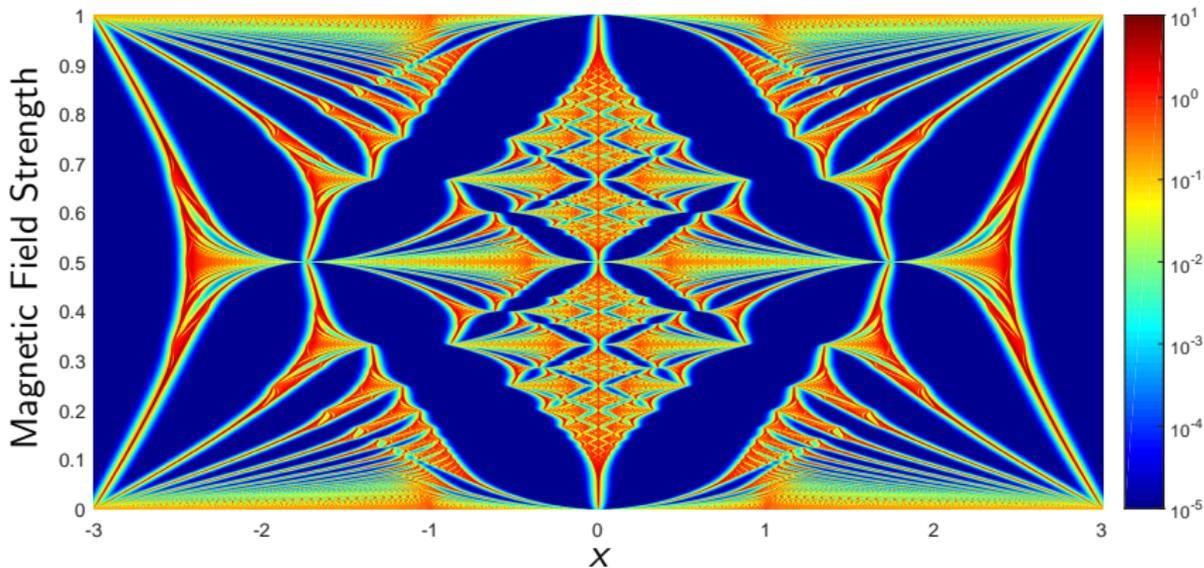


The Foundations of Infinite-Dimensional Spectral Computations

Matthew Colbrook

University of Cambridge



Spectral measure of magnetic graphene, computed to high precision (see log scale) using a method of this talk.

Collaborators

Jonathan Ben-Artzi
(Cardiff)



Anders Hansen
(Cambridge)



Andrew Horning
(Cornell)



Olavi Nevanlinna
(Aalto)



Bogdan Roman
(Cambridge)



Markus Seidel
(West Saxon)



Alex Townsend
(Cornell)



The infinite-dimensional problem

In discrete setting, operator acting on $\ell^2(\mathbb{N})$:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots \\ a_{21} & a_{22} & a_{23} & \dots \\ a_{31} & a_{32} & a_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad (Ax)_j = \sum_{k \in \mathbb{N}} a_{jk} x_k.$$

In cts setting, deal with differential operators, integral operators etc.

Finite Case	Infinite Case
Eigenvalues \Rightarrow	Spectrum $\text{Sp}(A) = \{z \in \mathbb{C} : A - z \text{ not bounded invertible}\}$
Eigenvectors \Rightarrow	Spectral Measure (normal case)

Goal: compute spectral properties of the operator from matrix elements, PDE coefficients, or other suitable information.

Many applications: quantum mechanics, chemistry, matter physics, statistical mechanics, optics, number theory, PDEs, mathematics of information etc.

Mathematicians and physicists contributing to computational spectral theory form a vast set including:

D. Arnold (Minnesota), W. Arveson (Berkeley), A. Böttcher (Chemnitz), W. Dahmen (South Carolina), E. B. Davies (King's College London), P. Deift (NYU), L. Demanet (MIT), C. Fefferman (Princeton), G. Golub (Stanford), A. Iserles (Cambridge), W. Schlag (Yale), E. Schrödinger (DIAS), J. Schwinger (Harvard), N. Trefethen (Oxford), V. Varadarajan (UCLA), S. Varadhan (NYU), J. von Neumann (IAS), M. Zworski (Berkeley),...

Can we do this for general classes of operators?

Can we do this for general classes of operators?

“Most operators that arise in practice are not presented in a representation in which they are diagonalized, and it is often very hard to locate even a single point in the spectrum... Thus, one often has to settle for numerical approximations [to the spectrum], and this raises the question of how to implement the methods of finite dimensional numerical linear algebra to compute the spectra of infinite dimensional operators. Unfortunately, there is a dearth of literature on this basic problem and, so far as we have been able to tell, there are no proven techniques.”

— W. Arveson, UC Berkeley (1994)

Can we do this for general classes of operators?

“Most operators that arise in practice are not presented in a representation in which they are diagonalized, and it is often very hard to locate even a single point in the spectrum... Thus, one often has to settle for numerical approximations [to the spectrum], and this raises the question of how to implement the methods of finite dimensional numerical linear algebra to compute the spectra of infinite dimensional operators. Unfortunately, there is a dearth of literature on this basic problem and, so far as we have been able to tell, there are no proven techniques.”

— W. Arveson, UC Berkeley (1994)

Partial answer: can compute spectra of general bounded operators on $\ell^2(\mathbb{N})$ (in Hausdorff metric) using three successive limits¹

$$\lim_{n_3 \rightarrow \infty} \lim_{n_2 \rightarrow \infty} \lim_{n_1 \rightarrow \infty} \Gamma_{n_3, n_2, n_1}(A) = \text{Sp}(A).$$

Turns out this is sharp! Hence impossible from numerical point of view.

¹Hansen. JAMS (2011)

Can we do this for general classes of operators?

“Most operators that arise in practice are not presented in a representation in which they are diagonalized, and it is often very hard to locate even a single point in the spectrum... Thus, one often has to settle for numerical approximations [to the spectrum], and this raises the question of how to implement the methods of finite dimensional numerical linear algebra to compute the spectra of infinite dimensional operators. Unfortunately, there is a dearth of literature on this basic problem and, so far as we have been able to tell, there are no proven techniques.”

— W. Arveson, UC Berkeley (1994)

Partial answer: can compute spectra of general bounded operators on $\ell^2(\mathbb{N})$ (in Hausdorff metric) using three successive limits¹

$$\lim_{n_3 \rightarrow \infty} \lim_{n_2 \rightarrow \infty} \lim_{n_1 \rightarrow \infty} \Gamma_{n_3, n_2, n_1}(A) = \text{Sp}(A).$$

Turns out this is sharp! Hence impossible from numerical point of view.

Q: What assumptions do we need to make it easier? What about other spectral problems?

¹Hansen. JAMS (2011)

Example: Bounded Diagonal Operators (Very Easy)

$$A = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & a_3 & \\ & & & \ddots \end{pmatrix}$$

If $\Gamma_n(A) = \{a_1, \dots, a_n\}$ then $\Gamma_n(A) \rightarrow \text{Sp}(A)$ in Hausdorff metric.

Also have $\Gamma_n(A) \subset \text{Sp}(A)$.

This is optimal from a foundations point of view.

Example: Compact Operators (Still Easy)

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots \\ a_{21} & a_{22} & a_{23} & \dots \\ a_{31} & a_{32} & a_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \text{ compact}$$

If $\Gamma_n(A) = \text{Sp}(P_n A P_n)$, then $\Gamma_n(A) \rightarrow \text{Sp}(A)$ in Hausdorff metric.

Known for decades.

Example: Compact Operators (Still Easy)

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots \\ a_{21} & a_{22} & a_{23} & \dots \\ a_{31} & a_{32} & a_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \text{ compact}$$

If $\Gamma_n(A) = \text{Sp}(P_n A P_n)$, then $\Gamma_n(A) \rightarrow \text{Sp}(A)$ in Hausdorff metric.

Known for decades.

Q: Can we gain error control as before?

Example: Compact Operators (Still Easy)

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots \\ a_{21} & a_{22} & a_{23} & \dots \\ a_{31} & a_{32} & a_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \text{ compact}$$

If $\Gamma_n(A) = \text{Sp}(P_n A P_n)$, then $\Gamma_n(A) \rightarrow \text{Sp}(A)$ in Hausdorff metric.

Known for decades.

Q: Can we gain error control as before?

No! No algorithm can gain error control on the whole class, even for self-adjoint compact operators.

What about Jacobi operators?

$$A = \begin{pmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & b_2 & a_3 & \ddots & \\ & & & \ddots & \ddots \end{pmatrix}$$

This problem has been open for decades.

What about Jacobi operators?

$$A = \begin{pmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & b_2 & a_3 & \ddots & \\ & & \ddots & \ddots & \ddots \end{pmatrix}$$

This problem has been open for decades.

What about sparse normal operators? Surely this is much harder?!

What about Jacobi operators?

$$A = \begin{pmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & b_2 & a_3 & \ddots & \\ & & & \ddots & \ddots \end{pmatrix}$$

This problem has been open for decades.

What about sparse normal operators? Surely this is much harder?!

New result: Large class Ω (covering arguably most applications and including sparse normal) such that we can compute $\Gamma_n(A) \rightarrow \text{Sp}(A)$ and $E_n(A) \downarrow 0$ for $A \in \Omega$ with

$$\text{dist}(z, \text{Sp}(A)) \leq E_n(A), \quad \forall z \in \Gamma_n(A).$$

What about Jacobi operators?

$$A = \begin{pmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & b_2 & a_3 & \ddots & \\ & & & \ddots & \ddots \end{pmatrix}$$

This problem has been open for decades.

What about sparse normal operators? Surely this is much harder?!

New result: Large class Ω (covering arguably most applications and including sparse normal) such that we can compute $\Gamma_n(A) \rightarrow \text{Sp}(A)$ and $E_n(A) \downarrow 0$ for $A \in \Omega$ with

$$\text{dist}(z, \text{Sp}(A)) \leq E_n(A), \quad \forall z \in \Gamma_n(A).$$

Paradox: Easier problem than compact operators!

What about Schrödinger operators?

Consider self-adjoint Schrödinger operators $-\nabla^2 + V$ on $L^2(\mathbb{R}^d)$.

Can we compute spectra from sampling V ?

What about Schrödinger operators?

Consider self-adjoint Schrödinger operators $-\nabla^2 + V$ on $L^2(\mathbb{R}^d)$.

Can we compute spectra from sampling V ?

NB: Open problem since Schwinger's work in the 1960s to do this for general Schrödinger operators (even without error control).

What about Schrödinger operators?

Consider self-adjoint Schrödinger operators $-\nabla^2 + V$ on $L^2(\mathbb{R}^d)$.

Can we compute spectra from sampling V ?

NB: Open problem since Schwinger's work in the 1960s to do this for general Schrödinger operators (even without error control).

New result: If V has locally bounded total variation and grows at most polynomially at infinity then yes! And with error control as before.

What about Schrödinger operators?

Consider self-adjoint Schrödinger operators $-\nabla^2 + V$ on $L^2(\mathbb{R}^d)$.

Can we compute spectra from sampling V ?

NB: Open problem since Schwinger's work in the 1960s to do this for general Schrödinger operators (even without error control).

New result: If V has locally bounded total variation and grows at most polynomially at infinity then yes! And with error control as before.

Paradox: Easier problem than compact operators!

What about Schrödinger operators?

Consider self-adjoint Schrödinger operators $-\nabla^2 + V$ on $L^2(\mathbb{R}^d)$.

Can we compute spectra from sampling V ?

NB: Open problem since Schwinger's work in the 1960s to do this for general Schrödinger operators (even without error control).

New result: If V has locally bounded total variation and grows at most polynomially at infinity then yes! And with error control as before.

Paradox: Easier problem than compact operators!

Can also be extended to many non-self-adjoint Schrödinger operators and singular potentials etc.

Foundations of Infinite-Dimensional Spectral Computations

How: Deal with operators directly, instead of previous 'truncate-then-solve'

⇒ Compute many spectral properties for the first time.

Framework: Classify problems in a computational hierarchy (the SCI hierarchy) measuring their difficulty and the optimality of algorithms.²

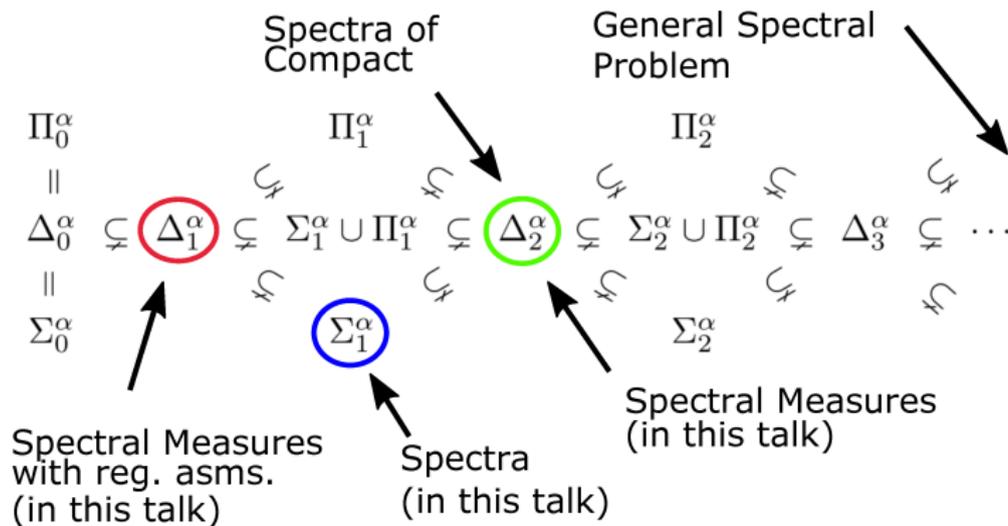
⇒ Algorithms that realise the boundaries of what computers can achieve.

NB: This holds regardless of model of computation - measure the intrinsic difficulty of these problems.

Common theme in examples of this talk: use the resolvent $(A - z)^{-1}$

²Also has deep connections with logic and descriptive set theory: C. Preprint (2019)

Structure of the Hierarchy



Example 1: Computing spectra with error control

Computing the resolvent norm $\gamma(z) := \|(A - z)^{-1}\|^{-1}$



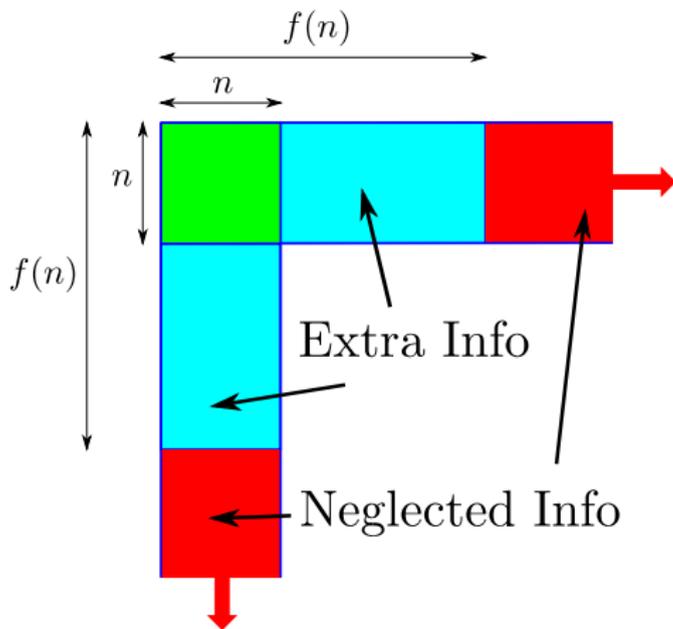
First algorithm that computes spectra of a very general class of operators.
Also does so with (rigorous provable) error control.

For simplicity, assume A acts on $l^2(\mathbb{N})$.

Definition (Dispersion: off-diagonal decay)

Dispersion of A bounded by function $f : \mathbb{N} \rightarrow \mathbb{N}$ and null sequence $\{c_n\}$ if

$$\max\{\|(I - P_{f(n)})AP_n\|, \|P_nA(I - P_{f(n)})\|\} \leq c_n.$$



Definition (Controlled growth of the resolvent: well-conditioned)

Continuous increasing function $g : [0, \infty) \rightarrow [0, \infty)$ with $g(x) \leq x$.
Controlled growth of the resolvent by g if

$$g(\text{dist}(z, \text{Sp}(A))) \leq \|(A - z)^{-1}\|^{-1} \quad \forall z \in \mathbb{C}.$$

- Measures conditioning of the problem through

$$\{z \in \mathbb{C} : \|(A - z)^{-1}\|^{-1} \leq \epsilon\} =: \text{Sp}_\epsilon(A) = \bigcup_{\|B\| \leq \epsilon} \text{Sp}(A + B).$$

- Normal operators (A commutes with A^*) well-conditioned with

$$\|(A - z)^{-1}\|^{-1} = \text{dist}(z, \text{Sp}(A)), \quad g(x) = x.$$

Definition (Controlled growth of the resolvent: well-conditioned)

Continuous increasing function $g : [0, \infty) \rightarrow [0, \infty)$ with $g(x) \leq x$.
Controlled growth of the resolvent by g if

$$g(\text{dist}(z, \text{Sp}(A))) \leq \|(A - z)^{-1}\|^{-1} \quad \forall z \in \mathbb{C}.$$

- Measures conditioning of the problem through

$$\{z \in \mathbb{C} : \|(A - z)^{-1}\|^{-1} \leq \epsilon\} =: \text{Sp}_\epsilon(A) = \bigcup_{\|B\| \leq \epsilon} \text{Sp}(A + B).$$

- Normal operators (A commutes with A^*) well-conditioned with

$$\|(A - z)^{-1}\|^{-1} = \text{dist}(z, \text{Sp}(A)), \quad g(x) = x.$$

Theorem (C., Roman, Hansen. PRL (2019))

Know $f, g \Rightarrow$ can compute $\text{Sp}(A)$ with error control.

Step 1: Approximate locally via smallest singular value:

$$\gamma_n(z) := \min\{\sigma_1(P_{f(n)}(A - z)P_n), \sigma_1(P_{f(n)}(A^* - \bar{z})P_n)\} + c_n.$$

This converges locally uniformly down to $\gamma(z) = \|(A - z)^{-1}\|^{-1}$.

Step 2: Bound the distance to the spectrum:

$$\gamma(z) \leq \text{dist}(z, \text{Sp}(A)) \leq g^{-1}(\gamma(z)) \leq g^{-1}(\gamma_n(z)).$$

Step 3: Find (almost) local minimisers and output $\Gamma_n(A)$ with

$$\Gamma_n(A) \rightarrow \text{Sp}(A), \quad \text{dist}(z, \text{Sp}(A)) \leq g^{-1}(\gamma_n(z)), \quad \sup_{z \in \Gamma_n(A)} g^{-1}(\gamma_n(z)) \rightarrow 0$$

NB: Without either f or g , constructing an algorithm convergent to the spectrum is impossible (even dropping error control).

Example: quartic potential on $L^2(\mathbb{R})$ using a Hermite basis

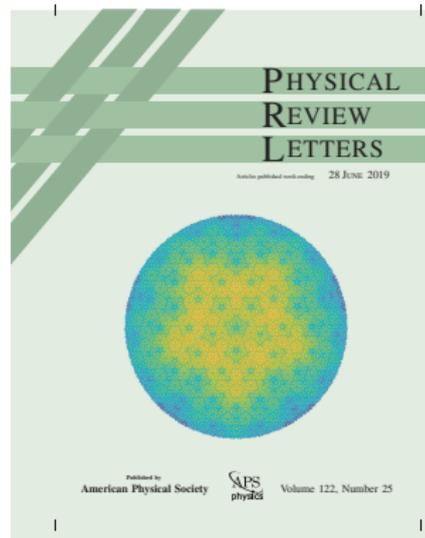
New exemplar of spectral computation

Method is:

- Local and parallelisable.
- Convergent for first time.
(e.g. no spectral pollution)
- Explicitly bounds the error:

$$\text{Error} \leq a_n \downarrow 0.$$

- Optimal from foundations point of view.



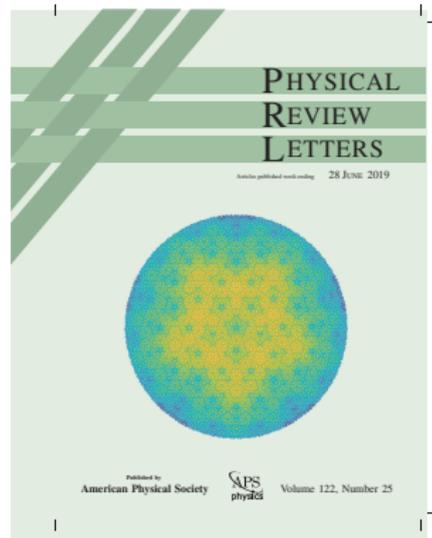
New exemplar of spectral computation

Method is:

- Local and parallelisable.
- Convergent for first time.
(e.g. no spectral pollution)
- Explicitly bounds the error:

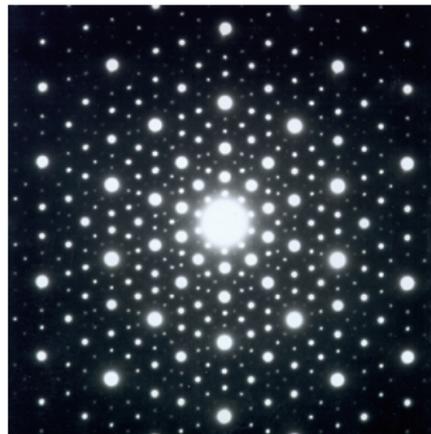
$$\text{Error} \leq a_n \downarrow 0.$$

- Optimal from foundations point of view.



Extends to unbounded operators and PDEs with coefficients of locally bounded total variation (e.g. algorithms point sample coefficients).

Example: Operators in condensed matter physics



Left: Dan Shechtman, **Nobel Prize in Chemistry 2011** for discovery of quasicrystal. Right: Diffraction pattern of a quasicrystal.

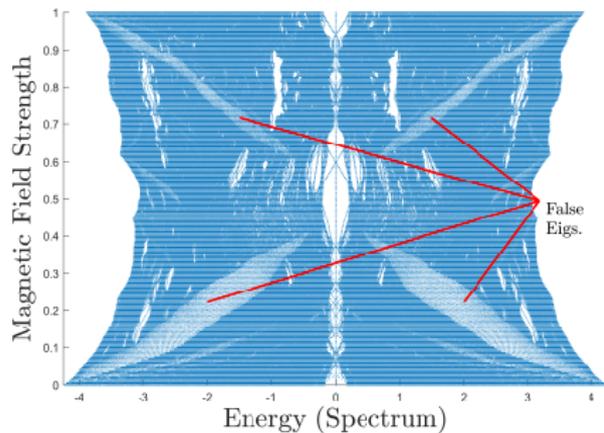
Magnetic properties of quasicrystal.

Hard problem - no previous method even converges to spectrum.

Example: Operators in condensed matter physics

Finite truncations

Edge states.



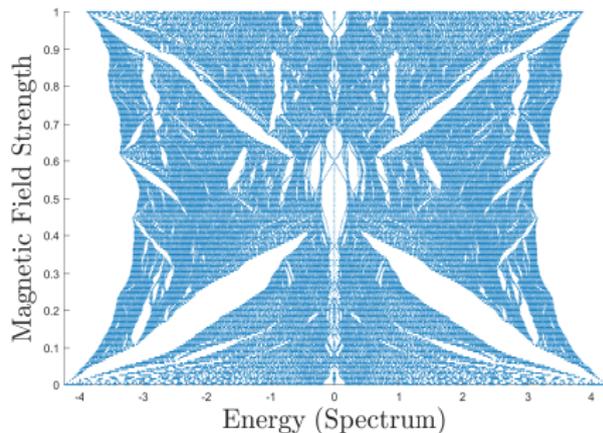
Unreliable

Does not converge

No error control

Infinite-dimensional techniques

First convergent computation.

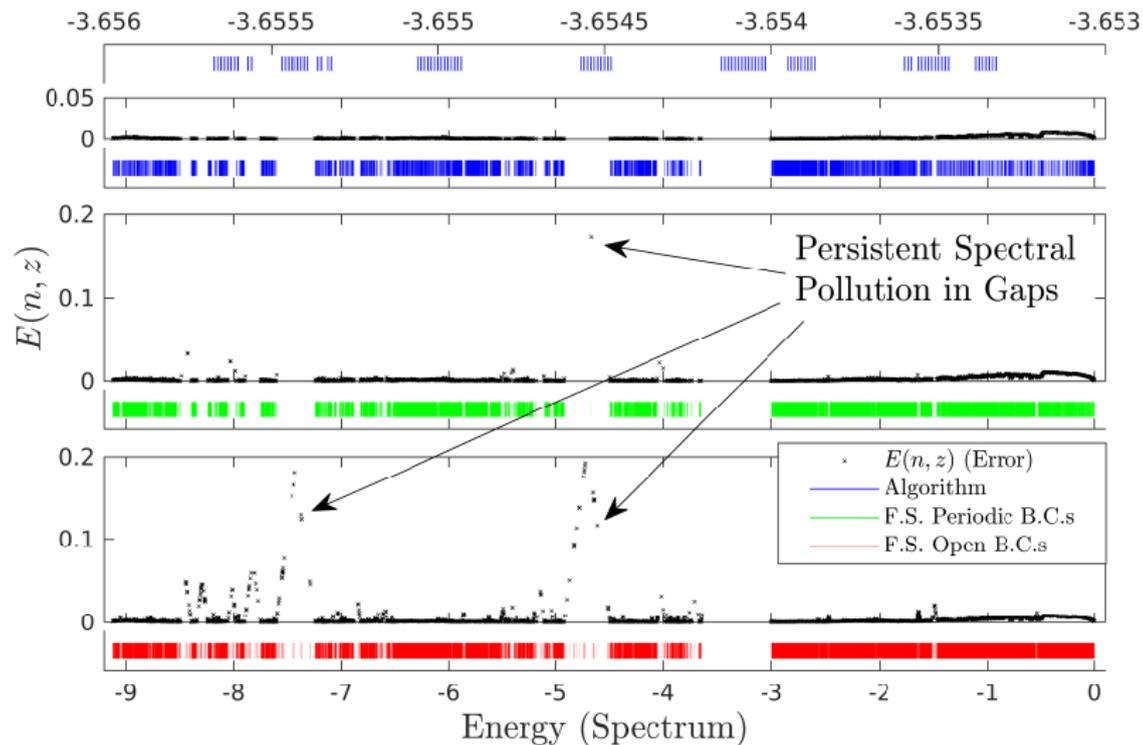


Reliable

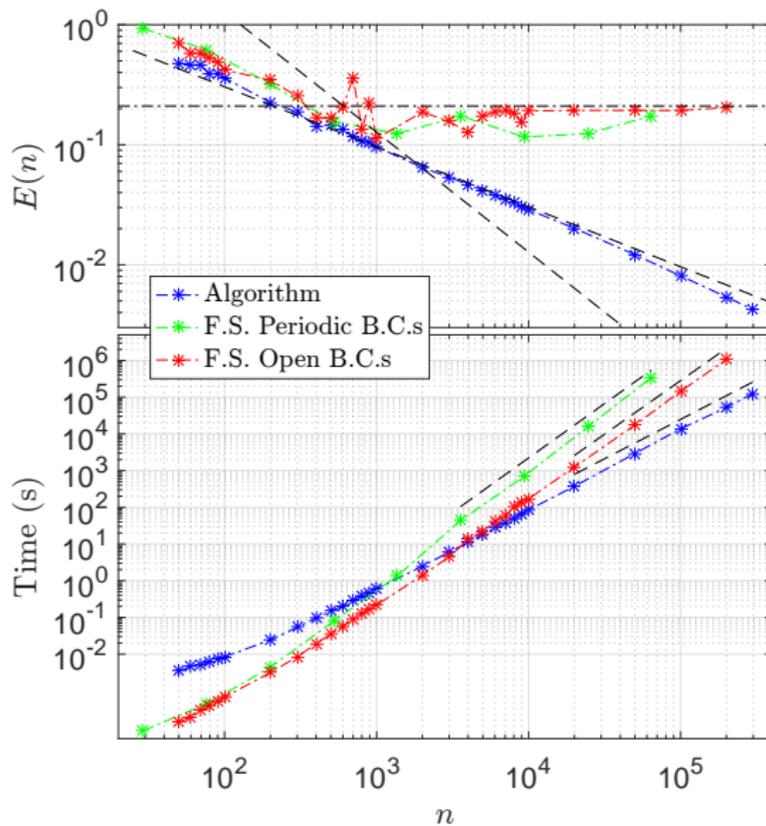
Converges

Error control

Example: Laplacian on Penrose tile



Example: Laplacian on Penrose tile



Example 2: Computing spectral measures

Compute the resolvent $(A - z)^{-1}\mathbf{x}$ ($\text{Im}(z) > 0$) with asymptotic error control as z approaches \mathbb{R} .



First algorithm that computes spectral measures of general self-adjoint operators - 'diagonalisation.'

Assume A is self-adjoint.

Spectral Measures

If A normal, associated projection-valued measure \mathcal{E} s.t.

$$A\mathbf{x} = \int_{\text{Sp}(A)} \lambda d\mathcal{E}(\lambda)\mathbf{x}, \quad \forall \mathbf{x} \in \mathcal{D}(A),$$

Allows computation of functional calculus, has interesting physics etc.

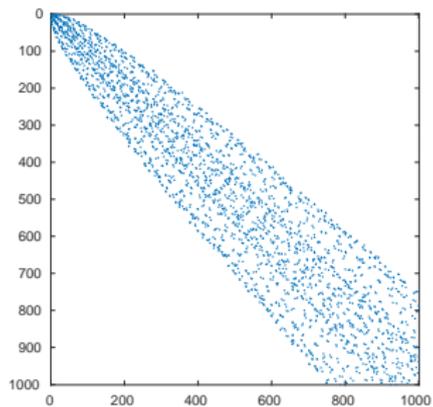
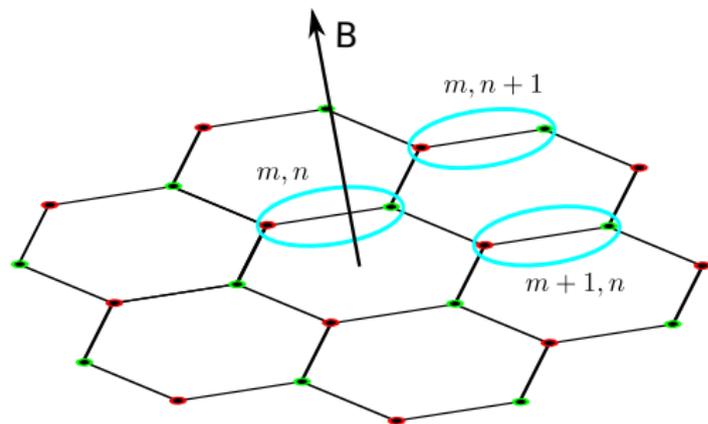
Idea: For $z = x + i\epsilon$, use

$$\frac{(A - z)^{-1} - (A - \bar{z})^{-1}}{2\pi i} = \frac{1}{\pi} \int_{\text{Sp}(A)} \frac{\epsilon}{(x - \lambda)^2 + \epsilon^2} d\mathcal{E}(\lambda).$$

Convolution with Poisson kernel: smoothed measure.

Converges weakly to measure as $\epsilon \downarrow 0$ (cf. Stone's formula).

Example: Magnetic Graphene



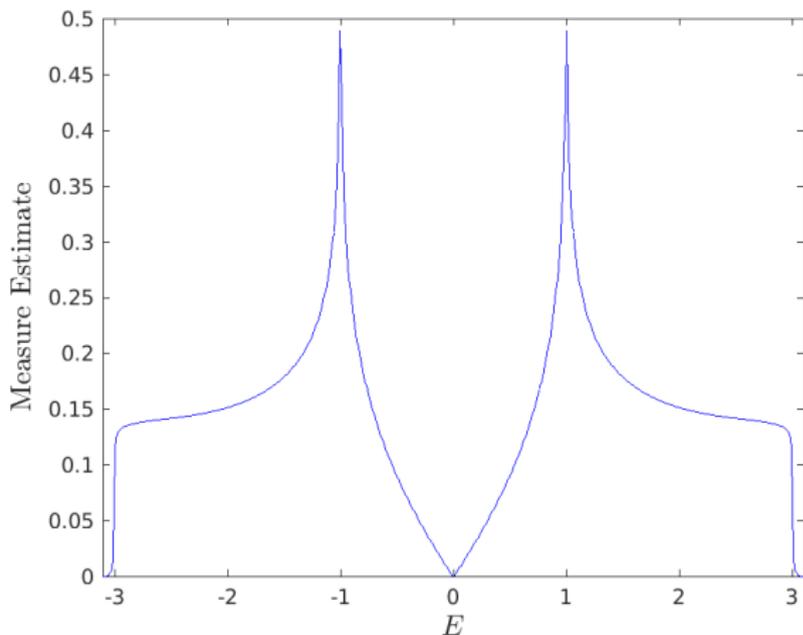
Example: Magnetic Graphene

Example: Magnetic Graphene

Theorem (C. Preprint (2019))

Know $f \Rightarrow$ can compute measure in one limit.

This is through a rectangular least squares type problem that computes $(A - z)^{-1}\mathbf{x}$ with (asymptotic) error control. $N(\epsilon)$ chosen **adaptively**.

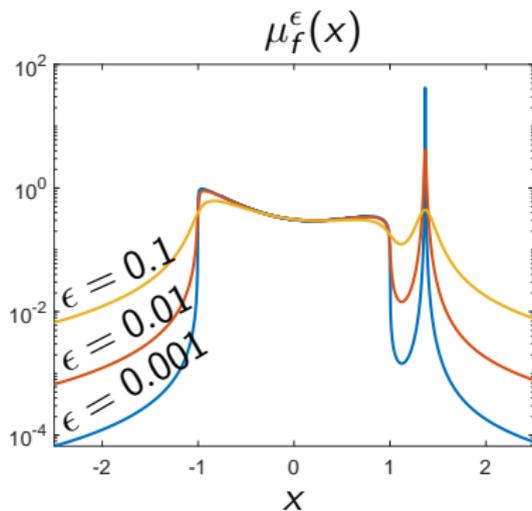


Example: Integral Operator

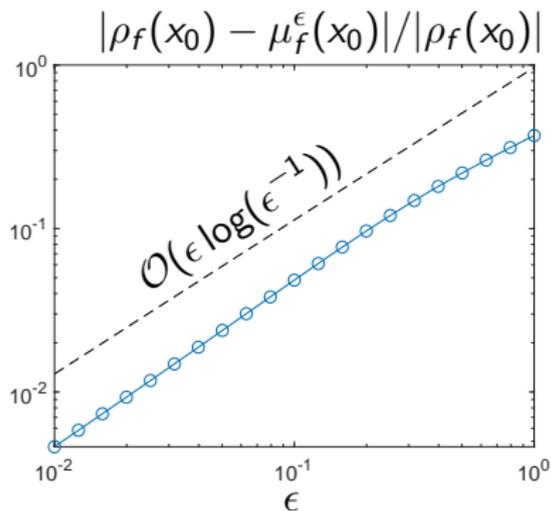
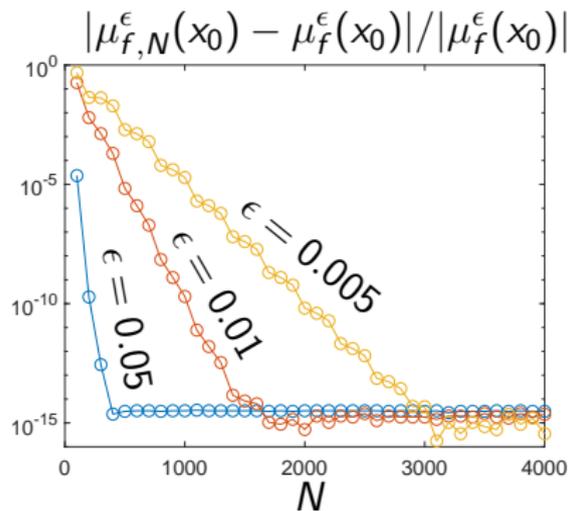
$$Au(x) = xu(x) + \int_{-1}^1 e^{-(x^2+y^2)} u(y) dy, \quad x \in [-1, 1].$$

Discretise using adaptive Chebyshev collocation method.

Look at $\mu_f : B \rightarrow \langle \mathcal{E}(B)f, f \rangle$. E.g. $f(x) = \sqrt{3/2} x$.



Example: Integral Operator



Converges like $\mathcal{O}(\epsilon \log(\epsilon^{-1}))$ and need $N \approx 20/\epsilon$.

\Rightarrow Infeasible to get more than five or six digits!

Q: Can we do better?

Accelerating convergence

Let $m \in \mathbb{N}$, $K \in L^1(\mathbb{R})$. We say K is an m th order kernel if:

- (i) Normalized: $\int_{\mathbb{R}} K(x) dx = 1$,
- (ii) Zero moments: $K(x)x^j$ is integrable and $\int_{\mathbb{R}} K(x)x^j dx = 0$ for $0 < j < m$, and
- (iii) Decay at $\pm\infty$: There is a constant C_K , independent of x , such that

$$|K(x)| \leq C_K(1 + |x|)^{-(m+1)}, \quad x \in \mathbb{R}.$$

Theorem (C., Horning, Townsend. Preprint (2020))

If K is m th order, $K_\epsilon(x) = \epsilon^{-1}K(x\epsilon^{-1})$ and μ_f locally absolutely continuous near x_0 with density ρ_f then

- Pointwise: If ρ_f locally $C^{n,\alpha}$ near x_0 then

$$|[K_\epsilon * \mu_f](x_0) - \rho_f(x_0)| = \mathcal{O}(\epsilon^{n+\alpha}) + \mathcal{O}(\epsilon^m \log(\epsilon^{-1}))$$

- L^p : If ρ_f locally $\mathcal{W}^{n,p}$ near x_0 ($1 \leq p < \infty$) then

$$\|[K_\epsilon * \mu_f] - \rho_f\|_{L^p_{loc}} = \mathcal{O}(\epsilon^n) + \mathcal{O}(\epsilon^m \log(\epsilon^{-1}))$$

Accelerating convergence with rational kernels

Idea: Replace Poisson kernel with rational kernel

$$K(x) = \frac{1}{2\pi i} \sum_{j=1}^{n_1} \frac{\alpha_j}{x - a_j} - \frac{1}{2\pi i} \sum_{j=1}^{n_2} \frac{\beta_j}{x - b_j}.$$

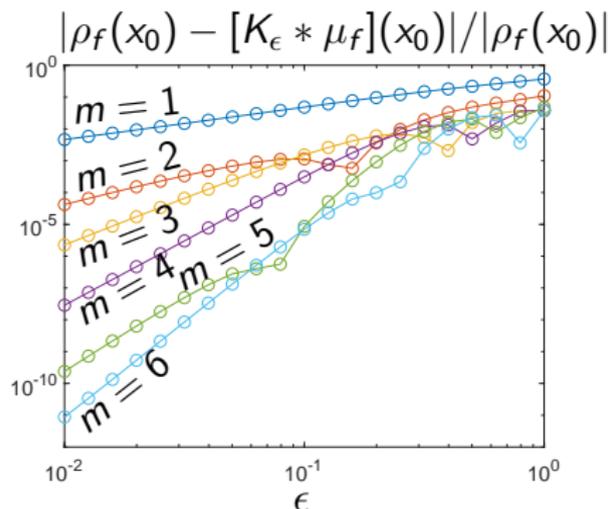
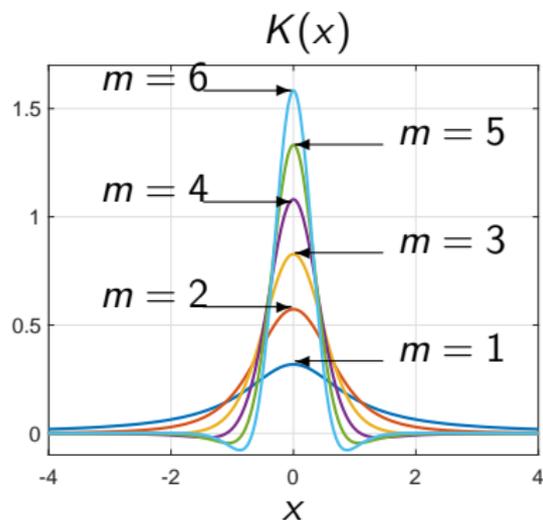
Can compute convolution with error control using resolvent

$$\begin{aligned} & [K_\epsilon * \mu_f](x) \\ &= \frac{1}{2\pi i} \left[\sum_{j=1}^{n_1} \alpha_j \langle (A - (x + \epsilon a_j))^{-1} f, f \rangle - \sum_{j=1}^{n_2} \beta_j \langle (A - (x + \epsilon b_j))^{-1} f, f \rangle \right]. \end{aligned}$$

Proposition (C., Horning, Townsend. Preprint (2020))

Let $n_1 = n_2 = m \in \mathbb{N}$ and fix a_j in UHP, b_j in LHP. Then there exists unique $\{\alpha_j, \beta_j\}$ such that K is an m th order kernel.

Integral operator revisited



See paper for general differential (even PDEs), integral and lattice operator examples - use sparse spectral methods for discretisation.

Eigenvalue hunting without spectral pollution

Can extend to compute decompositions (of measures and spectral sets).

Example: Dirac operator.

- Describes the motion of a relativistic spin-1/2 particle.
- Essential spectrum given by $\mathbb{R} \setminus (-1, 1) \Rightarrow$ spectral pollution!
- Consider radially symmetric potential, coupled system on half-line:

$$\mathcal{D}_V = \begin{pmatrix} 1 + V(r) & -\frac{d}{dr} + \frac{\kappa}{r} \\ \frac{d}{dr} + \frac{\kappa}{r} & -1 + V(r) \end{pmatrix}.$$

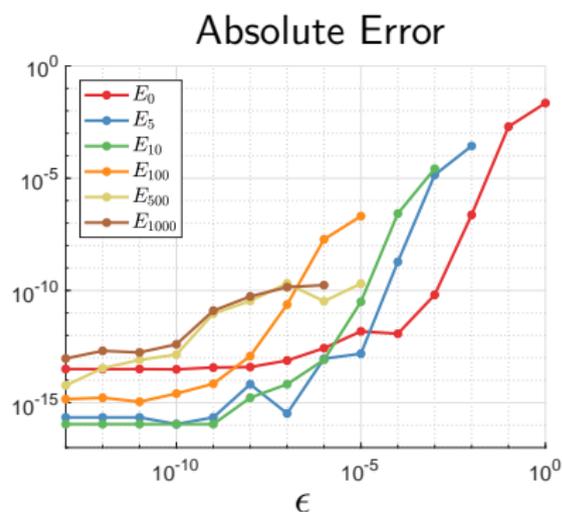
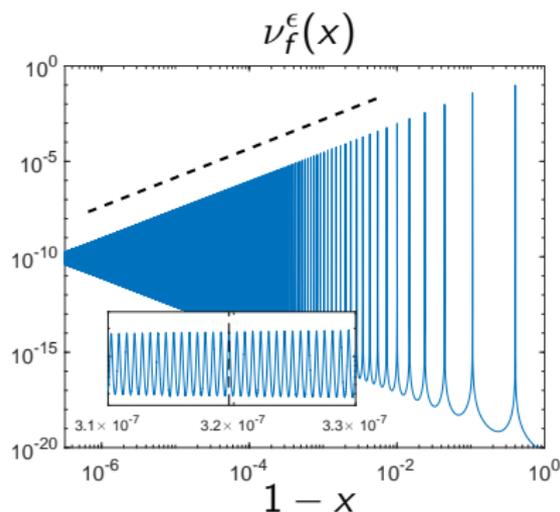
- Map to $[-1, 1]$ using

$$x = \frac{r - L}{r + L}$$

and compute $(A - z)^{-1}\mathbf{x}$ using the ultraspherical method.

Eigenvalue hunting without spectral pollution

$$\nu_f^\epsilon(x) := \epsilon \cdot \text{Im} \left(\langle (\mathcal{D}_V - (x + i\epsilon))^{-1} f, f \rangle \right) = \sum_{\lambda \in \Lambda^P(\mathcal{D}_V) \cap \{x\}} \langle \mathcal{P}_\lambda f, f \rangle + \mathcal{O}(\epsilon)$$



NB: Previous state-of-the-art achieves a few digits for a few excited states.

Results in this talk:

- Can compute spectra of a large class of operators with error control. New algorithm is fast, local and parallelisable.
- Can compute spectral measures and spectral decompositions through resolvent. Can be combined with state-of-the-art PDE methods.

Also have foundations for: spectral type (pure point, absolutely continuous, singularly continuous), Lebesgue measure and fractal dimensions of spectra, discrete spectra, essential spectra, eigenvectors + multiplicity, spectral radii, essential numerical ranges, geometric features of spectrum (e.g. capacity), spectral gap problem, ...

Ongoing and future work: foundations of computational PDEs, foundations of (stable) neural networks, and computer-assisted proofs.

New: high-performance numerical package with resolvent based algorithms for computing spectral measures (with Andrew Horning).

<https://github.com/ajhPHROS/SpecSolve>

References for algorithms in this talk

- M.J. Colbrook, B. Roman, and A.C. Hansen. “How to compute spectra with error control.” *Physical Review Letters* 122.25 (2019): 250201.
- J. Ben-Artzi, M.J. Colbrook, A.C. Hansen, O. Nevanlinna, and M. Seidel. “Computing Spectra - on the solvability complexity index hierarchy and towers of algorithms.” Submitted.
- M.J. Colbrook, A.C. Hansen. “On the infinite-dimensional QR algorithm.” *Numerische Mathematik* 143.1 (2019): 17-83.
- M.J. Colbrook. “Computing spectral measures and spectral types.” Submitted.
- M.J. Colbrook, A. Horning, and A. Townsend. “Computing spectral measures of self-adjoint operators.” Submitted.

For further papers in this program, classifications in the SCI hierarchy and numerical code: <http://www.damtp.cam.ac.uk/user/mjc249/home.html>