoretically and for practical applications. For example, spectral measures are intimately related to correlation functions in signal processing, resonance phenomena in scattering theory, and stability analysis for fluids. Moreover, the computation of E^T allows one to compute many additional objects, which we provide the first general algorithms for in this paper, such as the functional calculus (Theorem 4.1), the Radon–Nikodym derivative of the absolutely continuous component of the measure (Theorem 4.2), and the spectral measures and spectral set decompositions (Theorem 3.2 and Theorem 5.1). For instance, in Sect. 1.2.1 we show how our results allow the computation of spectral measures and the functional calculus of almost arbitrary self-adjoint partial differential operators on $L^2(\mathbb{R}^d)$. An important class of examples is given by solutions of evolution equations such as the Schrödinger equation on $L^2(\mathbb{R}^d)$ with a potential of locally bounded total variation. We prove that this is computationally possible even when an algorithm is only allowed to point sample the potential. A numerical example of fractional diffusion for a discrete quasicrystal is also given in Sect. 6.4.

Despite its importance, there has been no general method able to compute spectral measures of normal operators. Although there is a rich literature on the theory of spectral measures, most of the efforts to develop computational tools have focused on specific examples where analytical formulae are available (or perturbations thereof) or on classes of operators that carry a lot of structure. Indeed, apart from the work of Webb and Olver [151] (which deals with compact perturbations of tridiagonal Toeplitz operators) and methods for computing spectral density functions of Sturm–Liouville problems and other highly structured operators, there has been limited success in computing the measure E^T.[1] In some sense, this is not surprising given the difficulty in rigorously computing spectra. One can consider computing spectral measures/projections as the infinite-dimensional analogue of computing projections onto eigenspaces.[2] Thus, from a numerical/computational point of view, the current state of affairs in infinite-dimensional spectral computations is highly deficient, analogous in finite dimensions to being able to compute the location of eigenvalues but not eigenvectors! It has been unknown if the general computation of spectral measures is possible, even for simple subclasses such as discrete Schrödinger operators. In other words, the computational problem of "diagonalisation" through computing spectral measures remains an important and predominantly open problem.

In this paper, we solve this problem by providing the first set of algorithms for the computation of spectral measures for a large class of self-adjoint and unitary operators, namely, those whose matrix columns decay at a known asymptotic rate. This is a

---

[1] See Sect. 1.7 for connections with previous work.

[2] Of course eigenvectors exist in the infinite-dimensional case, but not all of the spectrum consists of eigenvalues. The projection-valued measure generalises the notion of projections onto eigenspaces.

very weak assumption and covers the majority of operators, even unbounded, found in applications. In particular, those whose representation is sparse (such as many of the graph or lattice operators typically dealt with in physics) and also partial differential operators, once a suitable basis has been selected (see Theorem 1.1 and Appendix B). We also show how to compute spectral measure decompositions, the functional calculus, the density of the absolutely continuous part of the measure (Radon–Nikodym derivative) and different types of spectra (pure point, absolutely continuous and singular continuous—these sets often characterise different physical properties in quantum mechanics [2,41,42,53,61,93,119,127]). A central ingredient of these new algorithms is the computation of the resolvent operator with error control through appropriate rectangular truncations (Theorem 2.1). Furthermore, we demonstrate the applicability of our algorithms. The algorithms are efficient and parallelisable, allowing large scale computations.

1.1. The solvability complexity index and classification of problems. A surprise thrown up by the infinite-dimensional spectral problem, which turns out to be quite generic, is the Solvability Complexity Index (SCI) [81]. The SCI provides a hierarchy for classifying the difficulty of computational problems. In classical numerical analysis, when computing spectra, one hopes to construct an algorithm, Γ_n, with one limit such that for an operator T,

$$\Gamma_n(T) \to \sigma(T), \quad \text{as } n \to \infty, \tag{1.2}$$

preferably with some form of error control or rate of convergence. However, this is not always possible. For example, when considering the class of bounded operators, the best possible alternative is an algorithm depending on three indices n_1, n_2, n_3 such that

$$\lim_{n_3 \to \infty} \lim_{n_2 \to \infty} \lim_{n_1 \to \infty} \Gamma_{n_3,n_2,n_1}(T) = \sigma(T).$$

Any algorithm with fewer than three limits will fail on some bounded operator, and neither error control nor convergence rates on any of the limits are possible since these would reduce the required number of limits. However, for self-adjoint operators, it is possible to reduce the number of limits to two, but not one [12,81]. With more structure (such as sparsity or column decay) it is possible to compute the spectrum in one limit with a certain type of error control [40]. Hence, the only way to characterise the computational spectral problem is through a hierarchy, classifying the difficulty of computing spectral properties of different subclasses of operators. The SCI classifies difficulty by considering the minimum number of limits that one must take to calculate the quantity of interest (see Appendix A for a full definition). The SCI has roots in the work of Smale [132,134], and his program on the foundations of computational mathematics and scientific computing, though it is quite distinct. The notions of Turing computability [145] and computability in the Blum–Shub–Smale (BSS) [17] sense become special cases, and impossibility results that are proven in the SCI hierarchy hold in all models of computation. The phenomenon of needing several limits also covers general numerical analysis problems, such as Smale's question on the existence of purely iterative algorithms for polynomial root finding [133]. As demonstrated by McMullen [101,102] and Doyle and McMullen [51], this is a case where several limits are needed in the computation, and their results become special cases of classification in the SCI hierarchy [12]. Extensions of the hierarchy to error control [33,36,37] also have potential applications

in the growing field of computer-assisted proofs, where one must perform a computation with absolute certainty. See, for example, the work of Fefferman and Seco on the Dirac–Schwinger conjecture [55,56] and Hales on Kepler's conjecture (Hilbert's 18th problem) [75], both of which implicitly provide classifications in the SCI hierarchy. As well as spectral problems [12,33,36–40,81], the SCI hierarchy has recently been used to solve problems related to computing resonances [13,14], computing solutions of semigroups and evolution PDEs [10,35], and computational barriers for stable and accurate neural networks [4].

An informal definition of the SCI hierarchy is as follows, with a detailed summary contained in Appendix A. The SCI hierarchy is based on the concept of a computational problem. This is described by a function

$$\Xi : \Omega \to \mathcal{M}$$

that we want to compute, where Ω is some domain, and $(\mathcal{M}, d)$ is a metric space. For example, we could take $\Xi(T) = \sigma(T)$ (the spectrum) for some class of operators Ω and $\mathcal{M}$ the collection of non-empty closed subsets of $\mathbb{C}$ equipped with the Attouch–Wets metric. The SCI of a computational problem is the smallest number of limits needed in order to compute the solution. For a given set of evaluation functions (the information our algorithm is allowed to read—in our case, Λ_1 or Λ_2 defined in (1.18)), class of objects (in our case, subclasses of operators acting on $l^2(\mathbb{N})$) and model of computation α (in this paper general, G, or arithmetic, A) we define:

(i) Δ_0^α is the set of problems that can be computed in finite time, the SCI $= 0$.
(ii) Δ_1^α is the set of problems that can be computed using one limit (the SCI $= 1$) with control of the error, i.e. $\exists$ a sequence of algorithms $\{\Gamma_n\}$ such that $d(\Gamma_n(A), \Xi(A)) \leq 2^{-n}$, $\forall A \in \Omega$.
(iii) Δ_2^α is the set of problems that can be computed using one limit (the SCI $= 1$) without error control, i.e. $\exists$ a sequence of algorithms $\{\Gamma_n\}$ such that $\lim_{n\to\infty} \Gamma_n(A) = \Xi(A)$, $\forall A \in \Omega$.
(iv) Δ_{m+1}^α, for $m \in \mathbb{N}$, is the set of problems that can be computed by using m limits, (the SCI $\leq m$), i.e. $\exists$ a family of algorithms $\{\Gamma_{n_m,\ldots,n_1}\}$ such that

$$\lim_{n_m\to\infty} \ldots \lim_{n_1\to\infty} \Gamma_{n_m,\ldots,n_1}(A) = \Xi(A), \ \forall A \in \Omega.$$

The class Δ_1 is, of course, a highly desired class; however, non-trivial spectral problems are higher up in the hierarchy. For example, the following classifications are known [12,81]:

(i) The general spectral problem is in $\Delta_4 \backslash \Delta_3$.
(ii) The self-adjoint spectral problem is in $\Delta_3 \backslash \Delta_2$.
(iii) The compact spectral problem is in $\Delta_2 \backslash \Delta_1$.

Here, the notation $\backslash$ indicates the standard "setminus". Hence, the computational spectral problem becomes an infinite classification theory to characterise the above hierarchy. In order to do so, there will, necessarily, have to be many different types of algorithms. Characterising the hierarchy will yield a myriad of different approaches, as different structures on the various classes of operators will require specific algorithms.

This paper provides classifications of spectral problems associated with E^T (such as decompositions of the measure and spectrum) in the SCI hierarchy, some of which can be computed in one limit. We provide algorithms for these problems, and one of the main tools used is the computation of the resolvent operator $R(z, T) := (T - zI)^{-1}$ with

error control (Theorem 2.1). This is possible through appropriate *rectangular* trunca-
tions of the infinite-dimensional operator. This approach differs from finite-dimensional
techniques, which typically consider square truncations.

Remark 1. (Recursivity and independence of the model of computation) The constructive
inclusion results we provide hold for arithmetic algorithms and the impossibility results
hold for general algorithms. We refer the reader to Appendix A for a detailed explanation.
Put simply, this means that the algorithms constructed can be recursively implemented
with inexact input and restrictions to arithmetic operations over the rationals (it is also
straightforward to implement them using interval arithmetic), whereas the impossibility
results hold in any model of computation (such as the Turing or BSS models).

1.2. Summary of the main results

1.2.1. Partial differential operators For $N \in \mathbb{N}$, consider the operator formally defined
on $L^2(\mathbb{R}^d)$ by

$$Lu(x) = \sum_{k \in \mathbb{Z}_{\geq 0}^d, |k| \leq N} a_k(x) \partial^k u(x), \tag{1.3}$$

where throughout we use multi-index notation with $|k| = \max\{|k_1|, \ldots, |k_d|\}$ and $\partial^k = \partial_{x_1}^{k_1} \partial_{x_2}^{k_2} \ldots \partial_{x_d}^{k_d}$. We will assume that the coefficients $a_k(x)$ are complex-valued measurable
functions on $\mathbb{R}^d$ and that L is self-adjoint. For dimension d and $r > 0$ consider the space

$$\mathcal{A}_r = \{f \in M([-r,r]^d) : \|f\|_\infty + \mathrm{TV}_{[-r,r]^d}(f) < \infty\}, \tag{1.4}$$

where $M([-r,r]^d)$ denotes the set of measurable functions on the hypercube $[-r,r]^d$
and $\mathrm{TV}_{[-r,r]^d}$ denotes the total variation norm in the sense of Hardy and Krause [103].
This space becomes a Banach algebra when equipped with the norm [18]

$$\|f\|_{\mathcal{A}_r} := \|f|_{[-r,r]^d}\|_\infty + (3^d + 1)\mathrm{TV}_{[-r,r]^d}(f).$$

Let Ω_{PDE} consist of all such L such that the following assumptions hold:

(1) The set $C_0^\infty(\mathbb{R}^d)$ of smooth, compactly supported functions forms a core of L.
(2) For each of the functions $a_k(x)$, there exists a positive constant A_k and an integer
B_k (both possibly unknown) such that

$$|a_k(x)| \leq A_k \left(1 + |x|^{2B_k}\right),$$

almost everywhere on $\mathbb{R}^d$, that is, we have at most polynomial growth of the coef-
ficients.
(3) The restrictions $a_k|_{[-r,r]^d} \in \mathcal{A}_r$ for all $r > 0$.

We consider the case where our algorithms can do the following:

1. Evaluate any coefficient $a_k(x)$ to a given precision at $x \in \mathbb{Q}^d$, where $\mathbb{Q}$ denotes the
field of rationals, and output an approximation in $\mathbb{Q} + i\mathbb{Q}$.

2. For each n, evaluate a positive number $b_n(L)$ such that the sequence $\{b_n(L)\}_{n \in \mathbb{N}}$ satisfies

$$\sup_{n \in \mathbb{N}} \max_{|k| \leq N} \frac{\|a_k\|_{\mathcal{A}_n}}{b_n(L)} < \infty. \tag{1.5}$$

In Appendix B, we prove (and state a more precise version of) the following theorem.

Theorem 1.1 (Spectral properties of self-adjoint partial differential operators can be computed). *Given the above set-up, there exist sequences of arithmetic algorithms that compute spectral measures, the functional calculus, and Radon–Nikodym derivatives of the absolutely continuous part of the measure over the class Ω_{PDE}. In other words, these objects can be computed in one limit with SCI = 1.*

Remark 2. As noted in Appendix B, we can extend this result to computing the decompositions into pure point, absolutely continuous and singular continuous parts of measures and spectra (with SCI > 1).

The above properties characterising Ω_{PDE} are deliberately very weak, and hence the class Ω_{PDE} is very large. For example, Schrödinger operators $L = -\Delta + V$ with polynomially bounded potentials of locally bounded total variation are a subclass of Ω_{PDE}. Hence, in this case, the theorem says that we can compute the spectral properties (measures, functional calculus etc.) of L by point sampling the potential V, if we have an asymptotic bound on the total variation of V over finite rectangles. In particular, we can solve the Schrödinger equation

$$\frac{du}{dt} = -iLu, \quad u_{t=0} = u_0 \tag{1.6}$$

by computing $\exp(-itL)u_0$ with guaranteed convergence. Applications of some of our results to semigroups (where error control can also be provided) are given in [35]. The proof of Theorem 1.1 can also be extended to other domains such as the half-line, and can be adapted to cope with other types of coefficients that are not of locally bounded total variation (for instance, Coulombic potentials for Dirac or Schrödinger operators). In order to prove Theorem 1.1, we prove theorems for operators on $l^2(\mathbb{N})$.

1.2.2. Operators on $l^2(\mathbb{N})$ We consider self-adjoint operators given as an infinite matrix T whose columns decay at a known asymptotic rate:

$$\|(P_{f(n)} - I)TP_n\| = O(\alpha_n) \tag{1.7}$$

for a sequence $\alpha_n \downarrow 0$ and function $f : \mathbb{N} \to \mathbb{N}$, where P_n denotes the orthogonal projection onto the linear span of the first n canonical basis vectors. This set-up is adapted in Appendix B to prove Theorem 1.1, where we make use of the following theorems proven for operators on $l^2(\mathbb{N})$.

We consider the problem of computing spectral measures. Specifically, for operators of the form (1.7) we develop the first algorithms and SCI classifications for:

- **Theorem** 3.1: Δ_2^A classification for the projection-valued spectral measure and, through taking inner products, the computation of the scalar spectral measures defined via

$$\mu_{x,y}^T(U) = \langle E^T(U)x, y \rangle.$$

This is done for open sets U and can be extended to other types of sets such as closed intervals or singletons (Theorem 5.2 shows that the problem $\notin \Delta_1^G$ in certain cases). These scalar-valued spectral measures play an important role in, for example, quantum mechanics, where they correspond to the distribution of the physical observable modelled by a Hamiltonian T.

- **Theorem** 3.2: $\Delta_3^A \backslash \Delta_2^G$ classification for the decompositions of the projection-valued and scalar-valued spectral measures into absolutely continuous, singular continuous and pure point parts. These decompositions often characterise different physical properties in quantum mechanics [2,41,42,53,61,93,119,127]

- **Theorem** 4.1: Δ_2^A classification for the functional calculus of operators, i.e. the computation of $F(T)$ for suitable functions F. For brevity, we consider functions that are bounded and continuous on the spectrum of T. However, the proof makes clear that wider classes can be dealt with. In some cases, Δ_1^A error control is possible, for instance, when considering the holomorphic functional calculus (see Sect. 6.4). A key application of this result is the computation of solutions of many evolution PDEs, such as the Schrödinger equation through the function $F(z) = \exp(-izt)$.

- **Theorem** 4.2: Δ_2^A classification for the Radon–Nikodym derivatives (densities) of the absolutely continuous parts of the scalar spectral measures with convergence in the L^1 sense on an open set. This requires a certain separation condition, without which our algorithm converges (Lebesgue) almost everywhere.

We also consider the computation of spectra as sets in the complex plane. Convergence is measured using the Hausdorff metric in the bounded case and using the Attouch–Wets metric in the unbounded case (i.e. uniform convergence on compact subsets of $\mathbb{C}$). Specifically, we provide the first algorithms computing these quantities and prove in **Theorem** 5.1 that:

- The absolutely continuous spectrum $\sigma_{ac}(T)$ can be computed in two limits but not one limit ($\Delta_3^A \backslash \Delta_2^G$ classification).
- The pure point spectrum $\sigma_{pp}(T)$ can be computed in two limits but not one limit ($\Delta_3^A \backslash \Delta_2^G$ classification).
- The singular continuous spectrum $\sigma_{sc}(T)$ can be computed in three limits (Δ_4^A classification). If $f(n) - n \geq \sqrt{2n} + 1/2$, then the computation cannot be done in two limits ($\notin \Delta_3^G$ classification). That is, if the local asymptotic bandwidth is allowed to grow sufficiently rapidly, three limits are needed, and this computational problem is exceedingly difficult. We do not know whether this growth condition on f can be dropped. However, without it, the problem still requires at least two limits ($\notin \Delta_2^G$ classification).

The main tool used to prove the above results is

- **Theorem** 2.1 **and Corollary** 2.2: The action of the resolvent $x \to R(z, T)x$ can be computed with *error control*. This also opens up potential applications in computer-assisted proofs.

We demonstrate that the "one-limit" algorithms constructed in this paper are implementable and efficient. These provide the first set of algorithms addressing these problems, and we have provided extensive numerical experiments in Sect. 6. This includes orthogonal polynomials on the real line and unit circle (where we also discuss acceleration through extrapolation), as well as fractional diffusion for a two-dimensional quasicrystal. Since writing the initial version of this paper, our algorithms have also been implemented in the software package SpecSolve of [39], which uses the machinery

of high-order rational kernels to accelerate computation of the Radon–Nikodym derivative of regular enough measures. In the current paper, we also show that in the case where the measure is regular enough, a global collocation method and local Richardson extrapolation for the computation of the Radon–Nikodym derivative can both accelerate convergence.

Finally, some brief remarks are in order.

(i) The impossibility results hold in general, even when restricted to tridiagonal operators. Furthermore, many of the impossibility results hold for structured operators such as bounded discrete Schrödinger operators. Our results (constructive algorithms and impossibility results) also carry over to a large class of normal operators, including unitary operators or skew-adjoint operators, both of which are important in applications. For the sake of clarity, we have stuck to the self-adjoint case in the statement of theorems and proofs. Numerical examples for unitary operators are given in Sect. 6.3.

(ii) The difficulty encountered when computing the singular continuous spectrum is partly due to the negative definition of the singular continuous part of a measure. It is the "leftover" part of the measure, the part that is not continuous with respect to Lebesgue measure and does not contain atoms. The challenge of studying σ_{sc} analytically also reflects this difficulty—singular continuous spectra were once thought to be rather rare or exotic. However, they are quite generic; see, for example, [128] for a precise topological statement to this effect.

(iii) One might at first expect computational results to be independent of the function f due to tridiagonalisation. However, the infinite-dimensional case is much more subtle than the finite-dimensional case. Using Householder transformations on a bounded sparse self-adjoint operator T leads to a tridiagonal operator, but, in general, this operator is T restricted to a *strict* subspace of $l^2(\mathbb{N})$. Part of the operator may be lost in the strong operator limit. Instead, one must consider a sum of possibly infinitely many tridiagonal operators (see [78, Ch. 2 & 8]). Hence some spectral problems may have different classifications for different f.

1.3. Open problems. The results of this paper form part of a program [33,34,36–40] on determining the foundations of computation (boundaries of what is possible) for infinite-dimensional spectral problems. As such, there remain many interesting open problems related to this paper, such as:

- **Computation of spectral measures for more general normal operators:** Proposition 2.4 demonstrates how the techniques of this paper can be generalised to certain classes of normal operators whose spectrum does not necessarily lie along a curve. However, it is not obvious how to extend the techniques to completely general normal operators. For example, the method of integrating the resolvent along a contour cannot be easily extended to interior points of the spectrum.
- **Brown measure for non-normal operators:** There is an extension of spectral measures to certain non-normal operators known as the Brown measure [25,71,72]. Computing spectra of non-normal operators is generally harder (in a sense made precise by the SCI hierarchy) than for normal operators (issues such as numerical stability are also present, even in the finite-dimensional case). It would be interesting to see if this phenomenon is also present for computing the Brown measure. Some works on approximating the Brown measure from matrix elements include [7,11,81].

1.4. A motivating example. As a motivating example, consider the case of a Jacobi operator with matrix

$$
J = \begin{pmatrix}
b_1 & a_1 & & \\
a_1 & b_2 & a_2 & \\
& a_2 & b_3 & \ddots \\
& & \ddots & \ddots
\end{pmatrix},
$$

where $a_j, b_j \in \mathbb{R}$ and $a_j > 0$. An enormous amount of work exists on the study of these operators, and the correspondence between bounded Jacobi matrices and probability measures with compact support [45,143]. The entries in the matrix provide the coefficients in the recurrence relation for the associated orthonormal polynomials. To study the canonical measure μ_J, one usually considers the principal resolvent function, which is defined on $\mathbb{C}\backslash\sigma(J)$ via

$$
G(z) := \langle R(z, J)e_1, e_1 \rangle = \int_{\mathbb{R}} \frac{d\mu_J(\lambda)}{\lambda - z}, \tag{1.8}
$$

and then takes z close to the real axis. The function G is also known in the differential equations and Schrödinger communities as the Weyl m-function [64,143] and one can develop the discrete analogue of what is known as Weyl–Titchmarsh–Kodaira theory for Sturm–Liouville operators. Going back to the work of Stieltjes [136] (see also [1,149]), there is a representation of G as a continued fraction:

$$
G(z) := \cfrac{1}{-z + b_1 - \cfrac{a_1^2}{-z+b_2-\dots}}. \tag{1.9}
$$

One can also approximate G via finite truncated matrices [143].

However, there are two major obstacles to overcome when using (1.9) and its variants as a means to compute measures. First of all, this representation of the principal resolvent function is structurally dependent. For example, (1.9) is valid for the restricted case of Jacobi operators and hence one is led to seeking different methods for different operators (such as tight-binding Hamiltonians on two-dimensional lattices which have a growing bandwidth when represented as operators on $l^2(\mathbb{N})$). Second, this would seem to give the wrong classification of the difficulty of the problem in the SCI hierarchy, giving rise to a tower of algorithms with two limits. One first takes a truncation parameter n to infinity to compute $G(z)$ for $\mathrm{Im}(z) > 0$, and then a second limit as z approaches the real axis. One of the main messages of this paper is that both of these issues can be overcome. Measures can be computed in one limit via an algorithm Γ_n and for a large class of operators. The only restriction is a known asymptotic decay rate of the off-diagonal entries. As a by-product, we compute the m-function of such operators with error control. Specific cases where this can be written explicitly do exist, such as periodic Jacobi matrices or perturbations of Toeplitz operators [52] (see also Sect. 1.7). However, there has been no general method proposed to compute the resolvent with error control. This consideration is crucial to allow the computation of measures in one limit.

To see how we might compute the measure using the resolvent, consider the Poisson kernels for the half-plane and the unit disk, defined respectively by

$$P_H(x, y) = \frac{1}{\pi} \frac{y}{x^2 + y^2} \text{ and } P_D(x, y) = \frac{1}{2\pi} \frac{1 - (x^2 + y^2)}{(x - 1)^2 + y^2}$$

$$= P_D(r, \theta) = \frac{1}{2\pi} \frac{1 - r^2}{1 - 2r\cos(\theta) + r^2}, \quad (1.10)$$

where (r, θ) denote the usual polar coordinates. Let T be a normal operator, then for $z \notin \sigma(T)$, we have from the functional calculus that

$$R(z, T) = \int_{\sigma(T)} \frac{1}{\lambda - z} dE^T(\lambda).$$

For self-adjoint T, $z = u + iv \in \mathbb{C}\backslash\mathbb{R}$ ($u, v \in \mathbb{R}$) and $x \in l^2(\mathbb{N})$, we define

$$K_H(z; T, x) := \frac{1}{2\pi i}[R(z, T) - R(\bar{z}, T)]x$$

$$= \frac{1}{2\pi i} \int_{-\infty}^{\infty} \left(\frac{1}{\lambda - z} - \frac{1}{\lambda - \bar{z}} \right) dE^T(\lambda)x \quad (1.11)$$

$$= \int_{-\infty}^{\infty} P_H(u - \lambda, v) dE^T(\lambda)x.$$

Similarly, if T is unitary, $z = r\exp(i\psi) \in \mathbb{C}\backslash\mathbb{T}$ (with $z \neq 0$) and $x \in l^2(\mathbb{N})$, we define

$$K_D(z; T, x) := \frac{1}{2\pi i}[R(z, T) - R(1/\bar{z}, T)]x$$

$$= \frac{1}{2\pi i} \int_{\mathbb{T}} \left[\frac{1}{\lambda - z} - \frac{1}{\lambda - 1/\bar{z}} \right] dE^T(\lambda)x. \quad (1.12)$$

We change variables $\lambda = \exp(i\theta)$ and, with an abuse of notation, write $dE^T(\lambda) = i\exp(i\theta)dE^T(\theta)$. A simple calculation then gives

$$K_D(z; T, x) = \int_0^{2\pi} P_D(r, \psi - \theta) dE^T(\theta)x. \quad (1.13)$$

Returning to our example, we see that the computation of the resolvent with error control allows the computation of $G(z)$ with error control through taking inner products. By considering $G(z) - G(\bar{z})$, this allows the computation of the convolution of the measure μ_J with the Poisson kernel P_H. In other words, we can compute a *smoothed version* of the measure μ_J with error control. We can then take the smoothing parameter to zero to recover the measure (one can show that these smoothed approximations converge weakly in the sense of measures). Fig. 1 demonstrates this for a typical example.

1.5. Functional analytic setup. We consider the canonical[3] separable Hilbert space $\mathcal{H} = l^2(\mathbb{N})$, the set of square summable sequences with canonical basis $\{e_n\}_{n \in \mathbb{N}}$. Let $\mathcal{C}(l^2(\mathbb{N}))$ be the set of closed densely defined linear operators T such that $\mathrm{span}\{e_n : n \in \mathbb{N}\}$ forms a core of T and T^*. The spectrum of $T \in \mathcal{C}(l^2(\mathbb{N}))$ will be denoted by $\sigma(T)$ and the point spectrum (the set of eigenvalues) by $\sigma_p(T)$. The latter set is not always closed and

---

[3] By a choice of basis our results extend to any separable Hilbert space.
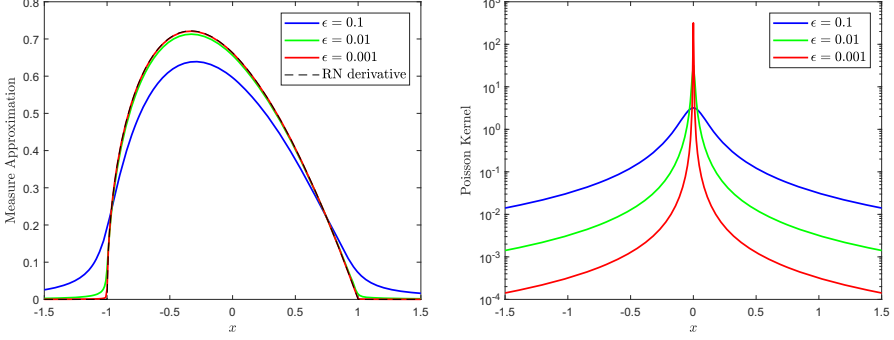
Fig. 1. Smoothed approximations of the Radon–Nikodym derivative for the Jacobi operator associated to Jacobi polynomials with $\alpha = 1, \beta = 1/2$. Here the measure is absolutely continuous and supported on $[-1, 1]$. Left: Computation of convolutions for different ϵ using the methods of this paper. Right: The associated Poisson kernel $\pi^{-1}\epsilon/(\epsilon^2 + x^2)$ which approaches a Dirac delta distribution as $\epsilon \downarrow 0$

in general the closure of a set S will be denoted by $\overline{S}$. The resolvent operator $(T - zI)^{-1}$ defined on $\mathbb{C}\backslash\sigma(T)$ will be denoted by $R(z, T)$.

This paper focusses on the subclass $\Omega_N \subset \mathcal{C}(l^2(\mathbb{N}))$ of normal operators, that is, operators for which $\mathcal{D}(T) = \mathcal{D}(T^*)$ and $\|Tx\| = \|T^*x\|$ for all $x \in \mathcal{D}(T)$. The subclasses $\subset \Omega_N$ of self-adjoint and unitary operators will be denoted by Ω_{SA} and Ω_U respectively. For $T \in \Omega_{SA}$ and $T \in \Omega_U$, $\sigma(T) \subset \mathbb{R}$ and $\sigma(T) \subset \mathbb{T}$ respectively, where $\mathbb{T}$ denotes the unit circle. Given $T \in \Omega_N$ and a Borel set B, E_B^T will denote the projection $E^T(B)$. Given $x, y \in l^2(\mathbb{N})$, we can define a bounded (complex-valued) measure $\mu_{x,y}^T$ via the formula

$$\mu_{x,y}^T(B) = \langle E_B^T x, y \rangle. \tag{1.14}$$

Via the Lebesgue decomposition theorem [76], the spectral measure $\mu_{x,y}^T$ can be decomposed into three parts

$$\mu_{x,y}^T = \mu_{x,y,\mathrm{ac}}^T + \mu_{x,y,\mathrm{sc}}^T + \mu_{x,y,\mathrm{pp}}^T, \tag{1.15}$$

the absolutely continuous part of the measure (with respect to the Lebesgue measure), the singular continuous part (singular with respect to the Lebesgue measure and atomless) and the pure point part. When considering Ω_{SA}, we will consider Lebesgue measure on $\mathbb{R}$ and let

$$\rho_{x,y}^T(\lambda) = \frac{d\mu_{x,y,\mathrm{ac}}^T}{dm}(\lambda), \tag{1.16}$$

the Radon–Nikodym derivative of $\mu_{x,y,\mathrm{ac}}^T$ with respect to Lebesgue measure. Of course this can be extended to the unitary (and, more generally, normal) case. This naturally gives a decomposition of the Hilbert space $\mathcal{H} = l^2(\mathbb{N})$. For $\mathcal{I} = \mathrm{ac}$, sc and pp, we let $\mathcal{H}_\mathcal{I}$ consist of vectors x whose measure $\mu_{x,x}^T$ is absolutely continuous, singular continuous and pure point respectively. This gives rise to the orthogonal (invariant subspace) decomposition

$$\mathcal{H} = \mathcal{H}_{\mathrm{ac}} \oplus \mathcal{H}_{\mathrm{sc}} \oplus \mathcal{H}_{\mathrm{pp}}, \tag{1.17}$$

whose associated projections will be denoted by P_{ac}^T, P_{sc}^T and P_{pp}^T respectively. These projections commute with T and the projections obtained through the projection-valued

measure. Of particular interest is the spectrum of T restricted to each $\mathcal{H}_\mathcal{I}$, which will be denoted by $\sigma_\mathcal{I}(T)$. These different sets and subspaces often, but not always, characterise different physical properties in quantum mechanics (such as the famous RAGE theorem [2,53,119]), where a system is modelled by some Hamiltonian $T \in \Omega_{SA}$ [41,42,61,93]. For example, pure point spectrum implies the absence of ballistic motion for many Schrödinger operators [127].

1.6. Algorithmic setup. Given an operator $T \in \mathcal{C}(l^2(\mathbb{N}))$, we can view it as an infinite matrix

$$T = \begin{pmatrix} t_{11} & t_{12} & t_{13} & \cdots \\ t_{21} & t_{22} & t_{23} & \cdots \\ t_{31} & t_{32} & t_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

through the inner products[4] $t_{ij} = \langle Te_j, e_i \rangle$. All of the algorithms constructed can also be adapted to operators on $l^2(\mathbb{Z})$, either through the use of a suitable re-ordering of the basis, or though considering truncations of matrices in two directions, which is useful numerically since it preserves bandwidth. To be precise about the information needed to compute spectral properties, we define two classes of evaluation functions as

$$\Lambda_1 = \{\langle Te_j, e_i \rangle : i, j \in \mathbb{N}\}, \quad \Lambda_2 = \{\langle Te_j, e_i \rangle, \langle T^*e_j, T^*e_i \rangle : i, j \in \mathbb{N}\}. \quad (1.18)$$

These can be understood as different sets of information our algorithms are allowed to access (see Appendix A for a precise meaning). All the results proven in this paper can be easily extended to the case of inexact input. This means replacing the evaluation functions by

$$f_{i,j,m}^{(1)}, f_{i,j,m}^{(2)} : \mathcal{C}(l^2(\mathbb{N})) \to \mathbb{Q} + i\mathbb{Q}$$

such that $|f_{i,j,m}^{(1)}(T) - \langle Te_j, e_i \rangle| \leq 2^{-m}$ and $|f_{i,j,m}^{(2)}(T) - \langle T^*e_j, T^*e_i \rangle| \leq 2^{-m}$. Hence, the existence results carry over to algorithms that are only allowed to perform arithmetic operations over $\mathbb{Q}$. This could be useful for rigorous bounds using interval arithmetic and computer-assisted proofs (for those familiar with the term, our algorithms are Turing recursive), though for brevity, we stick to Λ_1 and Λ_2 throughout. For discrete operators, the above information is often given to us, for example, in tight-binding models in physics or as a discretisation of a PDE, and hence it is natural to seek to compute spectral properties from matrix values. The set Λ_2 is motivated via variational problems. For partial differential operators, such information is often given through inner products with a suitable basis, and in this case, the inexact input model is needed due to approximating the integrals (see Appendix B). For the classes considered in this paper, the evaluation sets Λ_1 and Λ_2 are in general different, yet the classifications in the SCI remain the same.

We will be concerned operators whose matrix representation has a known *asymptotic rate of column/off-diagonal decay*. Namely, let $f : \mathbb{N} \to \mathbb{N}$ with $f(n) > n$ and let

---

[4] Our convention throughout will be that the inner product $\langle \cdot, \cdot \rangle$ is linear in the first component and conjugate-linear in the second.

$\alpha = \{\alpha_n\}_{n \in \mathbb{N}}$, $\beta = \{\beta_n\}_{n \in \mathbb{N}}$ be null sequences[5] of non-negative real numbers. We then define for X = SA or X = U,

$$
\begin{aligned}
\Omega_{f,\alpha,\beta}^{\mathrm{X}} = \{T \in \Omega_{\mathrm{X}} : \|(P_{f(n)} - I)T P_n\| = O(\alpha_n), \text{ as } n \to \infty\} \\
\times \{x \in l^2(\mathbb{N}) : \|P_n x - x\| = O(\beta_n), \text{ as } n \to \infty\},
\end{aligned} \tag{1.19}
$$

where P_n denotes the orthogonal projection onto span$\{e_1, \ldots, e_n\}$. In using the $O(\cdot)$ notation, the hidden constant is allowed to depend on the operator T or the vector x. We will also use

$$
\Omega_{f,\alpha}^{\mathrm{X}} = \{T \in \Omega_{\mathrm{X}} : \|(P_{f(n)} - I)T P_n\| = O(\alpha_n), \text{ as } n \to \infty\}. \tag{1.20}
$$

When discussing $\Omega_{f,\alpha,\beta}^{\mathrm{SA}}$ and $\Omega_{f,\alpha}^{\mathrm{SA}}$ we will use the notation $\Omega_{f,\alpha,\beta}$ and $\Omega_{f,\alpha}$. The collection of vectors in $l^2(\mathbb{N})$ satisfying $\|P_n x - x\| = O(\beta_n)$ will be denoted by V_β. Finally, when $\alpha_n \equiv 0$, we will abuse notation slightly in requiring the stronger condition that

$$
\|(P_{f(n)} - I)T P_n\| = 0.
$$

Thus $\Omega_{f,0}$ is the class of self-adjoint operators whose matrix sparsity structure is captured by the function f. For example, if $f(n) = n + 1$, we recover the class of self-adjoint tridiagonal matrices, the most studied class of infinite-dimensional operators. When discussing classes that include vectors x, we extend Λ_i to include pointwise evaluations of the coefficients of x. Other additions are sometimes needed, such as data regarding open sets as inputs for computations of measures, but this will always be made clear. When considering the general case of $\Omega_{f,\alpha}$, the function f and sequence α can also be considered as inputs to the algorithm—in other words, the same algorithm works for each class.

1.7. Connections with previous work. We have mentioned the literature on infinite-dimensional spectral problems and the SCI hierarchy. Computationally, our point of view in this paper is closest to the work of Olver, Townsend and Webb on practical infinite-dimensional linear algebra [104–107,151]. Their work includes efficient codes, such as the infinite-dimensional QL (IQL) algorithm [150] (see also [38] for the IQR algorithm, which has its roots in the work of Deift, Li and Tomei [46]), as well as theoretical results. A PDE version of the FEAST algorithm based on contour integration of the resolvent has recently been proposed by Horning and Townsend in [84], which computes discrete spectra. The set of algorithms we provide can be considered as a new member within the growing family of infinite-dimensional techniques.

A similar, though different, object studied in the mathematical physics literature, particularly when considering random Schrödinger operators, is the density of states [29, 89,91], which we mention for completeness and also to avoid potential confusion. This object is defined via the "thermodynamic limit", where instead of considering the infinite-dimensional operator T, one considers finite truncations, say $P_n T P_n$, and the limit $n \to \infty$ of the measure $\sum_{x_j \in \sigma(P_n T P_n)} \delta_{x_j}/n$. This measure is subtly different from the spectral measure of T (for instance, on the discrete spectrum). The density of states is an important quantity in quantum mechanics, and there is a large literature on its computation. We refer the reader to the excellent review article [95], which discusses the most common

---

[5] We use the term "null sequence" for a sequence converging to zero.

methods. The idea of using the resolvent to approximate the density of states of finite matrices can be found in the method of [82], which approximates the imaginary part of Trace $[R(z, P_n T P_n)]$ for $\text{Im}(z) > 0$. Similarly, in the random matrix literature, the connection is made through the Stieltjes transformation (see, for example, [9]). There are three immediate differences between our algorithms and those that compute the density of states. First, we seek to deal with the full, infinite-dimensional, operator directly to compute the spectral measure (and not the limit of increasing system sizes). Second, the object we are computing contains more refined spectral information of the operator and does not involve an averaging procedure. The density of states does not capture the full spectral information, such as the contribution of eigenvalues in the discrete spectrum, whereas the projection-valued spectral measure does. Third, there is a subtlety regarding the limits as $\text{Im}(z)$ goes to zero and the truncation parameter goes to infinity (a similar trade-off also occurs in random matrix literature when theoretically analysing the density of states—see [54]). In our case, appropriate *rectangular* truncations of the infinite-dimensional operator are required to compute the resolvent with error control (see Theorem 2.1). This approach differs from finite-dimensional techniques, which typically consider square truncations.

For the C^*-algebra viewpoint of the density of states, we refer the reader to the work of Areveson [7] and the references therein. Estimating the spectrum of T via $\sigma(P_n T P_n)$ is known as the finite section method. This has often been viewed in connection with Toeplitz theory. See, for example, the work by Böttcher [19,20], Böttcher and Silberman [23], Böttcher, Brunner, Iserles and Nørsett [21], Brunner, Iserles and Nørsett [27], Hagen, Roch and Silbermann [73], Lindner [96], Marletta [98], Marletta and Scheichl [99], and Seidel [121].

The study of spectral measures also has a rich history in the theory of orthogonal polynomials and quadrature rules for numerical integration going back to the work of Szegő [50,139], briefly touched upon in Sect. 1.4. In certain cases, one can recover a distribution function for the associated measure of the Jacobi operator as a limit of functions constructed using Gaussian quadrature [31, Ch. 2]. Our examples in Sect. 6.1 can be considered as a computational realisation of Favard's theorem.

There are several results in the literature considering the computation of spectral density functions for Sturm–Liouville problems. In the case of Sturm–Liouville problems, the spectral density function corresponds to the multiplicative version of the spectral theorem. This is subtly different from the measures we compute, which arise from the projection-valued measure version of the spectral theorem. A common approach to approximate spectral density functions associated with Sturm–Liouville operators on unbounded domains is to truncate the domain and use the Levitan–Levinson formula, as implemented in the software package SLEDGE [59,60,111]. This approach can be computationally expensive since the eigenvalues cluster as the domain size increases; often, hundreds of thousands of eigenvalues and eigenvectors need to be computed. More sophisticated methods avoiding domain truncation are considered for special cases in [57,58], and an application in plasma physics can be found in [152]. These make use of the additional structure present in Sturm–Liouville problems using results analogous to (1.9) in the continuous case. Our results, particularly Theorem 1.1, hold for partial differential operators much more complicated than Sturm–Liouville operators (see Appendix B).

Finally, we wish to highlight the work of Webb and Olver [151], which is of particular relevance to the present study. There the authors studied, through connection coefficients, Jacobi operators that arise as compact perturbations of Toeplitz operators.

Similar perturbations (where a stronger exponential decay of the perturbation is crucial for analyticity properties of the resolvent) were studied in the work of Bilman and Trogdon [16] in connection with the inverse scattering transform for the Toda lattice. (See also the work of Trogdon, Olver and Deconinck [144] for computations of spectral measures for inverse scattering for the KdV equation.) The results proven in [151] can be stated in terms of the SCI hierarchy:

- If the perturbation is finite rank (and known), the computation of σ_{pp} lies in Δ_1^G, and the computation of the μ_{ac} lies in Δ_0^G (note that σ_{ac} is known analytically).
- If the perturbation is compact with a known rate of decay at infinity, then the computation of the full spectrum σ lies in Δ_1^G.

The current paper complements the work of [151] by; considering operators much more general than tridiagonal compact perturbations of Toeplitz operators (we deal with arbitrary self-adjoint operators and assume we know f such that (1.20) holds) and partial differential operators, allowing operators to be unbounded, building algorithms that are arithmetic and can cope with inexact input, and considering computation of a wider range of spectral information. At the price of this greater generality, the objects we study are generally not computable with error control (unless one has local regularity assumptions on the measure—see [33, Ch. 4]), and some lead to computational problems higher up in the SCI hierarchy, though still computationally useful as we shall demonstrate. Our methods are also entirely different and rely on estimating the resolvent operator with error control (Theorem 2.1).

1.8. Organisation of the paper. The paper is organised as follows. In Sect. 2 we consider the computation of the resolvent with error control and generalisations of Stone's formula. The computation of measures, their various decompositions and projections are discussed in Sect. 3. We then discuss the functional calculus and density of measures in Sect. 4. The computation of the different types of spectra as sets in the complex plane is discussed in Sect. 5. We run extensive numerical tests in Sect. 6, where we also introduce a new collocation method for the computation of the Radon–Nikodym derivative. We find that increased rates of convergence can also be obtained through iterations of Richardson extrapolation. A summary of the SCI hierarchy can be found in Appendix A and a proof of Theorem 1.1 in Appendix B. Throughout, our theorems and proofs use the notation introduced in Sect. 1.5 and Sect. 1.6.

2. Preliminary Results

The algorithms built in this paper rely on the computation of the action of the resolvent operator $R(z, T) = (T - z)^{-1}$ for $z \notin \sigma(T)$ with (asymptotic) error control. Given this, one can compute the action of the projections E_S^T for a wide range of sets S (Theorem 3.1 and its generalisations), and hence the measures $\mu_{x,y}^T$. This section discusses the computation of the resolvent with error control and generalisations of Stone's formula, which relate the resolvent to the projection-valued measures.

2.1. Approximating the resolvent operator. The key theorem for computing the action of the resolvent operator is the following, where we use σ_1 to denote the injection modulus of an operator defined as

$$\sigma_1(T) := \min\{\|Tx\| : x \in \mathcal{D}(T), \|x\| = 1\}.$$

The proof of the following theorem boils down to a careful computation of a least-squares solution of a rectangular linear system.

Theorem 2.1. *Let $\alpha = \{\alpha_n\}_{n \in \mathbb{N}}$ and $\beta = \{\beta_n\}_{n \in \mathbb{N}}$ be null sequences, and $f : \mathbb{N} \to \mathbb{N}$ with $f(n) > n$. Define*

$$S_{f,\alpha,\beta} := \left\{ (T, x, z) \in \Omega^{\mathrm{N}}_{f,\alpha,\beta} \times \mathbb{C} : z \in \mathbb{C} \backslash \sigma(T) \right\}$$

and for $(T, x) \in \Omega^{\mathrm{N}}_{f,\alpha,\beta}$ (defined in (1.19)), let $C_1(T, x), C_2(T, x) \in \mathbb{R}_{\geq 0}$ be such that

1. $\|(I - P_{f(n)})T P_n\| \leq C_1 \alpha_n$,
2. $\|P_n x - x\| \leq C_2 \beta_n$,

where, for notational convenience, we drop the (T, x) dependence in the notation for C_1 and C_2. Then there exists a sequence of arithmetical algorithms

$$\Gamma_n : S_{f,\alpha,\beta} \to l^2(\mathbb{N}),$$

each of which use the evaluation functions in Λ_1 (defined in (1.18)), such that each vector $\Gamma_n(T, x, z)$ has finite support with respect to the canonical basis for each n, and

$$\lim_{n \to \infty} \Gamma_n(T, x, z) = R(z, T)x \text{ in } l^2(\mathbb{N}).$$

Moreover, the following error bound holds

$$
\begin{aligned}
&\|\Gamma_n(T, x, z) - R(z, T)x\| \\
&\leq \frac{C_2 \beta_{f(n)} + C_1 \alpha_n \|\Gamma_n(T, x, z)\| + \|P_{f(n)}(T - zI)\Gamma_n(T, x, z) - P_{f(n)}x\|}{\mathrm{dist}(z, \sigma(T))}.
\end{aligned}
$$

$$(2.1)$$

If a bound on C_1 and C_2 are known, this error bound can be computed to arbitrary accuracy using finitely many arithmetic operations and comparisons.

Proof. Let $(T, x, z) \in S_{f,\alpha,\beta}$. We have that $n = \mathrm{rank}(P_n) = \mathrm{rank}((T - zI)P_n) = \mathrm{rank}(P_{f(n)}(T - zI)P_n)$ for large n since $\sigma_1(T - zI) > 0$ and $\|(I - P_{f(n)})(T - zI)P_n\| \leq C_1 \alpha_n \to 0$ (recall that $z \notin \sigma(T)$). Hence we can define

$$
\widetilde{\Gamma}_n(T, x, z) := \begin{cases} 0 & \text{if } \sigma_1(P_n(T^* - \bar{z}I)P_{f(n)}(T - zI)P_n) \leq \frac{1}{n} \\ [P_n(T^* - \bar{z}I)P_{f(n)}(T - zI)P_n]^{-1} P_n(T^* - \bar{z}I)P_{f(n)}x & \text{otherwise.} \end{cases}
$$

Suppose that n is large enough so that $\sigma_1(P_n(T^* - \bar{z}I)P_{f(n)}(T - zI)P_n) > 1/n$. Then $\widetilde{\Gamma}_n(T, x, z)$ is a least-squares solution of the optimisation problem $\mathrm{argmin}_y \|P_{f(n)}(T - zI)P_n y - x\|$. The linear space $\mathrm{span}\{e_n : n \in \mathbb{N}\}$ forms a core of T and hence also of $T - zI$. It follows by invertibility of $T - zI$ that given any $\epsilon > 0$, there exists an $m = m(\epsilon)$ and a $y = y(\epsilon)$ with $P_m y = y$ such that

$$\|(T - zI)y - x\| \leq \epsilon.$$

It follows that for all $n \geq m$,

$$\|(T - zI)\widetilde{\Gamma}_n(T, x, z) - x\| \leq \|P_{f(n)}(T - zI)\widetilde{\Gamma}_n(T, x, z) - x\| + C_1\alpha_n\|\widetilde{\Gamma}_n(T, x, z)\|$$
$$\leq \|P_{f(n)}(T - zI)y - x\| + C_1\alpha_n\|\widetilde{\Gamma}_n(T, x, z)\|$$
$$\leq \|P_{f(n)}(T - zI)y - P_{f(n)}x\| + C_2\beta_{f(n)} + C_1\alpha_n\|\widetilde{\Gamma}_n(T, x, z)\|$$
$$\leq \epsilon + C_2\beta_{f(n)} + C_1\alpha_n\|\widetilde{\Gamma}_n(T, x, z)\|.$$

This implies that

$$\|\widetilde{\Gamma}_n(T, x, z) - R(z, T)x\| \leq \|R(z, T)\|\|(T - zI)\widetilde{\Gamma}_n(T, x, z) - x\|$$
$$\leq \|R(z, T)\| \left(\epsilon + C_2\beta_{f(n)} + C_1\alpha_n\|\widetilde{\Gamma}_n(T, x, z)\|\right).$$

In particular, since α and β are null, this implies that $\|\widetilde{\Gamma}_n(T, x, z)\|$ is uniformly bounded in n. Since $\epsilon > 0$ was arbitrary, we also see that $\widetilde{\Gamma}_n(T, x, z)$ converges to $R(z, T)x$.

Define the matrices

$$B_n = P_n(T^* - \overline{z}I)P_{f(n)}(T - zI)P_n, \quad C_n = P_n(T^* - \overline{z}I)P_{f(n)}.$$

Given the evaluation functions in Λ_1, we can compute the entries of these matrices to any given accuracy and hence also to arbitrary accuracy in the operator norm using finitely many arithmetic operations and comparisons (using the error in the Frobenius norm to bound the error in the operator norm). Denote approximations of B_n and C_n by $\widetilde{B}_n$ and $\widetilde{C}_n$ respectively and assume that

$$\|B_n - \widetilde{B}_n\| \leq u_n, \quad \|C_n - \widetilde{C}_n\| \leq v_n,$$

for null sequences $\{u_n\}$, $\{v_n\}$. Note that $\widetilde{B}_n^{-1}$ can be computed using finitely many arithmetic operations and comparisons. So long as u_n is small enough, the resolvent identity implies that

$$\|B_n^{-1} - \widetilde{B}_n^{-1}\| \leq \frac{\|\widetilde{B}_n^{-1}\|^2 u_n}{1 - u_n\|\widetilde{B}_n^{-1}\|} =: w_n.$$

By taking u_n and v_n smaller if necessary (so that the algorithm is adaptive and it is straightforward to bound the norm of a finite matrix from above), we can ensure that $\|\widetilde{B}_n^{-1}\|v_n \leq n^{-1}$ and $(\|\widetilde{C}_n\| + v_n)w_n \leq n^{-1}$. From Proposition A.7 and a simple search routine, we can also compute $\sigma_1(P_n(T^* - \overline{z}I)P_{f(n)}(T - zI)P_n)$ to arbitrary accuracy using finitely many arithmetic operations and comparisons. Suppose this is done to an accuracy $1/n^2$ and denote the approximation via τ_n. We then define

$$\Gamma_n(T, x, z) := \begin{cases} 0 & \text{if } \tau_n \leq \frac{1}{n} \\ \widetilde{B}_n^{-1}\widetilde{C}_n\widetilde{x}_n & \text{otherwise,} \end{cases}$$

where $\widetilde{x}_n = P_{f(n)}x$. It follows that $\Gamma_n(T, x, z)$ can be computed using finitely many arithmetic operations and, for large n,

$$\|\Gamma_n(T, x, z) - \widetilde{\Gamma}_n(T, x, z)\| \leq \left(\|\widetilde{B}_n^{-1}\|v_n + (\|\widetilde{C}_n\| + v_n)w_n\right)\|x\| \to 0,$$

so that $\Gamma_n(T, x, z)$ converges to $R(z, T)x$. By construction, $\Gamma_n(T, x, z)$ has finite support with respect to the canonical basis.

Furthermore, the following error bound holds (which also holds if $\tau_n \leq 1/n$)

$$\|\Gamma_n(T, x, z) - R(z, T)x\| \leq \|R(z, T)\| \|(T - zI)\Gamma_n(T, x, z) - x\|$$
$$\leq \frac{C_2\beta_{f(n)} + C_1\alpha_n\|\Gamma_n(T, x, z)\| + \|P_{f(n)}(T - zI)\Gamma_n(T, x, z) - P_{f(n)}x\|}{\text{dist}(z, \sigma(T))},$$

since T is normal so that $\|R(z, T)\| = \text{dist}(z, \sigma(T))^{-1}$. This bound converges to 0 as $n \to \infty$. If the C_1 and C_2 are known, the bound can be approximated to arbitrary accuracy using finitely many arithmetic operations and comparisons. □

Remark 3. If T corresponds to a choice of basis in a space of functions (for example when using a spectral method), there is often a link between the regularity of the functions x and the decay of the terms β_n. The bound (2.1) can then often be adapted to include such asymptotics, and hence indicate how large n needs to be to gain a given approximation.

Of course, a vast literature exists on computing $R(z, T)$, especially for infinite matrices with structure (such as being banded) and we refer the reader to [70,96,112,121] for a small sample. Note that if T is banded with bandwidth m, then we can take $f(n) = n + m$ and the above computation can be done in $O(nm^2)$ operations [66]. The following corollary of Theorem 2.1 will be used repeatedly in the following proofs.

Corollary 2.2. *There exists a sequence of arithmetic algorithms*

$$\Gamma_n : \Omega_{f,\alpha,\beta} \times \mathbb{C}\backslash\mathbb{R} \to l^2(\mathbb{N})$$

with the following properties:

1. *For all $(T, x) \in \Omega_{f,\alpha,\beta}$ and $z \in \mathbb{C}\backslash\mathbb{R}$, $\Gamma_n(T, x, z)$ has finite support with respect to the canonical basis and converges to $R(z, T)x$ in $l^2(\mathbb{N})$ as $n \to \infty$.*
2. *For any $(T, x) \in \Omega_{f,\alpha,\beta}$, there exists a constant $C(T, x)$ such that for all $z \in \mathbb{C}\backslash\mathbb{R}$,*

$$\|\Gamma_n(T, x, z) - R(z, T)x\| \leq \frac{C(T, x)}{|\text{Im}(z)|} [\alpha_n + \beta_n].$$

Proof. Let $\Gamma_n(T, x, z) = \widehat{\Gamma}_{m(n,T,x,z)}(T, x, z)$, where $\widehat{\Gamma}_k$ are the algorithms from the statement of Theorem 2.1 and $m(n, T, x, z)$ is a subsequence diverging to infinity as $n \to \infty$. Clearly statement (1) holds so we must show how to choose the sequence $m(n, T, x, z)$ such that (2) holds (and hence our algorithms will be adaptive). From (2.1), it is enough to show that $m = m(n, T, x, z)$ can be chosen such that

$$\beta_{f(m)} + \alpha_m\|\widehat{\Gamma}_m(T, x, z)\| + \|P_{f(m)}(T - zI)\widehat{\Gamma}_m(T, x, z) - P_{f(m)}x\| \lesssim \alpha_n + \beta_n.$$

The left-hand side can be approximated to arbitrary accuracy using finitely many arithmetic operations and comparisons. Hence by repeatedly computing approximations to within $\alpha_n + \beta_n$, we can choose the minimal m such that these approximate bounds are at most $2(\alpha_n + \beta_n)$. □

2.2. Stone's formula and Poisson kernels. Here we briefly discuss Stone's famous formula [32,113,137], which relates the convolution of spectral measures with Poisson kernels to the pointwise action of the projection-valued measures associated with an operator $T \in \Omega_{SA}$ as $\epsilon \downarrow 0$ (see Sect. 1.4). Stone's formula can also be generalised to unitary operators and a much larger class of normal operators (see Proposition 2.4). We include a short (and standard) proof of Proposition 2.3 for the benefit of the reader.

Proposition 2.3 (Stone's formula). *The following boundary limits hold:*

(i) Let $T \in \Omega_{SA}$. Then for any $-\infty \le a < b \le \infty$ and $x \in l^2(\mathbb{N})$,

$$\lim_{\epsilon \downarrow 0} \int_a^b K_H(u + i\epsilon; T, x)du = E_{(a,b)}^T x + \frac{1}{2} E_{\{a,b\}}^T x.$$

(ii) Let $T \in \Omega_U$. Then for any $0 \le a < b < 2\pi$ and $x \in l^2(\mathbb{N})$,

$$\lim_{\epsilon \downarrow 0} \int_a^b i \exp(i\psi) K_D((1 - \epsilon) \exp(i\psi); T, x)d\psi = E_{(a,b)_\mathbb{T}}^T x + \frac{1}{2} E_{\{\exp(ia),\exp(ib)\}}^T x,$$

where $(a, b)_\mathbb{T}$ denotes the image of (a, b) under the map $\theta \to \exp(i\theta)$.

Proof. To prove (i), we can apply Fubini's theorem to interchange the order of integration and arrive at

$$\int_a^b K_H(u + i\epsilon; T, x)du = \left[\int_{-\infty}^{\infty} \int_a^b P_H(u - \lambda, \epsilon)du \, dE^T(\lambda) \right] x$$

But

$$\int_a^b P_H(u - \lambda, \epsilon)du = \frac{1}{\pi} \left[\tan^{-1} \left(\frac{b - \lambda}{\epsilon} \right) - \tan^{-1} \left(\frac{a - \lambda}{\epsilon} \right) \right]$$

is bounded and converges pointwise as $\epsilon \downarrow 0$ to $\chi_{(a,b)}(\lambda) + \chi_{\{a,b\}}(\lambda)/2$, where χ_S denotes the indicator function of a set S. Part (i) now follows from the dominated convergence theorem.

To prove (ii), we apply Fubini's theorem again, now noting that

$$\int_a^b i \exp(i\psi) P_D((1 - \epsilon), \psi - \theta)d\psi = \frac{i \exp(i\theta)}{2\pi} \int_{a-\theta}^{b-\theta} \frac{(2\epsilon - \epsilon^2) \exp(i\psi)}{\epsilon^2 + 2(1 - \epsilon)(1 - \cos(\psi))}d\psi. \tag{2.2}$$

We can split the interval into small intervals of width ρ (where $0 < \rho < 1$) around each point where $\cos(\psi) = 1$, and a finite union of intervals on which $1 - \cos(\psi)$ is positive, bounded away from 0. On these later intervals, the limit vanishes as $\epsilon \downarrow 0$. Hence by periodicity and considering odd and even parts, we are left with considering

$$I_1(\rho, \epsilon) = \int_0^\rho \frac{(2\epsilon - \epsilon^2) \cos(\psi)}{\epsilon^2 + 2(1 - \epsilon)(1 - \cos(\psi))}d\psi,$$

$$I_2(\rho, \epsilon) = \int_0^\rho \frac{(2\epsilon - \epsilon^2) \sin(\psi)}{\epsilon^2 + 2(1 - \epsilon)(1 - \cos(\psi))}d\psi.$$
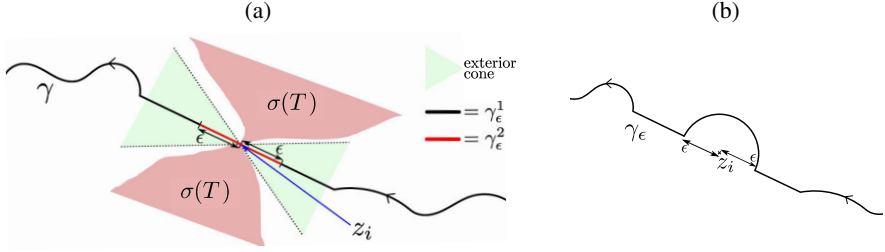
Fig. 2. Left: Exterior cone condition for Proposition 2.4. Right: Deformed contour γ_ϵ to compute $f_\epsilon(z_i)$

Explicit integration yields $I_2(\epsilon, \rho) = O(\epsilon \log(\epsilon))$ and hence the contribution vanishes in the limit. We also have

$$I_1(\rho, \epsilon) = \frac{(\epsilon^2 - 2\epsilon)\rho + 2(2 + \epsilon^2 - 2\epsilon)\tan^{-1}\left(\frac{(2-\epsilon)\tan\left(\frac{\rho}{2}\right)}{\epsilon}\right)}{2(1 - \epsilon)}.$$

This converges to π as $\epsilon \downarrow 0$. Considering the contributions of I_1 and I_2 in (2.2), we see that (2.2) converges pointwise as $\epsilon \downarrow 0$ to

$$i \exp(i\theta)\left\{\chi_{(a,b)}(\theta) + [\chi_{\{a\}}(\theta) + \chi_{\{b\}}(\theta)]/2\right\}.$$

Since the integral is also bounded, part (ii) now follows from the dominated convergence theorem and change of variables. $\square$

This type of construction can be generalised to $T \in \Omega_N$ whose spectrum lies on a regular enough curve. However, it is much more straightforward in the general case to use the analytic properties of the resolvent. The next proposition does this and also holds for operators whose spectrum does not necessarily lie along a curve.

Proposition 2.4 (Generalised Stone's formula). *Let $T \in \Omega_N$ and γ be a rectifiable positively oriented Jordan curve with the following properties. The spectrum $\sigma(T)$ intersects γ at finitely many points $z_1, \ldots, z_m$ and in a neighbourhood of each of the z_i, γ is formed of a line segment meeting $\sigma(T)$ only at z_i, at which point $\sigma(T)$ has a local exterior cone condition with respect to γ (see Fig. 2). Let $x \in l^2(\mathbb{N})$. Then we can define the Cauchy principal value integral of the resolvent $R(z, T)x$ along γ and have*

$$\frac{-1}{2\pi i}\text{PV}\int_\gamma R(z, T)x\,dz = E_{\sigma(T;\gamma)}^T x - \frac{1}{2}\left[\sum_{j=1}^m E_{\{z_j\}}^T x\right], \qquad (2.3)$$

where $\sigma(T; \gamma)$ is the closure of the intersection of $\sigma(T)$ with the interior of γ.

Proof. We will argue for the case $m = 1$, and the general case follows in exactly the same manner. Let $\epsilon > 0$ be small so that in a neighbourhood of the ϵ−ball around z_1, γ is given by a straight line. We then decompose γ into two disjoint parts

$$\gamma = \gamma_\epsilon^1 \cup \gamma_\epsilon^2,$$

where γ_ϵ^2 denotes the line segment of γ at most ϵ away from z_1 (as shown in Fig. 2). We set

$$F_\epsilon(x, T) = \int_{\gamma_\epsilon^1} R(z, T)x\,dz = \left[\int_{\sigma(T)} \int_{\gamma_\epsilon^1} \frac{1}{\lambda - z}\,dz\,dE^T(\lambda)\right]x.$$

We then consider the inner integral

$$f_\epsilon(\lambda) = \int_{\gamma_\epsilon^1} \frac{1}{\lambda - z}\,dz.$$

If λ is inside γ then $\lim_{\epsilon \downarrow 0} f_\epsilon(\lambda) = -2\pi i$ via Cauchy's residue theorem. Similarly, if λ is outside γ then $\lim_{\epsilon \downarrow 0} f_\epsilon(\lambda) = 0$. To calculate $f_\epsilon(z_1)$, consider the contour integral along γ_ϵ in Fig. 2. We see that

$$f_\epsilon(z_1) - i\pi = -2i\pi$$

and hence $f_\epsilon(z_1) = -i\pi$. We would like to apply the dominated convergence theorem. Clearly, away from z_1, f_ϵ is bounded as $\epsilon \downarrow 0$. Now let $0 < \delta < \epsilon$ then

$$f_\delta(\lambda) - f_\epsilon(\lambda) = \int_\delta^\epsilon \frac{1}{\frac{\lambda - z_1}{w} - s} + \frac{1}{\frac{\lambda - z_1}{w} + s}\,ds = \log\left(\frac{\epsilon + \frac{\lambda - z_1}{w}}{-\epsilon + \frac{\lambda - z_1}{w}}\right) - \log\left(\frac{\delta + \frac{\lambda - z_1}{w}}{-\delta + \frac{\lambda - z_1}{w}}\right)$$

for some $w \in \mathbb{T}$. Taking the pointwise limit $\delta \downarrow 0$, we see that $f_\epsilon(\lambda)$ is bounded for $\lambda \in \sigma(T)$ in a neighbourhood of z_1 as $\epsilon \downarrow 0$ if the same holds for

$$g_\epsilon(\lambda) = \log\left(\frac{\epsilon + \frac{\lambda - z_1}{w}}{-\epsilon + \frac{\lambda - z_1}{w}}\right).$$

By rotating and translating, we can assume that $w = 1$ and $z_1 = 0$ without loss of generality. Let $\lambda_1 = \mathrm{Re}(\lambda)$ and $\lambda_2 = \mathrm{Im}(\lambda)$. Using the cone condition gives $\alpha |\lambda_1| \leq |\lambda_2|$ for some $\alpha > 0$. Assume $\lambda_1 \neq 0$ then

$$\left|\frac{\epsilon + \lambda}{-\epsilon + \lambda}\right|^2 = \frac{(\epsilon + \lambda_1)^2 + \lambda_2^2}{(\epsilon - \lambda_1)^2 + \lambda_2^2} = 1 + \frac{4x}{(x - 1)^2 + y^2},$$

where $x = \epsilon/\lambda_1$ and $y = \lambda_2/\lambda_1$. Note that $y^2 \geq \alpha^2$ and without loss of generality we take $y \geq \alpha$. Define

$$h(x, y) = \frac{4x}{(x - 1)^2 + y^2}$$

Note that $h(x, y) \to 0$ as $|x|^2 + |y|^2 \to \infty$. We must show that $h(x, y)$ is bounded above -1 for $y \geq \alpha$. It is enough to consider points where $\partial h/\partial x = 0$ which occur when $x_\pm = \pm\sqrt{1 + y^2}$. We have

$$h(x_\pm, y) = \frac{\pm 2}{\sqrt{1 + y^2} \mp 1} \geq \frac{-2}{\sqrt{1 + \alpha^2} + 1} > -1,$$

and hence we have proved the required boundedness. We then define

$$\text{PV} \int_\gamma R(z, T)x\, dz = \lim_{\epsilon \downarrow 0} F_\epsilon(x, T).$$

The relation (2.3) now follows from the dominated convergence theorem. □

3. Computation of Measures

For the sake of brevity, the analysis in the rest of this paper will consider the self-adjoint case $T \in \Omega_{\text{SA}}$, which is the case most encountered in applications. However, the algorithms we build are based on Theorem 2.1 (and Corollary 2.2) and the link with Poisson kernels/Cauchy transforms. Given the relation (1.13) and Proposition 2.4, many of the results can be straightforwardly extended to the unitary case and more general cases where conditions similar to that of Proposition 2.4 hold. We consider examples of unitary operators in Sect. 6.3.

3.1. Full spectral measure. We start by considering the computation of $E_U^T x$, where $U \subset \mathbb{R}$ is a non-trivial open set. In other words, U is not the whole of $\mathbb{R}$ or the empty set. The collection of these subsets will be denoted by $\mathcal{U}$. To be precise, we assume that we have access to a finite or countable collection $a_m(U), b_m(U) \in \mathbb{R} \cup \{\pm\infty\}$ such that U can be written as a disjoint union

$$U = \bigcup_m (a_m(U), b_m(U)). \tag{3.1}$$

With an abuse of notation, we add this information as evaluation functions to Λ_i (defined in (1.18)).

Theorem 3.1 (Computation of measures on open sets). *Given the set-up in Sects. 1.5, 1.6 and the previous paragraph, consider the map*

$$\Xi_{\text{meas}} : \Omega_{f,\alpha,\beta} \times \mathcal{U} \to l^2(\mathbb{N})$$
$$(T, x, U) \to E_U^T x.$$

Then $\{\Xi_{\text{meas}}, \Omega_{f,\alpha,\beta} \times \mathcal{U}, \Lambda_1\} \in \Delta_2^A$. In other words, we can construct a convergent sequence of arithmetic algorithms for the problem.

Remark 4. Essentially, this theorem tells us that if we can compute the action of the resolvent operator with asymptotic error control near the real axis, then we can compute the spectral measures of open sets in one limit. In the unitary case, this can easily be extended to relatively open sets of $\mathbb{T}$ if we can evaluate the resolvent near the unit circle. For any $U \in \mathcal{U}$, the approximation of $E_U^T x$ has finite support, and hence we can take inner products to compute $\mu_{x,y}^T(U)$.

Remark 5. One may wonder whether it is possible to upgrade the convergence of the algorithm in Theorem 3.1 from Δ_2 to Δ_1. In other words, whether it is possible to compute the measure with error control. However, this is difficult because the measure may be singular. Theorem 5.2 shows this is impossible even for singleton sets and discrete Schrödinger operators acting on $l^2(\mathbb{N})$.

Proof of Theorem 3.1. Let $T \in \Omega_{SA}$ and $z_1, z_2 \in \mathbb{C} \backslash \mathbb{R}$. By the resolvent identity and self-adjointness of T,

$$\|R(z_1, T) - R(z_1, T)\| \leq |\text{Im}(z_1)|^{-1} |\text{Im}(z_2)|^{-1} |z_1 - z_2|.$$

Hence, for $z = u+i\epsilon$ with $\epsilon > 0$, the vector-valued function $K_H(u+i\epsilon; T, x)$ (considered with argument u) is Lipschitz continuous with Lipschitz constant bounded by $\epsilon^{-2}\|x\|/\pi$. Now consider the class $\Omega_{f,\alpha,\beta} \times \mathcal{U}$ and let $(T, x, U) \in \Omega_{f,\alpha,\beta} \times \mathcal{U}$. From Corollary 2.2, we can construct a sequence of arithmetic algorithms, $\widehat{\Gamma}_n$, such that

$$\|\widehat{\Gamma}_n(T, u, z) - K_H(u + i\epsilon; T, x)\| \leq \frac{C(T, x)}{\epsilon} (\alpha_n + \beta_n)$$

for all $(T, x) \in \Omega_{f,\alpha,\beta}$. It follows from standard quadrature rules and taking subsequences if necessary (using that $\{\alpha_n\}$ and $\{\beta_n\}$ are null), that for $-\infty < a < b < \infty$, the integral

$$\int_a^b K_H \left(u + \frac{i}{n}; T, x \right) du \tag{3.2}$$

can be approximated to an accuracy $\widehat{C}(T, x)/n$ using finitely many arithmetic operations and comparisons and the relevant set of evaluation functions Λ_1 (the constant C now becomes $\widehat{C}$ due to not knowing the exact value of $\|x\|$).

Recall that we assumed the disjoint union

$$U = \bigcup_m (a_m, b_m)$$

where $a_m, b_m \in \mathbb{R} \cup \{\pm\infty\}$ and the union is at most countable. Without loss of generality, we assume that the union is over $m \in \mathbb{N}$. We then let $a_{m,n}, b_{m,n} \in \mathbb{Q}$ be such that $a_{m,n} \downarrow a_m$ and $b_{m,n} \uparrow b_m$ as $n \to \infty$ with $a_{m,n} < b_{m,n}$ and hence $(a_{m,n}, b_{m,n}) \subset (a_m, b_m)$. Let

$$U_n = \bigcup_{m=1}^n (a_{m,n}, b_{m,n}),$$

then the proof of Stone's formula in Proposition 2.3 (essentially an application of the dominated convergence theorem) can be easily adapted to show that

$$\lim_{n\to\infty} \int_{U_n} K_H \left(u + \frac{i}{n}; T, x \right) du = E_U^T x.$$

Note that we do not have to worry about contributions from endpoints of the intervals (a_m, b_m) since we approximate strictly from within the open set U. To finish the proof, we simply let $\Gamma_n(T, x, U)$ be an approximation of the integral

$$\int_{U_n} K_H \left(u + \frac{i}{n}; T, x \right) du$$

with accuracy $\widehat{C}(T, x)/n$. By the above remarks, such an approximation can be computed using finitely many arithmetic operations and comparisons from the relevant set of evaluation functions Λ_1. $\square$

This theorem can clearly be extended to cover the more general case of Proposition 2.4 if γ is regular enough to allow approximation of

$$\mathrm{PV} \int_\gamma R(z, T)x\, dz,$$

given the ability to compute $R(z, T)x$ with asymptotic error control. Note that when it comes to numerically computing the integrals in Propositions 2.3 and 2.4, it is advantageous to deform the contour so that most of the contour lies far from the spectrum so that the resolvent has a smaller Lipschitz constant. The proof can also be adapted to compute $E_I x$, where $I = [a, b]$ is a closed interval, by considering intervals shrinking to $[a, b]$ (a, b finite). A special case of this is the computation of the spectral measure of singleton sets. However, for these it much easier to directly use the formulae

$$E^T_{\{u\}}x = \lim_{\epsilon \downarrow 0} \epsilon \pi K_H(u + i\epsilon; T, x),$$

$$E^T_{\{\exp(i\theta)\}}x = \lim_{\epsilon \downarrow 0} \epsilon \pi i \exp(i\theta) K_D((1 - \epsilon) \exp(i\theta); T, x),$$

for $T \in \Omega_{\mathrm{SA}}$ and $T \in \Omega_{\mathrm{U}}$ respectively.

3.2. Measure decompositions and projections.

Recall from Sect. 1.5 that $P^T_{\mathcal{I}}$ denotes the orthogonal projection onto the space $\mathcal{H}^T_{\mathcal{I}}$, where $\mathcal{I}$ denotes a generic type (ac, sc, pp, c or s). We have included the continuous and singular parts denoted by c or s which correspond to $\mathcal{H}_{\mathrm{ac}} \oplus \mathcal{H}_{\mathrm{sc}}$ and $\mathcal{H}_{\mathrm{sc}} \oplus \mathcal{H}_{\mathrm{pp}}$ respectively. These are often encountered in mathematical physics. As in Sect. 3.1, we assume the decomposition in (3.1) and add the $\{a_m, b_m\}$ as evaluation functions to Λ_i (defined in (1.18)). In this section, we prove the following theorem.

Theorem 3.2. *Given the set-up in Sects. 1.5, 1.6 and 3.1, consider the map*

$$\Xi_{\mathcal{I}} : \Omega_{f,\alpha,\beta} \times V_\beta \times \mathcal{U} \to \mathbb{C}$$
$$(T, x, y, U) \to \langle P^T_{\mathcal{I}} E^T_U x, y \rangle = \mu^T_{x,y,\mathcal{I}}(U),$$

for $\mathcal{I} =$ ac, sc, pp, c or s. Then for $i = 1, 2$

$$\Delta^G_2 \not\ni \{\Xi_{\mathcal{I}}, \Omega_{f,\alpha,\beta} \times V_\beta \times \mathcal{U}, \Lambda_i\} \in \Delta^A_3.$$

To prove this theorem, it is enough, by the polarisation identity, to consider $x = y$ (note that all the projections commute). We will split the proof into two parts: the Δ^A_3 inclusion, for which it is enough to consider Λ_1, and the Δ^G_2 exclusion, for which it is enough to consider Λ_2.

3.2.1. Proof of inclusion in Theorem 3.2

Proof of inclusion in Theorem 3.2. Since $P^T_{\mathrm{pp}} = I - P^T_{\mathrm{c}}$, $P^T_{\mathrm{ac}} = I - P^T_{\mathrm{s}}$ and $P^T_{\mathrm{sc}} = P^T_{\mathrm{s}} - P^T_{\mathrm{pp}}$, it is enough, by Theorem 3.1 and Remark 4, to consider only $\mathcal{I} =$ c and $\mathcal{I} =$ s.

Step 1: We first deal with $\mathcal{I} =$ c, where we shall use a similar argument to the proof of Theorem 4.1 (which is more general than what we need). We recall the RAGE theorem

[2,53,119] as follows. Let Q_n denote the orthogonal projection onto vectors in $l^2(\mathbb{N})$ with support outside the subset $\{1, \ldots, n\} \subset \mathbb{N}$. Then for any $x \in l^2(\mathbb{N})$,

$$\langle P_c^T E_U^T x, x \rangle = \|P_c^T E_U^T x\|^2 = \lim_{n \to \infty} \lim_{t \to \infty} \frac{1}{t} \int_0^t \left\| Q_n e^{-iTs} E_U^T x \right\|^2 ds$$

$$= \lim_{n \to \infty} \lim_{t \to \infty} \frac{1}{t} \int_0^t \left\| Q_n e^{-iTs} \chi_U(T) x \right\|^2 ds. \tag{3.3}$$

The proof of Theorem 4.1 is easily adapted to show that there exists arithmetic algorithms $\widetilde{\Gamma}_{n,m}$ using Λ_1 such that

$$\left\| Q_n e^{-iTs} \chi_U(T) x - \widetilde{\Gamma}_{n,m}(T, x, U, s) \right\| \leq \frac{C(T, x, U)}{m}$$

for all $(T, x, U, s) \in \Omega_{f,\alpha,\beta} \times \mathcal{U} \times \mathbb{R}$. Note that this bound can be made independent of s (as we have written above) by sufficiently approximating the function $\lambda \to \exp(-i\lambda s)\chi_U(\lambda)$ (it has known total variation for a given s and uniform bound). We now define

$$\Gamma_{n,m}(T, x, U) = \frac{1}{m^2} \sum_{j=1}^{m^2} \|\widetilde{\Gamma}_{m,n}(T, x, U, j/m)\|^2.$$

Using the fact that for $a, b \in l^2(\mathbb{N})$,

$$|\langle a, a \rangle - \langle b, b \rangle| \leq \|a - b\| (2\|a\| + \|a - b\|), \tag{3.4}$$

it follows that

$$\left| \|Q_n e^{-iTs} \chi_U(T) x\|^2 - \|\widetilde{\Gamma}_{n,m}(T, x, U, s)\|^2 \right| \leq \frac{C(T, x, U)}{m} \left(2\|x\| + \frac{C(T, x, U)}{m} \right).$$

Hence

$$\left| \Gamma_{n,m}(T, x, U) - \frac{1}{m} \int_0^m \left\| Q_n e^{-iTs} \chi_U(T) x \right\|^2 ds \right|$$

$$\leq \frac{1}{m^2} \sum_{j=1}^{m^2} \frac{C(T, x, U)}{m} \left(2\|x\| + \frac{C(T, x, U)}{m} \right)$$

$$+ \frac{1}{m^2} \sum_{j=1}^{m^2} \left| g_n(j/m) - m \int_{\frac{j-1}{m}}^{\frac{j}{m}} g_n(s) ds \right|,$$

where $g_n(s) = \|Q_n e^{-iTs} \chi_U(T) x\|^2$. Clearly the first term converges to 0 as $m \to \infty$, so we only need to consider the second. Using (3.4), it follows that for any $\epsilon > 0$

$$|g_n(s) - g_n(s + \epsilon)| \leq 4\|Q_n e^{-iTs}(e^{-iT\epsilon} - I)\chi_U(T) x\| \|x\|$$

$$\leq 4\|x\| \|(e^{-iT\epsilon} - I)\chi_U(T) x\|.$$

But $e^{-iT\epsilon} - I$ converges strongly to 0 as $\epsilon \downarrow 0$ and hence the quantity

$$\left| g_n(j/m) - m \int_{\frac{j-1}{m}}^{\frac{j}{m}} g_n(s) ds \right| \to 0$$

uniformly in j as $m \to \infty$. It follows that

$$\lim_{m \to \infty} \Gamma_{n,m}(T, x, U) = \lim_{t \to \infty} \frac{1}{t} \int_0^t \left\| Q_n e^{-iTs} E_U^T x \right\|^2 ds$$

and hence

$$\lim_{n \to \infty} \lim_{m \to \infty} \Gamma_{n,m}(T, x, U) = \langle P_c^T E_U^T x, x \rangle.$$

Step 2: Next we deal with the case $\mathcal{I} = $ s. Note that for $z \in \mathbb{C} \backslash \mathbb{R}$, $\langle R(z, T)x, x \rangle$ is simply the Stieltjes transform (also called the Borel transform) of the positive measure $\mu_{x,x}^T$

$$\langle R(z, T)x, x \rangle = \int_{\mathbb{R}} \frac{1}{\lambda - z} d\mu_{x,x}^T(\lambda).$$

The Hilbert transform of $\mu_{x,x}^T$ is given by the limit

$$H_{\mu_{x,x}^T}(t) = \frac{1}{\pi} \lim_{\epsilon \downarrow 0} \mathrm{Re} \left(\langle R(t + i\epsilon, T)x, x \rangle \right),$$

with the limit existing (Lebesgue) almost everywhere. This object was studied in [108, 109], where we shall use the result (since the measure is positive) that for any bounded continuous function f,[6]

$$\lim_{\theta \to \infty} \frac{\pi \theta}{2} \int_{\mathbb{R}} f(t) \chi_{\{w : |H_{\mu_{x,x}^T}(w)| \geq \theta\}}(t) dt = \int_{\mathbb{R}} f(t) d\mu_{x,x,s}^T(t). \tag{3.5}$$

Now let $(T, x, U) \in \Omega_{f,\alpha,\beta} \times \mathcal{U}$ with

$$U = \bigcup_m (a_m, b_m),$$

where $a_m, b_m \in \mathbb{R} \cup \{\pm \infty\}$ and the disjoint union is at most countable as in (3.1). Without loss of generality, we assume that the union is over $m \in \mathbb{N}$. Due to the possibility of point spectra at the endpoints a_m, b_m, we cannot simply replace f by χ_U in the above limit (3.5). However, this can be overcome in the following manner.

Let ∂U denote the boundary of U defined by $\overline{U} \backslash U$ and let ν denote the measure $\mu_{x,x}^T |_{\partial U}$. Let $\{f_l\}_{l \in \mathbb{N}}$ denote a pointwise increasing sequence of continuous functions, converging everywhere up to χ_U, such that the support of each f_l is contained in

$$[-l, l] \cap \left(\bigcup_{m=1}^{l} \left(a_m + 1/\sqrt{l}, b_m - 1/\sqrt{l} \right) \right).$$

Such a sequence exists (and can easily be explicitly constructed) precisely because U is open. We first claim that

$$\lim_{l \to \infty} \frac{\pi l}{2} \int_{\mathbb{R}} f_l(t) \chi_{\{w : |H_{\mu_{x,x}^T}(w)| \geq l\}}(t) dt = \mu_{x,x,s}^T(U). \tag{3.6}$$

---

[6] Note that this is stronger than weak* convergence which in this case means restricting to continuous functions vanishing at infinity. That the result holds for arbitrary bounded continuous functions is due to the tightness condition that the result holds for the function identically equal to 1.

To see this note that for any $k \in \mathbb{N}$, the following inequalities hold

$$\liminf_{l \to \infty} \frac{\pi l}{2} \int_{\mathbb{R}} f_l(t) \chi_{\{w:|H_{\mu_{x,x}^T}(w)| \geq l\}}(t) dt \geq \liminf_{l \to \infty} \frac{\pi l}{2} \int_{\mathbb{R}} f_k(t) \chi_{\{w:|H_{\mu_{x,x}^T}(w)| \geq l\}}(t) dt$$

$$= \int_{\mathbb{R}} f_k(t) d\mu_{x,x,\mathrm{s}}^T(t),$$

where the last equality is due to (3.5). Taking $k \to \infty$ yields

$$\liminf_{l \to \infty} \frac{\pi l}{2} \int_{\mathbb{R}} f_l(t) \chi_{\{w:|H_{\mu_{x,x}^T}(w)| \geq l\}}(t) dt \geq \mu_{x,x,\mathrm{s}}^T(U), \qquad (3.7)$$

so we are left with proving a similar bound for the limit supremum. Note that any point in the support of f_l is of distance at least $1/\sqrt{l}$ from ∂U. It follows that there exists a constant C independent of t such that for any $t \in \mathrm{supp}(f_l)$,

$$|H_\nu(t)| \leq C\sqrt{l}$$

Now let $\epsilon \in (0,1)$. Then, for large l, $l - C\sqrt{l} \geq (1-\epsilon)l$ and hence

$$\mathrm{supp}(f_l) \cap \{w : |H_{\mu_{x,x}^T}(w)| \geq l\} \subset \mathrm{supp}(f_l) \cap \{w : |H_{\mu_{x,x}^T - \nu}(w)| \geq (1-\epsilon)l\}. \quad (3.8)$$

Now let f be any bounded continuous function such that $f \geq \chi_U$. Then using (3.8),

$$\limsup_{l \to \infty} \frac{\pi l}{2} \int_{\mathbb{R}} f_l(t) \chi_{\{w:|H_{\mu_{x,x}^T}(w)| \geq l\}}(t) dt$$

$$\leq \limsup_{l \to \infty} \frac{1}{1-\epsilon} \frac{\pi(1-\epsilon)l}{2} \int_{\mathbb{R}} f_l(t) \chi_{\{w:|H_{\mu_{x,x}^T - \nu}(w)| \geq (1-\epsilon)l\}}(t) dt$$

$$\leq \limsup_{l \to \infty} \frac{1}{1-\epsilon} \frac{\pi(1-\epsilon)l}{2} \int_{\mathbb{R}} f(t) \chi_{\{w:|H_{\mu_{x,x}^T - \nu}(w)| \geq (1-\epsilon)l\}}(t) dt$$

$$= \frac{1}{1-\epsilon} \int_{\mathbb{R}} f(t) d([\mu_{x,x}^T - \nu]_{\mathrm{s}})(t).$$

Now we let $f \downarrow \chi_{\overline{U}}$, with pointwise convergence everywhere. This is possible since the complement of $\overline{U}$ is open. By the dominated convergence theorem, and since ϵ was arbitrary, this yields

$$\limsup_{l \to \infty} \frac{\pi l}{2} \int_{\mathbb{R}} f_l(t) \chi_{\{w:|H_{\mu_{x,x}^T}(w)| \geq l\}}(t) dt \leq [\mu_{x,x}^T - \nu]_{\mathrm{s}}(\overline{U}) = \mu_{x,x,\mathrm{s}}^T(U),$$

where the last equality follows from the definition of ν. The claim (3.6) now follows.

Let χ_n be a sequence of non-negative continuous piecewise affine functions on $\mathbb{R}$, bounded by 1 and such that $\chi_n(t) = 0$ if $t \leq n-1$ and $\chi_n(t) = 1$ if $t \geq n+1$. Consider the integrals

$$I(n,m) = \frac{\pi n}{2} \int_{\mathbb{R}} f_n(t) \chi_n(|F_m(t)|) dt,$$

where $F_m(t)$ is an approximation of

$$\frac{1}{\pi} \mathrm{Re}\left(\left\langle R\left(t + \frac{i}{m}, T\right) x, x\right\rangle\right)$$

to pointwise accuracy $O(m^{-1})$ over $t \in [-n, n]$. Note that a suitable piecewise affine function f_n can be constructed using Λ_1, as can suitable χ_n, and a suitable approximation function F_m can be pointwise evaluated using Λ_1 (again by Corollary 2.2). To see this, recall the definition of Λ_1 in (1.18) and that we added $\{a_m, b_m\}$ from (3.1) to Λ_1. To define f_n, we can define the function by suitable piecewise affine functions on each interval $[-n, n] \cap \left(a_m + 1/\sqrt{n}, b_m - 1/\sqrt{n}\right)$. It follows that there exists arithmetic algorithms $\Gamma_{n,m}(T, x, U)$ using Λ_1 such that

$$\left| I(n, m) - \Gamma_{n,m}(T, x, U) \right| \leq \frac{C(T, x, U)}{m}.$$

The dominated convergence theorem implies that

$$\lim_{m \to \infty} \Gamma_{n,m}(T, x, U) = \lim_{m \to \infty} I(n, m) = \frac{\pi n}{2} \int_{\mathbb{R}} f_n(t) \chi_n(|H_{\mu_{x,x}^T}(t)|) dt.$$

Note that continuity of χ_n is needed to gain convergence almost everywhere and prevent possible oscillations about the level set $\{H_{\mu_{x,x}^T}(t) = n\}$. We also have

$$\chi_{\{w : |H_{\mu_{x,x}^T}(w)| \geq n+1\}}(t) \leq \chi_n(|H_{\mu_{x,x}^T}(t)|) \leq \chi_{\{w : |H_{\mu_{x,x}^T}(w)| \geq n-1\}}(t)$$

The same arguments used to prove (3.6), therefore show that

$$\lim_{n \to \infty} \frac{\pi n}{2} \int_{\mathbb{R}} f_n(t) \chi_n(|H_{\mu_{x,x}^T}(t)|) dt = \mu_{x,x,s}^T(U).$$

Hence,

$$\lim_{n \to \infty} \lim_{m \to \infty} \Gamma_{n,m}(T, x, U) = \mu_{x,x,s}^T(U),$$

completing the proof of inclusion in Theorem 3.2. ☐

3.2.2. Proof of exclusion in Theorem 3.2 To prove the exclusion, we need two results which will also be used in Sect. 5. First, we consider a result connected to Anderson localisation (Theorem 3.3) and, second, we consider a result concerning sparse potentials of discrete Schrödinger operators (Theorem 3.4). The free Hamiltonian H_0 acts on $l^2(\mathbb{N})$ via the tridiagonal matrix representation

$$H_0 = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & \ddots \\ & & \ddots & \ddots \end{pmatrix}.$$

We define a Schrödinger operator acting on $l^2(\mathbb{N})$ to be an operator of the form

$$H_v = H_0 + v,$$

where v is a bounded (real-valued) multiplication operator with matrix $\mathrm{diag}(v(1), v(2), \ldots)$.

Since Anderson's introduction of his famous model 60 years ago [3], there has been a considerable amount of work by both physicists and mathematicians aiming to understand the suppression of electron transport due to disorder (Anderson localisation).

A full discussion of Anderson localisation is beyond the scope of this paper, and we refer the reader to [29, 42, 90] for broader surveys. When considering Anderson localisation, we will assume that $v = v_\omega = \{v(1), v(2), \ldots\}$ is a collection of independent identically distributed random variables. Following [67], we assume that the single-site probability distribution has a density $\rho \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ with $\|\rho\|_1 = 1$ (with respect to the standard Lebesgue measure). For such a potential, a measure of disorder is given by the quantity $\|\rho\|_\infty^{-1}$. The first result we need is the following theorem which follows straightforwardly from the technique of [67], and hence we have not provided a proof which would be almost verbatim to [67].

Theorem 3.3 (Anderson Localisation for Perturbed Operator [67]). *There exists a constant $C > 0$ such that if $\|\rho\|_\infty \leq C$ and ρ has compact support, then the operator $H_{v_\omega} + A$ has only pure point spectrum with probability 1 for any fixed self-adjoint operator A of the form*

$$A = \sum_{j=1}^{M} \alpha_j \left| x_{m_j} \right\rangle \left\langle x_{n_j} \right|. \tag{3.9}$$

In other words, the operator A's matrix with respect to the canonical basis has only finitely many non-zeros.

The second result we need is the following from [92].

Theorem 3.4 (Krutikov and Remling [92]). *Consider discrete Schrödinger operators acting on $l^2(\mathbb{N})$. Let v be a (real-valued and bounded) potential of the following form:*

$$v(n) = \sum_{j=1}^{\infty} g_j \delta_{n, m_j}, \quad m_{j-1} / m_j \to 0.$$

Then $[0, 4] \subset \sigma_{\mathrm{ess}}(H_v)$ and the following dichotomy holds:

(a) If $\sum_{j \in \mathbb{N}} g_j^2 < \infty$ then H_v is purely absolutely continuous on $(0, 4)$.
(b) If $\sum_{j \in \mathbb{N}} g_j^2 = \infty$ then H_v is purely singular continuous on $(0, 4)$.

Proof of exclusion in Theorem 3.2. Since $P_{\mathrm{pp}}^T = I - P_{\mathrm{c}}^T$, $P_{\mathrm{ac}}^T = I - P_{\mathrm{s}}^T$ and $P_{\mathrm{sc}}^T = P_{\mathrm{s}}^T - P_{\mathrm{pp}}^T$, it is enough, by Theorem 3.1 and Remark 4, to consider $\mathcal{I} = \mathrm{pp}$, ac and sc. We restrict the proof to considering bounded Schrödinger operators H_v acting on $l^2(\mathbb{N})$, which are clearly a subclass of $\Omega_{f,0}$ for $f(n) = n + 1$. Note that since the evaluation functions in Λ_2 can be recovered from those in Λ_1 in this special case, we can assume that we are dealing with Λ_1. We also set $x = e_1$, with the crucial properties that this vector is cyclic and hence $\mu_{e_1, e_1}^{H_v}$ has the same support as $\sigma(H_v)$, and that $x \in V_0$. Throughout, we also take $U = (0, 4)$.

Step 1: We begin with P_{pp}^T. Suppose for a contradiction that there does exist a sequence of general algorithms Γ_n such that

$$\lim_{n \to \infty} \Gamma_n(H_v) = \langle P_{\mathrm{pp}}^{H_v} E_{(0,4)}^{H_v} e_1, e_1 \rangle.$$

We take a general algorithm, denoted $\widehat{\Gamma}_n$, from Theorem 3.1 which has

$$\lim_{n \to \infty} \widehat{\Gamma}_n(H_v) = \mu_{e_1, e_1}^{H_v}((0, 4)).$$

Since e_1 is cyclic, this limit is non-zero if $(0, 4) \cap \sigma(H_v) \neq \emptyset$. We therefore define

$$\widetilde{\Gamma}_n(H_v) = \begin{cases} 0 & \text{if } \widehat{\Gamma}_n(H_v) = 0 \\ \frac{\Gamma_n(H_v)}{\widehat{\Gamma}_n(H_v)} & \text{otherwise.} \end{cases}$$

We will use Theorem 3.3 and the following well-known facts:

1. If for any $l \in \mathbb{N}$ there exists m_l such that $v(m_l + 1) = v(m_l + 2) = \cdots = v(m_l + l) = 0$, then $(0, 4) \subset \sigma(H_v)$.
2. If there exists $N \in \mathbb{N}$ such that $v(n)$ is 0 for $n \geq N$, then $\sigma_{\mathrm{pp}}(H_v) \cap (0, 4) = \emptyset$ [114], but $[0, 4] \subset \sigma(H_v)$ (the potential acts as a compact perturbation so the essential spectrum is $[0, 4]$).
3. If we are in the setting of Theorem 3.3, then the spectrum of $H_{v_\omega} + A$ is pure point almost surely. Moreover, if $\rho = \chi_{[-c,c]}/(2c)$ for some constant c, then $[-c, 4+c] \subset \sigma_{\mathrm{pp}}(H_{v_\omega} + A)$ almost surely.

The strategy will be to construct a potential v such that $(0, 4) \subset \sigma(H_v)$, yet $\widetilde{\Gamma}_n(H_v)$ does not converge. This is a contradiction since by our assumptions, for such a v we must have

$$\widetilde{\Gamma}_n(H_v) \rightarrow \frac{\langle P_{\mathrm{pp}}^{H_v} E_{(0,4)}^{H_v} e_1, e_1 \rangle}{\mu_{e_1,e_1}^{H_v}((0, 4))}.$$

To do this, choose $\rho = \chi_{[-c,c]}/(2c)$ for some constant c such that the conditions of Theorem 3.3 hold and define the potential v inductively as follows.

Let v_1 be a potential of the form v_ω (with the density ρ) such that $\sigma(H_{v_1})$ is pure point. Such a v_1 exists by Theorem 3.3 and we have $\langle P_{\mathrm{pp}}^{H_{v_1}} E_{(0,4)}^{H_{v_1}} e_1, e_1 \rangle = \mu_{e_1,e_1}^{H_{v_1}}((0, 4))$. Hence for large enough n it must hold that $\widetilde{\Gamma}_n(H_{v_1}) > 3/4$. Fix $n = n_1$ such that this holds. Then $\Gamma_{n_1}(H_{v_1})$ only depends on $\{v_1(j) : j \leq N_1\}$ for some integer N_1 by (i) of Definition A.2. Define the potential v_2 by $v_2(j) = v_1(j)$ for all $j \leq N_1$ and $v_2(j) = 0$ otherwise. Then by fact (2) above, $\langle P_{\mathrm{pp}}^{H_{v_2}} E_{(0,4)}^{H_{v_2}} e_1, e_1 \rangle = 0$ but $\mu_{e_1,e_1}^{H_{v_2}}((0, 4)) \neq 0$, and hence $\widetilde{\Gamma}_n(H_{v_2}) < 1/4$ for large n, say for $n = n_2 > n_1$. But then $\Gamma_{n_2}(H_{v_2})$ only depends on $\{v_2(j) : j \leq N_2\}$ for some integer N_2.

We repeat this process inductively switching between potentials which induce $\widetilde{\Gamma}_{n_k}(H_{v_k}) < 1/4$ for k even and potentials which induce $\widetilde{\Gamma}_{n_k}(H_{v_k}) > 3/4$ for k odd. Explicitly, if k is even then define a potential v_{k+1} by $v_{k+1}(j) = v_k(j)$ for all $j \leq N_k$ and $v_{k+1}(j) = v_\omega(j)$ (with the density ρ) otherwise such that the spectrum of H_{v_k} is pure point. Such a ω exists from Theorem 3.3 applied with the perturbation A to match the potential for $j \leq N_k$. If k is odd then we define v_{k+1} by $v_{k+1}(j) = v_k(j)$ for all $j \leq N_k$ and $v_{k+1}(j) = 0$ otherwise. We can then choose n_{k+1} such that the above inequalities hold and N_{k+1} such that $\Gamma_{n_{k+1}}(H_{v_{k+1}})$ only depends on $\{v_{k+1}(j) : j \leq N_{k+1}\}$. We also ensure that $N_{k+1} \geq N_k + k$.

Finally set $v(j) = v_k(j)$ for $j \leq N_k$. It is clear from (iii) of Definition A.2, that $\widetilde{\Gamma}_{n_k}(H_v) = \widetilde{\Gamma}_{n_k}(H_{v_k})$ and this implies that $\widetilde{\Gamma}_{n_k}(H_v)$ cannot converge. However, since $N_{k+1} \geq N_k + k$, for any k odd we have $v(N_k + 1) = v(N_k + 2) = \cdots = v(N_k + k) = 0$. Fact (1) implies that $(0, 4) \subset \sigma(H_v)$, hence $\mu_{e_1,e_1}^{H_v}((0, 4)) \neq 0$ and therefore $\widetilde{\Gamma}_n(H_v)$ converges. This provides the required contradiction.

Step 2: Next we deal with $\mathcal{I} = \mathrm{ac}$. To prove that one limit will not suffice, our strategy will be to reduce a certain decision problem to the computation of Ξ_{ac}. Let $(\mathcal{M}', d')$

be the discrete space $\{0, 1\}$, let Ω' denote the collection of all infinite sequence $\{a_j\}_{j \in \mathbb{N}}$ with entries $a_j \in \{0, 1\}$ and consider the problem function

$$\Xi'(\{a_j\}) : \text{ 'Does } \{a_j\} \text{ have infinitely many non-zero entries?'},$$

which maps to $(\mathcal{M}', d')$. In Appendix A, it is shown that $\text{SCI}(\Xi', \Omega')_G = 2$ (where the evaluation functions consist of component-wise evaluation of the array $\{a_j\}$). Suppose for a contradiction that Γ_n is a height one tower of general algorithms such that

$$\lim_{n \to \infty} \Gamma_n(H_v) = \langle P_{\text{ac}}^{H_v} E_{(0,4)}^{H_v} e_1, e_1 \rangle.$$

We will gain a contradiction by using the supposed tower to solve $\{\Xi', \Omega'\}$.

Given $\{a_j\} \in \Omega'$, consider the operator H_v, where the potential is of the following form:

$$v(m) = \sum_{k=1}^{\infty} a_k \delta_{m,k!}. \tag{3.10}$$

Then by Theorem 3.4, $\langle P_{\text{ac}}^{H_v} E_{(0,4)}^{H_v} e_1, e_1 \rangle = \mu_{e_1,e_1}^{H_v}((0,4))$ if $\sum_k a_k < \infty$ (that is, if $\Xi'(\{a_j\}) = 0$) and $\langle P_{\text{ac}}^{H_v} E_{(0,4)}^{H_v} e_1, e_1 \rangle = 0$ otherwise. Note that in either case we have $\mu_{e_1,e_1}^{H_v}((0,4)) \neq 0$. We follow Step 1 and take a general algorithm, denoted $\widehat{\Gamma}_n$, from Theorem 3.1 which has

$$\lim_{n \to \infty} \widehat{\Gamma}_n(H_v) = \mu_{e_1,e_1}^{H_v}((0,4)).$$

Since e_1 is cyclic, this limit is non-zero for H_v, where v is of the form (3.10). We therefore define

$$\widetilde{\Gamma}_n(H_v) = \begin{cases} 0 & \text{if } \widehat{\Gamma}_n(H_v) = 0 \\ \frac{\Gamma_n(H_v)}{\widehat{\Gamma}_n(H_v)} & \text{otherwise.} \end{cases}$$

It follows that

$$\lim_{n \to \infty} \widetilde{\Gamma}_n(H_v) = \begin{cases} 1 & \text{if } \Xi'(\{a_j\}) = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Given N, we can evaluate any matrix value of H using only finitely many evaluations of $\{a_j\}$ and hence the evaluation functions Λ_1 can be computed using component-wise evaluations of the sequence $\{a_j\}$. We now set

$$\overline{\Gamma}_n(\{a_j\}) = \begin{cases} 0 & \text{if } \Gamma_n(H_v) > \frac{1}{2} \\ 1 & \text{otherwise.} \end{cases}$$

The above comments show that each of these is a general algorithm and it is clear that it converges to $\Xi'(\{a_j\})$ as $n \to \infty$, the required contradiction.

Step 3: Finally, we must deal with $\mathcal{I} = \text{sc}$. The argument is the same as Step 2, replacing $\langle P_{\text{ac}}^{H_v} E_{(0,4)}^{H_v} e_1, e_1 \rangle$ with $\langle P_{\text{sc}}^{H_v} E_{(0,4)}^{H_v} e_1, e_1 \rangle$ and the resulting $\widetilde{\Gamma}_n(H_v)$ with $1 - \widetilde{\Gamma}_n(H_v)$. □

4. Two Important Applications

Two important applications of our techniques are the computation of the functional calculus and of the Radon-Nikodym derivative of $\mu_{x,y,\mathrm{ac}}^T$ with respect to Lebesgue measure, denoted by $\rho_{x,y}^T$. Both of these have applications throughout mathematics and the physical sciences, some of which are explored numerically in Sect. 6. For example, suppose that we wish to solve the Schrödinger equation

$$\frac{du}{dt} = -iHu, \quad u_{t=0} = u_0,$$

where H is some self-adjoint Hamiltonian. We can express the solution at time t as

$$u(t) = \exp(-iHt)u_0 = \left[\int_{\mathbb{R}} \exp(-i\lambda t) dE^H(\lambda) \right] u_0.$$

For example, the quantity

$$L(t) = \langle u(t), u_0 \rangle = \int_{\mathbb{R}} \exp(-i\lambda t) d\mu_{u_0,u_0}^H(\lambda),$$

known as the autocorrelation function [141], is simply the Fourier transform of the spectral measure $d\mu_{u_0,u_0}^H$. In particular, if $d\mu_{u_0,u_0}^H$ is absolutely continuous, then ρ_{u_0,u_0}^H and L form a Fourier transform pair. The computation of evolution generated by an operator is in some sense dual to the computation of the spectral measure. This interpretation of a time evolution can be adapted to describe many signals generated by PDEs [48,85,126] and stochastic processes [65,88] [118, Ch. 7]. In this section, we show how to compute the functional calculus and $\rho_{x,y}^T$.

4.1. Computation of the functional calculus. Recall that given a possibly unbounded complex-valued Borel function F, defined on $\sigma(T)$, and $T \in \Omega_{\mathrm{N}}$, $F(T)$ is defined by

$$F(T) = \int_{\sigma(T)} F(\lambda) dE^T(\lambda).$$

$F(T)$ is a densely defined closed normal operator with dense domain given by

$$\mathcal{D}(F(T)) = \left\{ x \in l^2(\mathbb{N}) : \int_{\sigma(T)} |F(\lambda)|^2 d\mu_{x,x}^T(\lambda) < \infty \right\}.$$

For simplicity, we will only deal with the case that F is a bounded continuous function on $\mathbb{R}$, that is, $F \in C_b(\mathbb{R})$. In this case $\mathcal{D}(F(T))$ is the whole of $l^2(\mathbb{N})$ (the variations $|\mu_{x,y}^T|$ are finite) and we can use standard properties of the Poisson kernel. We assume that given $F \in C_b(\mathbb{R})$, we have access to piecewise constant functions F_n supported in $[-n, n]$ such that $\|F - F_n\|_{L^\infty([-n,n])} \leq n^{-1}$. Clearly other suitable data also suffices and, as usual, we abuse notation slightly by adding this information to the evaluation sets Λ_i (recall that Λ_i are defined in (1.18)).

Theorem 4.1 (Computation of the functional calculus). *Given the set-up in Sects. 1.5 and 1.6, consider the map*

$$\Xi_{\mathrm{fun}} : \Omega_{f,\alpha,\beta} \times C_b(\mathbb{R}) \to l^2(\mathbb{N})$$
$$(T, x, F) \to F(T)x.$$

Then $\{\Xi_{\mathrm{fun}}, \Omega_{f,\alpha,\beta} \times C_b(\mathbb{R}), \Lambda_1\} \in \Delta_2^A$.

Proof. Let $(T, x, F) \in \Omega_{f,\alpha,\beta} \times C_b(\mathbb{R})$ then by Fubini's theorem,

$$\int_{-n}^{n} K_H(u + i/n; T, x) F_n(u) du = \left[\int_{-\infty}^{\infty} \int_{-n}^{n} P_H(u - \lambda, 1/n) F_n(u) du \, dE^T(\lambda) \right] x.$$

The inner integral is bounded since F is bounded and the Poisson kernel integrates to 1 along the real line. It also converges to $F(\lambda)$ everywhere. Hence by the dominated convergence theorem

$$\lim_{n \to \infty} \int_{-n}^{n} K_H(u + i/n; T, x) F_n(u) du = F(T)x.$$

We now use the same arguments used to prove Theorem 3.1. Using Corollary 2.2, together with $\|K_H(u + i/n; T, x)\|_{L^\infty(\mathbb{R})} \leq nC_1$ and the fact that $K_H(u + i/n; T, x)$ is Lipschitz continuous with Lipschitz constant $n^2 C_2$ for some (possibly unknown) constants C_1 and C_2, we can approximate this integral with an error that vanishes in the limit $n \to \infty$. □

If $\sigma(T)$ is bounded, then, with slightly more information available to our algorithms, a simpler proof holds using the Stone–Weierstrass theorem. Suppose that given x, the vectors $T^n x$ can be computed to arbitrary precision. There exists a sequence of polynomials $p_m(z)$ converging uniformly to $F(z)$ on $\sigma(T)$. Assuming such a sequence can be explicitly constructed (for example using Bernstein or Chebyshev polynomials), we can take $p_m(T)x$ as approximations of $F(T)x$. If we can bound $\|p_m(z) - F(z)\|_{\sigma(T)} \leq \epsilon_m$ with ϵ_m null, then the vector $F(T)x$ can be computed *with error control*. However, computing $T^n x$ for large n (even if $x = e_1$) may be computationally expensive as was found in the example in Sect. 6.4. We will also see in Sect. 6.4 that if $\sigma(T)$ is bounded and F is analytic in an open neighbourhood of $\sigma(T)$, then $F(T)x$ can be computed with error control by deforming the integration contour away from the spectrum. Such a deformation is useful since the resolvent does not blow up along such a contour, and we can bound its Lipschitz constant.

4.2. Computation of the Radon–Nikodym derivative. Recall the definition of the Radon–Nikodym derivative in (1.16) and note that $\rho_{x,y}^T \in L^1(\mathbb{R})$ for $T \in \Omega_{SA}$. We consider its computation in the L^1 sense in the following theorem, where, as before, we assume (3.1), adding the approximations of U to our evaluation set Λ_1 (defined in (1.18)), along with component-wise evaluations of a given vector y. However, we must consider the computation away from the singular part of the spectrum.

Theorem 4.2 (Computation of the Radon–Nikodym derivative)*. Given the set-up in Sects. 1.5 and 1.6, consider the map*

$$\Xi_{RN} : \Omega_{f,\alpha,\beta} \times l^2(\mathbb{N}) \times \mathcal{U} \to L^1(\mathbb{R})$$

$$(T, x, y, U) \to \rho_{x,y}^T|_U.$$

We restrict this map to the quadruples (T, x, y, U) such that U is strictly separated from $\text{supp}(\mu_{x,y,\text{sc}}^T) \cup \text{supp}(\mu_{x,y,\text{pp}}^T)$ *and denote this subclass by* $\widetilde{\Omega}_{f,\alpha,\beta}$. *Then* $\{\Xi_{RN}, \widetilde{\Omega}_{f,\alpha,\beta}, \Lambda_1\} \in \Delta_2^A$. *Furthermore, each output $\Gamma_n(T, x, y, U)$ of the algorithms constructed in the proof consists of a piecewise affine function, supported in U with rational knots and taking (complex) rational values at these knots.*

Remark 6. Essentially, this theorem tells us that if we can compute the action of the resolvent operator with asymptotic error control, then we can compute the Radon–Nikodym derivative of the absolutely continuous part of the measures on open sets which are a positive distance away from the singular support of the measure. The assumption that U is separated from $\operatorname{supp}(\mu^T_{x,y,\mathrm{sc}}) \cup \operatorname{supp}(\mu^T_{x,y,\mathrm{pp}})$ may seem unnatural but is needed to gain L^1 convergence of the approximation. However, without it, the proof still gives almost everywhere pointwise convergence.

Proof. Let $(T, x, y, U) \in \widetilde{\Omega}_{f,\alpha,\beta}$. For $u \in U$ we decompose as follows

$$
\langle K_H(u + i\epsilon; T, x), y \rangle = \frac{1}{\pi} \int_{\mathbb{R}} \frac{\epsilon}{(\lambda - u)^2 + \epsilon^2} \rho^T_{x,y}(\lambda) d\lambda
$$
$$
+ \frac{1}{\pi} \int_{\mathbb{R} \setminus U} \frac{\epsilon}{(\lambda - u)^2 + \epsilon^2} \left\{ d\mu^T_{x,y,\mathrm{sc}}(\lambda) + d\mu^T_{x,y,\mathrm{pp}}(\lambda) \right\}.
$$

(4.1)

The first term converges to $\rho^T_{x,y}|_U$ in $L^1(U)$ as $\epsilon \downarrow 0$ since $\rho^T_{x,y}|_U \in L^1(U)$. Since we assumed that U is separated from $\operatorname{supp}(\mu^T_{x,y,\mathrm{sc}}) \cup \operatorname{supp}(\mu^T_{x,y,\mathrm{pp}})$, it follows that the second term of (4.1) converges to 0 in $L^1(U)$ as $\epsilon \downarrow 0$. Hence we are done if we can approximate $\langle K_H(u + i/n; T, x), y \rangle$ in $L^1(U)$ with an error converging to zero as $n \to \infty$.

Recall that $K_H(u + i/n; T, x)$ is Lipschitz continuous with Lipschitz constant at most $n^2 \|x\|/\pi$. By assumption, and using Corollary 2.2, we can approximate $K_H(u + i/n; T, x)$ to asymptotic precision with vectors of finite support. Hence the inner product

$$
f_n(u) := \langle K_H(u + i/n; T, x), y \rangle
$$

can be approximated to asymptotic precision (now with a possibly unknown constant also depending on $\|y\|$) and f_n is Lipschitz continuous with Lipschitz constant at most $n^2 \|x\| \|y\|/\pi$.

Recall that U can be written as the disjoint union

$$
U = \bigcup_m (a_m, b_m),
$$

where $a_m, b_m \in \mathbb{R} \cup \{\pm\infty\}$ and the union is at most countable. Without loss of generality, we assume that the union is over $m \in \mathbb{N}$. Given an interval (a_m, b_m), let $a_m < z_{m,1,n} < z_{m,2,n} < \cdots < z_{m,r_m,n} < b_m$ such that $z_{m,j,n} \in \mathbb{Q}$ and $|z_{m,j,n} - z_{m,j+1,n}| \leq (b_m - a_m)^{-1} n^{-3} m^{-2}$ and $|a_m - z_{m,1,n}|, |b_m - z_{m,r_m,n}| \leq n^{-1}$. Let $f_{m,n}$ be a piecewise affine interpolant with knots $z_{m,1,n}, \ldots, z_{m,r_m,n}$ supported on $(z_{m,1,n}, z_{m,r_m,n})$ with the property that $|f_{m,n}(z_{m,j,n}) - f_n(z_{m,j,n})| < C(b_m - a_m)^{-1} n^{-1} m^{-2}$. Here C is some unknown constant which occurs from the asymptotic approximation of f_n that arises from Corollary 2.2 and we can always compute such $f_{m,n}$ in finitely many arithmetic operations and comparisons.

Let $\Gamma_n(T, x, y, U)$ be the function that agrees with $f_{m,n}$ on (a_m, b_m) for $m \leq n$ and is zero elsewhere. Clearly the nodes of $\Gamma_n(T, x, y, U)$ can be computed using finitely many arithmetic operations and comparisons and the relevant set of evaluation functions

Λ_1. A simple application of the triangle inequality implies that

$$
\int_U \left| \Gamma_n(T, U, x, y)(u) - \rho^T_{x,y}(u) \right| du
$$

$$
\leq \sum_{m>n} \int_{(a_m, b_m)} \left| \rho^T_{x,y}(u) \right| du + \sum_{m \leq n} \int_{(a_m, b_m) \setminus (z_{m,1,n}, z_{m,r_m,n})} \left| \rho^T_{x,y}(u) \right| du
$$

$$
+ \sum_{m \leq n} \int_{(z_{m,1,n}, z_{m,r_m,n})} \left| \rho^T_{x,y}(u) - f_n(u) \right| du + \frac{\widetilde{C}(x, y, T)}{n} \sum_{m \leq n} \frac{1}{m^2},
$$

where the last term arises due to the piecewise affine interpolant. The bound clearly converges to zero as required. ☐

5. Computing Spectra as Sets

We now turn to computing the different types of spectra as sets in the complex plane. Specifically, define the problem functions $\Xi^{\mathbb{C}}_{\mathcal{I}}(T) = \sigma_{\mathcal{I}}(T)$ for $\mathcal{I} = $ ac, sc or pp. Note also that $\sigma_{\mathrm{pp}}(T) = \overline{\sigma_{\mathrm{p}}(T)}$, the closure of the set of eigenvalues. Recalling the definition of a computational problem in Appendix A, we compute these quantities in a metric space $\mathcal{M}$ with metric $d_{\mathcal{M}}$. Since we wish to include unbounded operators, we use the *Attouch–Wets metric* defined by

$$
d_{\mathrm{AW}}(C_1, C_2) = \sum_{n=1}^{\infty} 2^{-n} \min \left\{ 1, \sup_{|x| \leq n} |\mathrm{dist}(x, C_1) - \mathrm{dist}(x, C_2)| \right\}, \qquad (5.1)
$$

for $C_1, C_2 \in \mathrm{Cl}(\mathbb{C})$, where $\mathrm{Cl}(\mathbb{C})$ denotes the set of closed non-empty subsets of $\mathbb{C}$. When considering bounded T, whose spectrum is necessarily a compact subset of $\mathbb{C}$, we let $(\mathcal{M}, d_{\mathcal{M}})$ be the set of all non-empty compact subsets of $\mathbb{C}$ provided with the *Hausdorff metric* $d_{\mathcal{M}} = d_{\mathrm{H}}$:

$$
d_{\mathrm{H}}(X, Y) = \max \left\{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \right\}, \qquad (5.2)
$$

where $d(x, y) = |x - y|$ is the usual Euclidean distance. Note that for compact sets (and hence for bounded operators), the topological notions of convergence according to d_{H} and d_{AW} coincide. To allow the possibility that the different spectral sets $\sigma_{\mathcal{I}}(T)$ are empty, we add the empty set to our metric space as an isolated point (the space remains metrisable).[7]

The main theorem of this section is the following:

Theorem 5.1. *Given the above setup and that in Sects. 1.5 and 1.6, for $i = 1, 2$ it holds that*

$$
\Delta^G_2 \not\ni \{\Xi^{\mathbb{C}}_{\mathrm{ac}}, \Omega_{f,\alpha}, \Lambda_i\} \in \Delta^A_3, \quad \Delta^G_2 \not\ni \{\Xi^{\mathbb{C}}_{\mathrm{sc}}, \Omega_{f,\alpha}, \Lambda_i\} \in \Delta^A_4,
$$

$$
\Delta^G_2 \not\ni \{\Xi^{\mathbb{C}}_{\mathrm{pp}}, \Omega_{f,\alpha}, \Lambda_i\} \in \Delta^A_3.
$$

If $f(n) - n \geq \sqrt{2n} + \frac{1}{2}$, then $\{\Xi^{\mathbb{C}}_{\mathrm{sc}}, \Omega_{f,0}, \Lambda_i\} \notin \Delta^G_3$ also holds.

---

[7] This simply means that $F_n \to \emptyset$ if and only if $F_n = \emptyset$ eventually.

In order to prove Theorem 5.1, we only need to prove the lower bounds for Λ_2 and the upper bounds for Λ_1 (recall that Λ_i are defined in (1.18)). These results show that despite the results of Sects. 2–4, in general it is very hard to compute the decomposition of the spectrum in the sense of (1.17). We also answer the question posed in Sect. 2.2 and prove that the spectral measures, while computable in one limit, cannot, in general, be computed with error control (see Theorem 5.2), unless one has additional regularity assumptions such as in [39] (computation with error control is made precise in [33, Ch. 4]).

5.1. Proof for point spectra. Proof that $\{\Xi_{\mathrm{pp}}^{\mathbb{C}}, \Omega_{f,\alpha}, \Lambda_2\} \notin \Delta_2^G$. To prove this, it is enough to consider bounded Schrödinger operators acting on $l^2(\mathbb{N})$ (defined in Sect. 3.2.2), which are clearly a subclass of $\Omega_{f,0}$ for $f(n) = n + 1$. Note that since the evaluation functions in Λ_2 can be recovered from those in Λ_1 in this special case, we can assume that we are dealing with Λ_1. Suppose for a contradiction that there does exist a sequence of general algorithms, Γ_n, with

$$\lim_{n\to\infty} \Gamma_n(H_v) = \Xi_{\mathrm{pp}}^{\mathbb{C}}(H_v).$$

We will construct a potential v such that $\Gamma_n(H_v)$ does not converge. To do this, choose $\rho = \chi_{[-c,c]}/(2c)$ for some constant c such that the conditions of Theorem 3.3 hold. We will use Theorem 3.3 and the following well-known facts:

1. If there exists $N \in \mathbb{N}$ such that $v(n)$ is 0 for $n \geq N$, then $\sigma_{\mathrm{pp}}(H_v) \cap (0, 4) = \emptyset$ [114], but $[0, 4] \subset \sigma(H_v)$ (the potential acts as a compact perturbation so the essential spectrum is $[0, 4]$).
2. If we are in the setting of Theorem 3.3, then the spectrum of $H_{v_\omega} + A$ is pure point almost surely. Moreover, if $\rho = \chi_{[-c,c]}/(2c)$ for some constant c, then $[-c, 4+c] \subset \sigma_{\mathrm{pp}}(H_{v_\omega} + A)$ almost surely.

We will define the potential v inductively as follows. Let v_1 be a potential of the form v_ω (with density ρ) such that $[-c, 4 + c] \subset \sigma(H_{v_1})$ and $\sigma(H_{v_1})$ is pure point. Such a v_1 exists by Theorem 3.3 and fact (2) above. Then for large enough n there exists $z_n \in \Gamma_n(H_{v_1})$ such that $|z_n - 2| \leq 1$. Fix $n = n_1$ such that this holds. Then $\Gamma_{n_1}(H_{v_1})$ only depends on $\{v_1(j) : j \leq N_1\}$ for some integer N_1 by (i) of Definition A.2. Define the potential v_2 by $v_2(j) = v_1(j)$ for all $j \leq N_1$ and $v_2(j) = 0$ otherwise. Then by fact (1) above, $\Gamma_n(H_{v_2}) \cap [1/2, 7/2] = \emptyset$ for large n, say for n_2. But then $\Gamma_{n_2}(H_{v_2})$ only depends on $\{v_2(j) : j \leq N_2\}$ for some integer N_2.

We repeat this process inductively switching between potentials which induce $\Gamma_{n_k}(H_{v_k}) \cap [1/2, 7/2] = \emptyset$ for k even and potentials which induce $\Gamma_{n_k}(H_{v_k}) \cap [1, 3] \neq \emptyset$ for k odd. Explicitly, if k is even then define a potential v_{k+1} by $v_{k+1}(j) = v_k(j)$ for all $j \leq N_k$ and $v_{k+1}(j) = v_\omega(j)$ (with the density ρ) otherwise such that $[-c, 4 + c] \subset \sigma(H_{v_{k+1}})$ and $\sigma(H_{v_{k+1}})$ is pure point. Such a ω exists from Theorem 3.3 and fact (2) above applied with the perturbation A to match the potential for $j \leq N_k$. If k is odd then we define v_{k+1} by $v_{k+1}(j) = v_k(j)$ for all $j \leq N_k$ and $v_{k+1}(j) = 0$ otherwise. We can then choose n_{k+1} such that the above intersections hold and N_{k+1} such that $\Gamma_{n_{k+1}}(H_{v_{k+1}})$ only depends on $\{v_{k+1}(j) : j \leq N_{k+1}\}$. Finally set $v(j) = v_k(j)$ for $j \leq N_k$. It is clear from (iii) of Definition A.2, that $\Gamma_{n_k}(H_v) = \Gamma_{n_k}(H_{v_k})$. But then this implies that $\Gamma_{n_k}(H_v)$ cannot converge, the required contradiction. $\square$

A similar argument gives the following theorem, where $\mathbb{V}$ is used to denote the set of bounded real-valued potentials on $\mathbb{N}$ and Λ_3 denotes the pointwise evaluations of such potentials.

Theorem 5.2 (Impossibility of computing spectral measures with error control). *Consider the problem function*

$$\widehat{\Xi} : \mathbb{V} \times \mathbb{N} \to \mathbb{R}_{\geq 0}$$
$$(v, j) \to \langle E^{H_v}_{\{1\}} e_j, e_j \rangle.$$

Then $\{\widehat{\Xi}, \mathbb{V} \times \mathbb{N}, \Lambda_3\} \in \Delta_2^A$ but $\{\widehat{\Xi}, \mathbb{V} \times \mathbb{N}, \Lambda_3\} \notin \Delta_1^G$. In other words, $\widehat{\Xi}$ can be computed in one limit, but it cannot be computed with error control.

Proof. The result $\{\widehat{\Xi}, \mathbb{V} \times \mathbb{N}, \Lambda_3\} \in \Delta_2^A$ follows directly from the remarks after Theorem 3.1 and Theorem 2.1. Suppose for a contradiction that $\{\widehat{\Xi}, \mathbb{V} \times \mathbb{N}, \Lambda_3\} \in \Delta_1^G$ and that Γ_n is a sequence of general algorithms solving the problem with error control. It follows that for each $j \in \mathbb{N}$, there exists a sequence of general algorithms Γ_n^j such that

$$\lim_{n \to \infty} \Gamma_n^j(v) = \begin{cases} 1, & \text{if } \widehat{\Xi}(v, j) > 0 \\ 0, & \text{otherwise} \end{cases}.$$

Informally, these are described as follows. Fix j and consider the lower bound on $\widehat{\Xi}(v, j)$ computed by $\{\Gamma_m(v, j) : m \leq n\}$. If this is greater than 0 then set $\Gamma_n^j(v) = 1$, otherwise set $\Gamma_n^j(v) = 0$. It follows that $\Gamma_n^j(v)$ also converges from below. It holds that $1 \in \sigma_p(H_v)$ if and only if $\widehat{\Xi}(v, j) > 0$ for some $j \in \mathbb{N}$. Now define

$$\widehat{\Gamma}_n(v) = \sup_{j \leq n} \Gamma_n^j(v).$$

It is clear that this is a general algorithm using Λ_3. Furthermore,

$$\lim_{n \to \infty} \widehat{\Gamma}_n(v) = \begin{cases} 1, & \text{if } 1 \in \sigma_p(H_v) \\ 0, & \text{otherwise} \end{cases},$$

with convergence from below.

Now we may choose a v such that $1 \in \sigma_p(H_v)$ (this can be achieved for example by taking a potential which induces pure point spectrum and shifting the operator accordingly). It follows that for large n, we have $\widehat{\Gamma}_n(v) = 1$. But the computation of $\widehat{\Gamma}_n(v)$ is only dependent on $v(j)$ for $j < N$ for some $N \in \mathbb{N}$. Define $v_0 \in \mathbb{V}$ by $v_0(j) = v(j)$ if $j < N$ and $v_0(j) = 0$ otherwise. It follows that $\widehat{\Gamma}_n(v_0) = 1$. But since the potential has compact support, $1 \notin \sigma_p(H_{v_0})$ and hence $\widehat{\Gamma}_n(v_0) = 0$, the required contradiction. $\square$

We now prove that $\Xi_{\text{pp}}^{\mathbb{C}}$ can be computed using a height two arithmetical tower. The first step is the following technical lemma, whose proof will also be used later when considering $\Xi_{\text{ac}}^{\mathbb{C}}$.

Lemma 5.3. *Let $a < b$ with $a, b \in \mathbb{R}$ and consider the decision problem*

$$\Xi_{a,b,\text{pp}} : \Omega_{f,\alpha} \to \{0, 1\}$$
$$T \to \begin{cases} 1 & \text{if } \sigma_{\text{pp}}(T) \cap [a, b] \neq \emptyset \\ 0 & \text{otherwise.} \end{cases}$$

Then there exists a height two arithmetical tower Γ_{n_2,n_1} (with evaluation functions Λ_1) for $\Xi_{a,b,\text{pp}}$. Furthermore, the final limit is from below in the sense that $\Gamma_{n_2}(T) := \lim_{n_1 \to \infty} \Gamma_{n_2,n_1}(T) \leq \Xi_{a,b,\text{pp}}(T)$.

Proof. Step 1 of the proof of Theorem 3.2 yields a height two arithmetical tower $\widehat{\Gamma}^j_{n_2,n_1}(T)$ for the computation of $\mu^T_{e_j,e_j,c}((a,b))$. Note that the final limit is from above and using the fact that $\mu^T_{e_j,e_j,c}(\{a,b\}) = 0$, we obtain a height two tower for $\mu^T_{e_j,e_j,c}([a,b])$. We can then use the height one tower for $\mu^T_{e_j,e_j}([a,b])$ constructed in Sect. 2.2, denoted by $\widetilde{\Gamma}^j_{n_1}(T)$, and define

$$a_{j,n_2,n_1}(T) = \widetilde{\Gamma}^j_{n_1}(T) - \widehat{\Gamma}^j_{n_2,n_1}(T).$$

This provides a height two arithmetical tower for $\mu^T_{e_j,e_j,pp}([a,b])$ with the final limit from below. Without loss of generality (by taking successive maxima), we can assume that these towers are non-decreasing in n_2. Now set

$$\Upsilon_{n_2,n_1}(T) = \max_{1 \leq j \leq n_2} a_{j,n_2,n_1}(T).$$

Then it is clear that the limit $\lim_{n_1 \to \infty} \Upsilon_{n_2,n_1}(T) = \Upsilon_{n_2}(T)$ exists. Furthermore, the monotonicity of $a_{j,n_2,n_1}(T)$ in n_2 implies that

$$\lim_{n_2 \to \infty} \Upsilon_{n_2}(T) = \sup_{n \in \mathbb{N}} \mu^T_{e_n,e_n,pp}([a,b]),$$

with monotonic convergence from below. This limiting value is zero if $\Xi_{a,b,pp}(T) = 0$, otherwise it is a positive finite number.

To convert this to a height two tower for the decision problem $\Xi_{a,b,pp}$, that maps to the discrete space $\{0,1\}$, we use the following trick. Consider the intervals $J^{n_2}_1 = [0, 1/n_2]$, and $J^{n_2}_2 = [2/n_2, \infty)$. Let $k(n_2,n_1) \leq n_1$ be maximal such that $\Upsilon_{n_2,n_1}(T) \in J^{n_2}_1 \cup J^{n_2}_2$. If no such k exists or $\Upsilon_{n_2,k}(T) \in J^{n_2}_1$ then set $\Gamma_{n_2,n_1}(T) = 0$. Otherwise set $\Gamma_{n_2,n_1}(T) = 1$. These can be computed using finitely many arithmetic operations and comparisons using Λ_1. The point of the intervals $J^{n_2}_1$ and $J^{n_2}_2$ is that we can show $\lim_{n_1 \to \infty} \Gamma_{n_2,n_1}(T) = \Gamma_{n_2}(T)$ exists. This is because $\lim_{n_1 \to \infty} \Upsilon_{n_2,n_1}(T) = \Upsilon_{n_2}(T)$ exists and hence we cannot oscillate infinitely often between the separated intervals $J^{n_2}_1$ and $J^{n_2}_2$. Now suppose that $\Xi_{a,b,pp}(T) = 0$, then $\lim_{n_1 \to \infty} \widehat{\Gamma}_{n_2,n_1}(T) = 0$ and hence $\lim_{n_1 \to \infty} \Gamma_{n_2,n_1}(T) = 0$ for all n_2. Now suppose that $\Xi_{a,b,pp}(T) = 1$, then for large enough n_2 we must have that $\Upsilon_{n_2}(T) > 2/n_2$ and hence $\Gamma_{n_2}(T) = 1$. Together, these prove the convergence and that $\Gamma_{n_2}(T) \leq \Xi_{a,b,pp}(T)$. □

Proof that $\{\Xi^{\mathbb{C}}_{pp}, \Omega_{f,\alpha}, \Lambda_1\} \in \Delta^A_3$. To construct a height two arithmetical tower for $\Xi^{\mathbb{C}}_{pp}$, we will use Lemma 5.3 repeatedly. Let $\widehat{\Gamma}_{n_2,n_1}(\cdot, I)$ denote the height two tower constructed in the proof of Lemma 5.3 for the closed interval I ($I = [a,b]$), where without loss of generality by taking successive maxima in n_2, we can assume that this tower is non-decreasing in n_2 (this is where we use convergence from below in the final limit in the statement of the lemma). For a given n_1 and n_2, we construct $\Gamma_{n_2,n_1}(T)$ as follows (we will use some basic terminology from graph theory).

Define the intervals $I^0_{n_2,n_1,j} = [j, j+1]$ for $j = -n_2, \ldots, n_2 - 1$ so that these cover the interval $[-n_2, n_2]$. Now suppose that $I^k_{n_2,n_1,j}$ are defined for $j = 1, \ldots, r_k(n_2, n_1, T)$. Compute each $\widehat{\Gamma}_{n_2,n_1}(T, I^k_{n_2,n_1,j})$ and if this is 1, bisect $I^k_{n_2,n_1,j}$ via its midpoint into two equal halves consisting of closed intervals. We then take all these bisected intervals and label them as $I^{k+1}_{n_2,n_1,j}$ for $j = 1, \ldots, r_{k+1}(n_2, n_1, k, T)$. This is repeated until we have

no further bisections or the intervals $I^{n_2}_{n_2,n_1,j}$ have been computed. By adding the interval $[-n_2, n_2]$ as a root with children $I^0_{n_2,n_1,j}$, this creates a finite tree structure where a non-root interval I is a parent of two intervals precisely if those two intervals are formed from its bisection and $\widehat{\Gamma}_{n_2,n_1}(T, I) = 1$. We then prune this tree by discarding all leaves I which have $\widehat{\Gamma}_{n_2,n_1}(T, I) = 0$ to form the tree $\mathcal{T}_{n_2,n_1}(T)$. Finally, we let $\Gamma_{n_2,n_1}(T)$ be the union of all the leaves of $\mathcal{T}_{n_2,n_1}(T)$. Clearly this can be computed using finitely many arithmetic operations and comparisons using Λ_1. The construction is shown visually in Fig. 3.

In the above construction, the number of intervals considered (including those not in the tree $\mathcal{T}_{n_2,n_1}(T)$) for a fixed n_2 is $n_2 2^{n_2+1} + 1$ and hence independent of n_1. It follows that $\mathcal{T}_{n_2,n_1}(T)$ and $\Gamma_{n_2,n_1}(T)$ are constant for large n_1 (due to the convergence of the $\widehat{\Gamma}_{n_2,n_1}(T, I)$ in $\{0, 1\}$). We denote these limiting values by $\mathcal{T}_{n_2}(T)$ and $\Gamma_{n_2}(T)$ respectively and also denote the corresponding intervals in the construction at the m−th level of this limit by $I^m_{n_2,j}$. Note also that if $\Xi^{\mathbb{C}}_{\mathrm{pp}}(T) = \emptyset$ then $\Gamma_{n_2}(T) = \emptyset$.

Now suppose that $z \in \Xi^{\mathbb{C}}_{\mathrm{pp}}(T)$, then there exists a sequence of nested intervals $I_m = I^m_{n_2,a_{m,n_2}}$ containing z for $m = 0, \ldots, n_2$, where these intervals are independent of n_2. Fix m, then for large n_2 we must have that $\widehat{\Gamma}_{n_2}(T, I_j) = 1$ for $j = 1, \ldots, m$. It follows that I_m has a descendent interval $I_{n_2,m}$ contained in $\Gamma_{n_2}(T)$ and hence we must have

$$\mathrm{dist}(z, \Gamma_{n_2}(T)) \leq 2^{-m}.$$

Since m was arbitrary it follows that $\mathrm{dist}(z, \Gamma_{n_2}(T))$ converges to 0 as $n_2 \to \infty$.

Conversely, suppose that $z_{m_j} \in \Gamma_{m_j}(T)$ with $m_j \to \infty$, then we must show that all limit points of $\{z_{m_j}\}$ lie in $\Xi^{\mathbb{C}}_{\mathrm{pp}}(T)$. Suppose this were false, then by taking a subsequence if necessary, we can assume that $z_{m_j} \to z$ and $\mathrm{dist}(z_{m_j}, \Xi^{\mathbb{C}}_{\mathrm{pp}}(T)) \geq \delta$ for some $\delta > 0$. We claim that it is sufficient to prove that the maximum length of the leaves of $\mathcal{T}_{n_2}(T)$ intersecting a fixed compact subset of $\mathbb{R}$, converges to zero as $n_2 \to \infty$. Suppose this has been shown, then $z_{m_j} \in I_{m_j}$ for some leaf I_{m_j} of $\mathcal{T}_{m_j}(T)$. It follows that $I_{m_j} \cap \Xi^{\mathbb{C}}_{\mathrm{pp}}(T) \neq \emptyset$ and $|I_{m_j}| \to 0$. But this contradicts z_{m_j} being positively separated from $\Xi^{\mathbb{C}}_{\mathrm{pp}}(T)$.

We are thus left with proving the claim regarding the lengths of leaves. Suppose this were false, then there exists a compact set $K \subset \mathbb{R}$ and leaves I_j in $\mathcal{T}_{b_j}(T)$ such that the lengths of I_j do not converge to zero and I_j intersect K. By taking subsequences if necessary, we can assume that the lengths of each I_j are constant. Then by the compactness of K and taking subsequences if necessary again, we can assume that each of the I_j are equal to a common interval I. It follows that $\widehat{\Gamma}_{b_j}(T, I) = 1$ but that $\widehat{\Gamma}_{b_j}(T, I_1) = \widehat{\Gamma}_{b_j}(T, I_2) = 0$ since I is a leaf, where I_1 and I_2 form the bisection of I. Taking $b_j \to \infty$, this implies that $I \cap \Xi^{\mathbb{C}}_{\mathrm{pp}}(T) \neq \emptyset$ but $I_1 \cap \Xi^{\mathbb{C}}_{\mathrm{pp}}(T) = I_2 \cap \Xi^{\mathbb{C}}_{\mathrm{pp}}(T) = \emptyset$ which is absurd. Hence we have shown the required contradiction, and we have finished the proof. □

5.2. Proof for absolutely continuous spectra.

To prove the lower bound (that one limit will not suffice), our strategy will be to reduce a certain decision problem to the computation of $\Xi^{\mathbb{C}}_{\mathrm{ac}}$. Let $(\mathcal{M}', d')$ be the discrete space $\{0, 1\}$, let Ω' denote the collection of all infinite sequences $\{a_j\}_{j\in\mathbb{N}}$ with entries $a_j \in \{0, 1\}$ and consider the problem function

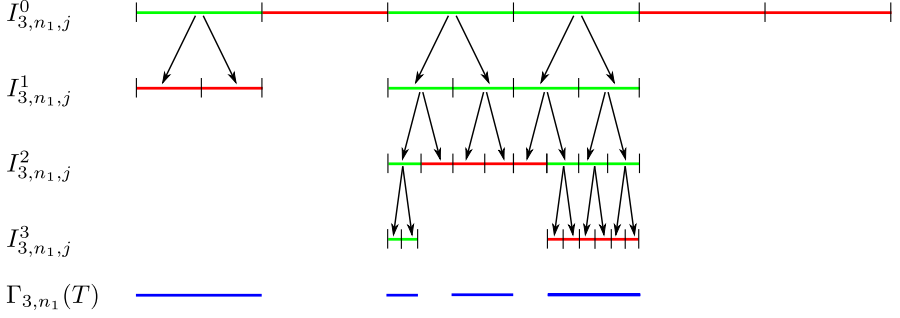$$\Xi'(\{a_j\}) : \text{ 'Does } \{a_j\} \text{ have infinitely many non-zero entries?',}$$

Fig. 3. Example of tree structure used to compute the point spectrum for $n_2 = 3$. Each tested interval is shown in green ($\widehat{\Gamma}_{n_2,n_1}(T, I) = 1$) or red ($\widehat{\Gamma}_{n_2,n_1}(T, I) = 0$). The arrows show the bisections and the final output is shown in blue

that maps to $(\mathcal{M}', d')$. In Appendix A, it is shown that $\mathrm{SCI}(\Xi', \Omega')_G = 2$ (where the evaluation functions consist of component-wise evaluations of the array $\{a_j\}$).

Proof that $\{\Xi_{\mathrm{ac}}^{\mathbb{C}}, \Omega_{f,\alpha}, \Lambda_2\} \notin \Delta_2^G$. We are done if we prove the result for $f(n) = n+1$ and $\alpha = 0$. In this case Λ_1 and Λ_2 are equivalent so we can restrict the argument to Λ_1. Suppose for a contradiction that Γ_n is a height one tower of general algorithms solving $\{\Xi_{\mathrm{ac}}^{\mathbb{C}}, \Omega_{f,0}, \Lambda_2\}$. We will gain a contradiction by using the supposed tower to solve $\{\Xi', \Omega'\}$.

Given $\{a_j\} \in \Omega'$, consider the operator $H = H_0 + v$, where the potential is of the following form:

$$v(m) = \sum_{k=1}^{\infty} a_k \delta_{m,k!}.$$

Then by Theorem 3.4, $[0, 4] \subset \sigma_{\mathrm{ac}}(H)$ if $\sum_k a_k < \infty$ (that is, if $\Xi'(\{a_j\}) = 0$) and $\sigma_{\mathrm{ac}}(H) \cap (0, 4) = \emptyset$ otherwise. Given N we can evaluate any matrix value of H using only finitely many evaluations of $\{a_j\}$ and hence the evaluation functions Λ_1 can be computed using component-wise evaluations of the sequence $\{a_j\}$. We now set

$$\widehat{\Gamma}_n(\{a_j\}) = \begin{cases} 0 & \text{if } \mathrm{dist}(2, \Gamma_n(H)) < 1 \\ 1 & \text{otherwise.} \end{cases}$$

The above comments show that each of these is a general algorithm, and it is clear that it converges to $\Xi'(\{a_j\})$ as $n \to \infty$, the required contradiction. $\qquad\square$

To construct the height two (arithmetical) tower for $\Xi_{\mathrm{ac}}^{\mathbb{C}}$, we require the following lemma.

Lemma 5.4. *Let* $a < b$ *with* $a, b \in \mathbb{R}$ *and consider the decision problem*

$$\Xi_{a,b,\mathrm{ac}} : \Omega_{f,\alpha} \to \{0, 1\}$$

$$T \to \begin{cases} 1 & \text{if } \sigma_{\mathrm{ac}}(T) \cap [a, b] \neq \emptyset \\ 0 & \text{otherwise.} \end{cases}$$

Then there exists a height two arithmetical tower Γ_{n_2,n_1} *(with evaluation functions* Λ_1*) for* $\Xi_{a,b,\mathrm{ac}}$*. Furthermore, the final limit is from below with* $\Gamma_{n_2}(T) := \lim_{n_1 \to \infty} \Gamma_{n_2,n_1}(T) \leq \Xi_{a,b,\mathrm{ac}}(T)$.

Proof. Fix such an a and b and let χ_n be a sequence of non-negative, continuous piecewise affine functions on $\mathbb{R}$, bounded by 1 and of compact support such that χ_n converge pointwise monotonically up to the constant function 1. Define the function

$$v_{m,n}(u, T) = \langle K_H(u + i/n, T, e_m), e_m \rangle$$

and set

$$a_{m,n_2,n_1}(T) = \int_a^b v_{m,n_1}(u, T) \chi_{n_2}\big(\big|v_{m,n_1}(u, T)\big|\big) du.$$

Since each χ_n is continuous and has compact support (which implies that the integrand is bounded for fixed n_2), and since $v_{m,n}(u, T)$ converges almost everywhere to $\rho^T_{e_m,e_m}(u)$ (the Radon–Nikodym derivative of the absolutely continuous part of the measure $\mu^T_{e_m,e_m}$), it follows by the dominated convergence theorem that

$$\lim_{n_1 \to \infty} a_{m,n_2,n_1}(T) =: a_{m,n_2}(T) = \int_a^b \rho^T_{e_m,e_m}(u) \chi_{n_2}(\rho^T_{e_m,e_m}(u)) du.$$

We now use the fact that the χ_n are increasing and the dominated convergence theorem again (with bounding integrable function $\rho^T_{e_m,e_m}$) to deduce that

$$\lim_{n_2 \to \infty} a_{m,n_2}(T) = \mu^T_{e_m,e_m,\mathrm{ac}}([a, b]),$$

with monotonic convergence from below.

Using Corollary 2.2 (and the now standard argument of Lipschitz continuity of the resolvent), we can compute approximations of $a_{m,n_2,n_1}(T)$ to accuracy $1/n_1$ in finitely many arithmetic operations and comparisons. Call these approximations $\widetilde{a}_{m,n_2,n_1}(T)$ and set

$$\Upsilon_{n_2,n_1}(T) = \max_{1 \le j \le n_2} \widetilde{a}_{j,n_2,n_1}(T).$$

The proof now follows that of Lemma 5.3 exactly. $\square$

Proof that $\{\Xi^{\mathbb{C}}_{\mathrm{ac}}, \Omega_{f,\alpha}, \Lambda_1\} \in \Delta^A_3$. This is exactly the same construction as in the above proof of the inclusion $\{\Xi^{\mathbb{C}}_{\mathrm{pp}}, \Omega_{f,\alpha}, \Lambda_1\} \in \Delta^A_3$. We simply replace the tower constructed in the proof of Lemma 5.3 by the tower constructed in the proof of Lemma 5.4. $\square$

5.3. Proof for singular continuous spectra.

We first prove the lower bound for the singular continuous spectrum via Theorem 3.4. Note that the impossibility result $\{\Xi^{\mathbb{C}}_{\mathrm{sc}}, \Omega_{f,\alpha}, \Lambda_2\} \notin \Delta^G_2$ follows from the same argument that was used to show $\{\Xi^{\mathbb{C}}_{\mathrm{ac}}, \Omega_{f,\alpha}, \Lambda_2\} \notin \Delta^G_2$. To show that two limits will not suffice for $f(n) - n \ge \sqrt{2n} + 1/2$, our strategy will be to again reduce a certain decision problem to the computation of $\Xi^{\mathbb{C}}_{\mathrm{sc}}$. Let $(\mathcal{M}', d')$ be the discrete space $\{0, 1\}$, let Ω' denote the collection of all infinite matrices $\{a_{i,j}\}_{i,j \in \mathbb{N}}$ with entries $a_{i,j} \in \{0, 1\}$ and consider the problem function

$\Xi'(\{a_{i,j}\})$: 'Does $\{a_{i,j}\}$ have a column containing infinitely many non-zero entries?',

that maps to $(\mathcal{M}', d')$. In [12], a Baire category argument was used to prove that $\mathrm{SCI}(\Xi', \Omega')_G = 3$ (where the evaluation functions consist of component-wise evaluations of the array $\{a_{i,j}\}$).

Proof that $\{\Xi^{\mathbb{C}}_{\mathrm{sc}}, \Omega_{f,\alpha}, \Lambda_2\} \notin \Delta^G_3$ *if* $f(n) - n \geq \sqrt{2n} + 1/2$. Assume that the function f satisfies $f(n) - n \geq \sqrt{2n} + 1/2$. The proof will use a direct sum construction. Given $\{a_{i,j}\} \in \Omega'$, consider the operators $H_j = H_0 + v_{(j)}$, where the potential is of the following form:

$$v_{(j)}(n) = \sum_{k=1}^{\infty} a_{k,j} \delta_{n,k!}.$$

Using Theorem 3.4, $[0, 4] \subset \sigma_{\mathrm{sc}}(H_j)$ if $\sum_k a_{k,j} = \infty$ (that is, if the j-th column has infinitely many 1s) and $\sigma_{\mathrm{sc}}(H_j) \cap (0, 4) = \emptyset$ otherwise. Now consider an effective bijection (with effective inverse) between the canonical bases of $l^2(\mathbb{N})$ and $\oplus_{j=1}^{\infty} l^2(\mathbb{N})$:

$$\phi : \{e_n : n \in \mathbb{N}\} \rightarrow \{e_{\mathbf{k}} : \mathbf{k} \in \mathbb{N}^{\mathbb{N}}, \|\mathbf{k}\|_0 = 1\}.$$

Set $H(\{a_{i,j}\}) = \oplus_{j=1}^{\infty} H_j$. Then through ϕ, we view $H = H(\{a_{i,j}\})$ as a self-adjoint operator acting on $l^2(\mathbb{N})$. Explicitly, we consider the matrix

$$H_{m,n} = \langle H e_{\phi(n)}, e_{\phi(m)} \rangle.$$

We choose the following bijection (where m lists the canonical basis in each Hilbert space):

$$
\begin{array}{cccc}
 & j = 1 & j = 2 & j = 3 & \cdots \\
m = 1 & \phi(1) & \phi(3) & \phi(6) & \cdots \\
m = 2 & \phi(2) & \phi(5) & \\
m = 3 & \phi(4) & & \\
\cdots & & &
\end{array}
$$

A straightforward computation shows that $H \in \Omega_{f,0}$. We also observe that if $\Xi'(\{a_{i,j}\}) = 1$ then $[0, 4] \subset \sigma_{\mathrm{sc}}(H)$, otherwise $\sigma_{\mathrm{sc}}(H) \cap (0, 4) = \emptyset$.

Suppose for a contradiction that Γ_{n_2,n_1} is a height two tower of general algorithms that solves $\{\Xi^{\mathbb{C}}_{\mathrm{sc}}, \Omega_{f,0}, \Lambda_1\}$. We will gain a contradiction by using the supposed height two tower to solve $\{\Xi', \Omega'\}$. We now set

$$\widehat{\Gamma}_{n_2,n_1}(\{a_{i,j}\}) = 1 - \min\{1, \mathrm{dist}(3, \Gamma_{n_2,n_1}(H(\{a_{i,j}\})))\},$$

where we use the convention $\mathrm{dist}(3, \emptyset) = 1$. The comments above show that each of these is a general algorithm. Furthermore, the convergence of Γ_{n_2,n_1} implies that

$$\lim_{n_2 \to \infty} \lim_{n_1 \to \infty} \widehat{\Gamma}_{n_2,n_1}(\{a_{i,j}\}) = 1 - \min\{1, \mathrm{dist}(3, \sigma_{\mathrm{sc}}(H(\{a_{i,j}\})))\} = \Xi'(\{a_{i,j}\}).$$

We are not quite done since the convergence here takes place on the interval $[0, 1]$ with the usual metric as opposed to $\{0, 1\}$ with the discrete metric. To get round this, we use the following, now familiar, trick.

Consider the intervals $J_1 = [0, 1/2]$, and $J_2 = [3/4, 1]$. Let $k(n_2, n_1) \leq n_1$ be maximal such that $\widehat{\Gamma}_{n_2,k}(\{a_{i,j}\}) \in J_1 \cup J_2$. If no such k exists or $\widehat{\Gamma}_{n_2,k}(\{a_{i,j}\}) \in J_1$ then

set $\Gamma'_{n_2,n_1}(\{a_{i,j}\}) = 0$. Otherwise set $\Gamma'_{n_2,n_1}(\{a_{i,j}\}) = 1$. As in the proof of Lemma 5.3, we can show $\lim_{n_1\to\infty} \Gamma'_{n_2,n_1}(\{a_{i,j}\}) = \Gamma'_{n_2}(\{a_{i,j}\})$ exists. If $\Xi'(\{a_{i,j}\}) = 0$, then for large n_2, $\lim_{n_1\to\infty} \widehat{\Gamma}_{n_2,k}(\{a_{i,j}\}) < 1/2$ and hence $\lim_{n_2\to\infty} \Gamma'_{n_2}(\{a_{i,j}\}) = 0$. Similarly, if $\Xi'(\{a_{i,j}\}) = 1$, then for large n_2 we must have $\lim_{n_1\to\infty} \widehat{\Gamma}_{n_2,k}(\{a_{i,j}\}) > 3/4$ and hence $\lim_{n_2\to\infty} \Gamma'_{n_2}(\{a_{i,j}\}) = 1$. Hence Γ'_{n_2,n_1} is a height two tower of general algorithms solving $\{\Xi', \Omega'\}$, a contradiction. □

Finally, we will use the following lemma to prove that the singular continuous spectrum can be computed in three limits.

Lemma 5.5. *Let $a < b$ with $a, b \in \mathbb{R}$ and consider the decision problem*

$$\Xi_{a,b,\mathrm{sc}} : \Omega_{f,\alpha} \to \{0, 1\}$$

$$T \to \begin{cases} 1 & \text{if } \sigma_{\mathrm{sc}}(T) \cap [a, b] \neq \emptyset \\ 0 & \text{otherwise.} \end{cases}$$

Then there exists a height three arithmetical tower Γ_{n_3,n_2,n_1} (with evaluation functions Λ_1) for $\Xi_{a,b,\mathrm{sc}}$. Furthermore, the final limit is from below with $\Gamma_{n_3}(T) := \lim_{n_2\to\infty} \lim_{n_1\to\infty} \Gamma_{n_3,n_2,n_1}(T) \leq \Xi_{a,b,\mathrm{sc}}(T)$.

Once this is proven, we can use the same construction that was used to prove $\{\Xi_{\mathrm{pp}}^{\mathbb{C}}, \Omega_{f,\alpha}, \Lambda_1\} \in \Delta_3^A$ and $\{\Xi_{\mathrm{ac}}^{\mathbb{C}}, \Omega_{f,\alpha}, \Lambda_1\} \in \Delta_3^A$ to show that $\{\Xi_{\mathrm{sc}}^{\mathbb{C}}, \Omega_{f,\alpha}, \Lambda_1\} \in \Delta_4^A$, but with an additional limit. Namely, we replace (n_2, n_1) by (n_3, n_2) in the proof and use the tower constructed in the proof of Lemma 5.4 instead of $\widehat{\Gamma}_{n_2,n_1}(T, I)$ for an interval I. We still gain the required convergence, since the only change is an additional limit in the finite number of decision problems that decide the appropriate tree.

Proof of Lemma 5.5. Note that we can write

$$\mu^T_{e_m,e_m,\mathrm{sc}}([a, b]) = \mu^T_{e_m,e_m}([a, b]) - \mu^T_{e_m,e_m,\mathrm{pp}}([a, b]) - \mu^T_{e_m,e_m,\mathrm{ac}}([a, b]).$$

From this and the proofs of Lemmas 5.3 and 5.4, it is clear that we can construct a height two arithmetical tower, $a_{m,n_2,n_1}(T)$, for $\mu^T_{e_m,e_m,\mathrm{sc}}([a, b])$, where the final limit is from above. Now set

$$\Upsilon_{n_3,n_2,n_1}(T) = \max_{1\leq j\leq n_3} a_{j,n_2,n_1}.$$

We see that each successive limit converges, with the second from above and the final from below. By taking successive maxima, minima of our base algorithms, we can assume that the second and final limits are monotonic and that $\Upsilon_{n_3,n_2,n_1}(T)$ is monotonic in both n_2 and n_3. Define $\Upsilon_{n_3,n_2}(T) = \lim_{n_1\to\infty} \Upsilon_{n_3,n_2,n_1}(T)$, $\Upsilon_{n_3}(T) = \lim_{n_2\to\infty} \Upsilon_{n_3,n_2}(T)$ and $\Upsilon(T) = \lim_{n_3\to\infty} \Upsilon_{n_3}(T)$. Then $\Upsilon(T)$ is zero if $\Xi_{a,b,\mathrm{sc}}(T) = 0$, otherwise it is a positive finite number.

With a slight change to the previous argument (the monotonicity in n_2 and n_3 is crucial for this to work), consider the intervals $J_1^m = [0, 1/m]$, and $J_2^m = [2/m, \infty)$. Let $k(m, n, n_1) \leq n_1$ be maximal such that $\Upsilon_{m,n,n_1}(T) \in J_1^m \cup J_2^m$. If no such k exists or $\Upsilon_{m,n,k}(T) \in J_1^m$ then set $\widehat{\Gamma}_{m,n,n_1}(T) = 0$. Otherwise set $\widehat{\Gamma}_{m,n,n_1}(T) = 1$. We then define

$$\Gamma_{n_3,n_2,n_1} = \max_{1\leq m\leq n_3} \min_{1\leq n\leq n_2} \widehat{\Gamma}_{m,n,n_1}(T).$$

These can be computed using finitely many arithmetic operations and comparisons using Λ_1, and, as before, the first limit exists with

$$\Gamma_{n_3,n_2}(T) = \lim_{n_1 \to \infty} \Gamma_{n_3,n_2,n_1}(T) = \max_{1 \le m \le n_3} \min_{1 \le n \le n_2} \widehat{\Gamma}_{m,n}(T).$$

Note also that the second and third sequential limits exist through the use of maxima and minima.

Now suppose that $\Xi_{a,b,\mathrm{sc}}(T) = 0$ and fix n_3. Then for large n_2, we must have that $\Upsilon_{m,n_2}(T) < 1/(2n_3)$ for all $m \le n_3$ due to the monotonic convergence of Υ_p as $p \to \infty$. It follows in this case that

$$\lim_{n_2 \to \infty} \Gamma_{n_3,n_2}(T) = 0, \quad \text{for all } n_3.$$

Now suppose that $\Xi_{a,b,\mathrm{sc}}(T) = 1$. It follows in this case that there exists $M \in \mathbb{N}$ such that if $m \ge M$ then $\Upsilon_m(T) > 3/m$. Due to the monotonic convergence of $\Upsilon_{m,p}$ as $p \to \infty$ it follows that for all p we must have $\Upsilon_{m,p} > 3/m$ and hence there exists $N(m, p) \in \mathbb{N}$ such that if $n_1 \ge N(m, p)$ then we must have $\Upsilon_{m,p,n_1} \ge 2/m$. It follows that if $n_3 \ge M$ then we must have $\widehat{\Gamma}_{n_3,p}(T) = 1$ for all p and hence that

$$\lim_{n_3 \to \infty} \Gamma_{n_3}(T) = 1.$$

The conclusion of the lemma now follows. $\quad \square$

6. Numerical Examples

We now demonstrate the applicability of the new algorithms. In particular, these are the first algorithms that compute their respective spectral properties for the whole class $\Omega_{f,\alpha}$, and even for the restricted case of tridiagonal self-adjoint matrices. The algorithms are also implicitly parallelisable, allowing large scale computations. We focus on discrete operators and the numerical implementation of the algorithms for ODEs, PDEs and integral operators will be the focus of a future software package.[8]

6.1. Jacobi matrices and orthogonal polynomials. For our first set of examples, we consider the natural link between the spectral measures of Jacobi operators and orthogonal polynomials on $\mathbb{R}$. Let J be a Jacobi matrix

$$J = \begin{pmatrix} b_1 & a_1 & & & \\ a_1 & b_2 & a_2 & & \\ & a_2 & b_3 & \ddots & \\ & & \ddots & \ddots & \end{pmatrix}$$

with $a_j, b_j \in \mathbb{R}$ and $a_j > 0$. In this case, under suitable conditions, the probability measure $\mu_J := \mu^J_{e_1,e_1}$ is the probability measure associated with the orthonormal polynomials defined by

$$x P_k(x) = a_{k+1} P_{k+1}(x) + b_{k+1} P_k(x) + a_k P_{k-1}(x),$$
$$P_{-1}(x) = 0, \quad P_0(x) = 1,$$

---

[8] Since writing the initial version of this paper, our algorithms have been implemented in the software package `SpecSolve` of [39], which treats general discrete, differential and integral operators.
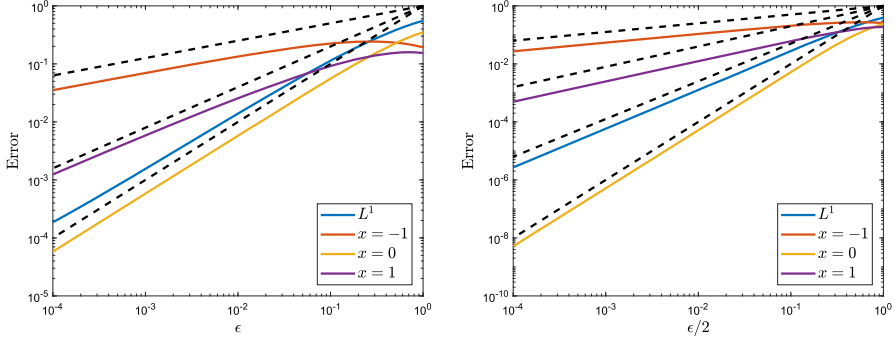
Fig. 4. Results for Jacobi polynomials with $\alpha = 0.7$ and $\beta = 0.3$. Left: Convergence in L^1 and at the points ± 1 and 0. The rates $O(\epsilon)$, $O(\epsilon^{0.7})$ and $O(\epsilon^{0.3})$ are also shown as dashed lines. Right: Convergence with Richardson extrapolation. The rates $O(\epsilon^2)$, $O(\epsilon^{1.3})$, $O(\epsilon^{0.7})$ and $O(\epsilon^{0.3})$ are also shown. As discussed in the text, the rates reflect the local smoothness properties of the density $f_{\alpha,\beta}$

and the spectral measure that appears in the multiplicative version of the spectral theorem (see, for example, [45,137,143]). Classically, one usually first considers the measure and then constructs the orthogonal polynomials and the corresponding J. In some sense, the algorithms constructed in this paper, and in particular Sect. 4, compute the inverse problem (note that $J \in \Omega_{n \to n+1,0}$). In other words, we compute the measure μ_J given the recurrence coefficients defining the orthogonal polynomials. This is a very difficult problem to study theoretically, and, until now, there has been no numerical method able to tackle the problem in this generality (see Sect. 1.7 for comments on methods that tackle compact perturbations of Toeplitz operators). To verify our method, we will consider problems where the measure μ_J is known analytically.

We begin with the well-known class of Jacobi polynomials defined for $\alpha, \beta > -1$ which have

$$a_k = 2\sqrt{\frac{k(k+\alpha)(k+\beta)(k+\alpha+\beta)}{(2k+\alpha+\beta-1)(2k+\alpha+\beta)^2(2k+\alpha+\beta+1)}},$$

$$b_k = \frac{\beta^2 - \alpha^2}{(2k+\alpha+\beta)(2k-2+\alpha+\beta)},$$

and measure on the interval $[-1, 1]$ given by

$$d\mu_J = \frac{(1-x)^\alpha (1+x)^\beta}{N(\alpha, \beta)} dx = f_{\alpha,\beta}(x)dx, \tag{6.1}$$

where $N(\alpha, \beta)$ is a normalising constant, ensuring the measure is a probability measure. To assess the convergence of the algorithm in Sect. 4.2 that approximates the Radon–Nikodym derivative $f_{\alpha,\beta}$, in this section we will plot various errors as a function of ϵ, the distance from the points at which we compute the resolvents to the real axis.

Figure 4 (left) shows a typical error plot for $\alpha = 0.7$ and $\beta = 0.3$. We plot both the L^1 error (computed using a large number of discrete points), and the pointwise errors at $-1, 0$ and 1. The procedure of Theorem 2.1 is used to determine adaptively how large our (rectangular) matrix truncations should be for a given ϵ. We see that both the L^1 error and error at 0 appear to decrease as $O(\epsilon),$[9] whereas the errors at -1 and

---

[9] This can be proven, though there is an additional $\log(\epsilon^{-1})$ factor. We shall omit such terms in the ensuing discussion.
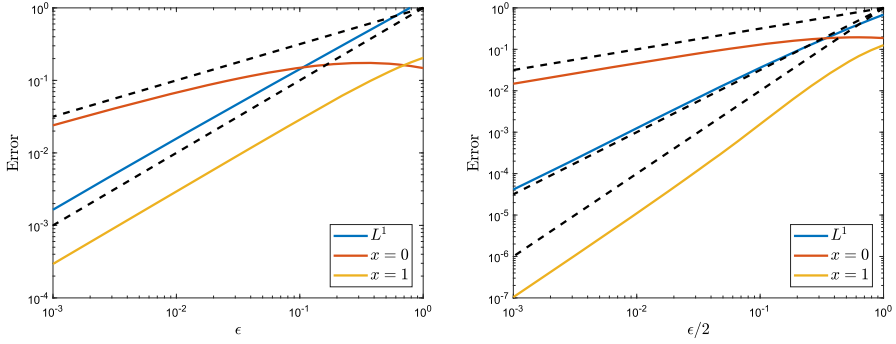
Fig. 5. Results for Laguerre polynomials with $\alpha = 0.5$. Left: Convergence in L^1 and at the points 0 and 1. The rates $O(\epsilon)$ and $O(\epsilon^{0.5})$ are also shown as dashed lines. Right: Convergence with Richardson extrapolation. The rates $O(\epsilon^2)$, $O(\epsilon^{1.5})$ and $O(\epsilon^{0.5})$ are also shown

+1 decrease as $O(\epsilon^\beta)$ and $O(\epsilon^\alpha)$ respectively, shown in the plot. This suggests using Richardson extrapolation to accelerate convergence [115]. This is shown in the right of Fig. 4, where the extrapolation was computed at distances ϵ and $\epsilon/2$ from the real axis. Now the error at 0 decreases as $O(\epsilon^2)$, whereas the L^1 error appears to decrease as $O(\epsilon^{1.3})$. We found similar results for different α and β. In general, interior points decrease at the rate $O(\epsilon)$ and then $O(\epsilon^2)$ after extrapolation. The left end point error decreases as $O(\epsilon^{\min\{1,\beta\}})$ and then $O(\epsilon^{\min\{2,\beta\}})$ after extrapolation. The right end point error decreases as $O(\epsilon^{\min\{1,\alpha\}})$ and then $O(\epsilon^{\min\{2,\alpha\}})$ after extrapolation. Finally, the L^1 error decreases as $O(\epsilon^{\min\{1,1+\alpha,1+\beta\}})$ and then $O(\epsilon^{\min\{2,1+\alpha,1+\beta\}})$ after extrapolation. These rates for pointwise and L^1 errors reflect the local Hölder exponent and integrability of the density and its first derivative respectively, and can be proven using a Taylor series argument for general measures.[10] Moreover, we found that increased rates of convergence could be obtained (and again proven) locally near smoother parts of the measure by using further iterates of extrapolation. Note also that we took a uniform value ϵ over the whole interval. However, ϵ could just have easily been a function of the position x, allowing it to be smaller for points where the resolvent is estimated more accurately for a given matrix size.

To demonstrate the algorithm on unbounded operators, we next consider the class of generalised Laguerre polynomials for $\alpha > -1$ which have

$$a_k = \sqrt{k(k+\alpha)}, \quad b_k = 2k + \alpha - 1,$$

and measure on the interval $[0, \infty)$ given by

$$d\mu_J = \frac{x^\alpha e^{-x}}{\Gamma(\alpha+1)} dx. \tag{6.2}$$

Results are shown in Fig. 5 for $\alpha = 0.5$, where we have plotted the (relative) L^1 error over the interval $[0, 1]$, as well as pointwise errors at 0 and 1. Similar conclusions can be drawn as before. Pointwise errors are also shown for this example and the Jacobi operator, but now using the 10th iterate of Richardson extrapolation, in Fig. 6. The errors decay at the expected rates (also shown), with $O(\epsilon^{10})$ convergence at smooth

---

[10] These results can be proven using the theory of high-order rational kernels developed in [39].
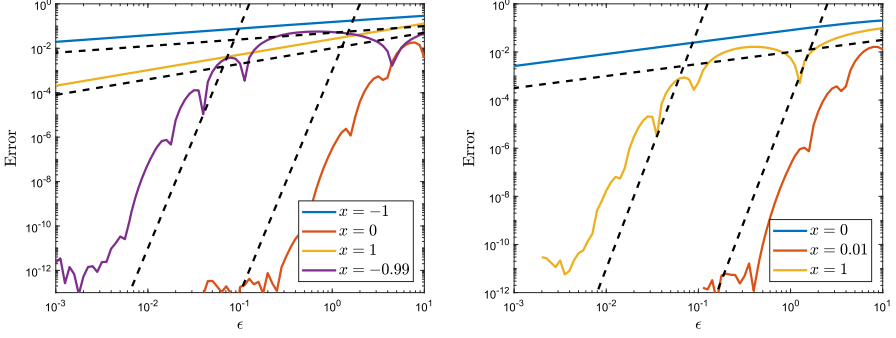
Fig. 6. Left: Pointwise errors for the Jacobi example ($\alpha = 0.7$ and $\beta = 0.3$) and (10th) iterated Richardson extrapolation. Right: Same but for the Laguerre example ($\alpha = 0.5$). The dotted lines show the expected rates of convergence, with $O(\epsilon^{10})$ convergence at smooth parts of the measure
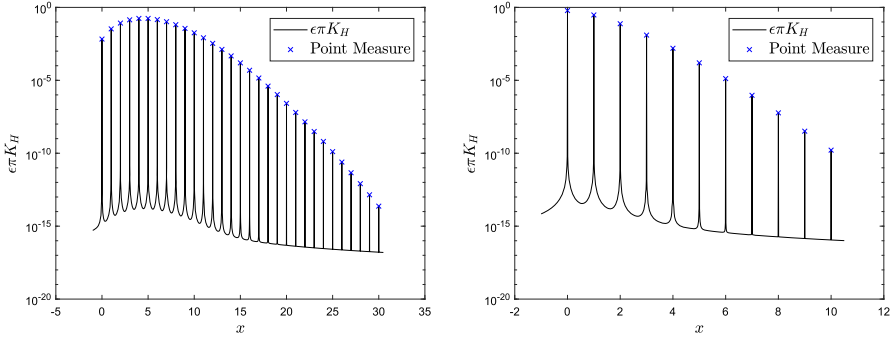

Fig. 7. Left: Computation of $\epsilon\pi\langle K_H(x + i\epsilon; T, e_1), e_1\rangle$ (denoted $\epsilon\pi\langle K_H\rangle$) for $\epsilon = 10^{-7}$ and $\alpha = 5$. Right: Same but for $\alpha = 0.5$. The blues crosses represent the weights of the atoms of the measure, corresponding to projections onto eigenspaces

parts of the measure. Near singular points (namely at $x = -0.99$ and $x = 0.01$ for the Jacobi and Laguerre cases, respectively), the prefactor in front of the $O(\epsilon^{10})$ term is larger, and smaller ϵ is needed before the expected rate kicks in.

Finally, we demonstrate the computation of measures for a Jacobi operator with non-empty discrete spectrum. The Charlier polynomials are generated by

$$a_k = \sqrt{\alpha k}, \quad b_k = k + \alpha - 1,$$

for $\alpha > 0$, and have measure

$$d\mu_J = \exp(-\alpha) \sum_{m=0}^{\infty} \frac{\alpha^m}{m!} \delta_m, \tag{6.3}$$

where δ_m denotes a Dirac measure located at the point m. Figure 7 shows plots of $\epsilon\pi\langle K_H(x + i\epsilon; T, e_1), e_1\rangle$ for $\epsilon = 10^{-7}$, computed using an $(n + 1) \times n$ matrix with $n = 1000$. The peaks clearly coincide with the atoms of the measure. The difference between the peak values and the weight $\exp(-\alpha)\alpha^m/m!$ was of the order 10^{-13} for both examples. This demonstrates that an effective way to compute eigenvalues (particularly

the challenging case of those in gaps of the essential spectrum, where spectral pollution occurs, or even those embedded in the essential spectrum) and projections onto eigenspaces may be to find local maxima or spikes of $\epsilon\pi\langle K_H(x+i\epsilon; T, e_1), e_1\rangle$. Such a routine would only require the solution of shifted linear systems (the resolvent), without diagonalisation, and could be executed rapidly in parallel.

6.2. A global collocation approach. Typically, the size of the linear system needed to approximate the resolvent accurately (Theorem 2.1) grows as $\epsilon \downarrow 0$ and we approach the spectrum. It is therefore beneficial to increase the rate of convergence of approximating the measures as $\epsilon \downarrow 0$. One local (in terms of $x \in \mathbb{R}$) approach, Richardson extrapolation, was used in Sect. 6.1. Here we briefly outline a different, more global approach.

Suppose that we know the spectral measure of an operator T has support included in $I \subset \mathbb{R}$ and is absolutely continuous. Alternatively, we may analytically know, or be able to estimate, the singular part of the measure and subtract this from what follows. In this case, a natural way to approximate the Radon–Nikodym derivative is through a formal basis expansion

$$\rho_{x,y}^T(\lambda) = \sum_{m=1}^{\infty} a_m \phi_m(\lambda),$$

where ϕ_m are functions with support I whose Cauchy's transforms are easy to compute. To approximate the coefficients a_m, we collocate in the complex plane as follows. Let $\mathcal{C}$ be a finite collection of complex points in the upper half-plane and truncate the approximation of $\rho_{x,y}^T$ to M terms. To generate a linear system for $\{a_m\}_{m=1}^M$, we evaluate the Cauchy transform at points $z \in \mathcal{C}$. The Cauchy transform satisfies

$$\int_{\mathbb{R}} \frac{\rho_{x,y}^T(\lambda)}{\lambda - z} d\lambda = \langle R(z, T)x, y\rangle,$$

which can be computed with error control using the results of Sect. 2.1. Call this approximation $b_{x,y}(z)$ and define

$$\widehat{\phi}_m(z) = \int_I \frac{\phi_m(\lambda)}{\lambda - z} d\lambda.$$

Then for each $z \in \mathcal{C}$, an approximate linear relation can be written as

$$\sum_{m=1}^{M} a_m \widehat{\phi}_m(z) = b_{x,y}(z).$$

Evaluating this relation at $\geq M$ points in $\mathcal{C}$ gives rise to a linear system, which can be inverted in the least-squares sense for the approximation of the coefficients $\{a_m\}_{m=1}^M$. If $x = y$, and the basis functions are real, then the coefficients are real. Hence, in this case, taking real and imaginary parts of the linear system gives further linear relations without having to compute further resolvents.

If our basis functions satisfy recursion relations of the form

$$\phi_{m+1}(\lambda) = \alpha_m \lambda \phi_m(\lambda) + \beta_m \phi_m(\lambda) + \gamma_m \phi_{m-1}(\lambda),$$
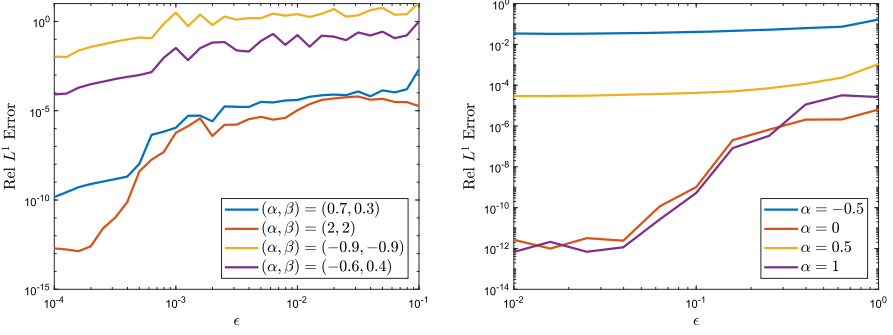
Fig. 8. Left: Convergence of collocation method using Chebyshev polynomials. Right: Convergence of collocation method using Laguerre functions on $[0, \infty)$

then

$$\widehat{\phi}_{m+1}(z) = \alpha_m \int_I \frac{\lambda \phi_m(\lambda)}{\lambda - z} d\lambda + \beta_m \widehat{\phi}_m(z) + \gamma_m \widehat{\phi}_{m-1}(z)$$

$$= \alpha_m \int_I \phi_m(\lambda) d\lambda + z \alpha_m \widehat{\phi}_m(z) + \beta_m \widehat{\phi}_m(z) + \gamma_m \widehat{\phi}_{m-1}(z).$$

Hence, given the values of the integrals

$$\int_I \phi_m(\lambda) d\lambda, \quad \widehat{\phi}_1(z), \tag{6.4}$$

we can compute $\widehat{\phi}_m(z)$ for all $m \in \mathbb{N}$. The integrals in (6.4) have simple forms for all the bases used in this paper. This method of computation of $\widehat{\phi}_m$ is fast, meaning that the most expensive part of the collocation method is the computation of the right hand side of the linear system, that is, computing the resolvent.

Figure 8 (left) revisits the Jacobi polynomials example in Sect. 6.1 and shows the collocation method in the case of using Chebyshev polynomials as the basis functions ϕ_m (taking the first 1500). For collocation points, we took M Chebyshev nodes offset by an additional ϵi to lie just above the interval $[-1, 1]$ in the complex plane. Note that the rate of convergence is faster than $O(\epsilon^{\min\{2, 1+\alpha, 1+\beta\}})$ achieved with the methods of Sect. 4 after extrapolation. There are at least two prices to pay, though. First, there is no current proof that collocation converges and, second, global regularity of the measure is needed. At the very least, we need to be able to subtract off the singular part of the measure. In practice, even if the measure is absolutely continuous, a large number of basis functions may be needed to capture the Radon–Nikodym derivative. Examples are given in Sect. 6.3, where collocation performs worse than the techniques of Sect. 4 due to either of these regularity conditions. Figure 8 (right) revisits the (generalised) Laguerre polynomials example in Sect. 6.1 and shows the collocation method in the case of using Laguerre functions (the polynomials multiplied by the square root of the weight function) as the basis functions with $M = 1000$. The collocation points were $\{1^2/M^2, 2^2/M^2, \ldots, 1\} + \epsilon i$. Again, this method converges with faster rates than that in Sect. 6.1.

6.3. CMV matrices and extensions to unitary operators. We now demonstrate that the algorithms extend to the unitary case through use of the functions K_D, namely, the convolution with the Poisson kernel of the unit disk. We will consider the class of CMV matrices (named after Cantero, Moral and Velázquez [28]) linked with orthogonal polynomials on the unit circle. A full discussion of this subject is beyond the scope of this paper, and we refer the reader to the monographs of Simon [130,131]. However, the background for this example is as follows. Given a probability measure μ on the unit circle $\mathbb{T}$, whose support is not a finite set, we can define a system of orthogonal polynomials $\{\Phi_n\}_{n=0}^{\infty}$ by applying the Gram–Schmidt process to $\{1, z, z^2, \ldots\}$. Given a polynomial $Q_n(z)$ of degree n, we define the reversed polynomial $Q_n^*(z)$ via $Q_n^*(z) = z^n \overline{Q_n(1/\overline{z})}$. Szegő's recurrence relation [139] is given by

$$\Phi_{n+1}(z) = z\Phi_n(z) - \overline{\alpha_n}\Phi_n^*(z), \tag{6.5}$$

where the α_n are known as Verblunsky coefficients [148] and satisfy $|\alpha_j| < 1$. Verblunsky's theorem [147] sets up a one-to-one correspondence between μ and the coefficients $\{\alpha_j\}_{j=0}^{\infty}$. Define also

$$\rho_j = \sqrt{1 - |\alpha_j|^2} > 0.$$

The CMV matrix associated with $\{\alpha_j\}_{j=0}^{\infty}$ is

$$C = \begin{pmatrix} \overline{\alpha_0} & \overline{\alpha_1}\rho_0 & \rho_1\rho_0 & 0 & 0 & \cdots \\ \rho_0 & -\overline{\alpha_1}\alpha_0 & -\rho_1\alpha_0 & 0 & 0 & \cdots \\ 0 & \overline{\alpha_2}\rho_1 & -\overline{\alpha_2}\alpha_1 & \overline{\alpha_3}\rho_2 & \rho_3\rho_2 & \cdots \\ 0 & \rho_2\rho_1 & -\rho_2\alpha_1 & -\overline{\alpha_3}\alpha_2 & -\rho_3\alpha_2 & \cdots \\ 0 & 0 & 0 & \overline{\alpha_4}\rho_3 & -\overline{\alpha_4}\alpha_3 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}. \tag{6.6}$$

This matrix is unitary and banded (and lies in $\Omega_{n \to n+2,0}^{\mathrm{U}}$). This last property does not hold for the so-called GGT representation [62,68,142], which has infinitely many non-zero entries in each row. The GGT representation uses the basis $\{\Phi_n\}_{n=0}^{\infty}$, whereas the CMV representation obtains a basis via applying Gram–Schmidt to $\{1, z, z^{-1}, z^2, z^2, \ldots\}$. The key result is that $\mu_C := \mu_{e_1,e_1}^C$ is precisely the measure μ on the unit circle. Hence our new algorithms can be considered as a computational tool for the correspondence

$$\{\alpha_j\}_{j=0}^{\infty} \to \mu,$$

in much the same way as for orthogonal polynomials on the real line in Sect. 6.1.

The first example we consider are the Rogers–Szegő polynomials [117] given by

$$\alpha_j = (-1)^j q^{(j+1)/2},$$

where $q \in (0, 1)$. In this case,

$$d\mu_C = \frac{1}{\sqrt{2\pi \log(q^{-1})}} \sum_{m \in \mathbb{Z}} \exp\left(-\frac{(\theta - 2\pi m)^2}{2\log(q^{-1})}\right) d\theta,$$

which can be expressed in terms of the theta function. Figure 9 (left) shows the convergence of the new algorithm for various q and we see algebraic convergence as before, with rates $O(\epsilon)$ and $O(\epsilon^2)$ before and after extrapolation respectively (here ϵ is such
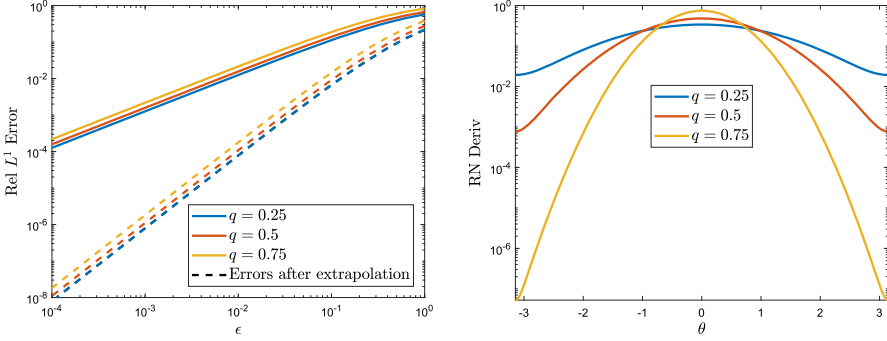
Fig. 9. Left: Convergence of algorithm for Rogers–Szegő polynomials. Right: Corresponding Radon–Nikodym derivatives (densities of the measure)

that we evaluate the convolutions with the Poisson kernel at radius $r = (1 + \epsilon)^{-1}$. Radon–Nikodym derivatives for larger values of q (shown in the right of Fig. 9) have larger derivatives and hence larger pre-factor in front of these rates. We can use a similar collocation method as in Sect. 6.2, using the standard Fourier basis $\{e^{im\theta}\}_{m \in \mathbb{Z}}$. Note that the relevant Cauchy transforms can be computed explicitly using Cauchy's residue theorem. Collocation points inside and outside the unit disk are needed (collocating inside the unit disk causes the Cauchy transforms of the basis functions with negative m to vanish). This achieved machine precision using just 41 basis functions and collocating at distance $\epsilon = 0.1$ from $\mathbb{T}$.

The next example we consider are the Geronimus polynomials [63], which have $\alpha_j = a$ with $|a| < 1$. In this case, for $|a + 1/2| \leq 1/2$,

$$d\mu_C = \chi_{|\theta| > \theta_a} \frac{\sqrt{\cos^2(\theta_a/2) - \cos^2(\theta/2)}}{2\pi \, |1 + a| \, |\sin((\theta - b)/2)|} d\theta,$$

where $\theta \in [-\pi, \pi]$, $b = 2\arg(1 + \overline{a})$ and $\theta_a = 2\sin^{-1}(|a|)$. Figure 10 (right) shows some typical examples, and note that these density functions are not smooth. When $|a + 1/2| > 1/2$, there is also a singular part of the measure (see [130] for the exact formula).

In the cases shown in Fig. 10, the collocation method struggles to gain an accuracy beyond 10^{-4} owing to discontinuity in the derivative of the Radon–Nikodym derivative and the algorithm based on convolutions is able to gain much more accurate results. Finally, a typical example is shown in Fig. 11 (left) for $a = 0.8$ (there is now a singular part located at $\theta = 0$), where we have shown the output of the algorithm and the exact convolution with the Poisson kernel for $r = 1/1.01$, as well as the collocation method using 21 basis functions and collocation points $\{(1 \pm 0.1)2\pi/21, (1 \pm 0.1)2\pi \cdot 2/21, \ldots, (1 \pm 0.1)2\pi\}$. Here, we see an exact agreement between the algorithm and convolution. Unsurprisingly in the presence of point spectra, the collocation method is unstable. Consistent with Theorem 4.2, the algorithm in Sect. 4.2 converges to the density over the portion of the spectrum, which is purely absolutely continuous. This is shown in the right of Fig. 11.

6.4. Fractional diffusion on a 2D quasicrystal. In this example, we demonstrate computation of the functional calculus and consider operators acting on the graph of a Penrose
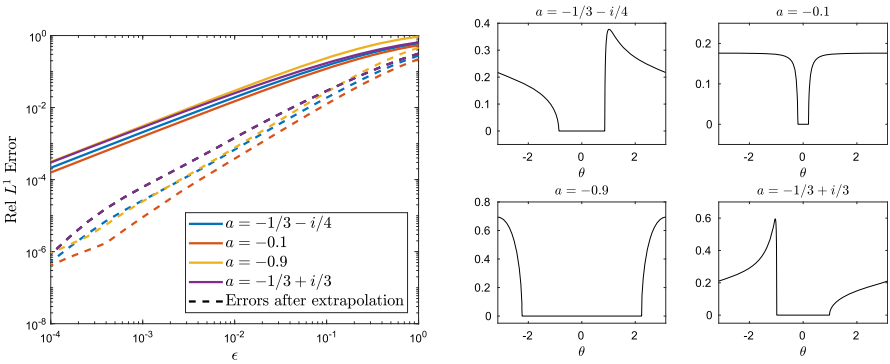
Fig. 10. Left: Convergence of algorithm for Geronimus polynomials. In this case the algorithm can obtain much more accurate results than collocation owing to the non-smoothness of the density functions. Right: Example density functions for the cases considered
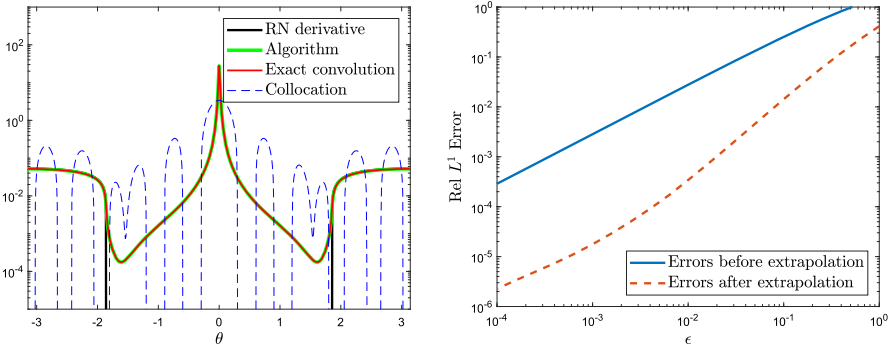

Fig. 11. Left: The smooth part of the density function (black) for the case of an additional point mass at $\theta = 0$. Note that the algorithm's output agrees almost perfectly with the exact convolution of the measure. We have also shown the output of the collocation method, which is unstable in the presence of the point mass. Right: Convergence of the algorithm from Theorem 4.2 over the portion of the spectrum which is purely absolutely continuous

tile—the canonical model of a two-dimensional quasicrystal [47,140,146] (aperiodic crystals which typically have anomalous spectra/transport properties). Quasicrystals were discovered in 1982 by Shechtman [125] who was awarded the 2011 Nobel Prize in Chemistry for his discovery. Since then, quasicrystals have generated considerable interest due to their often exotic physical properties [135], with a vast literature on the physics and spectral properties of such aperiodic systems.[11] Unlike the one-dimensional case, little is known about the spectral properties of two-dimensional quasicrystals. A finite portion of the infinite tile is shown in Fig. 12, and we consider the natural graph whose vertices are the vertices of the tiling and edges correspond to the edges of the rhombi. Such a graph posses no periodic structure, and it is generically impossible to study its spectral properties analytically with current methods. The free Hamiltonian H_0

---

[11] Recently, there has been renewed interest in aperiodic systems connected to graphene in the field of twistronics [30,97].
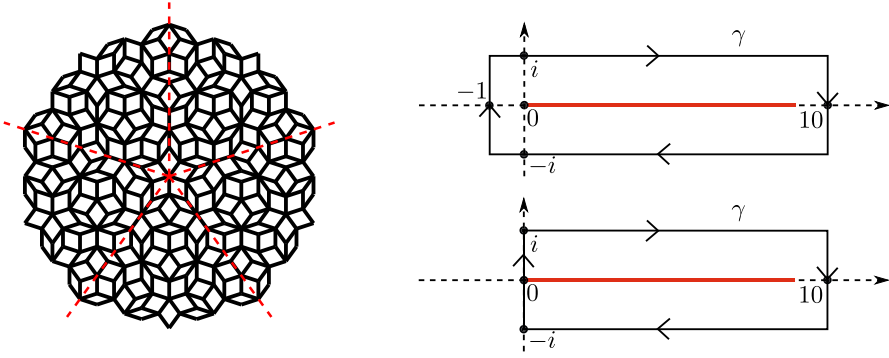
Fig. 12. Left: Finite portion of Penrose tile showing the fivefold rotational symmetry. We labelled vertices from the centre in a spiral outwards in increasing distance from the origin. Right: Contours used for the fractional diffusion on the Penrose tile ($\alpha \in \mathbb{N}$ top, $\alpha \notin \mathbb{N}$ bottom). The red line represents the interval containing the spectrum of $-H_0$, the branch cut for z^α is taken to be $\mathbb{R}_{\leq 0}$
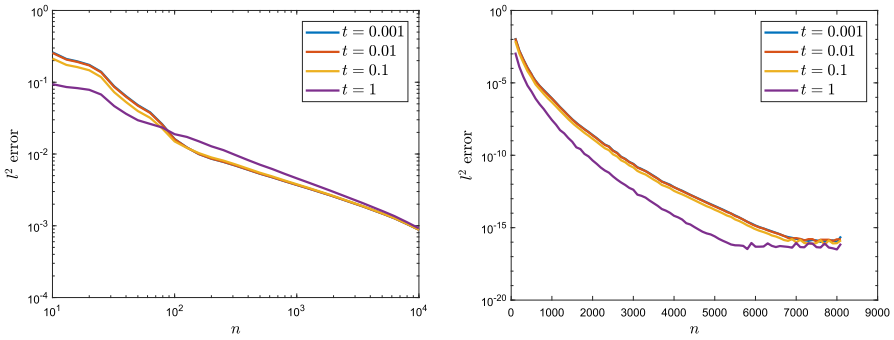

Fig. 13. Left: Convergence for $\alpha = 1/2$. Right: Convergence for $\alpha = 1$. We have plotted the errors as a function of the matrix size (number of matrix columns in the rectangular truncations) used

(Laplacian) is given by

$$(H_0\psi)_i = \sum_{i\sim j}\left(\psi_j - \psi_i\right), \tag{6.7}$$

with summation over nearest neighbour sites (vertices). The first rigorous computation of the spectrum of H_0 (also with additional error control) was performed in [40]. We chose a natural ordering of the vertices as in Fig. 12, which leads to an operator H_0 acting on $l^2(\mathbb{N})$. The local bandwidth grows for this operator (our ordering is asymptotically optimal) and hence computation of powers H_0^m is infeasible for $m \gtrsim 50$, rendering polynomial approximations of the functional calculus intractable. In the above notation, $H_0 \in \Omega_{f,0}$ with $f(n) - n = O(\sqrt{n})$, and so this example provides a demonstration of the algorithm for a non-banded matrix. Throughout, we take $u_0 = e_1$, though different initial conditions are handled in the same manner.

The ability to compute the functional calculus allows the solution of linear evolution equations. Given $A \in \Omega_N$, a function F (continuous and bounded on $\sigma(A)$) and $u_0 \in l^2(\mathbb{N})$, consider the evolution equation

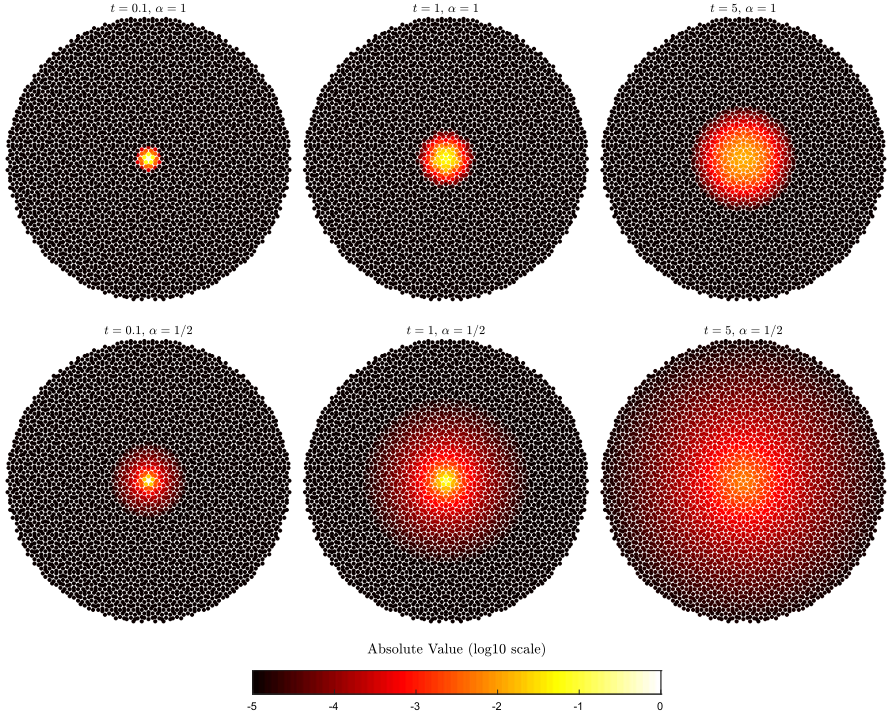$$\frac{du}{dt} = F(A)u, \quad u_{t=0} = u_0. \tag{6.8}$$

Fig. 14. Evolution of initial wavepacket under fractional diffusion

The solution of this equation is

$$u(t) = \exp(F(A)t)u_0$$

and can be computed via the algorithm outlined in Sect. 4.1.

We consider fractional diffusion governed by

$$\frac{du}{dt} = -(-H_0)^\alpha u, \quad u_{t=0} = u_0,$$

for $\alpha > 0$. If α is an integer, then the solution can be represented via contour deformation as

$$u(t) = \frac{1}{2\pi i} \int_\gamma \exp(-z^\alpha t) R(z, -H_0) u_0 dz, \tag{6.9}$$

where γ is a closed contour looping once around the spectrum. Typically we took the rectangular contour shown in Fig. 12 and approximated the integral via Gaussian quadrature. This allows us to compute the solution with error control (we known the minimal distance between γ and $\sigma(-H_0)$ so can bound the Lipschitz constant of the resolvent) and clearly, this holds for other functions F, holomorphic on a neighbourhood of $\sigma(-H_0)$. Note that other methods, such as direct diagonalisation of finite square truncations or discrete time stepping (which is difficult if $\alpha \notin \mathbb{N}$), do not give error control and are slower. In fact, for this example, direct diagonalisation was impractical.

When $\alpha \notin \mathbb{N}$, we can still deform the contour, but not at 0 since $0 \in \sigma(-H_0)$. Hence we deform the contour to that shown in Fig. 12. For a discussion of contour methods applied to finite matrices (in the case that the spectrum is strictly positive), we refer the reader to [74]. Unfortunately, such methods cannot be applied here since $0 \in \sigma(-H_0)$. Moreover, 0 appears to not be an isolated point of the spectrum, meaning that dealing with this point separately is also impossible.

Figure 13 shows the convergence of the algorithm for $\alpha = 1/2$ and $\alpha = 1$. For $\alpha = 1/2$, error control is not given by our algorithm, so we computed an error by comparing to a "converged" solution using larger n. The l^2 error refers to the error in $l^2(\mathbb{N})$. The method converges algebraically for $\alpha = 1/2$ (owing to the contour touching the spectrum at 0) but converges exponentially for $\alpha = 1$ with similar convergence observed over a large range of times t. Figure 14 shows the magnitude (log scale) of the computed solution for various times. Note that the techniques presented here can be applied to any evolution equation of the form (6.8) on *infinite-dimensional* Hilbert spaces. They may also be useful for splitting methods/exponential integrators, which require fast computation of matrix/operator exponentials (see [83,100] and the references therein) and more generally in the field of infinite-dimensional ODE/PDE systems. For methods that compute general semigroups on infinite-dimensional separable Hilbert spaces with error control, see [35].

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

A. The SCI Hierarchy: A Framework for Computation

Cornerstones in the SCI hierarchy are the definitions of a computational problem, a general algorithm and towers of algorithms. The basic objects in a computational problem are as follows:

(i) Ω is some set, called the *domain*.
(ii) Λ is a set of complex-valued functions on Ω called the *evaluation* set.
(iii) $\mathcal{M}$ is a metric space with metric $d_{\mathcal{M}}$.
(iv) $\Xi : \Omega \to \mathcal{M}$ is called the *problem function*.

The set Ω is the set of objects that give rise to our computational problems. The problem function $\Xi : \Omega \to \mathcal{M}$ is what we are interested in computing. Moreover, the set Λ is the collection of functions that provide us with the information we are allowed to read.

Remark 7. Throughout the paper we have relaxed the condition that $\lambda \in \Lambda$ maps into $\mathbb{C}$ by considering evaluation functions consisting of intervals exhausting a set, piecewise

constant functions of compact support etc. These seemingly more complicated objects can be effectively encoded by functions that map into $\mathbb{C}$. For example, when considering the decomposition (3.1):

$$U = \bigcup_m (a_m(U), b_m(U))$$

of an open set U, we consider $\lambda_{1,m}, \lambda_{2,m}$ with $\lambda_{1,m}(U) = a_m$ and $\lambda_{2,m}(U) = b_m$. For the sake of clarity of presentation of the proofs, such encodings are used implicitly throughout the paper.

This leads to the following definition.

Definition A.1 (*Computational problem*). Given a primary set Ω, an evaluation set Λ, a metric space $\mathcal{M}$ and a problem function $\Xi : \Omega \to \mathcal{M}$ we call the collection $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ a computational problem.

The goal is to find algorithms that approximate the function Ξ. More generally, we need the concept of a tower of algorithms, which is needed to describe problems that need several limits in the computation. However, first one needs the definition of a general algorithm.

Definition A.2 (*General algorithm*). Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$, a general algorithm is a mapping $\Gamma : \Omega \to \mathcal{M}$ such that for each $A \in \Omega$

(i) There exists a finite non-empty subset of evaluations $\Lambda_\Gamma(A) \subset \Lambda$,
(ii) The action of Γ on A only depends on $\{A_f\}_{f \in \Lambda_\Gamma(A)}$ where $A_f := f(A)$,
(iii) For every $B \in \Omega$ such that $B_f = A_f$ for every $f \in \Lambda_\Gamma(A)$, it holds that $\Lambda_\Gamma(B) = \Lambda_\Gamma(A)$.

Note that the definition of a general algorithm is more general than the definition of a Turing machine or a Blum–Shub–Smale (BSS) machine. A general algorithm has no restrictions on the operations allowed. The only restriction is that it can only take a finite amount of information, though it is allowed to *adaptively* choose the finite amount of information it reads depending on the input. Condition (iii) assures that the algorithm consistently reads the information. Note that the purpose of such a general definition is to get strong lower bounds. In particular, the more general the definition is, the stronger a lower bound will be.

With a definition of a general algorithm, we can define the concept of towers of algorithms.

Definition A.3 (*Tower of algorithms*). Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$, a tower of algorithms of height k for $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ is a family of sequences of functions

$$\Gamma_{n_k} : \Omega \to \mathcal{M}, \ \Gamma_{n_k, n_{k-1}} : \Omega \to \mathcal{M}, \ldots, \ \Gamma_{n_k, \ldots, n_1} : \Omega \to \mathcal{M},$$

where $n_k, \ldots, n_1 \in \mathbb{N}$ and the functions $\Gamma_{n_k, \ldots, n_1}$ at the "lowest level" of the tower are general algorithms in the sense of Definition A.2. Moreover, for every $A \in \Omega$,

$$\Xi(A) = \lim_{n_k \to \infty} \Gamma_{n_k}(A), \quad \Gamma_{n_k, \ldots, n_{j+1}}(A) = \lim_{n_j \to \infty} \Gamma_{n_k, \ldots, n_j}(A) \quad j = k-1, \ldots, 1.$$

In addition to a general tower of algorithms (defined above), we will focus on arithmetic towers. The definition of a general algorithm allows for strong lower bounds; however, to produce upper bounds we must add structure to the algorithm and towers of algorithms. An arithmetic tower allows for arithmetic operations and comparisons.

Definition A.4 (*Arithmetic towers*). Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$, where Λ is countable, we define the following: An arithmetic tower of algorithms of height k for $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ is a tower of algorithms where the lowest functions $\Gamma = \Gamma_{n_k,\ldots,n_1} : \Omega \to \mathcal{M}$ satisfy the following: For each $A \in \Omega$ the mapping $(n_k, \ldots, n_1) \mapsto \Gamma_{n_k,\ldots,n_1}(A) = \Gamma_{n_k,\ldots,n_1}(\{A_f\}_{f\in\Lambda})$ is recursive, the action of Γ on A consists of only finitely many arithmetic operations and comparisons on $\{A_f\}_{f\in\Lambda_\Gamma(A)}$, and $\Gamma_{n_k,\ldots,n_1}(A)$ is a finite string of complex numbers that can be identified with an element in $\mathcal{M}$.

Remark 8 (Recursiveness). By recursive we mean the following. If $f(A) \in \mathbb{Q}$ for all $f \in \Lambda$, $A \in \Omega$, and Λ is countable, then $\Gamma_{n_k,\ldots,n_1}(\{A_f\}_{f\in\Lambda})$ can be executed by a Turing machine [145], that takes $(n_k, \ldots, n_1)$ as input, and that has an oracle tape consisting of $\{A_f\}_{f\in\Lambda}$. If $f(A) \in \mathbb{R}$ (or $\mathbb{C}$) for all $f \in \Lambda$, then $\Gamma_{n_k,\ldots,n_1}(\{A_f\}_{f\in\Lambda})$ can be executed by a Blum–Shub–Smale machine [17] that takes $(n_k, \ldots, n_1)$, as input, and that has an oracle that can access any A_f for $f \in \Lambda$.

Given the above definitions, we can now define the Solvability Complexity Index:

Definition A.5 (*Solvability complexity index*). A computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ is said to have Solvability Complexity Index $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_\alpha = k$, with respect to a tower of algorithms of type α, if k is the smallest integer for which there exists a tower of algorithms of type α of height k. If no such tower exists then $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_\alpha = \infty$. If there exists a tower $\{\Gamma_n\}_{n\in\mathbb{N}}$ of type α and height one such that $\Xi = \Gamma_{n_1}$ for some $n_1 < \infty$, then we define $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_\alpha = 0$. We may sometimes write $\mathrm{SCI}(\Xi, \Omega)_\alpha$ to simplify notation when $\mathcal{M}$ and Λ are obvious.

The definition of the SCI immediately induces the SCI hierarchy:

Definition A.6 (*The solvability complexity index hierarchy*). Consider a collection $\mathcal{C}$ of computational problems and let $\mathcal{T}$ be the collection of all towers of algorithms of type α for the computational problems in $\mathcal{C}$. Define

$$\Delta_0^\alpha := \{\{\Xi, \Omega\} \in \mathcal{C} \mid \mathrm{SCI}(\Xi, \Omega)_\alpha = 0\}$$
$$\Delta_{m+1}^\alpha := \{\{\Xi, \Omega\} \in \mathcal{C} \mid \mathrm{SCI}(\Xi, \Omega)_\alpha \le m\}, \qquad m \in \mathbb{N},$$

as well as

$$\Delta_1^\alpha := \{\{\Xi, \Omega\} \in \mathcal{C} \mid \exists \{\Gamma_n\}_{n\in\mathbb{N}} \in \mathcal{T} \text{ s.t. } \forall A \, d(\Gamma_n(A), \Xi(A)) \le 2^{-n}\}.$$

In certain cases, the SCI hierarchy can also be refined with notions of error control—see [12,33,36,37]. We also need the following result.

Proposition A.7. *Given a matrix $B \in \mathbb{C}^{m\times n}$ and a number $\epsilon > 0$ one can test with finitely many arithmetic operations on the entries of B whether the smallest singular value $\sigma_1(B)$ of B is greater than ϵ.*

Proof. The matrix B^*B is self-adjoint and positive semidefinite, hence has its eigenvalues in $[0, \infty)$. The singular values of B are the square roots of these eigenvalues of B^*B. The smallest singular value is greater than ϵ if and only if the smallest eigenvalue of B^*B is greater than ϵ^2, which is the case if and only if $C := B^*B - \epsilon^2 I$ is positive definite. The matrix C is positive definite if and only if the pivots left after Gaussian elimination (without row exchange) are all positive. Thus, if C is positive definite, Gaussian elimination leads to pivots that are all positive, and this requires finitely many arithmetic

operations. If C is not positive definite, then at some point a pivot is zero or negative, at this point the algorithm aborts. An alternative is the Cholesky decomposition. Although forming the lower triangular $L \in \mathbb{C}^{n \times n}$ (if it exists) such that $C = LL^*$ requires the use of radicals, the existence of L can be determined using finitely many arithmetic operations. This follows from the standard Cholesky algorithm, and we omit the details.

<div align="right">□</div>

Finally, we also need the following result.

Proposition A.8. *Let $(\mathcal{M}', d')$ be the discrete space $\{0, 1\}$, let Ω' denote the collection of all infinite sequences $\{a_j\}_{j \in \mathbb{N}}$ with entries $a_j \in \{0, 1\}$, let Λ' consist of pointwise evaluations of the $\{a_j\}$ and consider the problem function*

$$\Xi'(\{a_j\}) : \ \text{'Does } \{a_j\} \text{ have infinitely many non-zero entries?'}$$

Then $\Delta_2^G \not\ni \{\Xi', \Omega', \Lambda'\} \in \Delta_3^A$

Proof. Consider the arithmetic tower of algorithms defined by

$$\Gamma_{m,n}(\{a_j\}) = \begin{cases} 1 & \text{if } \sum_{j=1}^{n} a_j > m, \\ 0 & \text{otherwise.} \end{cases}$$

This provides a height two arithmetic tower for Ξ' and hence $\{\Xi', \Omega', \Lambda'\} \in \Delta_3^A$.

To prove the lower bound, suppose for a contradiction that $\{\Gamma_n\}$ is a sequence of general algorithms, using Λ', such that

$$\lim_{n \to \infty} \Gamma_n(\{a_j\}) = \Xi'(\{a_j\}).$$

We will construct a sequence $\{a_j\}$ such that $\Gamma_n(\{a_j\})$ does not converge, providing the required contradiction.

Set $\{a_j^1\}_{j \in \mathbb{N}} = \{0, 0, \ldots\}$, then there exists $n_1 \in \mathbb{N}$ such that $\Gamma_{n_1}(\{a_j^1\}) = 0$. Moreover, $\Lambda_{\Gamma_{n_1}}(\{a_j^1\}) \subset \{\{a_j\}_{j \in \mathbb{N}} \to a_m : m \leq N_1\}$ for some integer N_1 by (i) of Definition A.2. We choose $\{a_j^2\}_{j \in \mathbb{N}}$ such that $a_j^2 = a_j^1$ for $j \leq N_1$ and $a_j^2 = 1$ otherwise. Then there exists $n_2 \in \mathbb{N}$ such that $n_2 > n_1$ and $\Gamma_{n_2}(\{a_j^2\}) = 1$. Moreover, $\Lambda_{\Gamma_{n_2}}(\{a_j^2\}) \subset \{\{a_j\}_{j \in \mathbb{N}} \to a_m : m \leq N_2\}$ for some integer $N_2 > N_1$ by (i) of Definition A.2. We now repeat this inductively. Explicitly, given the construction up to the kth step, we define $\{a_j^{k+1}\}_{j \in \mathbb{N}}$ by $a_j^{k+1} = a_j^k$ for all $j \leq N_k$ and $a_j^{k+1} = (1+(-1)^{k+1})/2$ otherwise. Then there exists $n_{k+1} \in \mathbb{N}$ such that $n_{k+1} > n_k$ and $\Gamma_{n_{k+1}}(\{a_j^{k+1}\}) = (1+(-1)^{k+1})/2$. Moreover, $\Lambda_{\Gamma_{n_{k+1}}}(\{a_j^{k+1}\}) \subset \{\{a_j\}_{j \in \mathbb{N}} \to a_m : m \leq N_{k+1}\}$ for some integer $N_{k+1} > N_k$ by (i) of Definition A.2.

Finally set $a_j = a_j^k$ for $j \leq N_k$. It is clear from (iii) of Definition A.2, that $\Gamma_{n_k}(\{a_j\}) = \Gamma_{n_k}(\{a_j^k\}) = (1+(-1)^k)/2$ and this implies that $\Gamma_n(\{a_j\})$ cannot converge, the required contradiction. □

B. Partial Differential Operators: Proof of Theorem 1.1

Before stating and proving a more mathematically precise version of Theorem 1.1, we need to be precise about the computational problems involved. Recall that we consider L formally defined on $L^2(\mathbb{R}^d)$ by

$$Lu(x) = \sum_{k \in \mathbb{Z}^d_{\geq 0}, |k| \leq N} a_k(x) \partial^k u(x)$$

with the class Ω_{PDE} consisting of self-adjoint L satisfying certain conditions given in Sect. 1.2.1. We take Λ_{PDE} to consist of

$$S_{k,q,m} : \Omega_{\text{PDE}} \to \mathbb{Q} + i\mathbb{Q},$$

where $|S_{k,q,m}(L) - a_k(q)| \leq 2^{-m}$ for all $q \in \mathbb{Q}$, together with

$$b_n : \Omega_{\text{PDE}} \to \mathbb{Q}_{>0}$$

such that

$$\sup_{n \in \mathbb{N}} \max_{|k| \leq N} \frac{\|a_k\|_{\mathcal{A}_n}}{b_n(L)} < \infty.$$

We also implicitly assume that for any given L, the dimension d and integer N (order) are known. As well as this, we need to consider the pairs of vectors (members of $L^2(\mathbb{R}^d)$) with which we compute the spectral properties. By the polarisation identity, we can consider equal functions of norm 1. We let V_{PDE} consist of all $f \in L^2(\mathbb{R}^d)$ of norm 1 such that

(1) There exists a positive constant C and an integer D (both possibly unknown) such that

$$|f(x)| \leq C \left(1 + |x|^{2D}\right),$$

almost everywhere on $\mathbb{R}^d$, that is, f is polynomially bounded.
(2) The restrictions $f|_{[-r,r]^d} \in \mathcal{A}_r$ for all $r > 0$.

We then add to Λ_{PDE} the evaluation functions

$$S_{q,m} : V_{\text{PDE}} \to \mathbb{Q} + i\mathbb{Q},$$

where $|S_{q,m}(f) - f(q)| \leq 2^{-m}$ for all $q \in \mathbb{Q}$, together with

$$c_n : V_{\text{PDE}} \to \mathbb{Q}_{>0}$$

such that

$$\sup_{n \in \mathbb{N}} \frac{\|f\|_{\mathcal{A}_n}}{c_n(f)} < \infty.$$

With an abuse of notation, we can consider the computational problem from Theorem 3.1

$$\Xi_{\text{meas}} : \Omega_{\text{PDE}} \times V_{\text{PDE}} \times \mathcal{U} \to L^2(\mathbb{R}^d)$$
$$(L, f, U) \to E_U^L f.$$

Recall that $\mathcal{U}$ is the collection of non-trivial open sets and we have access to a finite or countable collection $a_m(U), b_m(U) \in \mathbb{R} \cup \{\pm\infty\}$ such that $U \in \mathcal{U}$ can be written as a disjoint union

$$U = \bigcup_m (a_m(U), b_m(U)).$$

Similarly, from Theorem 4.1 we consider

$$\Xi_{\text{fun}} : \Omega_{\text{PDE}} \times V_{\text{PDE}} \times C_b(\mathbb{R}) \to L^2(\mathbb{R}^d)$$
$$(L, f, F) \to F(L)f.$$

We assume that given $F \in C_b(\mathbb{R})$, we have access to piecewise constant functions F_n supported in $[-n, n]$ such that $\|F - F_n\|_{L^\infty([-n,n])} \leq n^{-1}$. Finally, from Theorem 4.2 we consider

$$\Xi_{\text{RN}} : \Omega_{\text{PDE}} \times V_{\text{PDE}} \times \mathcal{U} \to L^1(\mathbb{R})$$
$$(L, f, U) \to \rho^L_{f,f}|_U.$$

We restrict this map to the quadruples (L, f, U) such that U is strictly separated from $\text{supp}(\mu^L_{f,f,\text{sc}}) \cup \text{supp}(\mu^L_{f,f,\text{pp}})$ and denote this subclass by $\widetilde{\Omega}_{\text{PDE}}$.

We can now state the precise form of Theorem 1.1.

Theorem B.1 (Precise form of Theorem 1.1). *Given the above set-up,*

$$\{\Xi_{\text{meas}}, \Omega_{\text{PDE}} \times V_{\text{PDE}} \times \mathcal{U}, \Lambda_{\text{PDE}}\} \in \Delta^A_2,$$
$$\{\Xi_{\text{fun}}, \Omega_{\text{PDE}} \times V_{\text{PDE}} \times C_b(\mathbb{R}), \Lambda_{\text{PDE}}\} \in \Delta^A_2,$$
$$\{\Xi_{\text{RN}}, \widetilde{\Omega}_{\text{PDE}}, \Lambda_{\text{PDE}}\} \in \Delta^A_2.$$

In other words, we can construct a convergent sequence of arithmetic algorithms for each problem.

Remark 9. The proof will make clear that we can assume different conditions on the operator L and function f. We simply choose an appropriate basis so that we can apply Theorem B.2. We can also extend this to scalar measures (through inner products) and also to the towers of algorithms (of height ≥ 2) used to compute the decompositions into pure point, absolutely continuous and singular continuous parts of measures and spectra.

To prove Theorem B.1, we need the following theorem, which is similar to the results of Sect. 2.1.

Theorem B.2. *Consider the class $\Omega_{\text{SA}} \times S_{l^2(\mathbb{N})}$, where $S_{l^2(\mathbb{N})}$ denotes the unit sphere (vectors of norm 1). Assume that for each $(T, x) \in \Omega_{\text{SA}} \times S_{l^2(\mathbb{N})}$ we have access to evaluation functions $\hat{\Lambda} = \{f^{(1)}_{i,j,m}, f^{(2)}_{i,j,m}, f^{(3)}_{i,m} : i, j, m \in \mathbb{N}\}$ with*

$$\left| f^{(1)}_{i,j,m}(T) - \langle Te_j, e_i \rangle \right|, \left| f^{(2)}_{i,j,m}(T) - \langle Te_j, Te_i \rangle \right|, \left| f^{(3)}_{i,m}(x) - \langle x, e_i \rangle \right|$$
$$= O(2^{-m}), \quad \forall i, j \in \mathbb{N},$$

where the hidden constant depends only on T and x. Then there exists a sequence of arithmetic algorithms using $\hat{\Lambda}$

$$\Gamma_N : \Omega_{\text{SA}} \times S_{l^2(\mathbb{N})} \times \mathbb{C}\backslash\mathbb{R} \to l^2(\mathbb{N})$$

with the following properties:

1. *For all $(T, x, z) \in \Omega_{SA} \times S_{l^2(\mathbb{N})} \times \mathbb{C}\backslash\mathbb{R}$, $\Gamma_N(T, x, z)$ has finite support with respect to the canonical basis and converges to $R(z, T)x$ in $l^2(\mathbb{N})$ as $N \to \infty$.*

2. *For any $(T, x) \in \Omega_{SA} \times S_{l^2(\mathbb{N})}$, there exists a constant $C(T, x)$ such that for all $z \in \mathbb{C}\backslash\mathbb{R}$,*

$$\|\Gamma_N(T, x, z) - R(z, T)x\| \leq \frac{C(T, x)}{N \, |\mathrm{Im}(z)|}. \tag{B.1}$$

Proof. The proof is similar to that of Theorem 2.1. Let $(T, x, z) \in \Omega_{SA} \times S_{l^2(\mathbb{N})} \times \mathbb{C}\backslash\mathbb{R}$. We have that $n = \mathrm{rank}(P_n) = \mathrm{rank}((T - zI)P_n)$ for large n since $\sigma_1(T - zI) > 0$ (recall that $z \notin \sigma(T)$). Hence we can define

$$\widetilde{\Gamma}_n(T, x, z) := \begin{cases} 0 & \text{if } \sigma_1(P_n(T^* - \overline{z}I)(T - zI)P_n) \leq \frac{1}{n} \\ [P_n(T^* - \overline{z}I)(T - zI)P_n]^{-1} P_n(T^* - \overline{z}I)x & \text{otherwise.} \end{cases}$$

Suppose that n is large enough so that $\sigma_1(P_n(T^* - \overline{z}I)(T - zI)P_n) > 1/n$. Then $\widetilde{\Gamma}_n(T, x, z)$ is a least-squares solution of the optimisation problem $\mathrm{argmin}_y \|(T-zI)P_n y - x\|$. The linear space $\mathrm{span}\{e_n : n \in \mathbb{N}\}$ forms a core of T and hence also of $T - zI$. It follows by invertibility of $T - zI$ that given any $\epsilon > 0$, there exists an $m = m(\epsilon)$ and a $y = y(\epsilon)$ with $P_m y = y$ such that

$$\|(T - zI)y - x\| \leq \epsilon.$$

It follows that for all $n \geq m$, $\|(T - zI)\widetilde{\Gamma}_n(T, x, z) - x\| \leq \|(T - zI)y - x\| \leq \epsilon$ and hence that

$$\|\widetilde{\Gamma}_n(T, x, z) - R(z, T)x\| \leq \frac{\epsilon}{|\mathrm{Im}(z)|}.$$

Since $\epsilon > 0$ was arbitrary, we see that $\widetilde{\Gamma}_n(T, x, z)$ converges to $R(z, T)x$.

For $n, m \in \mathbb{N}$, define the finite matrices

$$B_n = P_n(T^* - \overline{z}I)(T - zI)P_n, \quad C_{m,n} = P_n(T^* - \overline{z}I)P_m.$$

Given the evaluation functions in $\hat{\Lambda}$, we have access to the entries of these matrices to asymptotic accuracy (i.e. for a given diverging subsequence a_n, to precision $O(2^{-a_n})$). It follows that we have access to approximations of B_n and $C_{m,n}$ denoted $\widetilde{B}_n$ and $\widetilde{C}_{m,n}$ respectively with

$$\|B_n - \widetilde{B}_n\| = O(n^{-1}), \quad \|C_{m,n} - \widetilde{C}_{m,n}\| = O(n^{-1}).$$

Recall that the $O(\cdot)$ notation also means independently of z and other parameters (though it may depend on T and x). Note that $\widetilde{B}_n^{-1}$ can be computed using finitely many arithmetic operations and comparisons and the resolvent identity implies that

$$\|B_n^{-1} - \widetilde{B}_n^{-1}\| = O(n^{-1}).$$

From the proof of Proposition A.7 and a simple search routine, we can also compute $\sigma_1(\widetilde{B}_n)$ to accuracy n^{-2} using finitely many arithmetic operations and comparisons. Denote the approximation via τ_n. We then define

$$\Gamma_{m,n}(T, x, z) := \begin{cases} 0 & \text{if } \tau_n \leq \frac{1}{n} \\ \widetilde{B}_n^{-1}\widetilde{C}_{m,n}x_{(m)} & \text{otherwise,} \end{cases}$$

where $x_{(m)} = P_m x_{(m)}$ is an approximation of $P_m x$ to accuracy $O(m^{-1})$. It follows that $\Gamma_{m,n}(T, x, z)$ can be computed using finitely many arithmetic operations. We also have that

$$\|\Gamma_{m,n}(T, x, z) - \tilde{\Gamma}_n(T, x_{(m)}, z)\| \le \|B_n^{-1} - \tilde{B}_n^{-1}\| \|C_{m,n}\| \|x_{(m)}\| + \|\tilde{B}_n^{-1}\| \|C_{m,n} - \tilde{C}_{m,n}\| \|x_{(m)}\|,$$

so that $\Gamma_{m,n}(T, x, z)$ converges to $R(z, T)x_{(m)}$ as $n \to \infty$. By construction, $\Gamma_{m,n}(T, x, z)$ has finite support with respect to the canonical basis. Furthermore, since $P_m x_{(m)} = x_{(m)}$, the following error bound holds for any $l \ge m$

$$\begin{aligned} &\|\Gamma_{m,n}(T, x, z) - R(z, T)x_{(m)}\| \\ &\quad \le \|R(z, T)\| \|(T - zI)\Gamma_{m,n}(T, x, z) - x_{(m)}\| \\ &\quad \le \frac{\|(I - P_l)(T - zI)P_n\| \|\Gamma_{m,n}(T, x, z)\| + \|P_l(T - zI)\Gamma_{m,n}(T, x, z) - x_{(m)}\|}{|\text{Im}(z)|}. \end{aligned}$$

Since we have access to both $\langle Te_j, Te_j \rangle$ (norms of the columns of T) and $\langle Te_i, e_j \rangle$, we can estimate $\|(I - P_l)(T - zI)P_n\|$ to asymptotic accuracy $O(n^{-1})$. It follows that we can can compute, in finitely many arithmetic operations and comparisons, $l(n) \ge n$ such that

$$\|(I - P_{l(n)})(T - zI)P_n\| = O(n^{-1}).$$

Similarly, we can estimate $\|P_{l(n)}(T - zI)\Gamma_{m,l(n)}(T, x, z) - x_{(m)}\|$ to accuracy $O(n^{-1})$ and call this approximation $v_{m,n}$, and estimate $\|\Gamma_{m,l(n)}(T, x, z)\|$ to accuracy $O(n^{-1})$ and call this approximation $w_{m,n}$. It follows that we have

$$\|\Gamma_{m,n}(T, x, z) - R(z, T)x_{(m)}\| \le \frac{(w_{m,n} + O(n^{-1}))O(n^{-1}) + v_{m,n} + O(n^{-1})}{|\text{Im}(z)|}.$$

It follows that

$$\|\Gamma_{m,n}(T, x, z) - R(z, T)x\| \le \frac{\|x_{(m)} - x\| + (w_{m,n} + O(n^{-1}))O(n^{-1}) + v_{m,n} + O(n^{-1})}{|\text{Im}(z)|}.$$

For a fixed m, $v_{m,n} \to 0$ and $w_{m,n} \to \|R(z, T)x_{(m)}\|$ as $n \to \infty$. It follows that we can compute $n(m)$ (again in finitely many arithmetic operations and comparisons) such that

$$\|\Gamma_{m,n(m)}(T, x, z) - R(z, T)x\| \le \frac{\|x_{(m)} - x\| + O(m^{-1})}{|\text{Im}(z)|}. \tag{B.2}$$

We know that $\|x\|$ has norm 1 and hence we must have

$$\|x_{(m)} - x\|^2 \le \sum_{j=m+1}^{\infty} |x_j|^2 + O(m^{-2}) = 1 - \sum_{j=1}^{m} |x_{(m)}|^2 + O(m^{-2}).$$

However, we can compute $1 - \sum_{j=1}^{m} |x_{(m)}|^2$, which converges to zero as $m \to \infty$. For a given N, it follows that we can compute $m(N) \ge N$ such that

$$1 - \sum_{j=1}^{m} |x_{(m)}|^2 \le N^{-2}.$$

This implies that $\|x_{(m)} - x\| = O(N^{-1})$ and hence by setting

$$\Gamma_N(T, x, z) = \Gamma_{m(N),n(m(N))}(T, x, z)$$

we see that (B.2) implies (B.1). □

Proof of Theorem B.1. We choose an orthonormal basis of $L^2(\mathbb{R}^d)$ so that we can carry over the results for $l^2(\mathbb{N})$ proven in this paper. In [37] it was shown that (the orthonormal basis of) tensor products of Hermite functions form a core for any $L \in \Omega_{\text{PDE}}$. Namely for $d = 1$ we choose the Hermite functions

$$\psi_m(x) = (2^m m! \sqrt{\pi})^{-1/2} e^{-x^2/2} H_m(x), \, m \in \mathbb{Z}_{\geq 0},$$

where

$$H_n(x) = (-1)^n \exp(x^2) \frac{d^n}{dx^n} \exp(-x^2).$$

For $d > 1$ we abuse notation and write $\psi_m = \psi_{m_1} \otimes \cdots \otimes \psi_{m_d}$. The point of this is that by a suitable ordering of $\{\psi_m\}_{m \in \mathbb{N}^d} = \{\psi_{m(1)}, \psi_{m(2)}, \ldots\}$, any $L \in \Omega_{\text{PDE}}$ can be represented by $T \in \Omega_{\text{SA}}$ and $f \in V_{\text{PDE}}$ by $\hat{f} \in l^2(\mathbb{N})$ with the inner products

$$\langle T e_j, T e_i \rangle = \int_{\mathbb{R}^d} (L\psi_{m(j)}(x)) \overline{(L\psi_{m(i)}(x))} dx \tag{B.3}$$

$$\langle T e_j, e_i \rangle = \int_{\mathbb{R}^d} (L\psi_{m(j)}(x)) \psi_{m(i)} dx \tag{B.4}$$

$$\langle \hat{f}, e_i \rangle = \int_{\mathbb{R}^d} f(x) \psi_{m(i)}(x) dx. \tag{B.5}$$

In [37] it was shown, using the theory of quasi-Monte Carlo numerical integration, that the inner products in (B.3)–(B.5) can be computed using Λ_{PDE} to asymptotic error control. In other words, that we can compute any of the evaluation functions in $\hat{\Lambda}$ in the statement of Theorem B.2 in finitely many arithmetic operations and comparisons. The result now follows by using Theorem B.2 instead of Corollary 2.2 in simple adaptations of the proofs of Theorems 3.1, 4.1 and 4.2. □

References

1. Akhiezer, N.I.: The Classical Moment Problem and Some Related Questions in Analysis. Translated by N. Kemmer. Hafner Publishing, New York (1965)
2. Amrein, W.O., Georgescu, V.: On the characterization of bound states and scattering states in quantum mechanics. Helv. Phys. Acta, **46**, 635–658 (1973/74)
3. Anderson, P.W.: Absence of diffusion in certain random lattices. Phys. Rev. **109**(5), 1492 (1958)
4. Antun, V., Colbrook, M.J., Hansen, A.C.: Can stable and accurate neural networks be computed? On the barriers of deep learning and Smale's 18th problem. arXiv preprint arXiv:2101.08286 (2021)
5. Aronszajn, N.: Approximation methods for eigenvalues of completely continuous symmetric operators. In: Proceedings of the Symposium on Spectral Theory and Differential Problems, pp. 179–202. Oklahoma Agricultural and Mechanical College, Stillwater, Okla (1951)
6. Arveson, W.: Improper filtrations for C^*-algebras: spectra of unilateral tridiagonal operators. Acta Sci. Math. (Szeged) **57**(1–4), 11–24 (1993)
7. Arveson, W.: C^*-algebras and numerical linear algebra. J. Funct. Anal. **122**(2), 333–360 (1994)
8. Arveson, W.: The role of C^*-algebras in infinite-dimensional numerical linear algebra. In C^*-algebras: 1943–1993 (San Antonio, TX, 1993), volume 167 of Contemp. Math., pp. 114–129. Amer. Math. Soc., Providence, RI (1994)

9. Bai, Z., Silverstein, J.W.: Spectral Analysis of Large Dimensional Random Matrices, vol. 20. Springer, Berlin (2010)
10. Becker, S., Hansen, A.: Computing solutions of Schrödinger equations on unbounded domains—on the brink of numerical algorithms. arXiv preprint arXiv:2010.16347 (2020)
11. Bédos, E.: On Folner nets, Szegö's theorem and other eigenvalue distribution theorems. Exposit. Math. **15**(3), 193–228 (1997)
12. Ben-Artzi, J., Colbrook, M.J., Hansen, A.C., Nevanlinna, O., Seidel, M.: Computing Spectra—On the Solvability Complexity Index hierarchy and towers of algorithms. arXiv preprint arXiv:1508.03280v5 (2020)
13. Ben-Artzi, J., Marletta, M., Rösler, F.: Computing scattering resonances. arXiv:2006.03368 (2020)
14. Ben-Artzi, J., Marletta, M., Rösler, F.: Computing the sound of the sea in a seashell. arXiv:2009.02956 (2020)
15. Bender, C.M.: Making sense of non-Hermitian Hamiltonians. Rep. Prog. Phys. **70**(6), 947 (2007)
16. Bilman, D., Trogdon, T.: Numerical inverse scattering for the Toda lattice. Commun. Math. Phys. **352**(2), 805–879 (2017)
17. Blum, L., Cucker, F., Shub, M., Smale, S.: Complexity and Real Computation. Springer New York Inc, Secaucus (1998)
18. Blümlinger, M., Tichy, R.F.: Topological algebras of functions of bounded variation I. Manuscr. Math. **65**(2), 245–255 (1989)
19. Böttcher, A.: Pseudospectra and singular values of large convolution operators. J. Integr. Equ. Appl. **6**(3), 267–301 (1994)
20. Böttcher, A.: Infinite matrices and projection methods. In Lectures on operator theory and its applications (Waterloo, ON, 1994), volume 3 of Fields Inst. Monogr., pp. 1–72. Amer. Math. Soc., Providence, RI (1996)
21. Böttcher, A., Brunner, H., Iserles, A., Nørsett, S.P.: On the singular values and eigenvalues of the Fox-Li and related operators. N.Y. J. Math. **16**, 539–561 (2010)
22. Böttcher, A., Chithra, A.V., Namboodiri, M.N.N.: Approximation of approximation numbers by truncation. Integr. Equ. Oper. Theory **39**(4), 387–395 (2001)
23. Böttcher, A., Silbermann, B.: Introduction to Large Truncated Toeplitz Matrices. Universitext. Springer, New York (1999)
24. Boulton, L.: Projection methods for discrete Schrödinger operators. Proc. Lond. Math. Soc. (3) **88**(2), 526–544 (2004)
25. Brown, L.G.: Lidskii's theorem in the type II case, pp. 1–35. Proceedings US-Japan, Kyoto/Japan (1983)
26. Brown, N.P.: Quasi-diagonality and the finite section method. Math. Comput. **76**(257), 339–360 (2007)
27. Brunner, H., Iserles, A., Nørsett, S.P.: The computation of the spectra of highly oscillatory Fredholm integral operators. J. Integr. Equ. Appl. **23**(4), 467–519 (2011)
28. Cantero, M.J., Moral, L., Velázquez, L.: Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle. Linear Algebra Appl. **362**, 29–56 (2003)
29. Carmona, R., Lacroix, J.: Spectral Theory of Random Schrödinger Operators. Probability and its Applications. Birkhäuser Boston Inc, Boston (1990)
30. Chang, K.: A physics magic trick: Take 2 sheets of carbon and twist. The New York Times (Oct 2019)
31. Chihara, T.S.: An Introduction to Orthogonal Polynomials. Gordon and Breach Science, New York-London-Paris (1978)
32. Coddington, E.A., Levinson, N.: Theory of Ordinary Differential Equations. McGraw-Hill Book, New York-Toronto-London (1955)
33. Colbrook, M.: The Foundations of Infinite-Dimensional Spectral Computations. PhD thesis, University of Cambridge (2020)
34. Colbrook, M.J.: Pseudoergodic operators and periodic boundary conditions. Math. Comput. **89**(322), 737–766 (2020)
35. Colbrook, M.J.: Computing semigroups with error control. Preprint (2021)
36. Colbrook, M.J.: On the computation of geometric features of spectra of linear operators on Hilbert spaces. arXiv:1908.09598 (2021)
37. Colbrook, M.J., Hansen, A.C.: The foundations of spectral computations via the solvability complexity index hierarchy: Part I. arXiv:1908.09592 (2019)
38. Colbrook, M.J., Hansen, A.C.: On the infinite-dimensional QR algorithm. Numer. Mat. **143**(1), 17–83 (2019)
39. Colbrook, M.J., Horning, A., Townsend, A.: Computing spectral measures of self-adjoint operators. SIAM Rev. (to appear)
40. Colbrook, M.J., Roman, B., Hansen, A.C.: How to compute spectra with error control. Phys. Rev. Lett. **122**(25), 250201 (2019)

41. Combes, J.-M.: Connections between quantum dynamics and spectral properties of time-evolution operators. In Differential equations with applications to mathematical physics, volume 192 of Math. Sci. Engrg., pp. 59–68. Academic Press, Boston (1993)
42. Cycon, H.L., Froese, R.G., Kirsch, W., Simon, B.: Schrödinger operators with application to quantum mechanics and global geometry. Texts and Monographs in Physics, study edn. Springer, Berlin (1987)
43. Davies, E.B.: Linear Operators and Their Spectra. Cambridge Studies in Advanced Mathematics, vol. 106. Cambridge University Press, Cambridge (2007)
44. Dean, C.R., Wang, L., Maher, P., Forsythe, C., Ghahari, F., Gao, Y., Katoch, J., Ishigami, M., Moon, P., Koshino, M., et al.: Hofstadter's butterfly and the fractal quantum Hall effect in Moire superlattices. Nature **497**(7451), 598–602 (2013)
45. Deift, P.: Orthogonal polynomials and random matrices: a Riemann–Hilbert approach, volume 3 of Courant Lecture Notes in Mathematics. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI (1999)
46. Deift, P., Li, L.C., Tomei, C.: Toda flows with infinitely many variables. J. Funct. Anal. **64**(3), 358–402 (1985)
47. Della Villa, A., Enoch, S., Tayeb, G., Pierro, V., Galdi, V., Capolino, F.: Band gap formation and multiple scattering in photonic quasicrystals with a Penrose-type lattice. Phy. Rev. Lett. **94**(18), 183903 (2005)
48. Dell'Oro, F., Pata, V.: Second order linear evolution equations with general dissipation. Appl. Math. Opt. **2019**, 1–41 (2019)
49. Digernes, T., Varadarajan, V.S., Varadhan, S.R.S.: Finite approximations to quantum systems. Rev. Math. Phys. **6**(4), 621–648 (1994)
50. Dombrowski, J., Nevai, P.: Orthogonal polynomials, measures and recurrence relations. SIAM J. Math. Anal. **17**(3), 752–759 (1986)
51. Doyle, P., McMullen, C.: Solving the quintic by iteration. Acta Math. **163**(3–4), 151–180 (1989)
52. Dubbs, A., Edelman, A.: Infinite random matrix theory, tridiagonal bordered Toeplitz matrices, and the moment problem. Linear Algebra Appl. **467**, 188–201 (2015)
53. Enss, V.: Asymptotic completeness for quantum mechanical potential scattering. I. Short range potentials. Commun. Math. Phys. **61**(3), 285–291 (1978)
54. Erdos, L.: Universality for random matrices and log-gases. arXiv preprint arXiv:1212.0839 (2012)
55. Fefferman, C., Seco, L.: On the energy of a large atom. Bull. Am. Math. Soc. (N.S.) **23**(2), 525–530 (1990)
56. Fefferman, C., Seco, L.: Interval arithmetic in quantum mechanics. In: Applications of Interval Computations (El Paso, TX, 1995), volume 3 of Appl. Optim., pp. 145–167. Kluwer Acad. Publ., Dordrecht (1996)
57. Fulton, C.T., Pearson, D., Pruess, S.: Computing the spectral function for singular Sturm–Liouville problems. J. Comput. Appl. Math. **176**(1), 131–162 (2005)
58. Fulton, C.T., Pearson, D., Pruess, S.: New characterizations of spectral density functions for singular Sturm–Liouville problems. J. Comput. Appl. Math. **212**(2), 194–213 (2008)
59. Fulton, C.T., Pruess, S.: The computation of spectral density functions for singular Sturm–Liouville problems involving simple continuous spectra. ACM Trans. Math. Softw. **24**(1), 107–129 (1998)
60. Fulton, C.T., Pruess, S., Shoaff, W.: Parallel computation of Sturm–Liouville spectral density functions. Parallel Algorithms Appl. **4**(1–2), 41–51 (1994)
61. Geisel, T., Ketzmerick, R., Petschel, G.: New class of level statistics in quantum systems with unbounded diffusion. Phys. Rev. Lett. **66**(13), 1651 (1991)
62. Geronimus, J.: On polynomials orthogonal on the circle, on trigonometric moment-problem and on allied Carathéodory and Schur functions. Rec. Math. [Mat. Sbornik] N.S. **15**(57), 99–130 (1944)
63. Geronimus, J.: Certain limiting properties of orthogonal polynomials. Vestnik Kharkov. Gos. Univ. **1966**(14), 40–50 (1966)
64. Gesztesy, F., Simon, B.: m-functions and inverse spectral analysis for finite and semi-infinite Jacobi matrices. J. Anal. Math. **73**, 267–297 (1997)
65. Girardin, V., Senoussi, R.: Semigroup stationary processes and spectral representation. Bernoulli **9**(5), 857–876 (2003)
66. Golub, G.H., Van Loan, C.F.: Matrix Computations. Johns Hopkins Studies in the Mathematical Sciences, 4th edn. Johns Hopkins University Press, Baltimore, MD (2013)
67. Graf, G.M.: Anderson localization and the space-time characteristic of continuum states. J. Stat. Phys. **75**(1–2), 337–346 (1994)
68. Gragg, W.B.: Positive definite Toeplitz matrices, the Arnoldi process for isometric operators, and Gaussian quadrature on the unit circle. J. Comput. Appl. Math. **46**(1–2), 183–198 (1993)
69. Gray, R.M., et al.: Toeplitz and circulant matrices: a review. Found. Trends Commun. Inf. Theory **2**(3), 155–239 (2006)
70. Gröchenig, K., Klotz, A.: Norm-controlled inversion in smooth Banach algebras, I. J. Lond. Math. Soc. (2) **88**(1), 49–64 (2013)

71. Haagerup, U., Schultz, H.: Brown measures of unbounded operators affiliated with a finite von Neumann algebra. Mathematica Scandinavica, pp. 209–263 (2007)
72. Haagerup, U., Schultz, H.: Invariant subspaces for operators in a general II_1-factor. Publications Mathématiques de l'IHÉS **109**, 19–111 (2009)
73. Hagen, R., Roch, S., Silbermann, B.: C^*-algebras and numerical analysis. Monographs and Textbooks in Pure and Applied Mathematics, vol. 236. Marcel Dekker Inc., New York (2001)
74. Hale, N., Higham, N.J., Trefethen, L.N.: Computing A^α, $\log(A)$, and related matrix functions by contour integrals. SIAM J. Numer. Anal. **46**(5), 2505–2523 (2008)
75. Hales, T.C.: A proof of the Kepler conjecture. Ann. Math. (2) **162**(3), 1065–1185 (2005)
76. Halmos, P.R.: Measure Theory. D. Van Nostrand, New York (1950)
77. Halmos, P.R.: What does the spectral theorem say? Am. Math. Mon. **70**, 241–247 (1963)
78. Hansen, A.C.: On the approximation of spectra of linear Hilbert space operators. PhD thesis, University of Cambridge (2008)
79. Hansen, A.C.: On the approximation of spectra of linear operators on Hilbert spaces. J. Funct. Anal. **254**(8), 2092–2126 (2008)
80. Hansen, A.C.: Infinite-dimensional numerical linear algebra: theory and applications. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **466**(2124), 3539–3559 (2010)
81. Hansen, A.C.: On the solvability complexity index, the n-pseudospectrum and approximations of spectra of operators. J. Am. Math. Soc. **24**(1), 81–124 (2011)
82. Haydock, R., Heine, V., Kelly, M.J.: Electronic structure based on the local atomic environment for tight-binding bands. J. Phys. C: Solid State Phys. **5**(20), 2845 (1972)
83. Hochbruck, M., Ostermann, A.: Exponential integrators. Acta Numer. **19**, 209–286 (2010)
84. Horning, A., Townsend, A.: FEAST for differential eigenvalue problems. SIAM J. Numer. Anal. **58**(2), 1239–1262 (2020)
85. Hundertmark, D., Meyries, M., Machinek, L., Schnaubelt, R.: Operator semigroups and dispersive equations. In: 16th Internet Seminar on Evolution Equations (2013)
86. Kadison, R.V., Ringrose, J.R.: Fundamentals of the theory of operator algebras. Vol. I, volume 15 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI (1997)
87. Kadison, R.V., Ringrose, J.R.: Fundamentals of the theory of operator algebras. Vol. II, volume 16 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI (1997)
88. Kallianpur, G., Mandrekar, V.: Spectral theory of stationary H-valued processes. J. Multivar. Anal. **1**(1), 1–16 (1971)
89. Kirsch, W.: Random Schrödinger operators. A course. In: Schrödinger operators (Sønderborg, 1988), volume 345 of Lecture Notes in Phys., pp. 264–370. Springer, Berlin (1989)
90. Kirsch, W.: An invitation to random Schrödinger operators. arXiv:0709.3707, **25**, 1–119 (2007)
91. Kirsch, W., Metzger, B.: The integrated density of states for random Schrödinger operators. In Spectral theory and mathematical physics: a Festschrift in honor of Barry Simon's 60th birthday, volume 76 of Proc. Sympos. Pure Math., pp. 649–696. Amer. Math. Soc., Providence, RI (2007)
92. Krutikov, D., Remling, C.: Schrödinger operators with sparse potentials: asymptotics of the Fourier transform of the spectral measure. Commun. Math. Phys. **223**(3), 509–532 (2001)
93. Last, Y.: Quantum dynamics and decompositions of singular continuous spectra. J. Funct. Anal. **142**(2), 406–445 (1996)
94. Levitin, M., Shargorodsky, E.: Spectral pollution and second-order relative spectra for self-adjoint operators. IMA J. Numer. Anal. **24**(3), 393–416 (2004)
95. Lin, L., Saad, Y., Yang, C.: Approximating spectral densities of large matrices. SIAM Rev. **58**(1), 34–65 (2016)
96. Lindner, M.: Infinite Matrices and Their Finite Sections. Frontiers in Mathematics. Birkhäuser Verlag, Basel (2006)
97. Lu, X., Stepanov, P., Yang, W., Xie, M., Aamir, M.A., Das, I., Urgell, C., Watanabe, K., Taniguchi, T., Zhang, G., et al.: Superconductors, orbital magnets and correlated states in magic-angle bilayer graphene. Nature **574**(7780), 653–665 (2019)
98. Marletta, M.: Neumann–Dirichlet maps and analysis of spectral pollution for non-self-adjoint elliptic PDEs with real essential spectrum. IMA J. Numer. Anal. **30**(4), 917–939 (2010)
99. Marletta, M., Scheichl, R.: Eigenvalues in spectral gaps of differential operators. J. Spectral Theory **2**(3), 293–320 (2012)
100. McLachlan, R.I., Reinout, G., Quispel, W.: Splitting methods. Acta Numer. **11**, 341–434 (2002)
101. McMullen, C.: Families of rational maps and iterative root-finding algorithms. Ann. Math. (2) **125**(3), 467–493 (1987)
102. McMullen, C.: Braiding of the attractor and the failure of iterative algorithms. Invent. Math. **91**(2), 259–272 (1988)

103. Niederreiter, H.: Random number generation and quasi-Monte Carlo methods, volume 63 of CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (1992)

104. Olver, S.: ApproxFun.jl v0.8. github. https://github.com/JuliaApproximation/ApproxFun.jl (2018)

105. Olver, S., Townsend, A.: A fast and well-conditioned spectral method. SIAM Rev. **55**(3), 462–489 (2013)

106. Olver, S., Townsend, A.: A Practical Framework for Infinite-dimensional Linear Algebra. In: Proceedings of the 1st First Workshop for High Performance Technical Computing in Dynamic Languages, HPTCDL '14, pages 57–62, Piscataway, NJ, USA, 2014. IEEE Press

107. Olver, S., Webb, M.: SpectralMeasures.jl. github. https://github.com/JuliaApproximation/SpectralMeasures.jl (2018)

108. Poltoratski, A.G.: On the distributions of boundary values of Cauchy integrals. Proc. Am. Math. Soc. **124**(8), 2455–2463 (1996)

109. Poltoratski, A.G., Simon, B., Zinchenko, M.: The Hilbert transform of a measure. J. Anal. Math. **111**, 247–265 (2010)

110. Ponomarenko, L.A., Gorbachev, R.V., Yu, G.L., Elias, D.C., Jalil, R., Patel, A.A., Mishchenko, A., Mayorov, A.S., Woods, C.R., Wallbank, J.R., et al.: Cloning of Dirac fermions in graphene superlattices. Nature **497**(7451), 594–597 (2013)

111. Pruess, S., Fulton, C.T.: Mathematical software for Sturm–Liouville problems. ACM Trans. Math. Softw. **19**(3), 360–376 (1993)

112. Rabinovich, V., Roch, S., Silbermann, B.: Limit operators and their applications in operator theory. Operator Theory: Advances and Applications, vol. 150. Birkhäuser Verlag, Basel (2004)

113. Reed, M., Simon, B.: Methods of modern mathematical physics. I., 2nd edn. Academic Press [Harcourt Brace Jovanovich, Publishers], New York (1980)

114. Remling, C.: The absolutely continuous spectrum of one-dimensional Schrödinger operators with decaying potentials. Commun. Math. Phys. **193**(1), 151–170 (1998)

115. Richardson, L.F., Gaunt, J.A.: VIII. The deferred approach to the limit. Containing papers of a mathematical or physical character. Philos. Trans. R. Soc. Lond. Ser. A **226**(636–646), 299–361 (1927)

116. Riddell, R.C.: Spectral concentration for self-adjoint operators. Pac. J. Math. **23**, 377–401 (1967)

117. Rogers, L.J.: Second Memoir on the Expansion of certain Infinite Products. Proc. Lond. Math. Soc., **25**, 318–343 (1893/94)

118. Rosenblatt, M.: Stochastic curve estimation. IMS (1991)

119. Ruelle, D.: A remark on bound states in potential-scattering theory. Nuovo Cimento A **10**(61), 655–662 (1969)

120. Seidel, M.: On (N, ϵ)-pseudospectra of operators on Banach spaces. J. Funct. Anal. **262**(11), 4916–4927 (2012)

121. Seidel, M.: Fredholm theory for band-dominated and related operators: a survey. Linear Algebra Appl. **445**, 373–394 (2014)

122. Seidel, M., Silbermann, B.: Finite sections of band-dominated operators—norms, condition numbers and pseudospectra. In Operator theory, pseudo-differential equations, and mathematical physics, volume 228 of Oper. Theory Adv. Appl., pp. 375–390. Birkhauser/Springer Basel AG, Basel (2013)

123. Shargorodsky, E.: Geometry of higher order relative spectra and projection methods. J. Oper. Theory **44**(1), 43–62 (2000)

124. Shargorodsky, E.: On the limit behaviour of second order relative spectra of self-adjoint operators. J. Spectral Theory **3**(4), 535–552 (2013)

125. Shechtman, D., Blech, I., Gratias, D., Cahn, J.W.: Metallic phase with long-range orientational order and no translational symmetry. Phy. Rev. Lett. **53**, 1951–1953 (1984)

126. Simon, B.: Schrödinger semigroups. Bull. Am. Math. Soc. **7**(3), 447–526 (1982)

127. Simon, B.: Absence of ballistic motion. Commun. Math. Phys. **134**(1), 209–212 (1990)

128. Simon, B.: Operators with singular continuous spectrum. I. General operators. Ann. Math. (2) **141**(1), 131–145 (1995)

129. Simon, B.: The classical moment problem as a self-adjoint finite difference operator. Adv. Math. **137**(1), 82–203 (1998)

130. Simon, B.: Orthogonal polynomials on the unit circle. Part 1, volume 54 of American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI (2005)

131. Simon, B.: Orthogonal polynomials on the unit circle. Part 2, volume 54 of American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI (2005)

132. Smale, S.: The fundamental theorem of algebra and complexity theory. Bull. Am. Math. Soc. (N.S.) **4**(1), 1–36 (1981)

133. Smale, S.: On the efficiency of algorithms of analysis. Bull. Am. Math. Soc. (N.S.) **13**(2), 87–121 (1985)

134. Smale, S.: Complexity theory and numerical analysis. In Acta numerica, 1997, volume 6 of Acta Numer., pp. 523–551. Cambridge University Press, Cambridge (1997)

135. Stadnik, Z.M.: Physical Properties of Quasicrystals, vol. 126. Springer, Berlin (2012)

136. Stieltjes, T.J.: Recherches sur les fractions continues. Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys. **8**(4), J1–J122 (1894)
137. Stone, M.H.: Linear transformations in Hilbert space. American Mathematical Society Colloquium Publications, vol. 15. American Mathematical Society, Providence, RI (1990)
138. Szabo, A., Ostlund, N.S.: Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory. Courier Corporation, New York (2012)
139. Szegő, G.: Orthogonal Polynomials. American Mathematical Society, New York (1939)
140. Takemura, S., Takemori, N., Koga, A.: Valence fluctuations and electric reconstruction in the extended Anderson model on the two-dimensional Penrose lattice. Phys. Rev. B **91**(16), 165114 (2015)
141. Tannor, D.J.: Introduction to Quantum Mechanics: A Time-Dependent Perspective. University Science Books, New York (2007)
142. Teplyaev, A.V.: The pure point spectrum of random orthogonal polynomials on the circle. Dokl. Akad. Nauk SSSR **320**(1), 49–53 (1991)
143. Teschl, G.: Jacobi operators and completely integrable nonlinear lattices. Mathematical Surveys and Monographs, vol. 72. American Mathematical Society, Providence, RI (2000)
144. Trogdon, T., Olver, S., Deconinck, B.: Numerical inverse scattering for the Korteweg–de Vries and modified Korteweg–de Vries equations. Phys. D: Nonlinear Phenom. **241**(11), 1003–1025 (2012)
145. Turing, A.M.: On computable numbers, with an application to the Entscheidungsproblem. Proc. Lond. Math. Soc. (2) **42**(3), 230–265 (1936)
146. Vardeny, Z.V., Nahata, A., Agrawal, A.: Optics of photonic quasicrystals. Nat. Photonics **7**(3), 177–187 (2013)
147. Verblunsky, S.: On positive harmonic functions: a contribution to the algebra of Fourier series. Proc. Lond. Math. Soc. **2**(1), 125–157 (1935)
148. Verblunsky, S.: On positive harmonic functions. Proc. Lond. Math. Soc. **2**(1), 290–320 (1936)
149. Wall, H.S.: Analytic Theory of Continued Fractions. D. Van Nostrand, New York (1948)
150. Webb, M.: Isospectral algorithms, Toeplitz matrices and orthogonal polynomials. PhD thesis, University of Cambridge (2017)
151. Webb, M., Olver, S.: Spectra of Jacobi operators via connection coefficient matrices. Commun. Math. Phys. **2021**, 1–51 (2021)
152. Wilkening, J., Cerfon, A.: A spectral transform method for singular Sturm-Liouville problems with applications to energy diffusion in plasma physics. SIAM J. Appl. Math. **75**(2), 350–392 (2015)

Communicated by H. Yau