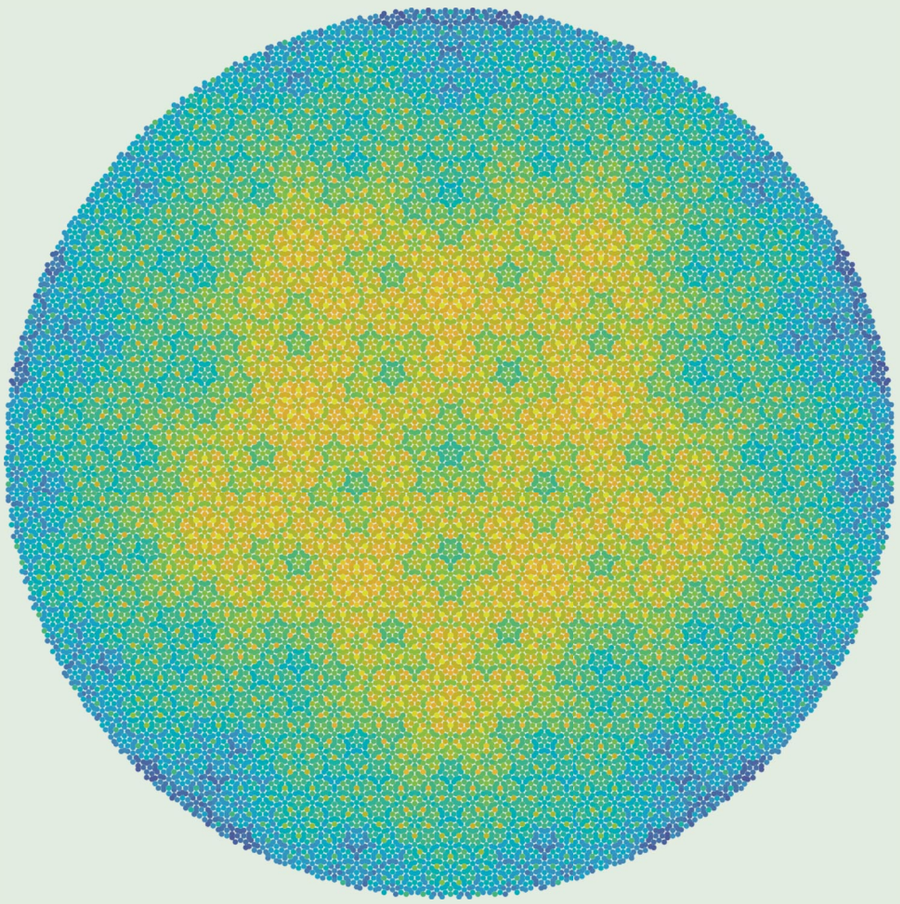
PHYSICAL REVIEW LETTERS

Articles published week ending 28 JUNE 2019



Published by
American Physical Society

APS physics

Volume 122, Number 25

How to Compute Spectra with Error Control

Matthew J. Colbrook, Bogdan Roman, and Anders C. Hansen
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom

(Received 28 November 2018; published 28 June 2019)

Computing the spectra of operators is a fundamental problem in the sciences, with wide-ranging applications in condensed-matter physics, quantum mechanics and chemistry, statistical mechanics, etc. While there are algorithms that in certain cases converge to the spectrum, no general procedure is known that (a) always converges, (b) provides bounds on the errors of approximation, and (c) provides approximate eigenvectors. This may lead to incorrect simulations. It has been an open problem since the 1950s to decide whether such reliable methods exist at all. We affirmatively resolve this question, and the algorithms provided are optimal, realizing the boundary of what digital computers can achieve. Moreover, they are easy to implement and parallelize, offer fundamental speed-ups, and allow problems that before, regardless of computing power, were out of reach. Results are demonstrated on difficult problems such as the spectra of quasicrystals and non-Hermitian phase transitions in optics.

DOI: 10.1103/PhysRevLett.122.250201

Introduction.—It is hard to overestimate the importance of computing the spectra of operators in mathematical physics, quantum chemistry, condensed-matter physics, statistical mechanics, Hermitian, as well as non-Hermitian, quantum mechanics, quasicrystals, optics, and many other fields. Motivated by the many applications, the topic has been intensely investigated, by both physicists [1–9] and mathematicians [10–17], since the 1950s. A reliable algorithm should converge and guarantee that any point of the output is close to the spectrum, up to a chosen arbitrary small error tolerance. A key question is whether such algorithms exist. Despite more than 90 years of quantum theory, the answer to this question has been unknown, even for Schrödinger operators.

The importance of this question is highlighted by the current interest in the spectral properties of systems with complicated spectra. The study of aperiodic systems, such as quasicrystals [18,19], often leads to complicated, even fractal-like spectra [20–24], which can make current methods of computation difficult. Another example is given by recent experimental breakthroughs in open systems in optics, which typically yield non-Hermitian Hamiltonians as there is no guaranteed energy preservation [25–29]. No previously known algorithm can handle non-Hermitian operators.

Questions on the foundations of computation and spectral computations have a rich history in physics. A recent example is the proof of the undecidability of the spectral gap [30]. Namely, one cannot construct an algorithm to determine whether a translationally invariant spin lattice system is gapped or gapless in the thermodynamic limit, a surprising result connected to seminal results in condensed-matter theory [31–33]. Hence, there are limitations to what a computer can achieve regarding limits of finite-dimensional systems.

In this Letter, we establish the boundaries for spectral problems in infinite dimensions. We show that it is *impossible* to design an algorithm for computing the spectra of Schrödinger operators which, given $\epsilon > 0$, halts and produces an output that is ϵ away from the true spectrum as measured in the Hausdorff metric. In other words, using information from a finite patch (truncation) of an operator A, it is impossible to produce an approximation $\Gamma(A)$ to the spectrum $\mathrm{Sp}(A)$, which satisfies the two inequalities (I) $\mathrm{dist}(z, \mathrm{Sp}(A)) \leq \epsilon$, for all $z \in \Gamma(A)$, and also (II) $\mathrm{dist}(w, \Gamma(A)) \leq \epsilon$, for all $w \in \mathrm{Sp}(A)$, simultaneously. However, we show that it is *possible* to create approximations, converging to the spectrum, that satisfy inequality (I). Indeed, we know the approximation is sound or reliable, but we do not know if we have got everything yet.

Namely, we provide an algorithm $\Gamma_n(\cdot)$, which both converges to the spectrum $\mathrm{Sp}(A)$ in the Hausdorff metric as $n \to \infty$ and also computes a local error bound function $E(n, z)$. Here, n is the size of the truncation of A used to compute $\Gamma_n(A)$, whereas $z \in \Gamma_n(A)$ approximates a member of $\mathrm{Sp}(A)$. For discrete lattice operators, n corresponds to considering the first n basis sites and their interactions, whereas, in general, n corresponds to the size of the truncation with respect to an orthonormal basis of the Hilbert space. The function $E(n, z)$ [defined in Eq. (4) in Ref. [34]] bounds the error of the approximations z through the inequality $\mathrm{dist}(z, \mathrm{Sp}(A)) \leq E(n, z)$. Moreover, we have

$$E_n(A) := \sup_{z \in \Gamma_n(A)} E(n, z) \to 0, \quad n \to \infty.$$

This means that $\Gamma_n(A)$ is contained in the true bulk spectrum of A up to the bound $E_n(A)$, regardless of how A extends outside the patch of size n used to compute $\Gamma_n(A)$. This is made possible through approximating the resolvent norm of A via using certain folded Hamiltonians, as outlined in the methods section.

The algorithm is applied to several problems known to be difficult in the literature and can also compute approximate states. The new algorithm has another key advantage over current methods—it is entirely local or parallelizable.

Statement of the results.—Consider a lattice model G with countably many vertices $V(G)$ (e.g., $\mathbb{Z}^d$) and write $v \sim_k w$ for $v, w \in V$ if there is a path of at most k edges connecting v and w. Let Ω_G be the class of normal finite range interaction Hamiltonians on $V(G)$, which are operators of the form

$$A = \sum_{v \sim_k w} \alpha(v, w) |v\rangle\langle w| \tag{1}$$

for some k (for more general operators, see [34]). The new algorithm computes $\mathrm{Sp}(A)$ and provides approximate eigenvectors (states). More precisely, (i) we prove that, given $\epsilon > 0$, there does not exist any algorithm Γ with input $A \in \Omega_G$ that halts and produces an output $\Gamma(A)$ with $d_H[\Gamma(A), \mathrm{Sp}(A)] \leq \epsilon$, where d_H denotes the Hausdorff metric. (ii) However, we provide an algorithm, $\Gamma_n(\cdot)$, that uses only the matrix elements $\alpha(v, w)$, which converges in the Hausdorff metric to $\mathrm{Sp}(A)$ for any $A \in \Omega_G$ as $n \to \infty$ and such that the local error bound function $E(n, z)$ satisfies $\mathrm{dist}(z, \mathrm{Sp}(A)) \leq E(n, z)$ with $E_n(A) \to 0$.

This has direct implications in computational boundaries in quantum mechanics [34]. With a minor modification, the algorithm also computes the pseudospectrum [15,42], $\mathrm{Sp}_\epsilon(A) = \{z : \|(A - zI)^{-1}\|^{-1} \leq \epsilon\}$, a generalization of the spectrum (and measure of its stability) which is popular for non-Hermitian problems, for which the above are still true. All cases of our algorithm(s) share several key features: (i) *Sharpness.*—They realize the boundary of what digital computers can achieve. (ii) *Known error of the output.*—The error bound $E_n(A)$ is computed with no added complexity. (iii) *Local computation of the spectrum.*—Computations are local and can be restricted to regions such as the extreme parts or parts in the middle of the spectrum (excited states). The algorithm can efficiently compute states corresponding to any chosen part of the output without diagonalizing the whole matrix. (iv) *Parallelization and speed.*—Local computation immediately implies parallelization. Every area where one wants to compute the spectrum can be divided into suitable subsets that can be handled individually. This allows for a substantial speed-up given access to several computer cores.

Methods.—The main ideas of the algorithm for computing the spectrum of $A \in \Omega_G$ are as follows (see [34] for generalizations, non-normal operators, proofs, and pseudocodes). By ordering the vertices or sites, we can consider A as an infinite matrix acting on $l^2(\mathbb{N})$, the space of square summable sequences. Given an integer n, consider the first n basis vectors $|e_1\rangle, \ldots, |e_n\rangle$ and let $f(n)$ be minimal such that $\langle e_i|A|e_j\rangle = 0$ if $j = 1, \ldots, n$ and $i > f(n)$. Such a $f(n)$ exists by the assumption of the finite range of interactions. We then consider the rectangular matrix $P_{f(n)}AP_n$, where P_n denotes orthogonal projection onto the span of the first n basis vectors. Physically, the rectangular matrix $P_{f(n)}AP_n$ contains all of the interactions of the first n sites without needing to apply boundary conditions (the range of the interactions controls precisely how rectangular the matrix should be).

The error bound function is then given as

$$E(n, z) \approx \min\{\|(A - zI)x\| : x \in \mathrm{span}\{|e_1\rangle, \ldots, |e_n\rangle\}\},$$

and we provide an efficient routine for its computation. This corresponds to an estimate of the distance of z to the spectrum and physically corresponds to approximating the square root of the ground state energy of the folded Hamiltonian $P_n(A - zI)^*(A - zI)P_n$ on the domain $\mathrm{span}\{|e_1\rangle, \ldots, |e_n\rangle\}$. We prove that our approximation converges uniformly to the resolvent norm $\|(A - zI)^{-1}\|^{-1} = \mathrm{dist}(z, \mathrm{Sp}(A))$, on compact subsets of the complex plane. The convergence is also from above, meaning that we gain the rigorous error bound $\mathrm{dist}(z, \mathrm{Sp}(A)) \leq E(n, z)$. It is precisely the use of the rectangular truncation $P_{f(n)}AP_n$ that leads to convergence from above, and, in general, taking a square truncation will not even converge.

Given a region $\mathcal{D} \subset \mathbb{C}$ of interest, the other ingredient of the algorithm is a search routine that seeks to approximate the spectrum locally on $\mathcal{D}$. We consider a grid of points $G_{\mathcal{D}}(n)$ of spacing $\delta(n) \to 0$ as $n \to \infty$. The resolution $\delta(n)^{-1}$ (which can be viewed as a discretization parameter) can be changed to allow one to vary the number of computed solutions. In our experiments, we chose $\delta(n)$ to ensure $\approx n$ solutions for fair comparisons with other methods. The first step is to compute $E(n, \cdot)$ over $G_{\mathcal{D}}(n)$, which can be done in parallel. Given $z \in G_{\mathcal{D}}(n)$, we let I_z be the points in $G_{\mathcal{D}}(n)$ at a distance most $E(n, z)$ away from z. We then let M_z be the minimizers of $E(n, \cdot)$ over the local set I_z. Since $E(n, \cdot)$ bounds the distance to the spectrum and converges to the true distance, M_z approximates the spectrum near the point z.

The negative result (i) we prove shows that it is, in general, impossible to know in finite time if we have computed all of the spectrum up to an arbitrary error parameter. However, the convergence of the algorithm [which depends on the use of the above local search radius $E(n, z)$] ensures that we will progressively obtain all of the spectrum through this search routine (and avoid solutions not in the bulk spectrum). The algorithm's output $\Gamma_n(A)$ is the union of all such M_z for z such that $E(n, z) \leq 1/2$, and

the error bound is $E_n(A) = \max\{E(n,z) : z \in \Gamma_n(A)\}$. The approximate states for $z \in \Gamma_n(A)$ then correspond to approximate states of the ground state energy of the folded Hamiltonian $(A - zI)^*(A - zI)$, which can be computed efficiently using an ldl decomposition and with residual error bounds. We also give a precise bound on the time needed for the computations based on the size of the region $G_{\mathcal{D}}(n)$, the resolution $\delta(n)^{-1}$, and n (Proposition 4 in Ref. [34]). $\delta(n)$ can also be chosen to vary across $\mathcal{D}$, which we predict will allow extremely fast computation of states corresponding to eigenvalues in gaps of the spectrum.

Results.—We now illustrate our algorithm(s) on difficult models arising in condensed-matter physics or quantum mechanics. In addition, an example considering the non-Hermitian Anderson model is included in Ref. [34].

Quasicrystals.—Quasicrystals, and more generally aperiodic systems, have generated considerable interest due to their often exotic physical or spectral properties [18,19]. We present the first rigorous spectral computational study with error bounds on a Penrose tile, the standard 2D model of a quasicrystal [22,43,44]. No previously known algorithm determines the spectrum with error bounds on the output.

The free Hamiltonian H_0 (Laplacian) is given by

$$(H_0\psi)_i = \sum_{i\sim j}(\psi_j - \psi_i), \tag{2}$$

with a summation over nearest neighbor sites (vertices). Previous numerical methods study the eigenvalues of the Hamiltonian restricted to a finite portion of the tiling with a choice of boundary conditions at the edges (finite section method) such as periodic approximants [45–49]. This can cause additional eigenvalues (spectral pollution or "edge states") [14] to appear which are not in the spectrum of H_0 acting on the infinite tiling. While edge states have a physical meaning and are important experimentally, it is often desirable to distinguish these from the "bulk" states of the operator. Our algorithm could have useful applications in the fast-growing research area of topological edge states [50,51].

Figure 1 shows the output of the algorithm for $n = 10^5$ and two finite section methods (open boundary conditions and periodic approximants). The error estimate, computed (using the algorithm) for both the algorithm and the finite section method, is displayed at each point. This error estimate converges uniformly to the true error on compact subsets of $\mathbb{R}$. Finite section methods produce spurious points in the gaps of the spectrum.

Figure 1 also shows the benefit of parallelization. The time taken for the algorithm (ran using 200 cores) and for the finite section methods (ran using four cores) to reach the final output suggests a speed-up of about 20 times. Moreover, the time for the finite section method appears to grow $\sim\mathcal{O}(n^{2.9})$ and $\mathcal{O}(n^{3.0})$ for open and periodic boundary conditions, respectively, whereas the time for the new algorithm grows $\sim\mathcal{O}(n^{2.1})$, predicting larger differences for larger n. The direct diagonalization approach is hard to parallelize and so will have difficulty competing with our method for large n. It is also possible to use the algorithm to locally compute approximate states corresponding to a given energy level without the need to diagonalize the whole system as shown in Fig. 2.

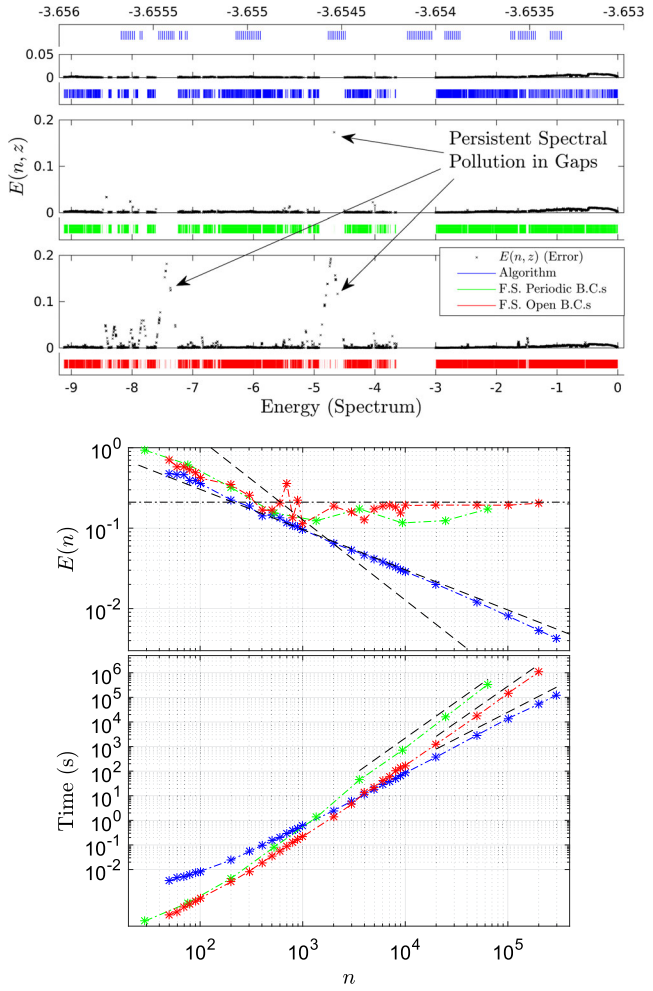


FIG. 1. Top: Large-scale experiment with $n = 10^5$ for the algorithm and finite section with open boundary conditions and periodic approximants ($n = 64079$). The top shows an enlarged section and the high resolution obtained. The approximation computed with the finite section methods produces spurious points in band gaps with large errors $\sim$0.2. Bottom: The maximum errors as well as time of outputs for the algorithm (blue) and finite section methods (red for open boundary conditions, green for periodic).

Finally, we consider a magnetic Hamiltonian [4,52–54]

$$(H\psi)_i = -\sum_{\langle i,j\rangle} e^{i\alpha_{ji}}\psi_j.$$

A constant perpendicular magnetic field $\mathbf{B} = B\mathbf{z}$ with potential $\mathbf{A} = (0, xB, 0)$ is applied, leading to the Peierls phase factor between sites i and j:

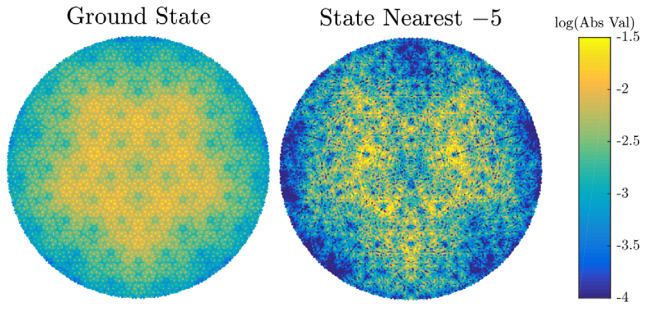


FIG. 2. The ground state for the Penrose Laplacian and a state corresponding to energy nearest −5. The algorithm allows us to choose which states to compute.

$\alpha_{ji} = (2\pi/\Phi_0)\int_{\mathbf{r}_j}^{\mathbf{r}_i} \mathbf{A}\cdot d\mathbf{l}$, where $\Phi_0 = hc/e$ is the flux quantum. Figure 3 shows the output for the finite section method and the algorithm for $n = 5000$ up to the first self-similar mode B_0. The absence of spectral pollution for the new algorithm is striking and agrees well the periodic approximant studied in Ref. [55].

Recently, Hofstadter's butterfly has been experimentally observed in graphene lattices [20,21,56]. Clearly, numerical methods that avoid spectral pollution, converge, and provide error bounds are needed to study such operators with fractal-like spectra. Although one may also study this type of problem via periodic approximants as before [52,55], this places restrictions on the value of B. One can increase the allowed values by the use of magnetic translations [57], but the total magnetic flux through a period cell must still be a multiple of the flux quantum. Such methods cannot be applied to problems with arbitrary (even nonconstant) magnetic fields nor models with large degrees of freedom, whereas the new algorithm can. Numerical difficulties have previously prevented

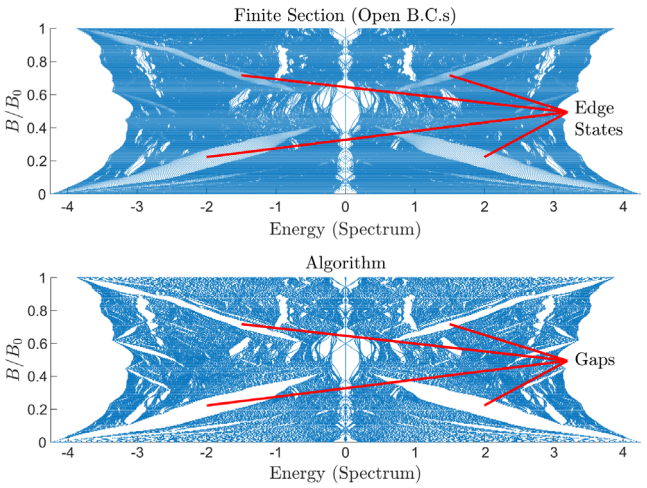


FIG. 3. Comparison of the finite section method and the algorithm for the magnetic Hamiltonian. The algorithm correctly leaves out the gaps and is able to capture the complicated structure with guaranteed error maximum 0.058 for $n = 5000$.

theoretical modeling of many experimental results of quasicrystals in higher dimensions. The new algorithm can tackle such models, and future work will study 3D systems.

Open systems in optics.—Open systems typically yield non-Hermitian Hamiltonians, as there is no guaranteed energy preservation. However, non-Hermitian Hamiltonians can possess real spectra when they respect parity-time (PT) symmetry [58–60]. Remarkably, many Hamiltonians undergo a phase transition to complex spectra if the imaginary part of their potential is increased beyond a certain threshold, known as symmetry breaking. Such systems are of wide interest [61–65] and can be realized in optics [25–29,66–69].

Detecting when symmetry breaking occurs poses a substantial challenge, since it is very sensitive to surface

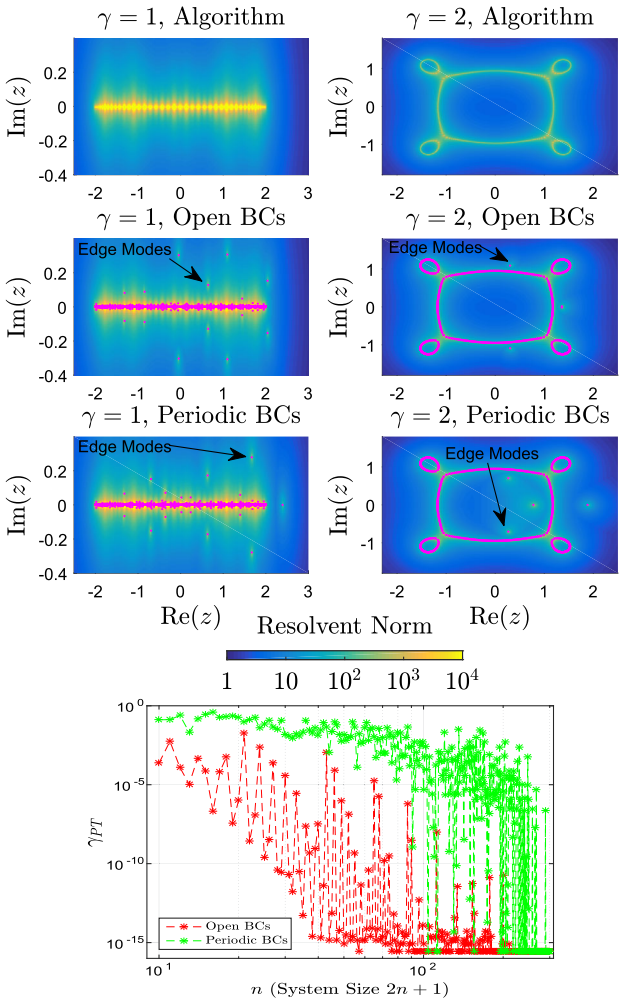


FIG. 4. Top: Pseudospectra of H computed with the proposed algorithm and finite sections (spectrum in magenta). Bottom: Fragile PT-symmetric phase as we increase the system size due to edge states (caused by spectral pollution) with complex eigenvalues, verifying the unsuitability of algorithms based on finite sections.

states arising from standard truncations. We discuss PT symmetry breaking for the case of an aperiodic potential on a discrete lattice:

$$(Hx)_n = x_{n-1} + x_{n+1} + V_n x_n$$

acting on $l^2(\mathbb{Z})$, where $V_n = \cos(n) + i\gamma \sin(n)$ and $\gamma \geq 0$. Here the aperiodicity occurs dues to the incommensurability of the potential and lattice. We stress that the algorithm can handle any type of potential (such as additional defects).

In the limit of an increasing system size, the critical parameter γ_{PT} depends on the boundary conditions imposed, often decreasing as the number of sites increases with a fragile PT-symmetric phase. This limit can differ from the value γ_{PT} on the infinite lattice due to surface or edge states [70]. Using the algorithm gives an estimate for γ_{PT} in the infinite lattice case avoiding this fragility, suggesting that symmetry breaking occurs at $\gamma_{PT} \approx 1 \pm 0.05$. This allows us to *rigorously detect edge states* (corresponding to spectral pollution) and the corresponding edge modes. Figure 4 shows pseudospectral plots generated by the algorithm for $\gamma = 1$, 2 as well as the corresponding plots for finite chains of length 2001 for open and periodic boundary conditions. We can easily use the algorithm to separate bulk states from edge states. We have also shown the values of γ_{PT} for the finite chains showing the fragility of the PT-symmetric phase.

Conclusion.—We have demonstrated the boundaries of what computers can do in spectral computations in large areas of applications by presenting the first algorithms that converge to the spectrum or pseudospectrum with error control on the output. In contrast, the state of the art finite section method, even for the cases where it does converge, does not provide error bounds. The related issue of spectral pollution has been well studied with previous results concentrating on the self-adjoint case [71], in particular, detecting spectral pollution in gaps of the essential spectrum. However, none of these methods provides convergence to the spectrum, as shown by Shargorodsky in Ref. [72]. The major difference between the algorithms presented here and those previously studied is the local approximation of the resolvent norm using uneven sections. Physically, for discrete systems, this corresponds to preserving the correct interactions of the first n sites and avoiding the use of imposed boundary conditions.

The presented method is very general and can be used in many diverse fields of applications. A significant advantage is the possibility of error control that will allow scientists to compare experiments with computational results that are now guaranteed to be correct up to an accuracy parameter set by the user. Moreover, the locality of the algorithms represents a considerable advantage if one is interested only in certain regions of the spectrum and makes computations extremely quick.

This work was supported by Engineering and Physical Sciences Research Council Grants No. EP/L016516/1, No. EP/R008272/1, No. EP/N014588/1, and No. EP/L003457/1, as well as a Royal Society University Research Fellowship.

[1] E. Schrödinger, in *Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences* (JSTOR, Dublin, 1940), Vol. 46, pp. 9–16.
[2] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).
[3] E. Balslev and J.-M. Combes, Commun. Math. Phys. **22**, 280 (1971).
[4] D. R. Hofstadter, Phys. Rev. B **14**, 2239 (1976).
[5] R. Haydock, Comput. Phys. Commun. **20**, 11 (1980).
[6] L. E. Gendenshtein, JETP Lett. **38**, 356 (1983).
[7] S. Ostlund, R. Pandit, D. Rand, H. J. Schellnhuber, and E. D. Siggia, Phys. Rev. Lett. **50**, 1873 (1983).
[8] A. Sütő, Commun. Math. Phys. **111**, 409 (1987).
[9] N. Laskin, Phys. Rev. E **66**, 056108 (2002).
[10] A. Böttcher and B. Silbermann, *Introduction to Large Truncated Toeplitz Matrices*, Universitext (Springer, New York, 2012).
[11] A. Böttcher, J. Integr. Equations Appl. **6**, 267 (1994).
[12] T. Digernes, V. S. Varadarajan, and S. R. S. Varadhan, Rev. Math. Phys. **06**, 621 (1994).
[13] A. C. Hansen, J. Am. Math. Soc. **24**, 81 (2011).
[14] M. Marletta, IMA J. Numer. Anal. **30**, 917 (2010).
[15] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators* (Princeton University, Princeton, NJ, 2005).
[16] E. B. Davies, *Linear Operators and Their Spectra* (Cambridge University Press, Cambridge, England, 2007), Vol. 106.
[17] W. Arveson, in *C^*-Algebras: 1943–1993 (San Antonio, TX, 1993)*, Contemporary Mathematics Vol. 167 (American Mathematical Society, Providence, RI, 1994), pp. 114–129.
[18] D. Shechtman, I. Blech, D. Gratias, and J. W. Cahn, Phys. Rev. Lett. **53**, 1951 (1984).
[19] Z. M. Stadnik, *Physical Properties of Quasicrystals* (Springer Science, New York, 2012), Vol. 126.
[20] B. Hunt, J. Sanchez-Yamagishi, A. Young, M. Yankowitz, B. J. LeRoy, K. Watanabe, T. Taniguchi, P. Moon, M. Koshino, P. Jarillo-Herrero *et al.*, Science **340**, 1427 (2013).
[21] C. Dean, L. Wang, P. Maher, C. Forsythe, F. Ghahari, Y. Gao, J. Katoch, M. Ishigami, P. Moon, M. Koshino *et al.*, Nature (London) **497**, 598 (2013).
[22] Z. V. Vardeny, A. Nahata, and A. Agrawal, Nat. Photonics **7**, 177 (2013).
[23] L. Levi, M. Rechtsman, B. Freedman, T. Schwartz, O. Manela, and M. Segev, Science **332**, 1541 (2011).
[24] A. Sütő, J. Stat. Phys. **56**, 525 (1989).
[25] A. Regensburger, C. Bersch, M.-A. Miri, G. Onishchukov, D. N. Christodoulides, and U. Peschel, Nature (London) **488**, 167 (2012).
[26] A. Guo, G. Salamo, D. Duchesne, R. Morandotti, M. Volatier-Ravat, V. Aimez, G. Siviloglou, and D. Christodoulides, Phys. Rev. Lett. **103**, 093902 (2009).

[27] C. E. Rüter, K. G. Makris, R. El-Ganainy, D. N. Christodoulides, M. Segev, and D. Kip, Nat. Phys. **6**, 192 (2010).

[28] L. Feng, Z. J. Wong, R.-M. Ma, Y. Wang, and X. Zhang, Science **346**, 972 (2014).

[29] H. Hodaei, M.-A. Miri, M. Heinrich, D. N. Christodoulides, and M. Khajavikhan, Science **346**, 975 (2014).

[30] T. S. Cubitt, D. Perez-Garcia, and M. M. Wolf, Nature (London) **528**, 207 (2015).

[31] T. D. Schultz, D. C. Mattis, and E. H. Lieb, Rev. Mod. Phys. **36**, 856 (1964).

[32] M. B. Hastings, Phys. Rev. B **69**, 104431 (2004).

[33] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, in *Condensed Matter Physics and Exactly Soluble Models* (Springer, New York, 1988), pp. 253–304.

[34] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.122.250201 for further details, including proofs of results, algorithmic detail, and remarks on efficient numerical implementation as well as implications in quantum mechanics and Refs. [35–41].

[35] A. M. Turing, Proc. London Math. Soc. **s2-42**, 230 (1937).

[36] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and Real Computation* (Springer-Verlag, New York, 1998).

[37] L. N. Trefethen and D. Bau III, *Numerical Linear Algebra* (Society for Industrial and Applied Mathematics, Philadelphia, 1997), Vol. 50.

[38] N. Hatano and D. R. Nelson, Phys. Rev. Lett. **77**, 570 (1996).

[39] S. Longhi, D. Gatti, and G. Della Valle, Sci. Rep. **5** (2015).

[40] D. R. Nelson and N. M. Shnerb, Phys. Rev. E **58**, 1383 (1998).

[41] I. Y. Goldsheid and B. A. Khoruzhenko, Phys. Rev. Lett. **80**, 2897 (1998).

[42] D. Krejčiřík, P. Siegl, M. Tater, and J. Viola, J. Math. Phys. (N.Y.) **56**, 103513 (2015).

[43] S. Takemura, N. Takemori, and A. Koga, Phys. Rev. B **91**, 165114 (2015).

[44] A. Della Villa, S. Enoch, G. Tayeb, V. Pierro, V. Galdi, and F. Capolino, Phys. Rev. Lett. **94**, 183903 (2005).

[45] H. Tsunetsugu, T. Fujiwara, K. Ueda, and T. Tokihiro, Phys. Rev. B **43**, 8879 (1991).

[46] M. A. Bandres, M. C. Rechtsman, and M. Segev, Phys. Rev. X **6**, 011016 (2016).

[47] E. C. Andrade, A. Jagannathan, E. Miranda, M. Vojta, and V. Dobrosavljević, Phys. Rev. Lett. **115**, 036403 (2015).

[48] A. Szallas and A. Jagannathan, Phys. Rev. B **77**, 104427 (2008).

[49] N. de Bruijn, Indagationes Math. (Proceedings) **84**, 39 (1981).

[50] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).

[51] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).

[52] D.-T. Tran, A. Dauphin, N. Goldman, and P. Gaspard, Phys. Rev. B **91**, 085125 (2015).

[53] T. Hatakeyama and H. Kamimura, Solid State Commun. **62**, 79 (1987).

[54] J. Vidal and R. Mosseri, J. Non-Cryst. Solids **334–335**, 130 (2004).

[55] T. Hatakeyama and H. Kamimura, J. Phys. Soc. Jpn. **58**, 260 (1989).

[56] L. Ponomarenko, R. Gorbachev, G. Yu, D. Elias, R. Jalil, A. Patel, A. Mishchenko, A. Mayorov, C. Woods, J. Wallbank *et al.*, Nature (London) **497**, 594 (2013).

[57] J.-N. Fuchs and J. Vidal, Phys. Rev. B **94**, 205437 (2016).

[58] C. M. Bender and S. Boettcher, Phys. Rev. Lett. **80**, 5243 (1998).

[59] S. Klaiman, U. Günther, and N. Moiseyev, Phys. Rev. Lett. **101**, 080402 (2008).

[60] C. M. Bender, Rep. Prog. Phys. **70**, 947 (2007).

[61] T. Gao, E. Estrecho, K. Bliokh, T. Liew, M. Fraser, S. Brodbeck, M. Kamp, C. Schneider, S. Höfling, Y. Yamamoto *et al.*, Nature (London) **526**, 554 (2015).

[62] M. Wimmer, A. Regensburger, M.-A. Miri, C. Bersch, D. N. Christodoulides, and U. Peschel, Nat. Commun. **6**, 77782 (2015).

[63] T. Eichelkraut, R. Heilmann, S. Weimann, S. Stützer, F. Dreisow, D. Christodoulides, S. Nolte, and A. Szameit, Nat. Commun. **4** (2013).

[64] H. Schomerus, Opt. Lett. **38**, 1912 (2013).

[65] S. Longhi, Phys. Rev. Lett. **103**, 123601 (2009).

[66] K. G. Makris, R. El-Ganainy, D. N. Christodoulides, and Z. H. Musslimani, Phys. Rev. Lett. **100**, 103904 (2008).

[67] J. Schindler, A. Li, M. C. Zheng, F. M. Ellis, and T. Kottos, Phys. Rev. A **84**, 040101 (2011).

[68] A. Regensburger, M.-A. Miri, C. Bersch, J. Näger, G. Onishchukov, D. N. Christodoulides, and U. Peschel, Phys. Rev. Lett. **110**, 223902 (2013).

[69] B. Zhen, C. W. Hsu, Y. Igarashi, L. Lu, I. Kaminer, A. Pick, S.-L. Chua, J. D. Joannopoulos, and M. Soljačić, Nature (London) **525**, 354 (2015).

[70] O. Bendix, R. Fleischmann, T. Kottos, and B. Shapiro, Phys. Rev. Lett. **103**, 030402 (2009).

[71] E. B. Davies, LMS J. Comput. Math. **1**, 42 (1998).

[72] E. Shargorodsky, J. Spectr. Theory **3**, 535 (2013).

Supplementary Material

Matthew J. Colbrook, Bogdan Roman, Anders C. Hansen

Contents

Overview

Here we provide statements and proofs of theorems as well as detailed explanations of the algorithm. Let $\mathcal{B}(l^2(\mathbb{N}))$ denote the set of bounded linear operators acting on $l^2(\mathbb{N})$. Given a class $\Omega \subset \mathcal{B}(l^2(\mathbb{N}))$ and $A \in \Omega$, we are interested in computing the following from the matrix values (with respect to the canonical basis) of A:

- $\mathrm{Sp}(A) := \{z \in \mathbb{C} : A - zI \text{ is not invertible}\}$ (spectrum),
- $\mathrm{Sp}_\epsilon(A) := \{z \in \mathbb{C} : \|(A - zI)^{-1}\|^{-1} \leq \epsilon\}$, with $\epsilon > 0$ (pseudospectrum),

where we interpret $\|B^{-1}\|^{-1}$ as 0 if the inverse B^{-1} does not exist. These are computed in the Hausdorff metric on non-empty compact subsets of $\mathbb{C}$. Recall that the Hausdorff metric is defined as

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

where $d(x, y) = |x-y|$ is the usual Euclidean distance. As will be mentioned later, the above problems easily generalise to arbitrary separable Hilbert spaces, in particular lattice models found in quantum mechanics and condensed matter physics. We have included a section which states these results explicitly.

For the classes Ω considered, there are two statements in the main manuscript (also for the pseudospectrum):

(i) There does not exists any algorithm that, given any $\delta > 0$ and $A \in \Omega$, halts and produces an output $\Gamma(A, \delta)$ that is no worse than δ away from the spectrum of A in the sense that

$$d_{\mathrm{H}}(\Gamma(A, \delta), \mathrm{Sp}(A)) \leq \delta. \tag{1}$$

(ii) However, we provide an algorithm, $\Gamma_n(\cdot)$, that uses only the matrix elements of A, which converges to the spectrum of any $A \in \Omega$ as $n \to \infty$. The algorithm also computes error bounds $E_n(A) \to 0$, as $n \to \infty$, such that for any $z \in \Gamma_n(A)$

$$\operatorname{dist}(z, \operatorname{Sp}(A)) \leq E_n(A).$$

We will prove these statements, and for the user's benefit, we have provided user-friendly pseudocode for the algorithm as well as discussing how to implement it efficiently. Other results such as approximate states and the implications for computational quantum mechanics are also discussed.

In order to be completely precise with statement (i) we need a (very broad) definition of an 'algorithm'. Suppose Ξ represents a mapping from Ω to some metric space that we want to compute (in our case the spectrum or pseudospectrum) and Λ denotes the family of evaluation functions $\Omega \to \mathbb{C}$ that our algorithms can access (in our case matrix elements of an operator). We call $\{\Xi, \Omega, \Lambda\}$ a computational problem.

Definition 1 (General Algorithm)**.** *Given a computational problem* $\{\Xi, \Omega, \Lambda\}$*, a* general algorithm *is a mapping* $\Gamma : \Omega \to \mathcal{M}$ *such that for each* $A \in \Omega$

(i) there exists a finite 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* $\Lambda_\Gamma(B) = \Lambda_\Gamma(A)$*.*

In other words, a general algorithm 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 reads the information in a consistent way. The purpose of such a broad definition is that the impossibility results become universal independent of the model of computation, and hence this leads to a stronger statement (i) compared to a statement based on a particular model such as the Turing model [1], or the Blum-Shub-Smale model [2]. However, all the algorithms we construct are recursive and can easily be adapted to work with inexact input/restrictions to arithmetic computations over $\mathbb{Q}$.

The Algorithm

Here we present in more detail the execution of our algorithm and prove our main results. We need to mention the subroutine `Grid`. This routine simply makes a grid in the complex plane in the area where one wants to compute the

Function `Grid`(m)
  Input : $m \in \mathbb{N}$
  Output: The intersection of $\frac{1}{m}(\mathbb{Z} + i\mathbb{Z})$ with the disc $B_m(0)$ centred at zero.
end

spectrum. Note that if we know an estimate $\|A\| \leq C$ of the norm of the operator, then obviously one will restrict the grid to the $B_C(0)$ disc, as this will save time substantially. However, the norm estimate is not needed for the algorithm to work.

Off-diagonal decay and bounded dispersion

In the algorithm presented in this paper, the function $f : \mathbb{N} \to \mathbb{N}$, that describes the off-diagonal decay of the infinite matrix, is crucial. More formally, let A be a bounded operator on $l^2(\mathbb{N})$ or $l^2(\mathbb{Z})$ interpreted as an infinite matrix with respect to the canonical basis $\{e_j\}$. We define for $f : \mathbb{N} \to \mathbb{N}$, $f(n) \geq n$

$$D_{f,m}(A) := \max\left\{\left\|(I - P_{f(m)})AP_m\right\|, \left\|P_m A(I - P_{f(m)})\right\|\right\}, \tag{2}$$

where P_n is the projection onto the span of $\{e_1, \ldots, e_n\}$ (or $\{e_{-n}, \ldots, e_n\}$ in the $l^2(\mathbb{Z})$ case), and assume that we have an estimate $D_{f,m}(A) \leq c_m$ for real numbers c_m, where $c_m \to 0$ as $m \to \infty$. We say that an operator that has this property has bounded dispersion. Note that for any infinite matrix A there is an f such that $c_m = 1/m$, the problem is to find it. Fortunately, in many applications this is quite easy.

Example 1 (Finding f)**.** If A is a banded matrix finding f is trivial. However, even for quasi-banded matrices, that have finitely many non-zeros in each column and row, this is quite easy. In particular, if $l(m)$ represents an integer such that the matrix elements $A_{m+j,m} = 0$ if $|j| > l(m)$ and $A_{m,m+k} = 0$ if $|k| > l(m)$ then we may take $f(m) = m + l(m) \Rightarrow D_{f,m}(A) = 0$. Note that all finite range interaction Hamiltonians on graphs will have this property as explained in Example 2 and Example 3.

The routine `DistSpec` takes f as an input and is used to estimate the distance of a point z from the spectrum of A.

Function `DistSpec`(A,n,$f(n)$,z)
  Input : $n \in \mathbb{N}$, $f(n) \in \mathbb{N}$, $A \in \mathcal{B}(l^2(\mathbb{N}))$, $z \in \mathbb{C}$
  Output: $y \in \mathbb{R}_+$, an approximation to the function $z \mapsto \left\|(A - zI)^{-1}\right\|^{-1}$
  $B = (A - zI)(1 : f(n), 1 : n)$
  $C = (A - zI)^*(1 : f(n), 1 : n)$
  $S = B^*B$
  $T = C^*C$
  $\nu = 1$, $l = 0$
  while $\nu = 1$ **do**
    $l = l + 1$
    $p =$ `IsPosDef`$(S - \frac{l^2}{n^2})$
    $q =$ `IsPosDef`$(T - \frac{l^2}{n^2})$
    $\nu = \min(p, q)$
  end
  $y = \frac{l}{n}$
end

The `IsPosDef` routine determines whether a matrix is positive definite (returns 1) or not (returns 0). This can be done by using, for example, a variant of the Cholesky decomposition (chosen for stability reasons and speed of computation). In practice, we also replace the while loop via an interval bisection method. The idea is to use `DistSpec` as an approximation to $z \mapsto \left\|(A - zI)^{-1}\right\|^{-1}$ which we convert to an approximation to the distance to the spectrum

The Main Routine

Let $g : \mathbb{R}_+ \to \mathbb{R}_+$ be a strictly increasing continuous function that vanishes only at 0 and tends to ∞ as x tends to ∞. Let Ω_2 denote the set of all operators with bounded dispersion bounded by f and with a known bound of the form

$$\|R(z, A)\|^{-1} \geq g(\mathrm{dist}(z, \mathrm{Sp}(A))),$$

for $z \in \mathbb{C}$. Note that such a g is always guaranteed to exist, however, the it is crucial for the algorithm that one knows an estimate for g. For example, in the Hermitian and normal case $g(x) = x$ is the trivial choice of g.

We begin by defining the subroutine `CompInvg` that computes an estimate of g^{-1} (which exists by assumptions on g). This subroutine is essential in the main routine `CompSpec`. Note that

$$g^{-1}(y) \leq \texttt{CompInvg}\,(n,y,g) \leq g^{-1}(y) + \frac{1}{n}, \tag{3}$$

thus, `CompInvg` approximates g^{-1} in a controlled way.

Function `CompInvg`(n,y,g)
  Input : $n \in \mathbb{N}$, $y \in \mathbb{R}_+$, $g : \mathbb{R}_+ \to \mathbb{R}_+$
  Output: $m \in \mathbb{R}_+$, an approximation to $g^{-1}(y)$
  $m = \min\{k/n : k \in \mathbb{N}, g(k/n) > y\}$
end

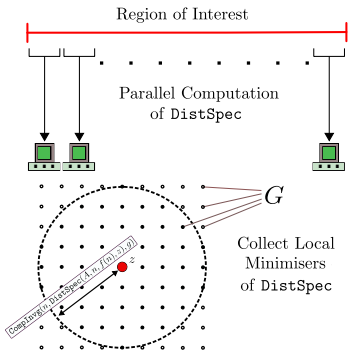


FIG. 1. Schematic of the method.

Moreover, in many cases, we can use an explicit form for g^{-1} and in that case one does not need `CompInvg`. For each output z we have an error estimate

$$E(n,z) = \texttt{CompInvg}\ (n,\texttt{DistSpec}\ (A,n,f(n),z)+c_n,\ g), \tag{4}$$

and this gives a guaranteed estimate which converges uniformly to $g^{-1}(\|R(z,A)\|^{-1})$. We can now give the main routine of the algorithm, shown in `CompSpec`.

Function `CompSpec`$(A,n,f(n),c_n,g)$
  Input : $n, f(n) \in \mathbb{N}$, $c_n \in \mathbb{R}_+$, $g : \mathbb{R}_+ \to \mathbb{R}_+$, $A \in \mathcal{B}(l^2(\mathbb{N}))$
  Output: $\Gamma \subset \mathbb{C}$, an approximation to $\mathrm{Sp}(A)$, $E \in \mathbb{R}_+$, the error estimate
  $G = \texttt{Grid}(n)$
  for $z \in G$ **do**
    $F(z) = \texttt{DistSpec}(A,n,f(n),z)$
    if $F(z) \leq 1/2$ **then**
      for $w_j \in B_{\texttt{CompInvg}(n,F(z),g)}(z) \cap G = \{w_1, ..., w_k\}$ **do**
        $F_j = \texttt{DistSpec}(A,n,f(n),w_j)$
      end
      $M_z = \{w_j : F_j = \min_q\{F_q\}\}$
    else
      $M_z = \emptyset$
    end
  end
  $\Gamma = \cup_{z\in G} M_z$
  $E = \max_{z\in\Gamma}\{\texttt{CompInvg}(n,\texttt{DistSpec}(A,n,f(n),z)+c_n,\ g)\}$
end

Algorithm 1: The subroutine `CompSpec` is the main routine in the algorithm.

Here, $B_\delta(z)$ denotes the closed disc of radius $\delta > 0$ centred at z. The simple idea of `CompSpec` is a local search routine. If z has $\texttt{DistSpec}(A,n,f(n),z) \leq 1/2$, we search in a radius $r = \texttt{CompInvg}(n, \texttt{DistSpec}(A,n,f(n),z), g)$ around z to minimise the approximated distance to the spectrum. This is our best guess of points in the spectrum near z. The output is then the collection of these local minimisers. This is shown in Figure 1.

There is also a physical interpretation of the method. If we are dealing with a finite range interaction Hamiltonian, then f can be chosen such that $A_{k,n} = 0$ for $k > f(n)$. In other words, the rectangular truncation $P_{f(n)}AP_n$ contains all of the interactions of the first n sites without needing to apply boundary conditions. The algorithm naturally uses just enough information to capture these interactions correctly.

Simplifications in the Normal/Hermitian Case

In the Hermitian case (the operator A being normal suffices), the algorithm is simplified as follows. We only need to check positive definiteness corresponding to one matrix in the routine `DistSpec`, and the altered version is shown as `DistSpecH`.

Function `DistSpecH`(A,n,$f(n)$,z)
  Input : $n \in \mathbb{N}$, $f(n) \in \mathbb{N}$, $A \in \mathcal{B}(l^2(\mathbb{N}))$, $z \in \mathbb{C}$
  Output: $y \in \mathbb{R}_+$, an approximation to the function $z \mapsto \mathrm{dist}(z, \mathrm{Sp}(A))$
  $B = (A - zI)(1 : f(n), 1 : n)$
  $\nu = 1$, $l = 0$
  while $\nu = 1$ **do**
    $l = l + 1$
    $\nu =$ `IsPosDef`$(B^*B - \frac{l^2}{n^2})$
  end
  $y = \frac{l}{n}$
end

`DistSpecH` has the added advantage of converging uniformly on compact subsets to the true error. Hence we can use it to gain a very good approximation of the error for outputs of other methods such as finite section. Furthermore, `DistSpecH` $+ c_n$ converges to the true error from above and hence we gain an error control on the output. We can also take the function $g(x) = x$ in the routine `CompSpec`, and, in the self-adjoint case, restrict everything to the real axis. The modified main routine will be denoted by `CompSpecH` with g suppressed in the notation.

Theorems on Classifications

Here we collect together the theorems and proofs of the classification of the spectral problem and prove our algorithm converges. We begin with the normal case and then extend the proof to the more general case. To state the theorem we recall that we have $f : \mathbb{N} \to \mathbb{N}$, $f(n) \geq n$ and $c = \{c_1, c_2, \ldots\}$, a sequence such that $c_n \to 0$ as $n \to \infty$. Moreover, we let Ω_1 denote the set of normal bounded operators A on $l^2(\mathbb{N})$ such that $D_{f,n}(A) \leq c_n$.

Theorem 2 (Hermitian and normal case)**.** *For $A \in \Omega_1$, let*

$$[\Gamma_n(A), E(n)] = \textit{ CompSpecH } (A,n,f(n),c_n).$$

Then $\Gamma_n(A) \to \mathrm{Sp}(A)$ and $E(n) \to 0$ as $n \to \infty$ and $\Gamma_n(A)$ is contained in the $E(n)$ neighbourhood of $\mathrm{Sp}(A)$*. Moreover, Γ_n can be implemented using finitely many arithmetic operations and comparisons on the matrix elements of A and the statement (i) on page 1 also holds for this class.*

Sketch of proof. From the definition of `CompSpecH` it is clear that this algorithm requires only finitely many arithmetic operations and comparisons except for the `IsPosDef` routine. Note, however, that if the `IsPosDef` routine is implemented similar to a Cholesky decomposition (which requires radicals), one does actually not need to form the full decomposition and can, therefore, execute the routine by using only arithmetic operations and comparisons. By considering diagonal operators, it is straightforward to see that (i) also holds.

The proof that $\Gamma_n(A) \to \mathrm{Sp}(A)$ and $E(n) \to 0$ as $n \to \infty$, and the fact that $\Gamma_n(A)$ is contained in the $E(n)$ neighbourhood of $\mathrm{Sp}(A)$ is close to the more general case in Theorem 3, so we use the Lemma 1 stated there, replacing h by the identity function. We also vary the definitions of γ_n and $\gamma_{m,n}$,

$$\gamma_{m,n}(z) := \sigma_1(P_n(A - zI)P_m), \qquad \gamma_n(z) := \sigma_1((A - zI)P_n),$$

where σ_1 denotes the smallest singular value (or injection modulus), $\sigma_1(T) = \inf\{\|T\psi\| : \|\psi\| = 1\}$. In the case of $T = (A - zI)P_n$ we take inf over all ψ of norm 1 in the span of $\{e_1, ..., e_n\}$. Now define $\gamma(z) := \sigma_1(A - zI)$. We can write

$$\texttt{DistSpecH}(A, n, f(n), z) = \min\{k/n : k \in \mathbb{Z}, k/n \geq \gamma_{n,f(n)}(z)\}$$

and hence we must have that $0 \leq \texttt{DistSpecH}(A,n,f(n),z) - \gamma_{n,f(n)}(z) \leq 1/n$. For any compact set K, we then have

$$\begin{aligned}\|\gamma - \texttt{DistSpecH}(A,n,f(n),\cdot)\|_{\infty,K} &\leq \|\gamma - \gamma_n\|_{\infty,K} + \left\|\gamma_n - \gamma_{n,f(n)}\right\|_{\infty,K} \\ &\quad + \left\|\gamma_{n,f(n)} - \texttt{DistSpecH}(A,n,f(n),\cdot)\right\|_{\infty,K} \\ &\leq \|\gamma - \gamma_n\|_{\infty,K} + \left\|\gamma_n - \gamma_{n,f(n)}\right\|_{\infty,K} + \frac{1}{n},\end{aligned} \tag{5}$$

where the subscript K on the norm denotes the restriction of the function to K. It is easy to see that γ_n is decreasing and converges pointwise to γ since $P_n \to I$ strongly. As γ is continuous, Dini's theorem then gives that $\|\gamma - \gamma_n\|_{\infty,K} \to 0$. By definition of f and the fact that $f(n) \geq n$ we have that

$$\|P_{f(n)}(A - zI)P_n - (A - zI)P_n\| = \|P_{f(n)}AP_n - AP_n\| \longrightarrow 0, \quad n \to \infty.$$

We then use the inequality $|\sigma_1(B + C) - \sigma_1(B)| \leq \|C\|$ to gain $\left\|\gamma_n - \gamma_{n,f(n)}\right\|_{\infty,K} \to 0$. Thus, for any $\epsilon > 0$, by Eq. 5, we can choose n large such that $\|\gamma - \texttt{DistSpecH}(A,n,f(n),\cdot)\|_{\infty,K} < \epsilon$. We can then use Lemma 1 to prove convergence in the same way as the proof of Theorem 3.

To get the error estimate, suppose that $z \in \Gamma_n(A)$ then we know that $\operatorname{dist}(z, \operatorname{Sp}(A)) = \gamma(z) \leq \gamma_n(z)$. In the case that the operator is quasi-banded with $f(m) = l(m) + m$ we have $\gamma_{n,f(n)} = \gamma_n$ and hence $\operatorname{dist}(z, \operatorname{Sp}(A)) \leq \texttt{DistSpecH}(A,n,f(n),z) =: E(n,z)$. Since $\gamma_n \downarrow \gamma = \operatorname{dist}(z, \operatorname{Sp}(A))$ uniformly and $\mathrm{d}_H(\Gamma_n(A), \operatorname{Sp}(A)) \to 0$, our error estimate converges uniformly to 0 over the output $\Gamma_n(A)$. If we know c_n then $\left\|\gamma_n - \gamma_{n,f(n)}\right\|_{\infty,K} \leq c_n$ and hence

$$\operatorname{dist}(z, \operatorname{Sp}(A)) \leq \texttt{DistSpecH}(A,n,f(n),z) + c_n =: E(n,z)$$

and we argue as before. Furthermore, in both cases we actually have that

$$E(n,z) \to \operatorname{dist}(z, \operatorname{Sp}(A))$$

uniformly over any compact set. Hence this error estimate gives us a good idea of the error of other methods such as finite section. □

Note that as well as uniform convergence, the proof makes it clear that, in the quasi-banded case, we have, up to an error of $1/n$, that $E(n,z)$ decreases monotonically from above down to the true error. This can be tweaked to gain accurate values of $\gamma_{n,f(n)}$ by increasing the search resolution in `DistSpecH`. We now turn to the general case. Recall that Ω_2 denotes the set of all operators with dispersion bounded by f and with a known bound of the form

$$\|R(z,A)\|^{-1} \geq g(\operatorname{dist}(z, \operatorname{Sp}(A))),$$

for $z \in \mathbb{C}$.

Theorem 3 (General case with controlled growth of the resolvent)**.** *Let $A \in \Omega_2$ and*

$$[\Gamma_n(A), E(n)] = \ \texttt{CompSpec}\ (A,n,f(n),c_n,g).$$

Then $\Gamma_n(A) \to \operatorname{Sp}(A)$ and $E(n) \to 0$ as $n \to \infty$ and $\Gamma_n(A)$ is contained in the $E(n)$ neighbourhood of $\operatorname{Sp}(A)$. Moreover, Γ_n can be implemented using finitely many arithmetic operations and comparisons on the matrix elements of A and the statement (i) on page 1 also holds for this class.

Sketch of proof. Again by considering diagonal operators we see that (i) holds and hence for the classification it is enough to prove the statements regarding $[\Gamma_n(A), E(n)]$. Let $A \in \Omega_2$ and define $\gamma(z) = \min\{\sigma_1(A - zI), \sigma_1(A^* - \bar{z}I)\}$, where again σ_1 denotes the injection modulus. It is well known (see [3] for example) that $\gamma(z) = \|R(z,A)\|^{-1} \leq \operatorname{dist}(z, \operatorname{Sp}(A))$.

For any $\zeta : \mathbb{C} \to \mathbb{R}_+$ define the algorithm `CompSpecII`$(A,n,f(n),c_n,g,\zeta)$ to be the subroutine `CompSpec`$(A,n,f(n),c_n,g)$, where `DistSpec`$(A,n,f(n),z)$ is replaced by $\zeta(z)$. Define

$$\begin{aligned}\gamma_{m,n}(z) &= \min\{\sigma_1(P_n(A - zI)P_m), \sigma_1(P_n(A^* - \bar{z}I)P_m)\}, \\ \gamma_n(z) &= \min\{\sigma_1((A - zI)P_n), \sigma_1((A^* - \bar{z}I)P_n)\},\end{aligned}$$

and note that in these definitions we need to include the adjoint A^* since our operators may no longer be normal. As before we can write

$$\texttt{DistSpec}(A,n,f(n),z) = \min\{k/n : k \in \mathbb{Z}, k/n \geq \gamma_{n,f(n)}(z)\}$$

with $0 \leq \texttt{DistSpec}(A,n,f(n),z) - \gamma_{n,f(n)}(z) \leq 1/n$. Thus, by arguing similarly to the proof of Theorem 2 we get that $\|\gamma - \texttt{DistSpec}(A,n,f(n),\cdot)\|_{\infty,K} \to 0$ as $n \to \infty$ for any compact set K. For simplicity let $h = g^{-1}$ and let $C > 0$ such that $\mathrm{Sp}(A) \subset B_C(0)$ (the closed ball of radius C about the origin). We then define

$$K_1 = B_{C+h(1/2)+1}(0), \quad K_2 = B_{C+h(1/2)+h(2C+h(1/2)+2)+2}(0).$$

Of course these are not known but they will be used to prove convergence of the algorithm via the following lemma. Note that $\texttt{DistSpec}(A,n,f(n),\cdot)$ satisfies the conditions of the lemma for large n.

Lemma 1. *Let K_1 and K_2 be as above and let $0 < \epsilon < 1/4$ and $n > 2/\epsilon$. Suppose there is a function ζ with $\|\zeta - \gamma\|_{\infty,K_2} < \epsilon$ and $\zeta(w) > 1/2$ for all $w \in \mathbb{C}\backslash K_1$. Then*

$$\mathrm{d_H}(\textit{CompSpecII}(A,n,f(n),c_n,g,\zeta), \mathrm{Sp}(A)) \leq u(\epsilon),$$

where $u(\eta) = \max\{h(3\eta + h(t+\eta) - h(t)) + \eta : t \in [0,1/2]\}$. Also, $u(\epsilon) \to 0$ as $\epsilon \to 0$.

To finish the proof we need the establish the remaining part of the theorem regarding the the error. Note that if $c_n = 0$ we have $\gamma_{n,f(n)} = \gamma_n$ and hence $g(\mathrm{dist}(z,\mathrm{Sp}(A))) \leq \texttt{DistSpec}(A,n,f(n),z)$. Now define $\tilde{E}(n,z)$ such that $\texttt{DistSpec}(A,n,f(n),z) =: g(\tilde{E}(n,z))$. Since $\gamma_n \downarrow \gamma = \|R(z,A)\|^{-1}$ uniformly and $\mathrm{d_H}(\Gamma_n(A),\mathrm{Sp}(A)) \to 0$ then our error estimate converges uniformly to 0 over the output $\Gamma_n(A)$. If we know $c_n \geq \left\|\gamma_n - \gamma_{n,f(n)}\right\|_{\infty,K}$ then we have

$$g(\mathrm{dist}(z,\mathrm{Sp}(A))) \leq \texttt{DistSpec}(A,n,f(n),z) + c_n =: g(\tilde{E}(n,z))$$

and we argue as before. Again, in both cases we have

$$\tilde{E}(n,z) \to h(\|R(z,A)\|^{-1}) \tag{6}$$

uniformly over any compact set. The final result follows from Eq. 3. □

Proof of Lemma 1. Let $z \in \texttt{Grid}(n)$ with $\zeta(z) \leq 1/2$ and $I_z = B_{\texttt{CompInvg}(n,\zeta(z),g)}(z) \cap \texttt{Grid}(n) \neq \emptyset$, the set we compute local minima over. It follows that $z \in K_1$. If $v \in I_z$ then

$$\begin{aligned}|z-v| &\leq r^n(\zeta(z)) \leq r^n(\gamma(z)+\epsilon) \text{ (as } h \text{ increasing)}\\ &\leq r^n(\mathrm{dist}(z,\mathrm{Sp}(A))+\epsilon) \leq h(\mathrm{dist}(z,\mathrm{Sp}(A))+\epsilon) + 1/n\\ &\leq h(2C + h(1/2) + 2) + 1,\end{aligned}$$

where we let $r^n(y) = \texttt{CompInvg}(n,y,g)$. This implies that $I_z \subset K_2$. Recall M_z from $\texttt{CompSpec}$ and suppose that $M_z \neq \emptyset$. By the growth condition, monotonicity of h and the compactness of $\mathrm{Sp}(A)$ there exists some $y \in \mathrm{Sp}(A)$ of minimal distance to z with $|z-y| \leq h(\gamma(z))$. By the assumptions on ζ we have $|z-y| \leq h(\zeta(z)+\epsilon)$. If $r^n(\zeta(z)) \leq 1/n$ then $|z-y| \leq h(\zeta(z)+\epsilon) - h(\zeta(z)) + 2/n$ and let $v = z$. If $r^n(\zeta(z)) > 1/n$ then a simple geometrical argument (which also works in $\mathbb{R}$) shows that there exists $v \in I_z$ with $|v-y| \leq h(\zeta(z)+\epsilon) - h(\zeta(z)) + 2/n$. Now $\gamma(v) \leq \mathrm{dist}(v,\mathrm{Sp}(A))$ and so, since $v \in K_2$, $\zeta(v) < \gamma(v) + \epsilon < h(\zeta(z)+\epsilon) - h(\zeta(z)) + 2\epsilon$ as $2/n < \epsilon$. By definition of M_z as a minimiser, this inequality must hold for all $w \in M_z$ and hence

$$\begin{aligned}\mathrm{dist}(w,\mathrm{Sp}(A)) &= h(g(\mathrm{dist}(w,\mathrm{Sp}(A)))) \leq h(\gamma(w))\\ &\leq h(\zeta(w)+\epsilon) \leq u(\epsilon),\end{aligned}$$

since $\zeta(z) \leq 1/2$. This holds for all points of $\texttt{CompSpecII}(A,n,f(n),c_n,g,\zeta)$ since $z \in \texttt{Grid}(n)$ was arbitrary.

Conversely, let $y \in \mathrm{Sp}(A)$ then there must exist some point $z \in \texttt{Grid}(n) \cap K_2$ with $|z-y| \leq 1/n$. It follows that $\zeta(z) < \gamma(z) + \epsilon \leq 2\epsilon < 1/2$. So M_z is not empty and if $w \in M_z$ then

$$\begin{aligned}|w-y| \leq& |w-z| + |z-y|\\ \leq& h(\zeta(z)) + 1/n + |z-y| \leq h(\zeta(z)) + \epsilon \leq u(\epsilon)\end{aligned}$$

where we use $\zeta(z) \leq 2\epsilon$ in the last line. The claim now follows. To finish the argument, let u be the maximum over t of $(t,\eta) \mapsto h(h(t+\eta) - h(t) + 3\eta) + \eta$ on $[0,1/2]^2$ which must be uniformly continuous. For fixed t, this function tends to 0 as $\eta \to 0$ and hence the convergence must be uniform in t. □

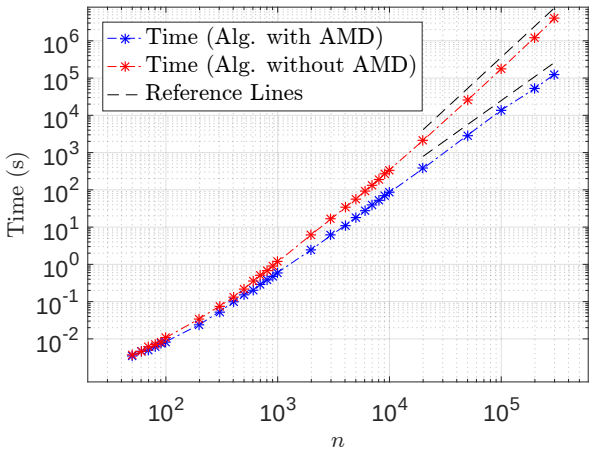


FIG. 2. Speed-up of method when we take advantage of the structure preservation under changes in z and use AMD ordering. The AMD ordering only needs to be calculated once for each n and can subsequently be used on all test points. Both these plots are for the operator negative Laplacian H_0 on the Penrose tile.

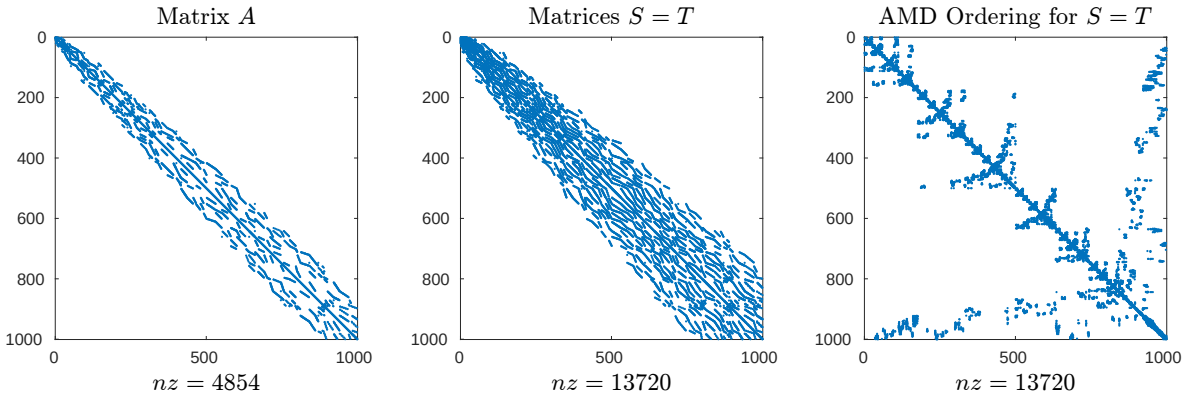


FIG. 3. Non-zero entries of the matrices $S = T$ with and without permutations of basis elements. The number of non-zero entries is denoted by nz.

Numerical Implementation and Speed-up

For sparse matrices A, the finite truncations $A(1 : f(n), 1 : n)$ will be sparse but the matrices S, T in `DistSpec` may be far from banded. To take advantage that S, T have the same shape as we vary our test point z, we can calculate a permutation of the indices corresponding to an Approximate Minimum Degree (AMD) ordering. This is a standard procedure to reduce the number of operations needed for Cholesky Decomposition or Gaussian Elimination. This can be computed with the Matlab commands `[~,~,Q1]=chol(S-speye(m))` and `[~,~,Q2]=chol(T-speye(m))`, where S, T are $m \times m$. We can then replace S, T by $S(Q1, Q1)$ and $T(Q2, Q2)$ in subsequent calculations. As shown in Figure 2 this offers considerable speed-up, especially in two-dimensional models where the initial matrix A is not banded. For the case considered in the text, the time taken was of order $\sim \mathcal{O}(n^{2.1})$ and $\sim \mathcal{O}(n^{2.8})$ for large n with and without the AMD ordering respectively (shown as reference lines). For the case of self-adjoint operators and real z, $S = T$ and Figure 3 shows the non-zero entry patterns of these matrices and their permutations for $n = 1000$ and the operator H_0 on the Penrose tile. Of course, for large sparse rectangular truncations $P_n(A - zI)P_m$, there exist efficient iterative methods to approximate the smallest singular value. We found the partial Cholesky approach (with interval bisection and AMD ordering) slightly faster for the examples in this paper, but note that the user can easily use different subroutines for the computation of the smallest singular value. For the case of computing pseudospectra (e.g. Fig. 4), using the partial Cholesky positive definite test is more efficient since we can test levels of the resolvent norm on a logarithmic scale and we found it to be more stable for non-normal A. It is also easier to implement the incomplete Cholesky approach when using interval arithmetic, allowing completely rigorous error bounds.

The number of operations, pre-AMD ordering can be bounded in the following way.

Proposition 4. *Let A be a bounded operator and suppose that for large n, $f(n) - n \sim Cn^{\alpha}$, where f is the dispersion function, C a constant and $\alpha \in [0, 1)$. Suppose that f is non-decreasing and also describes the off-diagonal sparsity*

structure of A in the sense that $A_{n,k}, A_{k,n} = 0$ if $k > f(n)$. If we use $m_1(n)$ test points and an accuracy of $1/m_2(n)$ for `DistSpec`*, then the proposed algorithm can be computed in*

$$\mathcal{O}(m_1(n)n^{(\alpha+1)\alpha+1}\log(m_2(n)))$$

operations.

Proof. We will show first that testing positive definiteness of $S - \epsilon$ in `DistSpec` can be achieved in $\mathcal{O}(n^{(\alpha+1)\alpha+1})$ operations. Is is then clear that the computations of all the estimates using `DistSpec` can be achieved in $\mathcal{O}(m_1(n)n^{(\alpha+1)\alpha+1}\log(m_2(n)))$ operations if we use a binary search routine. It is easily seen that the rest of `CompSpec` can be executed in $\mathcal{O}(m_1(n)n)$ operations, yielding the result.

To test positive definiteness (`IsPos`) we checked whether a Cholesky decomposition of the matrix $S-\epsilon$ was possible. One can see that S also has a dispersion function $\tilde{f}(n)-n \sim \tilde{C}n^{\alpha}$ and hence without loss of generality we can assume $f = \tilde{f}$. Furthermore, $S - \epsilon$ is sparse with f describing its sparsity structure. We refer the reader to [4] Chapter 23 where Cholesky factorisation is explained. Following the notation there, one computes (assuming $S > \epsilon$)

$$S - \epsilon = R_1^* ... R_m^* R_m ... R_1$$

with $R = R_m ... R_1$ upper triangular. Using the fact that f is non-decreasing with $f(n) \geq n$ it is straightforward to prove that all R_i's used to compute R have the same sparsity/dispersion function f. A simple operation count gives complexity of order

$$\sum_{k=1}^{n}\sum_{j=k+1}^{f(k)}(f(j)-j) \lesssim \sum_{k=1}^{n}\sum_{j=k+1}^{f(k)} j^{\alpha} \lesssim \sum_{k=1}^{n}(f(k)^{\alpha+1} - k^{\alpha+1}) \lesssim \sum_{k=1}^{n} k^{(\alpha+1)\alpha} \lesssim n^{(\alpha+1)\alpha+1}$$

and we get the result. □

Remark 5. If we are studying a finite range Hamiltonian on the lattice $l^2(\mathbb{Z}^d)$ then one can choose $\alpha = (d-1)/d$ and in the general case of such Hamiltonians this is easily seen to be optimal. If $m_1 = Ln, m_2 = n$ then in two dimensions for a constant L this reduces to $n^{2.75}\log(n)$ which is the slope in Figure 2.

Approximate States

In this section we mention how the proposed algorithm can be used to gain states corresponding to elements in the spectrum in addition to the spectrum itself. For simplicity we will consider an operator on $l^2(\mathbb{N})$. For such an operator, not all of the spectrum is composed of eigenvalues. In the normal/Hermitian case, given $z \in \mathrm{Sp}(A)$, there exists a sequence of unit vectors $x_n \in l^2(\mathbb{N})$ such that

$$\left\|P_{f(n)}(A - zI)P_n x_n\right\| \to 0 \text{ as } n \to \infty.$$

Such a sequence is known as an approximate eigenvector sequence or an approximate eigenstate sequence. In the non-normal case, one only has the existence of x_n such that at least one of the quantities $\left\|P_{f(n)}(A - zI)P_n x_n\right\|$ and $\left\|P_{f(n)}(A^* - \overline{z}I)P_n x_n\right\|$ converge to zero. The question is whether given a z in the output $\Gamma_n(A)$ of the algorithm and an approximation of $\gamma_{n,f(n)}(z)$, we can find a x_n satisfying

$$\min\{\|(A - zI)x_n\|, \|(A^* - \overline{z}I)x_n\|\} \leq \gamma_{n,f(n)}(z).$$

The convergence proof of the algorithm shows that such a sequence will be an approximate eigenvector sequence.

Theorem 6 (Approximate States)**.** *Suppose A is a bounded operator with dispersion bound f. Given any $z \in \Gamma_n(A)$ with computed bound on $\gamma_{n,f(n)}(z)$, we can compute a corresponding vector x_n satisfying*

$$\min\{\left\|P_{f(n)}(A - zI)P_n x_n\right\|, \left\|P_{f(n)}(A^* - \overline{z}I)P_n x_n\right\|\} \leq \gamma_{n,f(n)}(z)$$

in finitely many arithmetic and square root operations without any iterative procedure.

Function `ApproxState`$(A,n,f(n),z,\epsilon)$
  Input : $n \in \mathbb{N}$, $f(n) \in \mathbb{N}$, $A \in \mathcal{B}(l^2(\mathbb{N}))$, $z \in \mathbb{C}$, $\epsilon = E(n,z)^2$
  Output: $x_n \in \mathbb{C}^n$, a vector satisfying $\left\|P_{f(n)}(A - zI)P_n x_n\right\| \leq E(n,z)$
  $B = [(A - zI)(1 : f(n), 1 : n)]^*[(A - zI)(1 : f(n), 1 : n)] - \epsilon I$
  $[L, D] =$ `ldl`(B)
  if *D is diagonal* **then**
    Find i with $D(i,i) \leq 0$
    $y = e_i$
  else
    Find y eigenvector of D with eigenvalue ≤ 0
  end
  Solve upper triangular system for x_n with $y = L^* x_n$.
end

Algorithm 2: The subroutine `ApproxState` computes approximate eigenvalues/states. The `ldl` denotes a standard subroutine that computes the LDL^* decomposition.

Proof. We will deal with the normal case and note that dealing with the general case is simply a matter of applying the following argument to $(A^*, \overline{z})$ as well as (A, z). Let $\epsilon = (\gamma_{n,f(n)}(z))^2$ and consider the matrix

$$B = [(A - zI)(1 : f(n), 1 : n)]^*[(A - zI)(1 : f(n), 1 : n)] - \epsilon I$$

then B is a Hermitian matrix but not positive definite. It follows that B can be put into the form

$$PBP^T = LDL^*,$$

where L is lower triangular with 1's along its diagonal, D is block diagonal with block sizes 1 or 2 and P is a permutation matrix. This can be computed in finitely many arithmetic operations. Without loss of generality we assume that $P = I$. Let x be an eigenvector of B with non-positive eigenvalue then set $y = L^* x$. Such an x exists by assumption. Note that

$$\langle y, Dy \rangle = \langle L^* x, DL^* x \rangle = \langle x, Bx \rangle \leq 0.$$

It follows that there exists a unit vector y_n with $\langle y_n, Dy_n \rangle \leq 0$. Such a vector is easy to spot by either considering 1 blocks or 2 blocks (where we need to extract square roots) in the block diagonal matrix D. L^* is invertible and upper triangular so we can efficiently solve for $\tilde{x}_n = (L^*)^{-1} y_n$ and then normalise to get x_n. Finally note that

$$\left\|P_{f(n)}(A - zI)P_n x_n\right\|^2 = \langle x_n, Bx_n \rangle + \epsilon = \frac{1}{\|\tilde{x}_n\|^2} \langle y_n, Dy_n \rangle + \epsilon \leq \epsilon.$$

□

The upshot of this is that the algorithm not only computes $\Gamma_n(A)$ converging to the spectrum of A, but it also computes approximating eigenvector sequences for the spectrum. Since not all of the spectrum is necessarily composed of eigenvectors in the infinite dimensional case, this is the best any algorithm can hope to achieve in generality. Furthermore, even in the finite dimensional case this is the best possible owing to numerical errors due to round-off and finite precision. Finally, we note that this method is very quick and can be efficiently implemented. The algorithm is shown in the subroutine `ApproxState` for the Hermitian case which takes as input A, n, $f(n)$, z and the corresponding error bound $\epsilon = E(n,z)^2$ and computes the corresponding approximate state x_n.

Computing Pseudospectra

Here we state and prove the theorem on the use of the algorithm to compute pseudospectra, as well as show that the algorithm produces output without error in this case. Recall that if A is a bounded linear operator on a Hilbert space $\mathcal{H}$ and $\epsilon > 0$, then the ϵ-pseudospectrum of A is defined as

$$\mathrm{Sp}_\epsilon(A) = \{z : \left\|(A - zI)^{-1}\right\| \geq \epsilon^{-1}\}.$$

Function `PseudoSpec`$(A,n,f(n),c_n,\epsilon)$
  Input : $n, f(n) \in \mathbb{N}$, $c_n \in \mathbb{R}_+^{\mathbb{N}}$, $A \in \mathcal{B}(l^2(\mathbb{N}))$, $\epsilon > 0$
  Output: $\Gamma \subset \mathbb{C}$, an approximation to $\mathrm{Sp}_\epsilon(A)$
  $G = \mathtt{Grid}(n)$
  $m = \min\{k \geq n \,|\, c_k < \epsilon\}$
  for $z \in G$ **do**
    $B = (A - zI)(1 : f(m), 1 : m)$
    $C = (A - zI)^*(1 : f(m), 1 : m)$
    $S = B^*B$
    $T = C^*C$
    $p = \mathtt{IsPosDef}(S - (\epsilon - c_m)^2)$
    $q = \mathtt{IsPosDef}(T - (\epsilon - c_m)^2)$
    $\nu(z) = \min(p, q)$
  end
  $\Gamma = \bigcup\{z \in G \,|\, \nu(z) = 0\}$
end

Algorithm 3: The `PseudoSpec` routine computes an approximation to the true pseudospectrum of the operator with the property that it is always inside the true pseudospectrum, and of course converges as n becomes large.

We now let Ω_3 be the set of all (bounded) operators A on $l^2(\mathbb{N})$ such that $D_{f,n}(A) \leq c_n$. Consider first $\psi_n(A, z) = \min\{\sigma_1(P_{f(n)}(A - zI)P_n), \sigma_1(P_{f(n)}(A^* - \bar{z}I)P_n)\}$ where σ_1 is the usual injection modulus and define

$$\tilde{\Gamma}_n^\epsilon(A) = \{z \in \mathtt{Grid}(m) : \psi_n(A, z) \leq \epsilon\}. \tag{7}$$

In the case that A is quasi-banded with $f(m) = m + l(m)$ then it is clear that the output is guaranteed to be inside the pseudospectrum, since the resolvent norm estimate can only increase as we increase n (see the proof of Theorem 3). If we know that $D_{f,n}(A) \leq c_n$, we can use the following trick even with no known behaviour of $\|R(z, A)\|$ between level sets. Recall that $|\sigma_1(B + C) - \sigma_1(B)| \leq \|C\|$ for bounded operators B, C and let n be large such that $c_n < \epsilon$. Replace $\tilde{\Gamma}_n^\epsilon(A)$ by $\tilde{\Gamma}_m^{\epsilon - c_m}(A)$ with $m \geq n$ such that $c_m < \epsilon$.

Theorem 7 (Pseudospectra)**.** *Let $A \in \Omega_3$ and*

$$\Gamma_n(A) = \mathtt{PseudoSpec}(A, n, f(n), c_n, \epsilon).$$

Then $\Gamma_n(A) \to \mathrm{Sp}_\epsilon(A)$ as $n \to \infty$ and $E(n) = 0$. Moreover, Γ_n can be implemented using finitely many arithmetic operations and comparisons on the matrix elements of A and the statement (i) on page 1 also holds for this class.

Proof. If D is a diagonal operator and M the closure of its diagonal matrix elements then we have $\mathrm{Sp}_\epsilon = M + B_\epsilon(0)$. Extending the usual argument for the spectrum, it is clear that (i) holds.

To show the convergence result, first we consider the case when $c_m = 0$ for all m. We use the definition of γ_n, γ as in the proof of Theorem 3. Recall that $f(m) \geq m$ and we have

$$\begin{aligned} \psi_n(A, z) &= \min\left\{\sigma_1(P_{f(n)}(A - zI)P_n), \sigma_1(P_{f(n)}(A^* - \overline{z}I)P_n)\right\}, \\ \Gamma_n(A) = \tilde{\Gamma}_n^\epsilon(A) &= \{z \in \mathtt{Grid}(n) : \psi_n(A, z) \leq \epsilon\}. \end{aligned}$$

Note that $\Gamma_n(A)$ relies on arithmetic operations on only finitely many computations of A's matrix elements. Let $|z| > \|A\|$ then we know via a standard series argument that $A - zI$ is invertible with

$$\|R(z, A)\| \leq \frac{1}{|z| - \|A\|}.$$

Consider $\tilde{R}_n(z) := R(z, P_{f(n)}AP_{f(n)}) \in \mathcal{B}(P_{f(n)}\mathcal{H})$ and note that for any operator B, $\sigma_1(BP_{m_1}) \geq \sigma_1(BP_{m_2})$ for $m_1 \leq m_2$. By considering the Banach algebra $\mathcal{B}(P_{f(n)}\mathcal{H})$ we have

$$\psi_n(A, z) \geq \frac{1}{\left\|\tilde{R}_n(z)\right\|} \geq |z| - \left\|P_{f(n)}AP_{f(n)}\right\| \geq |z| - \|A\|.$$

Similarly $\gamma_n(z) \geq |z| - \|A\|$ for all such z. It follows that, given $\epsilon > 0$, we can choose a compact subset $K \subset \mathbb{C}$ with $\gamma_n(z) > 2\epsilon$ and $\psi_n(A,z) > 2\epsilon$ for all $z \in \mathbb{C}\backslash K$ and for all n. We have already seen that $\|\gamma_n - \psi_n\|_{\infty,K} \to 0$ as $n \to \infty$. As proven before, $\gamma_n \downarrow \gamma$ locally uniformly as $n \to \infty$ and hence $\psi_n \to \gamma$ uniformly on K as $n \to \infty$. By the uniform convergence and the Arzelá-Ascoli theorem we can choose $\delta_n \downarrow 0$ such that for all n,

$$|\psi_n(A,z) - \psi_n(A,w)| < \delta_n \text{ for all } z,w \in K \text{ with } |z-w| < 1/n \tag{8}$$

and such that $|\psi_n - \gamma_n| < \delta_n$ on K. Now define the sets

$$\Psi_{\epsilon,n} = \{z \in \mathbb{C} : \psi_n(A,z) \leq \epsilon\}, \quad \Upsilon_{\epsilon,n} = \{z \in \mathbb{C} : \gamma_n(z) \leq \epsilon\}.$$

The above shows that we have for $n \geq k$

$$\Upsilon_{\epsilon-\delta_k,n} \subset \Upsilon_{\epsilon-\delta_n,n} \subset \Psi_{\epsilon,n} \subset \Upsilon_{\epsilon+\delta_n,n} \subset \Upsilon_{\epsilon+\delta_k,n}. \tag{9}$$

We have that $\Gamma_n(A) \subset \Psi_{\epsilon,n}(A)$ and we assume n is large so that $K \subset Q_n(0)$, where $Q_n(0)$ denotes the closed square centred at 0 with side length $2n$. We can write $\texttt{Grid}(n) = (n^{-1}(\mathbb{Z} + i\mathbb{Z})) \cap Q_n(0)$. If this holds and $z_1 \in \Psi_{\epsilon-\delta_n,n}$ then there exists some $z_2 \in \texttt{Grid}(n)$ with $|z_1 - z_2| < 1/n$ and hence $|\psi_n(A,z_1) - \psi_n(A,z_2)| < \delta_n$. It follows that $z_2 \in \Gamma_n(A)$ and hence

$$\Psi_{\epsilon-\delta_n,n}(A) \subset \Gamma_n(A) + B_{1/n}(0).$$

Putting these together and using that $\delta_n \downarrow 0$, we get for $n \geq k$

$$\Psi_{\epsilon-\delta_k,n}(A) \subset \Psi_{\epsilon-\delta_n,n}(A) \subset \Gamma_n(A) + B_{1/n}(0) \subset \Psi_{\epsilon,n}(A) + B_{1/n}(0). \tag{10}$$

As $\gamma_n \downarrow \gamma$ it is clear that $\Upsilon_{\epsilon,n} \subset \mathrm{Sp}_\epsilon(A)$. A simple proof by contradiction using compactness of K and the fact that $\gamma(z) \geq 2\epsilon$ on $\mathbb{C}\backslash K$ shows that we must have $\lim_{n\to\infty} \mathrm{d_H}(\Upsilon_{\epsilon,n}, \mathrm{Sp}_\epsilon(A)) = 0$. Using Eq. 9 and the fact that $\mathrm{Sp}_\epsilon(A)$ is continuous in ϵ (with respect to the Hausdorff metric) we must have $\lim_{n\to\infty} \mathrm{d_H}(\Psi_{\epsilon,n}, \mathrm{Sp}_\epsilon(A)) = 0$. Similarly, letting $n \to \infty$ and then $k \to \infty$ in Eq. 10 we deduce that $\lim_{n\to\infty} \mathrm{d_H}(\Gamma_n, \mathrm{Sp}_\epsilon(A)) = 0$.

To prove that the algorithm converges with non-zero c_m's, let $\delta < \epsilon$. For large m with $c_m < \delta$ we have that

$$\tilde{\Gamma}_m^{\epsilon-\delta}(A) \subset \tilde{\Gamma}_m^{\epsilon-c_m}(A) \subset \mathrm{Sp}_\epsilon(A).$$

Hence, using the convergence proven above we have

$$\limsup_{m\to\infty} \mathrm{d_H}(\tilde{\Gamma}_m^{\epsilon-c_m}(A), \mathrm{Sp}_\epsilon(A)) \leq \mathrm{d_H}(\mathrm{Sp}_{\epsilon-\delta}(A), \mathrm{Sp}_\epsilon(A)).$$

We then note that $\Gamma_n(A) = \tilde{\Gamma}_{m(n)}^{\epsilon-c_{m(n)}}(A)$ with $m(n) \geq n$. Letting $\delta \downarrow 0$ and using continuity of the pseudospectrum we get the result. □

To produce diagrams like those shown in the main text, we choose a grid, corresponding to the wanted resolution in the diagram, and compute the approximation of the resolvent norm at each point. One may also choose the accuracy of the approximation wanted for a given n, which corresponds to a choice of different ϵ or level sets of the resolvent norm. In all cases, the output is guaranteed to be inside the pseudospectrum, i.e. we under-estimate the resolvent norm. Keeping the ϵ constant and increasing n, we have an approximation that fills up each pseudospectrum.

Non-Hermitian Anderson model

Here we give another example of the algorithm in action. Hatano and Nelson initiated the study of the non-Hermitian Anderson model in the context of vortex pinning in type-II superconductors [5]. Their model showed that an imaginary gauge field in a disordered one-dimensional lattice can induce a delocalisation transition. Whilst synthesising such an imaginary vector potential is a challenge in condensed-matter physics, this phenomenon has been investigated in the field of optics [6]. From a computational point of view, non-Hermitian Hamiltonians pose a serious challenge as no previous algorithm is known to work for computing pseudospectra of non-Hermitian operators and providing error bounds. The operator on $l^2(\mathbb{Z})$ can be written as

$$(Hx)_n = e^{-g}x_{n-1} + e^{g}x_{n+1} + V_n x_n$$

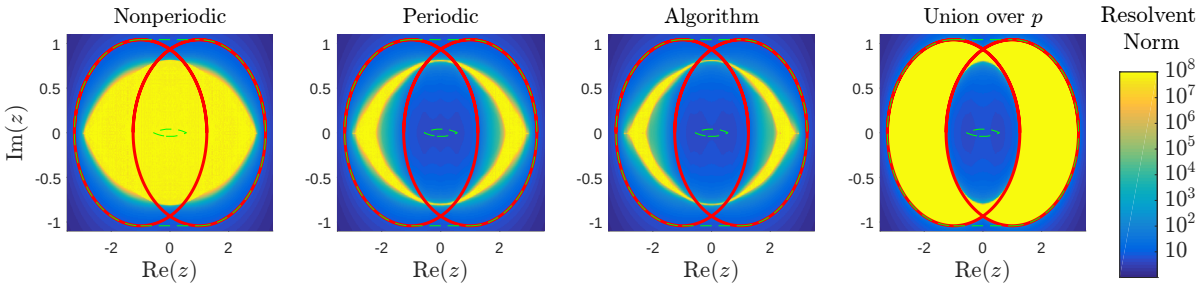


FIG. 4. Pseudospectra of the finite section method with non-periodic boundary conditions shown as contours of the resolvent norm $\|(H_n - zI)^{-1}\|$ for $n = 10^6$. Similar plots for periodic boundary conditions, the new algorithm with and without varying p. Bounds on the spectrum are shown in green and the set $E + M$ in red.

where $g > 0$ and V is a random potential and also has applications in population biology [7].

Spectral computations of H are delicate. Once truncated to a finite lattice of size n, the spectrum and pseudospectrum of the finite section H_n depend on the boundary conditions imposed. Non-periodic boundary conditions (standard finite-section) yield an entirely real spectrum, completely different to the complex spectrum of H. Hatano and Nelson argued that a more physical model would be periodic boundary conditions. However, there is no guarantee that this is correct. In our case, periodic boundary conditions lead to spectra that converge to a curve in the complex plane *strictly* contained in the spectrum [8].

If $(V_n)_{n\in\mathbb{Z}}$ are i.i.d. random variables, then $\mathrm{Sp}(H)$ and $\mathrm{Sp}_\epsilon(H)$ only depend on M, the support of the potential, almost surely. We consider the Bernoulli case $M = \{\pm 1\}$ where $V_n = 1$ with probability $p \in (0, 1)$. This choice ensures the spectrum has a hole in it by a standard series argument. Defining the ellipse $E = \{e^{g+\mathrm{i}\theta} + e^{-g-\mathrm{i}\theta} : \theta \in [0, 2\pi)\}$, we also have $E \pm 1 \subset \mathrm{Sp}(H)$ which is contained in the convex hull of $E + [-1, 1]$. Fig. 4 shows the result of the finite section i.e. the pseudospectra of H_n for $n = 10^6$ (corresponding to a matrix size of $2n + 1$) and the new algorithm with $g = 1/2$ and $p = 1/2$. Non-periodic boundary conditions give the wrong set in the limit $n \to \infty$, filling the hole in the spectrum and converging to the interval $[-3, 3]$ (this can be proven).

We can take advantage of the fact that, up to round off errors, the new algorithm has zero error in its output and that the pseudospectrum is invariant under changes in $p \in (0, 1)$. Thus, we have also shown the output over a union of varying p. This gives a very good estimate of the spectrum and the pseudospectrum given that the computation is guaranteed without error.

Computational Boundaries in Quantum Mechanics

In this section we extend our results to the problem of computing spectra and pseudospectra of Hamiltonians on graphs. Consider a connected, undirected graph G, such that the degree of each vertex is finite and such that the set of vertices $V(G)$ is countably infinite. Throughout this section we consider an arbitrary, but fixed, such G. We may then pick any vertex v and label it e_1. We can enumerate $v's$ neighbours (including itself) as $S_1 = \{e_1, e_2, ..., e_{q_1}\}$ for some finite q_1. Consider the set of neighbours of these vertices and label these as $S_2 = \{e_1, ..., e_{q_2}\}$ for some finite q_2 where we continue the enumeration of S_1. Continue this process inductively enumerating S_m. By the connectivity assumption and countability of V, we may identify $\{e_1, e_2, ...\}$ with an orthonormal basis of $l^2(V(G))$. In what follows, we write $v \sim_k w$ for two vertices $v, w \in V$ if there is a path of at most k edges in E (the set of edges) connecting v and w. In other words v, w are k-th nearest neighbours. We will also use the abuse of notation of identifying each $v \in V$ with its corresponding e_i in $l^2(V(G))$, including in the use of the inner product $\langle\cdot,\cdot\rangle$ on $l^2(V(G))$.

Now let Ω_4 denote the set of all bounded operators A on $l^2(V(G)) \cong l^2(\mathbb{N})$ such that A has the property that the set $S(v) := \{w \in V : \langle w, Av\rangle \neq 0 \text{ or } \langle w, A^*v\rangle \neq 0\}$ is finite for any $v \in V$. This is the analogue of quasi-bandedness and covers a large range of operators in mathematical physics. We also assume we can control the resolvent of A by a known function g as before (for instance, if the A is self-adjoint or normal). It is clear that, with respect to the chosen basis, the matrix of A is quasi-banded. If we know $S(v)$ for all $v \in V$ then it is also clear that the above basis construction also allows us to find an $f : \mathbb{N} \to \mathbb{N}$ such that $A_{j,m} = 0$ if $|j| > f(m)$ and $A_{m,k} = 0$ if $|k| > f(m)$, where $\{A_{ij}\}$ are the matrix elements of A with respect to the basis $\{e_1, e_2, ...\}$. Hence, we assume the knowledge of this f and that Λ contains the functions providing the matrix elements $\{A_{i,j}\}$. Note that finding such an f is straightforward for Hamiltonians as the next example demonstrates.

Example 2. Suppose that the bounded operator A can be written (using the Dirac notation $|v\rangle$ to represent the vertex v) as

$$A = \sum_{v \sim_k w} \alpha(v, w) \, |v\rangle \, \langle w| \tag{11}$$

for some $k \in \mathbb{N}$ and $\alpha : V \times V \to \mathbb{C}$ (i.e. A only involves k-th nearest neighbour interactions). Suppose also that the vertex degree of G is bounded by M. We must have $e_n \in S_n$ and $\{w \in V : v \sim_k w\} \subset S_{n+k}$. Inductively we have $|S_m| \leq (M+1)^m$ and hence we may take the upper bound

$$f(n) = (M+1)^{n+k}.$$

In most cases a smaller f can be used and explicitly computed such as in the construction of the Penrose tile. Note also that the form of A in (11) includes all finite range interaction Hamiltonians on the discrete lattice G.

Example 3. Consider a nearest neighbour operator ($k = 1$ in Eq. 11) on $l^2(\mathbb{Z}^d)$. We have $|S_m| \sim \mathcal{O}(m^d)$ whilst $|S_{m+1} - S_m| \sim \mathcal{O}(m^{d-1})$ (think of the S as radial spheres). It is easy to see that we can choose a suitable f such that

$$f(n) - n \sim \mathcal{O}(n^{\frac{d-1}{d}}),$$

i.e. f grows at most linearly.

Now consider the case that the vertex degree is bounded, we can then compute the spectrum of any Schrödinger operator on any such graph. Furthermore, statements (i) and (ii) hold for this class also. Recall that the (negative) Laplacian or free Hamiltonian H_0 acts on $\psi \in l^2(V(G)) \cong l^2(\mathbb{N})$ by

$$(H_0\psi)(x) = \sum_{y \sim x} (\psi(y) - \psi(x)) \, . \tag{12}$$

The vertex degree being bounded is equivalent to H_0 being a bounded operator. We define a Schrödinger operator on G to be an operator of the form

$$H = H_0 + V,$$

where V is a bounded multiplication operator

$$(V\psi)(x) = V(x)\psi(x).$$

Denote this smaller class by Ω_5 then we have the following.

Theorem 8. *Statements (i) and (ii) (page 1) hold for the classes Ω_4 and Ω_5 when considering $\mathrm{Sp}(\cdot)$ or $\mathrm{Sp}_\epsilon(\cdot)$.*

Proof. The existence of algorithms (part (ii)) follow immediately from the above remarks and Theorems 2, 3 and 7. For statement (i), we can argue as before for the class Ω_4. We will prove statement (i) for Ω_5 when considering the spectrum $\mathrm{Sp}(\cdot)$. The proof for the pseudospectrum follows from similar reasoning, and we omit the details.

Let $\{e_j\}$ be a basis generated as described above. We argue by contradiction and assume that given $\delta > 0$ there does exist an algorithm Γ such that given $A \in \Omega_5$ we have

$$d_{\mathrm{H}}(\Gamma(A), \mathrm{Sp}(A)) \leq \delta.$$

Now let $A = H_0$ as defined in Eq. 12 and let $C = \|H_0\|$. Then there exists a $N(\delta, A) \in \mathbb{N}$ such that $\Gamma(A)$ only depends on $\langle Ae_j, e_i \rangle$ with $i, j \leq N(\delta, A)$. We may then choose a potential V_δ by

$$V_\delta(e_m) = \begin{cases} 0, & m \leq N(\delta, H_0) \\ 3C, & m > N(\delta, H_0) \end{cases}$$

and set $T_\delta = H_0 + V_\delta$. By the definition of a General Algorithm, in particular by (iii) in Definition 1, we must have that $\Gamma(T_\delta) = \Gamma(A)$. Now V_δ and T_δ are both normal and hence

$$\mathrm{d_H}(\mathrm{Sp}(V_\delta), \mathrm{Sp}(T_\delta)) \leq \|V_\delta - T_\delta\| = C.$$

It follows that

$$\mathrm{d_H}(\mathrm{Sp}(V_\delta), \mathrm{Sp}(A)) \leq \mathrm{d_H}(\mathrm{Sp}(V_\delta), \mathrm{Sp}(T_\delta)) + \mathrm{d_H}(\mathrm{Sp}(T_\delta), \Gamma(T_\delta)) + \mathrm{d_H}(\Gamma(T_\delta), \mathrm{Sp}(A)) \leq C + 2\delta.$$

But by construction, $\mathrm{Sp}(V_\delta) = \{0, 3C\}$ and $\mathrm{Sp}(A) \subset B_0(C)$ so we gain a contradiction for small enough δ. □

[1] A. M. Turing, Proc. London Math. Soc. **S2-42**, 230 (1936).
[2] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and Real Computation* (Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1998).
[3] A. C. Hansen, J. Amer. Math. Soc. **24**, 81 (2011).
[4] L. N. Trefethen and D. Bau III, *Numerical linear algebra*, Vol. 50 (Siam, 1997).
[5] N. Hatano and D. R. Nelson, Phys. Rev. Lett. **77**, 570 (1996).
[6] S. Longhi, D. Gatti, and G. Della Valle, Sci. Rep. **5** (2015).
[7] D. R. Nelson and N. M. Shnerb, Phys. Rev. E **58**, 1383 (1998).
[8] I. Y. Goldsheid and B. A. Khoruzhenko, Phys. Rev. Lett. **80**, 2897 (1998).