A New Class of Algorithms for Computing Spectra with **Error Control**

Matthew Colbrook

3rd May 2017

Background

- Hilbert space $l^2(\mathbb{N})$ with $\|x\|_2 = \sqrt{\sum_{j=1}^{\infty} |x_j|^2}$, $\langle x, y \rangle = \sum_{j=1}^{\infty} x_j \bar{y}_j$
- Bounded linear operator $A : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ realised as matrix

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

  Denote these by $\mathcal{B}(l^2(\mathbb{N}))$.
- Want to compute spectrum (generalistion of eigenvalues)

$$\mathrm{Sp}(A) := \{z \in \mathbb{C} : A - zI \text{ not invertible}\}.$$

  from the matrix elements.

Motivation

- Quantum mechanics, quasicrystals

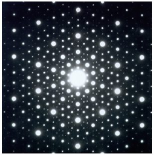


Figure: Left: Dan Shechtman, Nobel Prize in Chemistry 2011. Right: Electron diffraction pattern of quasicrystal.

- Intensely investigated since the 1950s, still very active today.

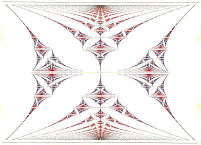


Figure: Left: Artur Avila, Fields Medal 2014. Right: Hofstadter butterfly.

Our computational problem

- Convergence in $\mathcal{M}$, set of all compact subsets of $\mathbb{C}$ provided with the Hausdorff metric $d = d_{\mathrm{H}}$

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

Our computational problem

- Convergence in $\mathcal{M}$, set of all compact subsets of $\mathbb{C}$ provided with the Hausdorff metric $d = d_{\mathrm{H}}$

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

- Allowed to use entries of the matrix representation of A. Want to compute spectra of a class $\Omega \subset \mathcal{B}(l^2(\mathbb{N}))$.

Our computational problem

- Convergence in $\mathcal{M}$, set of all compact subsets of $\mathbb{C}$ provided with the Hausdorff metric $d = d_{\mathrm{H}}$

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

- Allowed to use entries of the matrix representation of A. Want to compute spectra of a class $\Omega \subset \mathcal{B}(l^2(\mathbb{N}))$.

What attributes should an "algorithm" $\Gamma_n(A)$ possess?

Our computational problem

- Convergence in $\mathcal{M}$, set of all compact subsets of $\mathbb{C}$ provided with the Hausdorff metric $d = d_{\mathrm{H}}$

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

- Allowed to use entries of the matrix representation of A. Want to compute spectra of a class $\Omega \subset \mathcal{B}(l^2(\mathbb{N}))$.

What attributes should an “algorithm” $\Gamma_n(A)$ possess?

1. Should converge!

Our computational problem

- Convergence in $\mathcal{M}$, set of all compact subsets of $\mathbb{C}$ provided with the Hausdorff metric $d = d_{\mathrm{H}}$

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

- Allowed to use entries of the matrix representation of A. Want to compute spectra of a class $\Omega \subset \mathcal{B}(l^2(\mathbb{N}))$.

What attributes should an "algorithm" $\Gamma_n(A)$ possess?

1. Should converge!
2. Given index n, should only use finite amounts of information.

Our computational problem

- Convergence in $\mathcal{M}$, set of all compact subsets of $\mathbb{C}$ provided with the Hausdorff metric $d = d_{\mathrm{H}}$

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

- Allowed to use entries of the matrix representation of A. Want to compute spectra of a class $\Omega \subset \mathcal{B}(l^2(\mathbb{N}))$.

What attributes should an "algorithm" $\Gamma_n(A)$ possess?

1. Should converge!
2. Given index n, should only use finite amounts of information.
3. Should read and use this information in a consistent way.

Our computational problem

- Convergence in $\mathcal{M}$, set of all compact subsets of $\mathbb{C}$ provided with the Hausdorff metric $d = d_{\mathrm{H}}$

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

- Allowed to use entries of the matrix representation of A. Want to compute spectra of a class $\Omega \subset \mathcal{B}(l^2(\mathbb{N}))$.

What attributes should an "algorithm" $\Gamma_n(A)$ possess?

1. Should converge!
2. Given index n, should only use finite amounts of information.
3. Should read and use this information in a consistent way.
4. Error control?

Hierarchy of complexity

Definition 1 (Tower of Algorithms)

A *tower of algorithms of height k* is a family of sequences of functions

$$\Gamma_{n_k,\ldots,n_1} : \Omega \to \mathcal{M},$$

where $n_k, \ldots, n_1 \in \mathbb{N}$ and $\Gamma_{n_k,\ldots,n_1}$ are "algorithms". Moreover,

$$\mathrm{Sp}(A) = \lim_{n_k\to\infty} \ldots \lim_{n_1\to\infty} \Gamma_{n_k,\ldots,n_1}(A).$$

Hierarchy of complexity

Definition 1 (Tower of Algorithms)

A *tower of algorithms of height k* is a family of sequences of functions

$$\Gamma_{n_k,\ldots,n_1} : \Omega \to \mathcal{M},$$

where $n_k, \ldots, n_1 \in \mathbb{N}$ and $\Gamma_{n_k,\ldots,n_1}$ are "algorithms". Moreover,

$$\mathrm{Sp}(A) = \lim_{n_k \to \infty} \ldots \lim_{n_1 \to \infty} \Gamma_{n_k,\ldots,n_1}(A).$$

Definition 2 (Solvability Complexity Index (SCI))

Solvability Complexity Index, $\mathrm{SCI}(\mathrm{Sp}, \Omega)$ is the smallest integer k for which there exists a tower of algorithms of height k. If no such tower exists then $\mathrm{SCI}(\mathrm{Sp}, \Omega) = \infty$.

Problem: $\mathrm{SCI}(\mathrm{Sp},\mathcal{B}(l^2(\mathbb{N})))=3$

Problem: $\mathrm{SCI}(\mathrm{Sp}, \mathcal{B}(l^2(\mathbb{N}))) = 3$

Solution: restrict to smaller subclass. Large subclass $\Omega \subset \mathcal{B}(l^2(\mathbb{N}))$ with $\mathrm{SCI}(\mathrm{Sp}, \Omega) = 1$. What's more can gain error control, an algorithm such that

$$z \in \Gamma_n(A) \Rightarrow \mathrm{dist}(z, \mathrm{Sp}(A)) < 2^{-n}.$$

Call this class Σ_1 - "output is reliable".

Problem: $\mathrm{SCI}(\mathrm{Sp}, \mathcal{B}(l^2(\mathbb{N}))) = 3$

Solution: restrict to smaller subclass. Large subclass $\Omega \subset \mathcal{B}(l^2(\mathbb{N}))$ with $\mathrm{SCI}(\mathrm{Sp}, \Omega) = 1$. What's more can gain error control, an algorithm such that

$$z \in \Gamma_n(A) \Rightarrow \mathrm{dist}(z, \mathrm{Sp}(A)) < 2^{-n}.$$

Call this class Σ_1 - "output is reliable". A **stricter** condition is that

$$z \in \mathrm{Sp}(A) \Rightarrow \mathrm{dist}(z, \Gamma_n(A)) < 2^{-n}$$

as well - "output is reliable and captures nearly everything". Call this class Δ_1

Problem: $\mathrm{SCI}(\mathrm{Sp}, \mathcal{B}(l^2(\mathbb{N}))) = 3$

Solution: restrict to smaller subclass. Large subclass $\Omega \subset \mathcal{B}(l^2(\mathbb{N}))$ with $\mathrm{SCI}(\mathrm{Sp}, \Omega) = 1$. What's more can gain error control, an algorithm such that

$$z \in \Gamma_n(A) \Rightarrow \mathrm{dist}(z, \mathrm{Sp}(A)) < 2^{-n}.$$

Call this class Σ_1 - "output is reliable". A **stricter** condition is that

$$z \in \mathrm{Sp}(A) \Rightarrow \mathrm{dist}(z, \Gamma_n(A)) < 2^{-n}$$

as well - "output is reliable and captures nearly everything". Call this class Δ_1

Some notation:

- If $z \notin \mathrm{Sp}(A)$ write

$$R(z, A) := (A - zI)^{-1} \in \mathcal{B}(l^2(\mathbb{N})).$$

- Operator norm: $\|A\| := \sup_{x:\|x\|=1} \|Ax\|$.

Main Result

Theorem 3

$$\mathrm{SCI}(\mathrm{Sp},\Omega) = 1$$

The problem is Σ_1 **not** Δ_1. *What's more we can build* Σ_1 *algorithms.*

Main Result

Theorem 3

$$\mathrm{SCI}(\mathrm{Sp}, \Omega) = 1$$

The problem is Σ_1 **not** Δ_1*. What's more we can build* Σ_1 *algorithms.*

Corollary 4

Let G *be a countable, locally finite, connected graph and* Ω_G *be class of finite range interaction Hamiltonians on vertices of* G*.*

$$\mathrm{SCI}(\mathrm{Sp}, \Omega_G) = 1$$

The problem is Σ_1 **not** Δ_1*. What's more we can build* Σ_1 *algorithms.*

Application: Laplacian on Quasi-crystals

G a graph, V set of vertices. $x \sim y$ if sites x, y connected by edge. Laplacian H_0 acts on $\psi \in l^2(V) \cong l^2(\mathbb{N})$ by

$$(H_0\psi)(x) = \sum_{y \sim x} \left(\psi(y) - \psi(x)\right).$$

Application: Laplacian on Quasi-crystals

G a graph, V set of vertices. $x \sim y$ if sites x, y connected by edge. Laplacian H_0 acts on $\psi \in l^2(V) \cong l^2(\mathbb{N})$ by

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

- Spectrum corresponds to the single-electron energy (electronic transportation).

Application: Laplacian on Quasi-crystals

G a graph, V set of vertices. $x \sim y$ if sites x, y connected by edge. Laplacian H_0 acts on $\psi \in l^2(V) \cong l^2(\mathbb{N})$ by

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

- Spectrum corresponds to the single-electron energy (electronic transportation).
- Calculating the spectrum currently an open problem in the quasicrystal community - 2D cases very hard!

Application: Laplacian on Quasi-crystals

G a graph, V set of vertices. $x \sim y$ if sites x, y connected by edge. Laplacian H_0 acts on $\psi \in l^2(V) \cong l^2(\mathbb{N})$ by

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

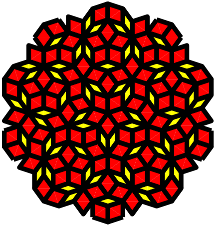
- Spectrum corresponds to the single-electron energy (electronic transportation).
- Calculating the spectrum currently an open problem in the quasicrystal community - 2D cases very hard!
- Vast physics literature on Hamiltonians on aperiodic structures.

Applications of Quasicrystals

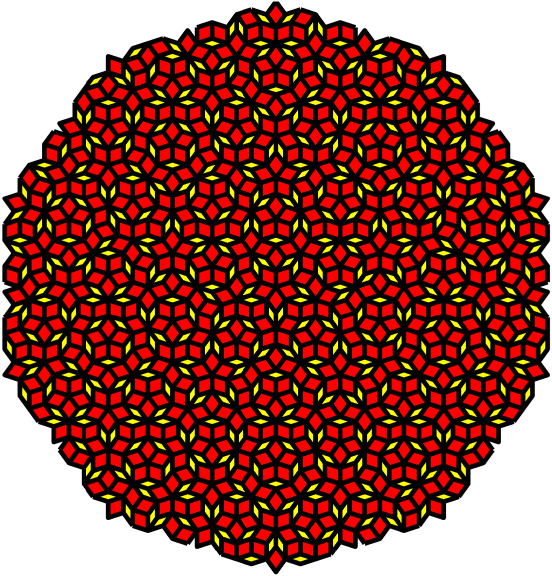
- Reinforce steel via coating - e.g. machinery, surgical instruments...
- Heat insulation
- LEDs
- Solar absorbers
- Unique electrical properties, optical properties, hardness and nonstick properties...

Constructing Penrose Tile

Constructing Penrose Tile



Constructing Penrose Tile



Naïve Approximations

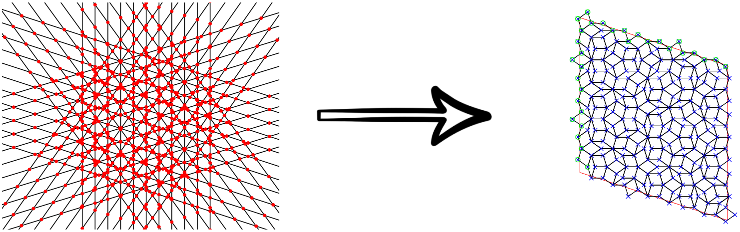
1. Finite section with open boundary conditions: compute eigenvalues of **truncated matrix** $P_n H_0 P_n$ for large n. Similar "Galerkin" methods - suffer from spectral pollution.

Naïve Approximations

1. Finite section with open boundary conditions: compute eigenvalues of **truncated matrix** $P_n H_0 P_n$ for large n. Similar "Galerkin" methods - suffer from spectral pollution.
2. Can construct Penrose tile via "Pentagrid" $\rightsquigarrow$ "Periodic Approximants"

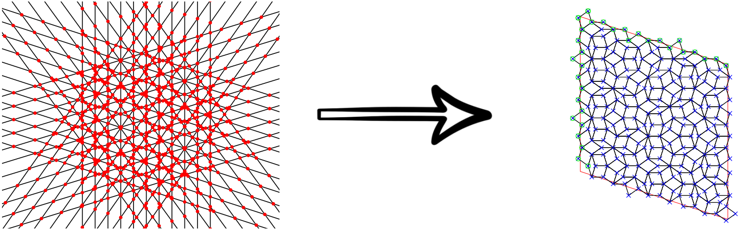
Naïve Approximations

1. Finite section with open boundary conditions: compute eigenvalues of **truncated matrix** $P_n H_0 P_n$ for large n. Similar "Galerkin" methods - suffer from spectral pollution.
2. Can construct Penrose tile via "Pentagrid" $\rightsquigarrow$ "Periodic Approximants"



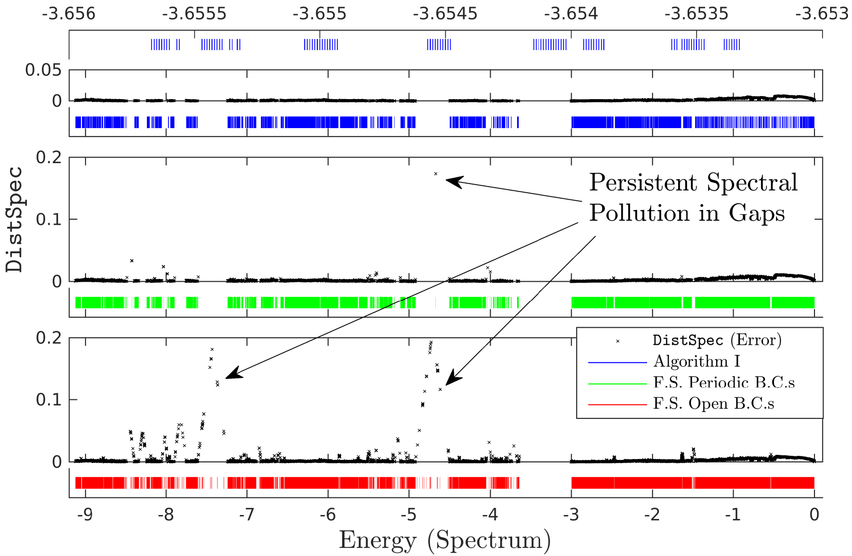
Naïve Approximations

1. Finite section with open boundary conditions: compute eigenvalues of **truncated matrix** $P_n H_0 P_n$ for large n. Similar "Galerkin" methods - suffer from spectral pollution.
2. Can construct Penrose tile via "Pentagrid" $\rightsquigarrow$ "Periodic Approximants"



These represent state of art in literature. Can we beat this?

Numerical Results



Advantages of Method

1. Gain error control - no other general method can do this!

Advantages of Method

1. Gain error control - no other general method can do this!
2. Completely local - we can calculate spectrum in any neighbourhood.

Advantages of Method

1. Gain error control - no other general method can do this!
2. Completely local - we can calculate spectrum in any neighbourhood.
3. Parallelisable.

Advantages of Method

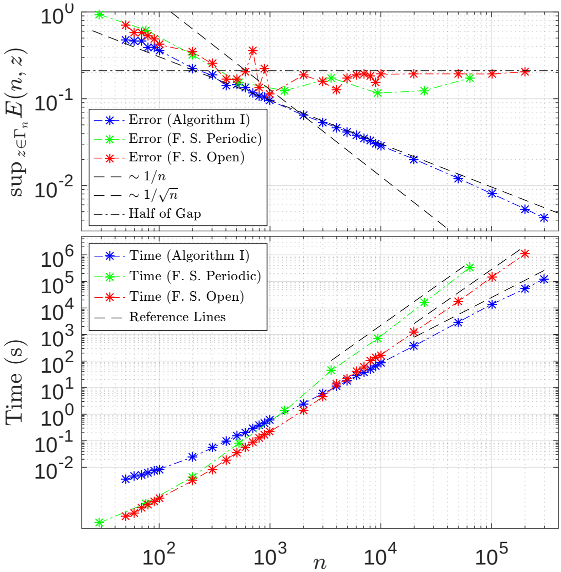
1. Gain error control - no other general method can do this!
2. Completely local - we can calculate spectrum in any neighbourhood.
3. Parallelisable.
4. Numerically stable.

Advantages of Method

1. Gain error control - no other general method can do this!
2. Completely local - we can calculate spectrum in any neighbourhood.
3. Parallelisable.
4. Numerically stable.

First algorithm that realises the sharp Σ_1 classification - logically impossible to do better.

Error and Speed Results



Thanks for Listening!