On spectral inclusion sets and computing the spectra and pseudospectra of bounded linear operators

Simon Chandler-Wilde, Ratchanikorn Chonchaiya, and Marko Lindner

Abstract. In this paper, we derive novel families of inclusion sets for the spectra and pseudospectra of large classes of bounded linear operators, and establish convergence of particular sequences of these inclusion sets to the spectrum or pseudospectrum, as appropriate. Our results apply, in particular, to bounded linear operators on a separable Hilbert space that, with respect to some orthonormal basis, have a representation as a bi-infinite matrix that is banded or band-dominated. More generally, our results apply in cases where the matrix entries themselves are bounded linear operators on some Banach space. In the scalar matrix entry case, we show that our methods, given the input information we assume, lead to a sequence of approximations to the spectrum, each element of which can be computed in finitely many arithmetic operations, so that, with our assumed inputs, the problem of determining the spectrum of a band-dominated operator has solvability complexity index one in the sense of Ben-Artzi et al. (2020). As a concrete and substantial application, we apply our methods to the determination of the spectra of non-self-adjoint bi-infinite tridiagonal matrices that are pseudoergodic in the sense of Davies [Commun. Math. Phys. 216 (2001), 687–704].

Dedicated to Prof. E. Brian Davies on the occasion of his 80th birthday

Contents

1. Introduction and overview

The determination of the spectra of bounded linear operators is a fundamental prob-
lem in functional analysis with applications across science and engineering (e.g.,
[22, 26, 33, 56, 85]). In this paper, we establish novel families of inclusion sets for
the spectrum and pseudospectrum of large classes of bounded linear operators, and
establish convergence of particular sequences of these inclusion sets to the spectrum
or pseudospectrum, as appropriate. Our results apply, in particular, to bounded linear
operators A on a separable Hilbert space Y that, with respect to some orthonormal
basis $(e_i)_{i \in \mathbb{Z}}$ for Y, have a matrix representation $[a_{i,j}]_{i,j \in \mathbb{Z}}$, where $a_{i,j} := (Ae_j, e_i)$
with $(\cdot, \cdot)$ the inner product on Y, that is banded or band-dominated (as defined in
Section 1.3 below). More generally, our results apply in the case that Y is the Banach
space $Y = \ell^2(\mathbb{Z}, X)$, where X is a Banach space, and $A : Y \to Y$ is a bounded lin-
ear operator that has the banded or band-dominated matrix representation $[a_{ij}]_{i,j \in \mathbb{Z}}$,
where each $a_{i,j} \in L(X)$, the space of bounded linear operators on X.

1.1. The main ideas, their provenance and significance

Let us sketch the key ideas of the paper, deferring a more detailed account of the main
results to Section 1.4. To explain these ideas, focusing first on the case that

$$A = [a_{i,j}]_{i,j \in \mathbb{Z}}$$

is a bi-infinite tridiagonal matrix, let $A(i : j, k : \ell)$ denote the $(j+1-i) \times (\ell+1-k)$
submatrix (or finite section) of A consisting of the elements in rows i through j and
columns k through ℓ. Abbreviate the $n \times n$ principal submatrix $A(k+1 : k+n, k+1 :
k+n)$ as $A_{n,k}$, and the $(n+2) \times n$ submatrix $A(k : k+n+1, k+1 : k+n)$ as $A_{n,k}^+$ (see
Figure 1.1). Further, let $A_n := A_{2n+1,-n-1}$ and $A_n^+ := A_{2n+1,-n-1}^+$, so that A_n^+ con-
tains all the non-zero elements in columns $-n$ through n. Similarly, for $x = (x_i)_{i \in \mathbb{Z}}$,
let $x_{n,k}$ denote the (column) vector of length n consisting of entries $k + 1$ through
$k + n$ of x.

Suppose that we wish to compute an approximation to Spec A, the spectrum of
a bi-infinite tridiagonal matrix A acting as a bounded linear operator on $\ell^2(\mathbb{Z})$. It is
well known that if $\lambda \in$ Spec A then, for every $\varepsilon > 0$, there exists $x \in \ell^2(\mathbb{Z})$ such that x
is an ε-*pseudoeigenvector of A for* λ, meaning that $x \neq 0$ and $\|(A - \lambda I)x\| \leq \varepsilon \|x\|$,

or there exists an ε-pseudoeigenvector of A^* for λ, where A^* is the matrix that is the transpose of A. Our first key observation is the following:

If x is an ε-pseudoeigenvector of A for λ then, for each $n \in \mathbb{N}$, there is a computable $\varepsilon_n > 0$, with $\varepsilon_n \to 0$ as $n \to \infty$, such that $x_{n,k}$ is an $(\varepsilon + \varepsilon_n)$-pseudoeigenvector of $A_{n,k}$ for λ for some $k \in \mathbb{Z}$.

This observation implies that, for some $k \in \mathbb{Z}$, $\mathrm{Spec}\, A \subset \mathrm{Spec}_\varepsilon A_{n,k}$, the (closed) ε-pseudospectrum of the matrix $A_{n,k}$ (see Section 1.3 for definitions). Indeed, this inclusion holds for all $\varepsilon > 0$, which we will see implies that

$$\mathrm{Spec}\, A \subset \Sigma_{\varepsilon_n}^n (A) := \overline{\bigcup_{k \in \mathbb{Z}} \mathrm{Spec}_{\varepsilon_n} A_{n,k}} \tag{1.1}$$

for some positive null sequence $(\varepsilon_n)_{n \in \mathbb{N}}$. Optimising the above idea (see Section 1.4.1) leads, concretely, to the conclusion that (1.1) holds with $\varepsilon_n = r(A)\eta_n$, where

$$r(A) := \sup_{i \in \mathbb{Z}} |a_{i,i+1}| + \sup_{i \in \mathbb{Z}} |a_{i,i-1}| \quad \text{and} \quad 2\sin\left(\frac{\pi}{4n+2}\right) \le \eta_n \le 2\sin\left(\frac{\pi}{2n+4}\right). \tag{1.2}$$

In terms of provenance, the above observation and (1.1) are reminiscent of the classical Gershgorin theorem [45, 86] that provides an enclosure for the eigenvalues of an arbitrary $n \times n$ matrix. Indeed, in the simplest case $n = 1$, $A_{n,k} = [a_{k+1,k+1}]$ is the $(k+1)$th diagonal entry of the matrix A, the above inequality implies that $\eta_n = 1$, and $\mathrm{Spec}_{\varepsilon_n} A_{n,k}$ is just $a_{k+1,k+1} + r(A)\overline{\mathbb{D}}$, the closed disc of radius $r(A)$ centred on $a_{k+1,k+1}$. Thus, (1.1) in the case $n = 1$ is simply

$$\mathrm{Spec}\, A \subset \overline{\bigcup_{k \in \mathbb{Z}} (a_{k,k} + r(A)\mathbb{D})}, \tag{1.3}$$

an enclosure of $\mathrm{Spec}\, A$ by discs centred on the diagonal entries of A that is an immediate consequence of (an infinite matrix version of) the classical Gershgorin theorem[1]. So, our enclosures (1.1) can be seen as an extension of Gershgorin's theorem to create a whole family of inclusion sets for the spectra of bi-infinite tridiagonal matrices.

An attraction of a sequence $(\Sigma_{\varepsilon_n}^n (A))_{n \in \mathbb{N}}$ of inclusion sets is that one can think of taking the limit. We will see significant examples below, including well-studied examples of tridiagonal pseudoergodic matrices in the sense of Davies [31], for which $\Sigma_{\varepsilon_n}^n (A) \xrightarrow{H} \mathrm{Spec}\, A$ as $n \to \infty$ ($\xrightarrow{H}$ is the Hausdorff convergence that we recall in

---

[1] An infinite matrix version of the Gershgorin theorem that implies (1.3) is proved as [20, Theorem 2.50]. Alternatively, the inclusion (1.3) follows by applying (2.3) below, with B the diagonal of A, noting that $\|A - B\| \le r(A)$ and $\mathrm{Spec}_{r(A)} B = \mathrm{Spec}\, B + r(A)\overline{\mathbb{D}}$.

Section 1.3). But, in general, the sequence $(\Sigma^n_{\varepsilon_n}(A))_{n\in\mathbb{N}}$, based on computing spectral quantities of square matrix finite sections of A, suffers from *spectral pollution* [34]: there are limit points of the sequence $(\Sigma^n_{\varepsilon_n}(A))$ that are not in Spec A.

The second key idea, dating back to Davies and Plum [34, pp. 423–434] (building on Davies [29]) for the self-adjoint case, and Hansen [55, Theorem 48] (proved in [56]) and [56] for the general case (see also [2,25,26,61]), is to avoid spectral pollution by working with rectangular rather than square finite sections. In our context this means to replace the $n \times n$ matrix $A_{n,k}$ by the $(n + 2) \times n$ matrix $A^+_{n,k}$, containing all the non-zero entries of A in columns $k + 1$ to $k + n$. It is a standard characterisation of the pseudospectrum (see (1.10) and (7.2), or [85, Section I.2]) that

$$\mathrm{Spec}_{\varepsilon_n} A_{n,k} = \left\{\lambda \in \mathbb{C} : \min(s_{\min}((A - \lambda I)_{n,k}), s_{\min}((A^* - \lambda I)_{n,k})) \le \varepsilon_n\right\},$$

where $s_{\min}$ denotes the smallest singular value. This implies (with a little work, see Proposition 3.1) that

$$\Sigma^n_{\varepsilon_n}(A) = \left\{\lambda \in \mathbb{C} : \inf_{k\in\mathbb{Z}} \min(s_{\min}((A - \lambda I)_{n,k}), s_{\min}((A^* - \lambda I)_{n,k})) \le \varepsilon_n\right\}.$$

Replacing square with rectangular finite sections in the above expression gives an alternative sequence of approximations to Spec A, viz. $(\Gamma^n_{\varepsilon''_n}(A))_{n\in\mathbb{N}}$, where

$$\Gamma^n_{\varepsilon''_n}(A) := \left\{\lambda \in \mathbb{C} : \inf_{k\in\mathbb{Z}} \min(s_{\min}((A - \lambda I)^+_{n,k}), s_{\min}((A^* - \lambda I)^+_{n,k})) \le \varepsilon''_n\right\}. \quad (1.4)$$

Again, Spec $A \subset \Gamma^n_{\varepsilon''_n}(A)$ for each n (equation (1.30)), provided ε''_n is large enough. (We show in Section 5 that if $\varepsilon''_n := r(A)\eta''_n$ then $\eta''_n = 2\sin(\pi/(2n + 2))$ is the smallest choice of η''_n that maintains this inclusion for all tridiagonal A.) Crucially, and this is the benefit of replacing square by rectangular finite sections, also $\Gamma^n_{\varepsilon''_n}(A) \xrightarrow{H} \mathrm{Spec}\, A$ as $n \to \infty$ (Theorem 1.2), for every tridiagonal A.

The third idea is to realise that the entries in the tridiagonal matrix can themselves be matrices (or indeed operators on some Banach space). This extends the above results to general banded matrices (see Section 1.4 below). Further, by perturbation arguments, we can obtain inclusion sets, and a sequence of approximations converging to the spectrum, also for the band-dominated case (Section 1.4.3).

The final main idea of the paper is that, through a further perturbation argument (see Section 1.4.4), the union and infimum over all $k \in \mathbb{Z}$, in (1.1) and (1.4), can be replaced by a union and infimum over $k \in K_n$, where $K_n \subset \mathbb{Z}$ is a finite, A- and n-dependent subset. Thus, provided that one has access through some oracle to the sets K_n, for $n \in \mathbb{N}$, a sequence of approximations to Spec A is obtained that can be computed in a finite number of arithmetic operations. In the language of the solvability complexity index (SCI) of [2, 3, 22, 23, 25] these results (see Section 1.4.5 and Section 7) lead to a proof that the computational problem of determining the spectrum

of a band-dominated operator has SCI equal to one, if one chooses an evaluation set (in the sense of [2, 22, 25]) that provides access to appropriately chosen submatrices, i.e., to $A_{n,k}^+$ and $(A^*)_{n,k}^+$, for $n \in \mathbb{N}$ and $k \in K_n$. To showcase this result we provide a concrete realisation of the associated algorithm (concrete in the sense that the sets K_n can be constructed) for the class of tridiagonal matrices that are pseudoergodic in the sense of Davies [31] (see Section 8.4), illustrating this algorithm by numerical results taken from [14].

To clarify the significance of this combination of ideas, let us make comparison to the classical idea of approximating the spectrum or pseudospectrum of A by those of a single large finite section of A. To avoid spectral pollution we take a rectangular finite section: precisely (cf. (1.4)), consider the approximation

$$\mathcal{S}_\varepsilon^n(A) := \left\{ \lambda \in \mathbb{C} : \min(s_{\min}((A - \lambda I)_n^+), s_{\min}((A^* - \lambda I)_n^+)) \leq \varepsilon \right\}, \qquad (1.5)$$

defined in terms of smallest singular values of $2n+3 \times 2n+1$ finite sections of A and A^*, which is a particular instance of the approximation sequence proposed by Hansen [55]. Then, it follows from [55, Theorem 48] (the proof deferred to [56]) that, for every $\varepsilon > 0$, $\mathcal{S}_\varepsilon^n(A) \xrightarrow{H} \mathrm{Spec}_\varepsilon A$ as $n \to \infty$. Since also $\mathrm{Spec}_\varepsilon A \xrightarrow{H} \mathrm{Spec}\, A$ as $\varepsilon \to 0$, we have that

$$\mathrm{Spec}\, A = \lim_{\varepsilon \to 0} \lim_{n \to \infty} \mathcal{S}_\varepsilon^n(A).$$

This result trivially implies that there exists a positive null sequence $(\varepsilon_n)_{n \in \mathbb{N}}$ such that

$$\mathcal{S}_{\varepsilon_n}^n(A) \xrightarrow{H} \mathrm{Spec}\, A \quad \text{as } n \to \infty.$$

But, in contrast to the new approximation families that we introduce in this paper, this null sequence (ε_n) depends on A in an ill-defined, and unquantifiable way. Indeed, the SCI result arguments of [2], that we discuss in Section 1.4.5 and Section 1.6, imply that there exists no sequence of universal algorithms, indexed by $n \in \mathbb{N}$, operating on the set of all tridiagonal matrices, such that the nth algorithm takes just finitely many values of the matrix entries $a_{i,j}$ as input and gives ε_n as output, in such a way that $\mathcal{S}_{\varepsilon_n}^n(A) \xrightarrow{H} \mathrm{Spec}\, A$ for all tridiagonal A [23]. In the language of [2], the SCI of computing the spectrum of a tridiagonal operator is two [23] if the evaluation set is restricted to the mappings $A \mapsto a_{i,j}$, for $i, j \in \mathbb{Z}$ (i.e., to only *local information* about the matrix A). In contrast, our results in this paper show that the SCI is one if we expand the evaluation set to include also the mappings $A \mapsto K_n$ for $n \in \mathbb{N}$ (in some sense *global information* about the matrix A). We are hopeful that changing the evaluation set in analogous ways, to include global as well as local information, should lead to reductions in SCI for a broad range of problems, in particular, for the computation of a range of other spectral quantities (cf. [2, 22, 25, 26]).

1.2. The structure of the paper

In the next subsection, we briefly recall key notations and definitions necessary to read our main results. In Section 1.4, we give an overview of these main results. We state our results on inclusion sets for the spectrum and pseudospectrum in the case that the matrix representation of A is banded in Section 1.4.1, prove our results on convergence of our inclusion set sequences to the spectrum and pseudospectrum for the banded case in Section 1.4.2, detail the extension to the band-dominated case in Section 1.4.3, and summarise our results on computability and the solvability complexity index in Section 1.4.4 and Section 1.4.5. In Section 1.5, we give a first example (the shift operator) that serves to illustrate our inclusions (1.22), (1.25), and (1.30) below and demonstrate their sharpness, and in Section 1.6, we survey related work, expanding on the discussion above.

In Section 2, we detail further properties of pseudospectra needed for our later arguments and introduce the so-called Globevnik property. Sections 3–5 prove the inclusions (1.22), (1.25), and (1.30), respectively. In Section 6, we prove Theorem 1.9 that we state in Section 1.4.4, that justifies replacement of the infinite set of submatrices that are part of the definitions of our inclusion sets with a finite subset. This result is extended to the general band-dominated case in Theorem 6.1. In Section 7, we give details of implementation, leading to our solvability complexity index results. We discuss concrete applications of our general results in Section 8, notably to the case where A is tridiagonal and pseudoergodic in the sense of Davies [31]. In Section 9, we indicate open problems and directions for further work.

1.3. Key notations

As usual, we write $\mathbb{Z}, \mathbb{R}, \mathbb{C}$ for the integer, real, and complex numbers, $\mathbb{N} \subset \mathbb{Z}$ is the set of positive integers, $\mathbb{D} \subset \mathbb{C}$ the open unit disc, $\mathbb{T} \subset \mathbb{C}$ the unit circle, and, for $n \in \mathbb{N}$, $\mathbb{T}_n \subset \mathbb{T}$ denotes the set of nth roots of unity, $\mathbb{T}_n = \{z \in \mathbb{C} : z^n = 1\}$. $\bar{S}$ denotes the closure of a set $S \subset \mathbb{C}$ and $\bar{z}$ denotes the complex conjugate of $z \in \mathbb{C}$. Throughout, X will denote some Banach space: at particular points in the text we will make clear where we are specialising in the case that X is a Hilbert space or is finite-dimensional. Further,

$$E := \ell^2(\mathbb{Z}, X) \tag{1.6}$$

will denote the space of bi-infinite X-valued sequences $x = (x_j)_{j \in \mathbb{Z}}$ that are square-summable, a Banach space with the norm $\| \cdot \|$ defined by $\|x\| = (\sum_{j \in \mathbb{Z}} |x_j|_X^2)^{1/2}$. We will abbreviate $\ell^2(\mathbb{Z}, \mathbb{C})$ as $\ell^2(\mathbb{Z})$.

The spectrum, pseudospectrum and lower norm. Let Y be a Banach space (throughout the paper our Banach and Hilbert spaces are assumed complex). Given $A \in L(Y)$,

the space of bounded linear operators on Y, we denote the spectrum of A by Spec A. With the convention that, for $B \in L(Y)$, $\|B^{-1}\| := \infty$ if B is not invertible, and that $1/\infty := 0$, and where I is the identity operator on Y, we define the *closed ε-pseudospectrum*, $\mathrm{Spec}_\varepsilon A$, by

$$\mathrm{Spec}_\varepsilon A := \{\lambda \in \mathbb{C} : 1/\|(A - \lambda I)^{-1}\| \leq \varepsilon\}, \quad \varepsilon > 0. \tag{1.7}$$

Clearly, Spec $A \subset \mathrm{Spec}_\varepsilon A \subset \mathrm{Spec}_{\varepsilon'} A$, for $0 < \varepsilon < \varepsilon'$, and

$$\mathrm{Spec}\, A = \bigcap_{\varepsilon > 0} \mathrm{Spec}_\varepsilon A. \tag{1.8}$$

To write certain inclusion statements more compactly, we will define $\mathrm{Spec}_0 A := \mathrm{Spec}\, A$. For $\varepsilon > 0$ we define also the *open ε-pseudospectrum*[2], $\mathrm{spec}_\varepsilon A$, given by (1.7) with the $\leq$ replaced by $<$; the above inclusions and (1.8) hold also with $\mathrm{Spec}_\varepsilon A$ replaced by $\mathrm{spec}_\varepsilon A$. As a consequence of standard perturbation arguments (e.g., [33, Theorem 1.2.9]), $\mathrm{Spec}_\varepsilon A$ is closed, for $\varepsilon \geq 0$, and $\mathrm{spec}_\varepsilon A$ open, for $\varepsilon > 0$.

Given Banach spaces X and Y and $B \in L(X, Y)$, the space of bounded linear operators from X to Y, one can study, as a counterpart to the operator norm $\|B\| := \sup_{\|x\|=1} \|Bx\|$, the quantity

$$\nu(B) := \inf_{\|x\|=1} \|Bx\|$$

that is sometimes (by abuse of notation) called the *lower norm* (and sometimes, e.g., in [78], the *injection modulus*) of B. In the case that Y is a Hilbert space and $A \in L(Y)$, $\|A\|$ is the largest singular value of A and $\nu(A)$ is the smallest[3]. For every Banach space Y and $A \in L(Y)$ the equality

$$1/\|A^{-1}\| = \min(\nu(A), \nu(A^*)) =: \mu(A), \tag{1.9}$$

holds, where A^* is the Banach space adjoint (see below). In particular, A is invertible if and only if $\nu(A)$ and $\nu(A^*)$ are both nonzero, i.e., if and only if $\mu(A) \neq 0$, in which

---

[2]We refer readers unfamiliar with the pseudospectrum, its properties and applications, to Section 2 below, and to Böttcher and Lindner [9], Davies [33], Trefethen and Embree [85], and Hagen, Roch, Silbermann [51]. We note that the ε-pseudospectrum, as defined in [33, 85] and most recent literature, is our open ε-pseudospectrum, but in part of the literature, including [51], ε-pseudospectrum means our closed ε-pseudospectrum. As we recall in Section 2, if the Banach space X has the so-called *Globevnik property*, which holds in particular, if X is a Hilbert space or is finite-dimensional, these two pseudospectra are related simply by $\mathrm{Spec}_\varepsilon(B) = \overline{\mathrm{spec}_\varepsilon(B)}$, for $\varepsilon > 0$ and $B \in L(X)$.

[3]Recall that the singular values of $A \in L(Y)$, defined whenever Y is a Hilbert space, are the points in the spectrum of $(A'A)^{1/2}$, where A' is the Hilbert space adjoint.

case $\nu(A) = \nu(A^*) = \mu(A)$. Further, if A is Fredholm of index zero, in particular, if Y is finite-dimensional, then $\nu(A) = 0$ if and only if $\nu(A^*) = 0$, so that $\mu(A) = \nu(A)$.

From the definitions of $\mathrm{spec}_\varepsilon A$ and $\mathrm{Spec}_\varepsilon A$ and (1.9) it follows that

$$
\begin{aligned}
\mathrm{spec}_\varepsilon A &= \{\lambda \in \mathbb{C} : \mu(A - \lambda I) < \varepsilon\}, \quad \varepsilon > 0, \\
\mathrm{Spec}_\varepsilon A &= \{\lambda \in \mathbb{C} : \mu(A - \lambda I) \le \varepsilon\}, \quad \varepsilon \ge 0.
\end{aligned}
\tag{1.10}
$$

Simple but important properties of the lower norm are that, if $A, B \in L(X, Y)$ and Z is a closed subspace of X, then

$$
\nu(A|_Z) \ge \nu(A) \quad \text{and} \quad |\nu(A) - \nu(B)| \le \|A - B\|;
\tag{1.11}
$$

see, e.g., [64, Lemma 2.38] for the second of these properties. It is easy to see (since $\mu(A) = \nu(A)$ if $\mu(A) \ne 0$) that also

$$
|\mu(A) - \mu(B)| \le \|A - B\|.
\tag{1.12}
$$

Hausdorff distance and notions of set convergence. Let $\mathbb{C}^B$, $\mathbb{C}^C$ denote, respectively, the sets of bounded and compact non-empty subsets of $\mathbb{C}$. For $a \in \mathbb{C}$ and $S \in \mathbb{C}^B$, let $\mathrm{dist}(a, S) := \inf_{b \in S} |a - b|$. For $S_1, S_2 \subset \mathbb{C}^B$ let

$$
d_H(S_1, S_2) := \max\left(\sup_{s_1 \in S_1} \mathrm{dist}(s_1, S_2), \sup_{s_2 \in S_2} \mathrm{dist}(s_2, S_1) \right),
$$

which is termed the *Hausdorff distance* between S_1 and S_2. It is well known (e.g., [51, 60]) that $d_H(\cdot, \cdot)$ is a metric on $\mathbb{C}^C$, the so-called *Hausdorff metric*. For $S_1, S_2 \in \mathbb{C}^B$ it is clear that $d_H(S_1, S_2) = d_H(\overline{S_1}, \overline{S_2})$, so that $d_H(S_1, S_2) = 0$ if and only if $\overline{S_1} = \overline{S_2}$, and $d_H(\cdot, \cdot)$ is a pseudometric on $\mathbb{C}^B$. For a sequence $(S_n) \subset \mathbb{C}^B$ and $S \in \mathbb{C}^B$ we write $S_n \xrightarrow{H} S$ if $d_H(S_n, S) \to 0$. This limit is in general not unique: if $S_n \xrightarrow{H} S$ and $T \in \mathbb{C}^B$ then $S_n \xrightarrow{H} T$ if and only if $\overline{S} = \overline{T}$.

Simple results that we will use extensively are that if $(S_n) \subset \mathbb{C}^C$ and $S_1 \supset S_2 \supset \cdots$, in which case $S := \cap_{n=1}^\infty S_n$ is non-empty, then $S_n \xrightarrow{H} S$, and that if $S \in \mathbb{C}^B$, $(S_n), (T_n) \subset \mathbb{C}^B$, $S \subset T_n \subset S_n$ and $S_n \xrightarrow{H} S$, then $T_n \xrightarrow{H} S$. (These results can be seen directly, or via the equivalence of Hausdorff convergence with other notions of set convergence, see [60, p. 171] or [51, Proposition 3.6].)

Band and band-dominated operators. Let $A \in L(E)$, in which case (e.g., [64, Section 1.3.5]) A has a matrix representation $[A] = [a_{i,j}]$, that we will denote again by A, with each $a_{i,j} \in L(X)$, such that, for every $x = (x_j)_{j \in \mathbb{Z}} \in E$ with finitely many non-zero entries, $Ax = y = (y_i)_{i \in \mathbb{Z}}$, where

$$
y_i = \sum_{j \in \mathbb{Z}} a_{i,j} x_j, \quad i \in \mathbb{Z}.
\tag{1.13}
$$

The main body of our results are for the case where $A \in \mathrm{BO}(E)$, where $\mathrm{BO}(E)$ denotes the linear subspace of those $A \in L(E)$ whose matrix representation is *banded* with some *band-width* w, meaning that $a_{i,j} = 0$ for $|i - j| > w$, in which case (1.13) holds (and is a finite sum) for all $x \in E$. By perturbation arguments we also extend our results to the case where $A \in \mathrm{BDO}(E)$, where $\mathrm{BDO}(E)$, the space of *band-dominated operators*, is the closure of $\mathrm{BO}(E)$ in $L(E)$ with respect to the operator norm (see, e.g., [64, Section 1.3.6]).

Adjoints. In many of our arguments, we require the (Banach space) adjoint A^* of A which we think of as an operator on $E^* := \ell^2(\mathbb{Z}, X^*)$, where X^* is the dual space of X. This has the matrix representation $[a_{j,i}^*]$, where $a_{j,i}^* \in L(X^*)$ is the Banach space adjoint of $a_{j,i} \in L(X)$. In the case that $X = \mathbb{C}^n$, equipped with the ℓ^p norm, with $1 \leq p \leq \infty$, we will identify X^* with $\mathbb{C}^n$ equipped with the ℓ^q norm, where q is the usual conjugate index $(1/p + 1/q = 1)$, so that $a_{j,i} \in \mathbb{C}^{n \times n}$ is just an $n \times n$ matrix with scalar entries, and $a_{j,i}^*$ is just $a_{j,i}^T$, the transpose of the matrix $a_{j,i}$.

1.4. Our main results

Having introduced key notations and definitions, let us now state our main results.

Reduction to the tridiagonal case. A simple but important observation is that, to construct inclusion sets for the spectrum and pseudospectrum of $A \in \mathrm{BO}(E)$ it is enough to consider the case when the matrix representation has band-width one, so that the matrix is tridiagonal[4]. For suppose that A is banded with band-width $w \in \mathbb{N}$ and, for $b \in \mathbb{N}$, let $\mathcal{I}_b : E \to E_b := \ell^2(\mathbb{Z}, X^b)$ be the mapping

$$x = (x_i)_{i \in \mathbb{Z}} \mapsto \xi = (\xi_i)_{i \in \mathbb{Z}}$$

with $\xi_i = (x_{bi+1}, \dots, x_{b(i+1)}) \in X^b$. Then, $\mathcal{I}_b$ is an isomorphism, indeed an isometric isomorphism if we equip the product space X^b with the norm

$$\|(\xi_1, \dots, \xi_b)\|_{X^b} := \left(\sum_{j=1}^{b} \|\xi_j\|_X^2 \right)^{1/2} \tag{1.14}$$

(which we assume hereafter). Thus, $A \in L(E)$ and $A_b \in L(E_b)$, given by $A_b = \mathcal{I}_b A \mathcal{I}_b^{-1}$, share the same spectrum and pseudospectra and, provided $b \geq w$, the matrix representation of A_b is tridiagonal.

---

[4]In the case $X = \mathbb{C}$, in which case $E = \ell^2(\mathbb{Z})$ is a Hilbert space, and if A is self-adjoint, reduction of computation of Spec A (and the pseudospectra of A) to consideration of the tridiagonal case is discussed in another sense in [55, Corollary 26], viz. the sense of reduction to tridiagonal form by application of a sequence of Householder transformations (which preserve spectrum and pseudospectra).

Thus, in the statement of our main results we focus on the case where $A \in L(E)$ has a matrix representation that is tridiagonal, so that, introducing the abbreviations $\alpha_j := a_{j+1,j}, \beta_j := a_{j,j}, \gamma_j := a_{j-1,j}$,

$$
A = \begin{pmatrix}
\ddots & \ddots & & & & & \\
\ddots & \beta_{-2} & \gamma_{-1} & & & & \\
 & \alpha_{-2} & \beta_{-1} & \gamma_0 & & & \\
 & & \alpha_{-1} & \boxed{\beta_0} & \gamma_1 & & \\
 & & & \alpha_0 & \beta_1 & \gamma_2 & \\
 & & & & \alpha_1 & \beta_2 & \ddots \\
 & & & & & \ddots & \ddots
\end{pmatrix},
\tag{1.15}
$$

where the box encloses $\beta_0 = a_{0,0}$. As a special case of (1.13), $y := Ax \in E$ has ith entry $y_i \in X$ given by

$$
y_i = \alpha_{i-1} x_{i-1} + \beta_i x_i + \gamma_{i+1} x_{i+1}, \quad i \in \mathbb{Z}.
$$

It is easy to see that the mapping A given by this rule is bounded, i.e., $A \in L(E)$, if and only if (α_i), (β_i), and (γ_i) are bounded sequences in $L(X)$, i.e., are sequences in $\ell^\infty(\mathbb{Z}, L(X))$, a Banach space equipped with the norm $\|\cdot\|_\infty$ defined by $\|x\|_\infty := \sup_{i \in \mathbb{Z}} \|x_i\|_{L(X)}$, in which case

$$
\|A\| \le r(A) + \|\beta\|_\infty, \quad \text{where } r(A) := \|\alpha\|_\infty + \|\gamma\|_\infty.
\tag{1.16}
$$

Conversely, noting that $\|a_{i,j}\|_{L(X)} \le \|A\|$ for every $i, j \in \mathbb{Z}$,

$$
r(A) = \|\alpha\|_\infty + \|\gamma\|_\infty \le 2\|A\|.
\tag{1.17}
$$

1.4.1. Our spectral inclusion sets for the tridiagonal case: The τ, π, and τ_1 methods.

Since to derive spectral inclusion sets for $A \in \mathrm{BO}(E)$ it is enough to study the tridiagonal case, we restrict to this case in this subsection (and in much of the rest of the paper); the action of A is that of multiplication by the tridiagonal matrix (1.15). In this paper, we derive for this tridiagonal case three different families of spectral inclusion sets that enclose the spectrum or pseudospectra of the operator A. Each spectral inclusion set is defined as the union of the pseudospectra (or related sets) of certain finite matrices. These finite matrices are either principal submatrices of the infinite matrix (1.15) (we call the corresponding approximation method the "τ method", where τ is for *truncation*), or are circulant-type modifications of these submatrices (the "π method", π for *periodised truncation*), or they are connected to infinite submatrices with finite row- or column number (this is our "τ_1 method", τ_1 for *one-sided truncation*).

In each method, as discussed below (1.23), there is a "penalty term", ε_n, ε_n', and ε_n'', for the τ, π, and τ_1 methods, respectively, which arises from replacing infinite matrices by finite matrices. As will be demonstrated by the simple example where A is the shift operator in Section 1.5, our values for ε_n, ε_n', and ε_n'' are optimal; there is at least one tridiagonal $A \in L(E)$ such that, if we make any of these values smaller, the claimed inclusions fail. In this limited sense our inclusion sets are sharp.

To define these methods, for $k \in \mathbb{Z}$ and $n \in \mathbb{N}$, let $P_{n,k} : E \to E$ be the projection operator given by

$$(P_{n,k}x)_j = \begin{cases} x_j, & j \in \{k+1, k+2, \ldots, k+n\}, \\ 0, & \text{otherwise,} \end{cases} \tag{1.18}$$

and let $E_{n,k} := P_{n,k}(E)$ be the range of $P_{n,k}$. In the τ and π methods, we construct inclusion sets for the spectrum and pseudospectrum of A from those of the finite section operators

$$A_{n,k} := P_{n,k} A P_{n,k}|_{E_{n,k}} : E_{n,k} \to E_{n,k}$$

and their "periodised" versions $A_{n,k}^{\mathrm{per},t} : E_{n,k} \to E_{n,k}$, with $t \in \mathbb{T}$, so that, for $n \in \mathbb{N}$, $k \in \mathbb{Z}$, $A_{n,k}$ acts by multiplication by the $n \times n$ principal submatrix

$$A_{n,k} = \begin{pmatrix} \beta_{k+1} & \gamma_{k+2} & & & \\ \alpha_{k+1} & \beta_{k+2} & \gamma_{k+3} & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_{k+n-2} & \beta_{k+n-1} & \gamma_{k+n} \\ & & & \alpha_{k+n-1} & \beta_{k+n} \end{pmatrix} \tag{1.19}$$

of A and $A_{n,k}^{\mathrm{per},t}$ acts by multiplication by the matrix
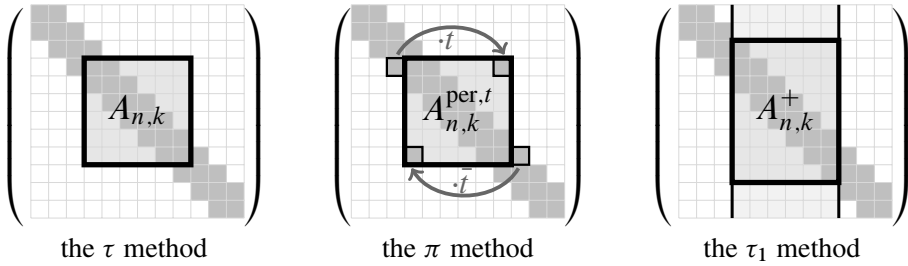
$$A_{n,k}^{\mathrm{per},t} := A_{n,k} + B_{n,k}^t,$$

where $B_{n,k}^t$ is the $n \times n$ matrix whose entry in row i, column j is

$$\delta_{i,1}\delta_{j,n}t\alpha_k + \delta_{i,n}\delta_{j,1}\bar{t}\gamma_{k+n+1},$$

where δ_{ij} is the Kronecker delta, so that, for $n \geq 3$,

$$A_{n,k}^{\mathrm{per},t} = \begin{pmatrix} \beta_{k+1} & \gamma_{k+2} & & & t\alpha_k \\ \alpha_{k+1} & \beta_{k+2} & \gamma_{k+3} & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_{k+n-2} & \beta_{k+n-1} & \gamma_{k+n} \\ \bar{t}\gamma_{k+n+1} & & & \alpha_{k+n-1} & \beta_{k+n} \end{pmatrix} \tag{1.20}$$

(see Figure 1.1 for visualisations of these finite section matrices).

Figure 1.1. The finite section matrices $A_{n,k}$, $A_{n,k}^{\text{per},t}$, and $A_{n,k}^+$, given by (1.19), (1.20), and (1.32), respectively, that arise in the τ, π, and τ_1 methods, respectively.

The τ method. Defining[5], for $n \in \mathbb{N}$,

$$\sigma_\varepsilon^n(A) := \bigcup_{k \in \mathbb{Z}} \text{spec}_\varepsilon A_{n,k}, \quad \varepsilon > 0 \quad \text{and} \quad \Sigma_\varepsilon^n(A) := \overline{\bigcup_{k \in \mathbb{Z}} \text{Spec}_\varepsilon A_{n,k}}, \quad \varepsilon \geq 0, \quad (1.21)$$

the τ-method inclusions sets are the right-hand sides of the following inclusions (recall that $\text{Spec}_0 A := \text{Spec } A$ so that taking $\varepsilon = 0$ in the first of these inclusions provides an inclusion for $\text{Spec } A$):

$$\boxed{\text{Spec}_\varepsilon A \subset \Sigma_{\varepsilon+\varepsilon_n}^n(A), \quad \varepsilon \geq 0 \quad \text{and} \quad \text{spec}_\varepsilon A \subset \sigma_{\varepsilon+\varepsilon_n}^n(A), \quad \varepsilon > 0.} \quad (1.22)$$

These inclusions (proved as Theorem 3.6) hold for all $n \in \mathbb{N}$, the second inclusion with no constraint on the Banach space X, the first with the constraint that X satisfies Globevnik's property (see Section 2), which holds, in particular, if X is finite-dimensional or a Hilbert space. The term ε_n, defined in Theorem 3.6 (and see Corollary 3.7 for the special case that A is bidiagonal), enters (1.22) as a "truncation penalty" when passing from the infinite matrix A to a finite matrix of size $n \times n$. (Similar comments apply to the terms ε_n' and ε_n'' in the π and τ_1 methods below.) Corollary 3.8 provides an upper bound for ε_n that implies that

$$0 \leq \varepsilon_n \leq 2(\|\alpha\|_\infty + \|\gamma\|_\infty) \sin \frac{\pi}{2(n+2)}$$
$$\leq (\|\alpha\|_\infty + \|\gamma\|_\infty) \frac{\pi}{n+2}. \quad (1.23)$$

---

[5]In the following equation $\text{spec}_\varepsilon A_{n,k}$ denotes the open ε-pseudospectrum of the operator $A_{n,k} : E_{n,k} \to E_{n,k}$, where $E_{n,k} \subset E$ is equipped with the norm of E or, what is the same thing, the open ε-pseudospectrum of the matrix (1.19), where the matrix norm is that induced by the norm (1.14) on X^n. A similar comment holds for the other ε-pseudospectra in (1.21) and (1.24).

The π method. Similarly, defining, for $n \in \mathbb{N}$ and $t \in \mathbb{T}$,

$$\pi_\varepsilon^{n,t}(A) := \bigcup_{k \in \mathbb{Z}} \mathrm{spec}_\varepsilon A_{n,k}^{\mathrm{per},t}, \quad \varepsilon > 0 \quad \text{and} \quad \Pi_\varepsilon^{n,t}(A) := \overline{\bigcup_{k \in \mathbb{Z}} \mathrm{Spec}_\varepsilon A_{n,k}^{\mathrm{per},t}}, \quad \varepsilon \geq 0, \tag{1.24}$$

we prove in Corollary 4.5, for $n \in \mathbb{N}$ and $t \in \mathbb{T}$, the π-method inclusions

$$\boxed{\mathrm{Spec}_\varepsilon A \subset \Pi_{\varepsilon+\varepsilon_n'}^{n,t}(A), \quad \varepsilon \geq 0 \quad \text{and} \quad \mathrm{spec}_\varepsilon A \subset \pi_{\varepsilon+\varepsilon_n'}^{n,t}(A), \quad \varepsilon > 0,} \tag{1.25}$$

where

$$\varepsilon_n' := 2(\|\alpha\|_\infty + \|\gamma\|_\infty) \sin \frac{\pi}{2n}, \tag{1.26}$$

and the second inclusion in (1.25) holds whatever the Banach space X, while the first requires that X has the Globevnik property.

Our motivation for the π method and the definitions (1.24) comes from consideration of the case when the diagonals of A are constant, i.e., A is a so-called *block-Laurent* matrix. If A is tridiagonal and block-Laurent and X is a finite-dimensional Hilbert space, then $\mathrm{Spec}_\varepsilon A = \cup_{t \in \mathbb{T}} \Pi_\varepsilon^{n,t}(A)$, for $\varepsilon \geq 0$, and $\mathrm{spec}_\varepsilon A = \cup_{t \in \mathbb{T}} \pi_\varepsilon^{n,t}(A)$, for $\varepsilon > 0$ (see Theorem 8.1).

The τ_1 method. The τ_1 method modifies these constructions using ideas from [29, 34, 55] (see the discussion in Section 1.1). To see the similarities but distinction between the τ and τ_1 methods, for $B \in L(E)$ and $n \in \mathbb{N}$ let

$$\mu_n^\dagger(B) := \inf_{k \in \mathbb{Z}} \mu(P_{n,k} B|_{E_{n,k}}) = \inf_{k \in \mathbb{Z}} \min\left(\nu(P_{n,k} B|_{E_{n,k}}), \nu(P_{n,k} B^*|_{E_{n,k}})\right). \tag{1.27}$$

Then, see Proposition 3.1,

$$\sigma_\varepsilon^n(A) = \left\{\lambda \in \mathbb{C} : \mu_n^\dagger(A - \lambda I) < \varepsilon\right\} \quad \text{and} \quad \Sigma_\varepsilon^n(A) = \left\{\lambda \in \mathbb{C} : \mu_n^\dagger(A - \lambda I) \leq \varepsilon\right\}, \tag{1.28}$$

for $\varepsilon > 0$, the first of these identities holding for every Banach space X, the second if X has the Globevnik property. The operators $P_{n,k} B|_{E_{n,k}}$ are two-sided truncations of B, each corresponding to an $n \times n$ matrix. By contrast, in the τ_1 method we make one-sided truncations, dropping the $P_{n,k}$'s in (1.27), and so, replacing $\mu_n^\dagger(B)$ by

$$\mu_n(B) := \inf_{k \in \mathbb{Z}} \min\left(\nu(B|_{E_{n,k}}), \nu(B^*|_{E_{n,k}})\right).$$

Precisely, for $n \in \mathbb{N}$ let

$$\begin{aligned}
\gamma_\varepsilon^n(A) &:= \left\{\lambda \in \mathbb{C} : \mu_n(A - \lambda I) < \varepsilon\right\}, \quad \varepsilon > 0, \\
\Gamma_\varepsilon^n(A) &:= \left\{\lambda \in \mathbb{C} : \mu_n(A - \lambda I) \leq \varepsilon\right\}, \quad \varepsilon \geq 0.
\end{aligned} \tag{1.29}$$

Then, the τ_1 method inclusions are

$$
\boxed{
\begin{aligned}
\Gamma_\varepsilon^n(A) &\subset \operatorname{Spec}_\varepsilon A \subset \Gamma_{\varepsilon+\varepsilon_n''}^n(A), & \varepsilon &\geq 0, \\
\gamma_\varepsilon^n(A) &\subset \operatorname{spec}_\varepsilon A \subset \gamma_{\varepsilon+\varepsilon_n''}^n(A), & \varepsilon &> 0,
\end{aligned}
}
\tag{1.30}
$$

which hold for all $n \in \mathbb{N}$ (whatever the Banach space X), where

$$
\varepsilon_n'' := 2(\|\alpha\|_\infty + \|\gamma\|_\infty) \sin \frac{\pi}{2(n+1)}.
\tag{1.31}
$$

In contrast to (1.22) and (1.25), (1.30) provides two-sided inclusions, which is significant for the convergence of the τ_1 method inclusion sets in the limit $n \to \infty$ (see Theorem 1.2). We will establish the inclusions from the right in Section 5.1 (see Corollary 5.4), but the inclusions from the left, i.e., that $\Gamma_\varepsilon^n(A) \subset \operatorname{Spec}_\varepsilon A$, for $\varepsilon \geq 0$, and $\gamma_\varepsilon^n(A) \subset \operatorname{spec}_\varepsilon A$, for $\varepsilon > 0$, are immediate consequences of the definitions and the first inequality in (1.11).

Like the τ and π method inclusion sets (1.21) and (1.24), also the τ_1 method sets (1.29) can be expressed in terms of finite submatrices. For $n \in \mathbb{N}$ and $k \in \mathbb{Z}$, $(A - \lambda I)|_{E_{n,k}}$ corresponds to the $\infty \times n$ matrix consisting of columns $k+1$ to $k+n$ of A. Let $A_{n,k}^+$ denote the matrix consisting of the $n+2$ rows of this $\infty \times n$ matrix that are non-zero, by the tridiagonal structure of A, and let I_n^+ denote the corresponding submatrix of the identity operator, so that $A_{n,k}^+$ and I_n^+ are the $(n+2) \times n$ matrices

$$
A_{n,k}^+ := \left(
\begin{array}{ccccc}
\gamma_{k+1} & 0 & \cdots & 0 & 0 \\
\hline
& & A_{n,k} & & \\
\hline
0 & 0 & \cdots & 0 & \alpha_{k+n}
\end{array}
\right),
$$

$$
I_n^+ := \left(
\begin{array}{ccc}
0 & \cdots & 0 \\
\hline
& I_n & \\
\hline
0 & \cdots & 0
\end{array}
\right), \quad \text{where } I_n := \begin{pmatrix} I_X & & \\ & \ddots & \\ & & I_X \end{pmatrix}
\tag{1.32}
$$

is an $n \times n$ identity matrix, and I_X is the identity operator on X (see Figure 1.1 for a visualisation of $A_{n,k}^+$). Then, for $n \in \mathbb{N}$, $k \in \mathbb{Z}$, and $\lambda \in \mathbb{C}$, $\nu((A - \lambda I)|_{E_{n,k}}) = \nu(A_{n,k}^+ - \lambda I_n^+)$ and $\nu((A - \lambda I)^*|_{E_{n,k}}) = \nu((A^*)_{n,k}^+ - \lambda I_n^+)$, so that

$$
\mu_n(A - \lambda I) = \inf_{k \in \mathbb{Z}} \min\left(\nu(A_{n,k}^+ - \lambda I_n^+), \nu((A^*)_{n,k}^+ - \lambda I_n^+) \right).
\tag{1.33}
$$

1.4.2. Our convergence results for the banded case. The inclusions (1.22), (1.25), and (1.30), for the τ, π, and τ_1 methods, are main results of the paper. A further key result, a corollary of the two-sided inclusion (1.30), is the convergence result, stated as Theorem 1.2 below, that the τ_1 method inclusion sets, $\Gamma_{\varepsilon+\varepsilon_n''}^n(A)$ and $\gamma_{\varepsilon+\varepsilon_n''}^n(A)$, converge to the spectral sets that they include as $n \to \infty$.

The one-sided inclusions for the τ and π methods, (1.22) and (1.25), do not imply convergence of the corresponding inclusion sets to the spectral sets that they include; indeed we will exhibit examples where this convergence is absent (see Example 1.11, Section 8.1, and Section 8.3). But these methods are convergent if they do not suffer from spectral pollution for the particular A in the sense of the following definition[6] (cf. [34]). This is a strong assumption; indeed, to say that the τ method does not suffer from spectral pollution is equivalent to saying that, for some positive null sequence η_n, $\sigma_\varepsilon^n(A) \subset \mathrm{spec}_{\varepsilon+\eta_n} A$, for $\varepsilon > 0$, and similarly for the π method, in which inclusion takes the place, for the τ/π methods, of the left-hand inclusions in (1.30), for the τ_1 method. Nevertheless, we will see below, in Section 8.2 and Section 8.4, important applications where this absence-of-spectral-pollution assumption is satisfied.

Definition 1.1 (Absence of spectral pollution). We say that the τ method (the π method for a particular $t \in \mathbb{T}$) *does not suffer from spectral pollution* for a particular tridiagonal $A \in L(E)$ if there exists a positive null sequence $(\eta_n)_{n\in\mathbb{N}}$ such that, for every $\varepsilon > 0$, $\mathrm{spec}_\varepsilon A_{n,k} \subset \mathrm{spec}_{\varepsilon+\eta_n} A$ ($\mathrm{spec}_\varepsilon A_{n,k}^{\mathrm{per},t} \subset \mathrm{spec}_{\varepsilon+\eta_n} A$) for $n \in \mathbb{N}$ and $k \in \mathbb{Z}$.

The following three theorems are our main convergence results for the τ_1, τ, and π methods. These are theorems for the case that $A \in L(E)$ is tridiagonal, but recall from Section 1.4 that every $A \in \mathrm{BO}(E)$ can be written in this tridiagonal form. The proofs of these results, in which $\xrightarrow{H}$ denotes the standard Hausdorff convergence of sets introduced in Section 1.3, are so short that we include them here.

Theorem 1.2 (Convergence of the τ_1 method). *Suppose that the matrix representation of $A \in L(E)$ is tridiagonal. Then, as $n \to \infty$,*

$$\Gamma_{\varepsilon+\varepsilon_n''}^n(A) \xrightarrow{H} \mathrm{Spec}_\varepsilon A \quad \text{for } \varepsilon \geq 0, \quad \text{in particular,} \quad \Gamma_{\varepsilon_n''}^n(A) \xrightarrow{H} \mathrm{Spec}\, A. \quad (1.34)$$

If X has Globevnik's property (see Section 2) then also

$$\gamma_{\varepsilon+\varepsilon_n''}^n(A) \xrightarrow{H} \mathrm{spec}_\varepsilon A \quad \text{for } \varepsilon > 0. \quad (1.35)$$

---

[6]The standard notion of spectral pollution, e.g., [34], is that a sequence of linear operators (A_n) approximating A *suffers from spectral pollution* if there exists a sequence $z_n \in \mathrm{Spec}\, A_n$ such that $z_n \to z$ and $z \notin \mathrm{Spec}\, A$. Spectral pollution is said to be absent if this does not hold. One might, alternatively, require absence of spectral pollution to mean that also every subsequence of (A_n) does not suffer from spectral pollution. Absence of spectral pollution in this stronger sense is equivalent to a requirement that, for some positive null sequence (η_n), $\mathrm{Spec}\, A_n \subset \mathrm{spec}_{\eta_n} A$ for each n. Our notion of absence of spectral pollution is a version of this definition that is concerned with pseudospectra not just spectra, and is uniform with respect to the parameter $k \in \mathbb{Z}$.

Proof. It follows from (1.30) that, for $\varepsilon \geq 0$,

$$\mathrm{Spec}_\varepsilon A \subset \Gamma^n_{\varepsilon+\varepsilon''_n}(A) \subset \mathrm{Spec}_{\varepsilon+\varepsilon''_n} A. \tag{1.36}$$

Since $\varepsilon''_n \to 0$ as $n \to \infty$, (1.34) follows from (2.8) and the observation at the end of the Hausdorff convergence discussion in Section 1.3. By (1.30), for $\varepsilon > 0$ the inclusions (1.36) hold also with $\mathrm{Spec}_\varepsilon$ replaced by $\mathrm{spec}_\varepsilon$ and $\Gamma^n_{\varepsilon+\varepsilon''_n}(A)$ replaced by $\gamma^n_{\varepsilon+\varepsilon''_n}(A)$. Further, if X has Globevnik's property then

$$\mathrm{spec}_{\varepsilon+\varepsilon''_n} A \xrightarrow{H} \mathrm{spec}_\varepsilon A$$

(see the comment below (2.8)), and (1.35) follows. ∎

Theorem 1.3 (Convergence of the τ method). *Suppose that the matrix representation of $A \in L(E)$ is tridiagonal, the τ method does not suffer from spectral pollution for A, and X has Globevnik's property. Then, as $n \to \infty$,*

$$\Sigma^n_{\varepsilon+\varepsilon_n}(A) \xrightarrow{H} \mathrm{Spec}_\varepsilon A \quad \textit{for } \varepsilon \geq 0, \quad \textit{in particular,} \quad \Sigma^n_{\varepsilon_n}(A) \xrightarrow{H} \mathrm{Spec}\, A, \tag{1.37}$$

and

$$\sigma^n_{\varepsilon+\varepsilon_n}(A) \xrightarrow{H} \mathrm{spec}_\varepsilon A \quad \textit{for } \varepsilon > 0. \tag{1.38}$$

Proof. It follows from (1.22) and that the τ method does not suffer from spectral pollution for A that, for some positive null sequence (η_n),

$$\mathrm{Spec}_\varepsilon A \subset \Sigma^n_{\varepsilon+\varepsilon_n}(A) \subset \sigma^n_{\varepsilon+1/n+\varepsilon_n} A \subset \mathrm{spec}_{\varepsilon+1/n+\varepsilon_n+\eta_n} A \subset \mathrm{Spec}_{\varepsilon+1/n+\varepsilon_n+\eta_n} A \tag{1.39}$$

for $\varepsilon \geq 0$, so that (1.37) follows from (2.8). Similarly, for $\varepsilon > 0$,

$$\mathrm{spec}_\varepsilon A \subset \sigma^n_{\varepsilon+\varepsilon_n}(A) \subset \mathrm{spec}_{\varepsilon+\varepsilon_n+\eta_n} A.$$

So, (1.38) follows from (2.8). ∎

Arguing similarly, but using (1.25) instead of (1.22), we have an analogous result for the π method.

Theorem 1.4 (Convergence of the π method). *Suppose that the matrix representation of $A \in L(E)$ is tridiagonal, that the π method does not suffer from spectral pollution for A for some $t \in \mathbb{T}$, and that X has Globevnik's property. Then, as $n \to \infty$,*

$$\Pi^{n,t}_{\varepsilon+\varepsilon'_n}(A) \xrightarrow{H} \mathrm{Spec}_\varepsilon A \quad \textit{for } \varepsilon \geq 0, \quad \textit{in particular,} \quad \Pi^{n,t}_{\varepsilon_n}(A) \xrightarrow{H} \mathrm{Spec}\, A, \tag{1.40}$$

and

$$\pi^{n,t}_{\varepsilon+\varepsilon_n}(A) \xrightarrow{H} \mathrm{spec}_\varepsilon A \quad \textit{for } \varepsilon > 0. \tag{1.41}$$

Remark 1.5 (Estimation of ε_n, ε_n', and ε_n''). It is clear, inspecting the proofs, that Theorems 1.2-1.4 remain valid if ε_n, ε_n', and ε_n'' are replaced in the statements of the theorems by upper bounds η_n, η_n', and η_n'', as long as these upper bounds tend to zero as $n \to \infty$. In particular, Theorem 1.3 remains true with ε_n, as defined in Theorem 3.6, replaced by the explicit upper bounds for ε_n obtained in Corollaries 3.8 and 3.9.

1.4.3. The band-dominated case. The above results for banded operators can be extended to the band-dominated case by perturbation arguments, precisely by using (2.4) below, a simple perturbation inclusion for pseudospectra. Perhaps unexpectedly, this leads not just to inclusion sets but also to convergent sequences of approximations, even for the spectrum.

To see how this works, suppose that $A \in \mathrm{BDO}(E)$, and let us focus for brevity on the τ_1 method and the case that X is finite-dimensional or a Hilbert space (we prove a closely related result for the general Banach space case in Section 6 as Theorem 6.1). Then, by definition, there exists a sequence $(A^{(n)})_{n \in \mathbb{N}} \subset \mathrm{BO}(E)$ such that $\delta_n := \|A - A^{(n)}\| \to 0$ as $n \to \infty$. Let w_n be the bandwidth of $A^{(n)}$ and let $A_n := \mathcal{I}_{w_n} A^{(n)} \mathcal{I}_{w_n}^{-1}$, where $\mathcal{I}_b$ is defined, for $b \in \mathbb{N}$, as below (1.14). Then, as discussed in Section 1.4, the matrix representation of A_n is tridiagonal. We will apply the inclusion (1.30) to A_n. Let us write ε_n'' in that inclusion as $\varepsilon_n''(A)$ to indicate explicitly its dependence on A, so that

$$\varepsilon_n''(A) = 2r(A) \sin \frac{\pi}{2(n+1)}, \tag{1.42}$$

where $r(A)$ is defined by (1.2). Then, for $\varepsilon \geq 0$ and $n \in \mathbb{N}$, using (1.30) and (2.5), we see that

$$\begin{aligned}
\mathrm{Spec}_\varepsilon A &\subset \mathrm{Spec}_{\varepsilon+\delta_n} A^{(n)} = \mathrm{Spec}_{\varepsilon+\delta_n} A_n \\
&\subset \Gamma_{\varepsilon+\delta_n+\varepsilon_n''(A_n)}^n(A_n) \subset \mathrm{Spec}_{\varepsilon+\delta_n+\varepsilon_n''(A_n)}(A_n) \\
&= \mathrm{Spec}_{\varepsilon+\delta_n+\varepsilon_n''(A_n)}(A^{(n)}) \subset \mathrm{Spec}_{\varepsilon+2\delta_n+\varepsilon_n''(A_n)}(A).
\end{aligned}$$

Now, $\varepsilon_n''(A_n) \to 0$ as $n \to \infty$, since, by (1.17), $r(A_n) \leq 2\|A_n\| = 2\|A^{(n)}\| \leq 2(\|A\| + \delta_n)$. Thus, by (2.8), $\mathrm{Spec}_{\varepsilon+2\delta_n+\varepsilon_n''(A_n)}(A) \xrightarrow{H} \mathrm{Spec}_\varepsilon A$ as $n \to \infty$, which gives immediately the following result.

Theorem 1.6. *Suppose that $A \in \mathrm{BDO}(E)$ and X is finite-dimensional or a Hilbert space, and let $(A^{(n)})_{n \in \mathbb{N}} \subset \mathrm{BO}(E)$ be such that $\delta_n := \|A - A^{(n)}\| \to 0$ as $n \to \infty$, and let $A_n := \mathcal{I}_{w_n} A^{(n)} \mathcal{I}_{w_n}^{-1}$, where w_n is the band-width of $A^{(n)}$, so that A_n is $A^{(n)}$ written in tridiagonal form. Then, for $\varepsilon \geq 0$ and $n \in \mathbb{N}$,*

$$\mathrm{Spec}_\varepsilon A \subset \Gamma_{\varepsilon+\delta_n+\varepsilon_n''(A_n)}^n(A_n) \subset \mathrm{Spec}_{\varepsilon+2\delta_n+\varepsilon_n''(A_n)}(A),$$

so that $\Gamma_{\varepsilon+\delta_n+\varepsilon_n''(A_n)}^n(A_n) \xrightarrow{H} \mathrm{Spec}_\varepsilon A$ as $n \to \infty$, in particular, $\Gamma_{\delta_n+\varepsilon_n''(A_n)}^n(A_n) \xrightarrow{H} \mathrm{Spec}\, A$.

Remark 1.7 (Construction of banded approximations and estimation of δ_n). It is easy to see that the statements in the above theorem remain true (cf. Remark 1.5) if δ_n is replaced throughout with any upper bound $\eta_n \geq \delta_n := \|A - A^{(n)}\|$, as long as $\eta_n \to 0$ as $n \to \infty$; in particular, $\mathrm{Spec}_\varepsilon A \subset \Gamma^n_{\varepsilon + \eta_n + \varepsilon_n''(A_n)}(A_n)$ and $\Gamma^n_{\varepsilon + \eta_n + \varepsilon_n''(A_n)}(A_n) \xrightarrow{H} \mathrm{Spec}_\varepsilon A$ as $n \to \infty$. But to compute the inclusion set $\Gamma^n_{\varepsilon + \eta_n + \varepsilon_n''(A_n)}(A_n)$ one needs both a concrete banded approximation $A^{(n)}$ to A and the upper bound $\eta_n \geq \delta_n$. Concretely, one can take, for some $\mathfrak{p} \geq 0$, $A^{(n)} \in \mathrm{BO}(E)$ to be the matrix with band-width n and entries $[a_{i,j}^{(n)}]$ given by

$$a_{i,j}^{(n)} := \begin{cases} \left(1 - \frac{|i-j|}{n+1}\right)^{\mathfrak{p}} a_{i,j}, & |i-j| \leq n, \\ 0, & |i-j| > n, \end{cases} \tag{1.43}$$

for $n \in \mathbb{N}$. Provided $\mathfrak{p} > 0$, it holds that $\delta_n = \|A - A^{(n)}\| \to 0$ as $n \to \infty$ (see the proof that (e)$\Rightarrow$(a) in [74, Theorem 2.1.6]). For $\mathfrak{p} = 0$, $A^{(n)}$ is the obvious approximation to A with band-width n obtained by simply discarding the matrix entries a_{ij} with $|i - j| > n$. In many cases this approximation is adequate, but it does not hold in this case that $\|A - A^{(n)}\| \to 0$ as $n \to \infty$ for all $A \in \mathrm{BDO}(E)$ (see the example in [64, Remark 1.40]).

One case where the simple choice $\mathfrak{p} = 0$ is effective is where $A \in \mathcal{W}(E)$, the so-called *Wiener algebra* (e.g., [64, Definition 1.43]). In terms of the matrix representation $[a_{i,j}]$, $A \in \mathcal{W}(A)$ precisely when there exists $\kappa = (\kappa_i)_{i \in \mathbb{Z}} \in \ell^1(\mathbb{Z}, \mathbb{R})$ such that $\|a_{i,j}\|_{L(X)} \leq \kappa_{i-j}$, for $i, j \in \mathbb{Z}$. In that case it is clear that (see [64, equation (1.27)]), defining $A^{(n)}$ by (1.43) with $\mathfrak{p} = 0$,

$$\|A - A^{(n)}\| \leq \eta_n := \sum_{|k|>n} \kappa_k, \quad n \in \mathbb{N},$$

with $\eta_n \to 0$ as $n \to \infty$.

Remark 1.8 (Proving invertibility of operators). Let us flag one significant application of Theorem 1.6 and our other convergence theorems, Theorems 1.2–1.4, when coupled with the inclusion results (1.22), (1.25), and (1.30). Suppose that $A \in \mathrm{BDO}(E)$ and that $(U_n)_{n \in \mathbb{N}} \subset \mathbb{C}^{\mathbb{C}}$ is a sequence of compact sets with the following properties: (a) $\mathrm{Spec}\, A \subset U_n$, for $n \in \mathbb{N}$; (b) $U_n \xrightarrow{H} \mathrm{Spec}\, A$ as $n \to \infty$. Then, it is easy to see that the following claim holds for any $\lambda \in \mathbb{C}$:

$$\lambda I - A \text{ is invertible} \iff \lambda \notin U_n \quad \text{for some } n \in \mathbb{N}. \tag{1.44}$$

In particular, A is invertible if and only if $0 \notin U_n$, for some $n \in \mathbb{N}$.

Of course, by (1.30) and Theorem 1.6, if X is finite-dimensional or a Hilbert space, a sequence $(U_n)_{n \in \mathbb{N}}$ with these properties, for any $A \in \mathrm{BDO}(E)$, is the sequence

$U_n = \Gamma^n_{\varepsilon''_n}(A)$, $n \in \mathbb{N}$. We will see below in Theorem 7.1 that, in the case that $X = \mathbb{C}$, another sequence with properties (a) and (b) is the sequence $U_n = \widehat{\Gamma}^n_{\text{fin}}(A)$, $n \in \mathbb{N}$. The attraction of this alternative choice for $(U_n)_{n \in \mathbb{N}}$ is that one can determine whether $\lambda \in U_n$ in only finitely many arithmetic operations (see Theorem 7.1). Thus, if $\lambda I - A$ is invertible, this can be proved in finite time, by checking whether $\lambda \in U_n$, successively for $n = 1, 2, \ldots$, stopping when a U_n is found such that $\lambda \notin U_n$. (The existence of such a U_n is guaranteed by (1.44).)

At the end of Section 8.4.1 we use (1.44), with $U_n = \Sigma^n_{\varepsilon_n}(A)$, for an operator A, the so-called Feinberg–Zee operator, for which this choice of U_n satisfies (a) and (b), to show that a particular $\lambda \notin \text{Spec } A$, proving that $\text{Spec } A$ is a strict subset of the closure of its numerical range.

1.4.4. Computability of our inclusion sets: Reduction to finitely many matrices.

A criticism of the inclusion sets that we have introduced in Section 1.4.1 from the perspective of practical computation is that to determine whether or not a particular $\lambda \in \mathbb{C}$ is in one of these inclusion sets, one has to make a computation for infinitely many finite matrices, indexed by $k \in \mathbb{Z}$.

In certain cases this infinite collection of finite matrices is in fact a finite collection. For the inclusion sets with parameter $n \geq 2$, this is the case precisely when the tridiagonal matrix A has only finitely many distinct entries. (If $n = 1$ it is enough, for the τ method, if the set $\{\beta_k : k \in \mathbb{Z}\}$ is finite.)

When A has infinitely many distinct entries, provided the set of matrix entries is a relatively compact subset of $L(X)$, which is the case, in particular, if X is finite-dimensional, we can approximate the infinite collection arbitrarily closely by a finite subset. We spell out the details in the following theorem for the τ_1 method under the assumption that X is finite-dimensional or a Hilbert space; similar results hold for the τ and π methods. The point of this theorem is that, for every $\varepsilon \geq 0$, it can be established whether or not $\lambda \in \Gamma^{n,\text{fin}}_\varepsilon(A)$ by determining whether or not the lower norms of finitely many finite matrices are $\leq \varepsilon$. Note that the definition of $\Gamma^{n,\text{fin}}_\varepsilon(A)$ in (1.47) is identical to the definition (1.29) of $\Gamma^n_\varepsilon(A)$, except that μ^n, defined by taking an infimum over $\mathbb{Z}$, is replaced by μ^{fin}_n, defined by taking a minimum over the finite set K_n. Thus, for $\varepsilon \geq 0$,

$$\Gamma^{n,\text{fin}}_\varepsilon(A) \subset \Gamma^n_\varepsilon(A), \text{ indeed also } \Gamma^n_\varepsilon(A) \subset \Gamma^{n,\text{fin}}_{\varepsilon+1/n}(A), \tag{1.45}$$

as a consequence of (1.11), if (1.46) holds. This result, indeed an extended version of this result that applies in the general band-dominated case and for every Banach space X, is proved in Section 6 as Theorem 6.1. A version of this theorem for the τ and π methods, under the assumption of absence of spectral pollution (Definition 1.1), is proved as Theorem 6.3.

Theorem 1.9. *Suppose that X is finite-dimensional or a Hilbert space, that $A \in L(E)$ is tridiagonal, and that $\{A_{ij} : i, j \in \mathbb{Z}\} \subset L(X)$ is relatively compact. Then, for every $n \in \mathbb{N}$ there exists a finite set $K_n \subset \mathbb{Z}$ such that*

$$\forall k \in \mathbb{Z} \, \exists j \in K_n \text{ such that } \|A_{n,k}^+ - A_{n,j}^+\| \leq 1/n \text{ and } \|(A^*)_{n,k}^+ - (A^*)_{n,j}^+\| \leq 1/n. \tag{1.46}$$

Further, for $\varepsilon \geq 0$ and $n \in \mathbb{N}$,

$$\Gamma_\varepsilon^{n,\text{fin}}(A) \subset \text{Spec}_\varepsilon A \subset \Gamma_{\varepsilon+\varepsilon_n''+1/n}^{n,\text{fin}}(A),$$

where

$$\begin{aligned} \Gamma_\varepsilon^{n,\text{fin}}(A) &:= \{\lambda \in \mathbb{C} : \mu_n^{\text{fin}}(A - \lambda I) \leq \varepsilon\}, \\ \mu_n^{\text{fin}}(B) &:= \min_{k \in K_n} \min\left(\nu(B|_{E_{n,k}}), \nu(B^*|_{E_{n,k}})\right) \end{aligned} \tag{1.47}$$

for $B \in L(E)$, so that

$$\mu_n^{\text{fin}}(A - \lambda I) = \min_{k \in K_n} \min\left(\nu(A_{n,k}^+ - \lambda I_n^+), \nu((A^*)_{n,k}^+ - \lambda I_n^+)\right) \quad \text{for } \lambda \in \mathbb{C}. \tag{1.48}$$

Further, for $\varepsilon \geq 0$,

$$\Gamma_{\varepsilon+\varepsilon_n''+1/n}^{n,\text{fin}}(A) \xrightarrow{H} \text{Spec}_\varepsilon A$$

as $n \to \infty$, in particular

$$\Gamma_{\varepsilon_n''+1/n}^{n,\text{fin}}(A) \xrightarrow{H} \text{Spec } A.$$

Of course, a practical issue for computation is how to determine the finite set K_n. As a concrete, nontrivial example, we demonstrate how to reduce from an infinite set to a finite set for the class of tridiagonal operators that are pseudoergodic in the sense of Davies [31] in Section 8.4.2.

1.4.5. An algorithm for computing the spectrum in the case $X = \mathbb{C}^p$ and its solvability complexity index.

In the finite-dimensional case that $X = \mathbb{C}^p$, for some $p \in \mathbb{N}$, equipped with the usual Euclidean norm, a sequence of approximations converging to Spec A, in the case when A is band-dominated, can be realised with each member of the sequence computed in finitely many operations, provided we have available as inputs to the computation sufficient information about A. We will demonstrate this for the general band-dominated case in Section 7. For the tridiagonal case, with A given by (1.15) and $X = \mathbb{C}^p$ (recall that we noted in Section 1.4 that every $A \in \text{BO}(E)$ with $X = \mathbb{C}$ can be written in this form), the τ_1 version of the algorithm proceeds as follows. We use in this definition the notation (cf. (1.58))

$$\text{Grid}(n, r) := \frac{1}{n}(\mathbb{Z} + i\mathbb{Z}) \cap r\mathbb{D}, \quad n \in \mathbb{N}, \, r > 0. \tag{1.49}$$

Let Ω_T^p denote the set of all tridiagonal matrices (1.15), with $\alpha, \beta, \gamma \in \ell^\infty(\mathbb{Z}, X)$ and $X = \mathbb{C}^p$, for some $p \in \mathbb{N}$. Where, as usual, $P(S)$ denotes the power set of a set S, the inputs we need are the following.

(1) A mapping $\mathcal{A}_p : \Omega_T^p \times \mathbb{N} \to P(\mathbb{Z})$, $(A, n) \mapsto K_n$, where $K_n \subset \mathbb{Z}$ is a finite set such that (1.46) holds.

(2) A mapping $\mathcal{B}_p : \Omega_T^p \to \mathbb{R}^3$, $A \mapsto (\alpha_{\max}, \beta_{\max}, \gamma_{\max})$, such that $\alpha_{\max} \geq \|\alpha\|_\infty$, $\beta_{\max} \geq \|\beta\|_\infty$, $\gamma_{\max} \geq \|\gamma\|_\infty$.

(3) A mapping $\mathcal{C}_p : \Omega_T^p \times \mathbb{Z} \times \mathbb{N} \to (X^{(n+2)\times n})^2$, $(A, n, k) \mapsto (A_{n,k}^+, (A^*)_{n,k}^+)$.

It is not clear to us how to construct a map $\mathcal{A}_p$ such that $\mathcal{A}_p(A, n)$ can be explicitly computed for general $A \in \Omega_T^p$. Indeed, establishing existence of a map $\mathcal{A}_p : \Omega_T^p \times \mathbb{N} \to P(\mathbb{Z})$ with the required properties may require an application of the axiom of choice. But, as a non-trivial example, we will construct in Section 8.4.2 a version of the mapping $\mathcal{A}_p$ for the important subset $\Omega_\Psi \subset \Omega_T^1$ of operators that are pseudoergodic in the sense of Davies [31].

The sequence of approximations to Spec A, each element of which can be computed in finitely many arithmetical operations, given finitely many evaluations of the above input maps, is the sequence $(\Gamma_{\mathrm{fin}}^n(A))_{n\in\mathbb{N}}$, defined using the notation (1.47) by

$$\Gamma_{\mathrm{fin}}^n(A) := \Gamma_{\varepsilon_n^*+3/n}^{n,\mathrm{fin}}(A) \cap \mathrm{Grid}(n, R), \quad n \in \mathbb{N}, \tag{1.50}$$

with (cf. (1.31))

$$\varepsilon_n^* := (\alpha_{\max} + \gamma_{\max}) \frac{22}{7(n+1)} \geq 2(\alpha_{\max} + \gamma_{\max}) \sin \frac{\pi}{2(n+1)} \geq \varepsilon_n'', \tag{1.51}$$

$(\alpha_{\max}, \beta_{\max}, \gamma_{\max}) := \mathcal{B}_p(A)$, $R := \alpha_{\max} + \beta_{\max} + \gamma_{\max}$ (so that R is an upper bound for $\|A\|$), and K_n, in the definition (1.47) and (1.48), given by $K_n := \mathcal{A}_p(A, n)$.

Proposition 1.10. *For $A \in \Omega_T^p$ and $n \in \mathbb{N}$, $\Gamma_{\mathrm{fin}}^n(A)$ can be computed in finitely many arithmetic operations, given finitely many evaluations of the functions $\mathcal{A}_p$, $\mathcal{B}_p$, and $\mathcal{C}_p$, namely, the evaluations: $K_n := \mathcal{A}_p(A, n)$; $(A_{n,k}^+, (A^*)_{n,k}^+) := \mathcal{C}_p(A, k, n)$, for $k \in K_n$; and $(\alpha_{\max}, \beta_{\max}, \gamma_{\max}) := \mathcal{B}_p(A)$. Further, $\Gamma_{\mathrm{fin}}^n(A) \xrightarrow{H} \mathrm{Spec}\, A$ as $n \to \infty$, and also*

$$\widehat{\Gamma}_{\mathrm{fin}}^n(A) := \Gamma_{\mathrm{fin}}^n(A) + \frac{2}{n}\overline{\mathbb{D}} \xrightarrow{H} \mathrm{Spec}\, A$$

as $n \to \infty$, with $\mathrm{Spec}\, A \subset \widehat{\Gamma}_{\mathrm{fin}}^n(A)$ for each $n \in \mathbb{N}$.

We give the proof of the above result, and a version for $A \in \mathrm{BDO}(E)$ in Section 7.

The above result can be interpreted as a result relating to the solvability complexity index (SCI) of [2, 25]. Let us abbreviate Ω_T^1, $\mathcal{A}_1$, $\mathcal{B}_1$, and $\mathcal{C}_1$ as Ω_T, $\mathcal{A}$, $\mathcal{B}$, and

$\mathcal{C}$, respectively, so that Ω_T is the set of tridiagonal matrices with complex number entries. With the *evaluation set* (in the sense of [2, 25])[7]

$$\Lambda := \big\{\mathcal{A}(\cdot, n), \mathcal{B}, \mathcal{C}(\cdot, k, n) : k \in \mathbb{Z}, n \in \mathbb{N}\big\}, \qquad (1.52)$$

where we equip $\mathbb{C}^C$, the set of compact subsets of $\mathbb{C}$, with the Hausdorff metric (see Section 1.3), the mappings

$$\Omega_T \to \mathbb{C}^C, \quad A \mapsto \Gamma^n_{\text{fin}}(A), \quad A \mapsto \widehat{\Gamma}^n_{\text{fin}}(A),$$

are *general algorithms* in the sense of [2, 25] for each $n \in \mathbb{N}$. Further, $\widehat{\Gamma}^n_{\text{fin}}(A)$ can be computed in finitely many arithmetic operations and specified using finitely many complex numbers (the elements of $\Gamma^n_{\text{fin}}(A)$ and the value of n). Thus, where $\Xi : \Omega_T \to \mathbb{C}^C$ is the mapping given by $\Xi(A) := \text{Spec } A$, for $A \in \Omega_T$, the *computational problem* $\{\Xi, \Omega_T, \mathbb{C}^C, \Lambda\}$ has arithmetic SCI, in the sense of [2, 25], equal to one; more precisely, since also $\text{Spec } A \subset \widehat{\Gamma}^n_{\text{fin}}(A)$, for each $n \in \mathbb{N}$ and $A \in \Omega_T$, this computational problem is in the class Π^A_1, as defined in [2, 25].

The same observations on SCI classification hold true for the corresponding computational problem with Ω_T replaced by $\Omega := \text{BDO}(E)$, that we consider in Section 7; again this has arithmetic SCI equal to one. In contrast, as we recall in Section 1.6, the computational problem of determining the spectrum of a band-dominated operator, indeed even the restricted problem of determining the spectrum of a tridiagonal matrix, has been shown in [23], adapting the arguments of [2], to have an SCI of two when the evaluation set, the set of allowed inputs to the computation, is restricted to the mappings $\Omega_T \to \mathbb{C}$, $A \mapsto a_{i,j}$, for $i, j \in \mathbb{Z}$, providing evaluation of the matrix elements.

1.5. A first example demonstrating the sharpness of our inclusion sets

We will explore the utility and properties of these new families of inclusion sets in a range of examples in Section 8. But, to aid comprehension of the τ, π, and τ_1 inclusion sets introduced in Section 1.4.1, and demonstrate their sharpness, let us pause to illustrate them as applied to one of the simplest scalar examples, an example with $X = \mathbb{C}$ so that the matrix entries are just complex numbers. We will return to this example in Section 8.1.

---

[7]Strictly, to fit with the definition in [25, Section 2.1], each element of the evaluation set Λ should be a complex-valued function on Ω_T. But it is easy to express our evaluation set in this form, expressing each of our functions in terms of a finite number of complex-valued functions. In particular, where $M := |K_n|$, the set of integers $K_n = \{k_1, \ldots, k_M\}$, with $k_1 < k_2 < \cdots < k_M$, which is the output of the mapping $\mathcal{A}(\cdot, n)$ applied to A, might be encoded as the complex number $x + iy$, with $x := k_1$ and $y := \Pi^M_{i=1} p_i^{k_i - k_1}$, where p_i is the ith prime number.

Example 1.11 (The shift operator). Let V denote the shift operator on $\ell^2(\mathbb{Z})$, taking $x = (x_j)_{j \in \mathbb{Z}} \in \ell^2(\mathbb{Z})$ to $y = (y_j)_{j \in \mathbb{Z}} \in \ell^2(\mathbb{Z})$ with $x_j = y_{j+1}$ for all $j \in \mathbb{Z}$. Its matrix is of the form (1.15) with $\alpha_j = 1$, $\beta_j = 0$ and $\gamma_j = 0$ for all $j \in \mathbb{Z}$. For $A := V$, the matrices (1.19) and (1.20) are given by

$$A_{n,k} = V_n := \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & \ddots & \ddots & & \\ & & 1 & 0 \end{pmatrix}_{n \times n},$$

$$A_{n,k}^{\text{per},t} = V_n^{\text{per},t} := \begin{pmatrix} 0 & & & & t \\ 1 & 0 & & & \\ & \ddots & \ddots & & \\ & & & 1 & 0 \end{pmatrix}_{n \times n}$$

(1.53)

for all $n \in \mathbb{N}$, independent of $k \in \mathbb{Z}$. The circulant matrix $V_n^{\text{per},1}$ is normal and has spectrum spec $V_n^{\text{per},1} = \mathbb{T}_n = \{z \in \mathbb{C} : z^n = 1\}$ (e.g., [85, Theorem 7.1]). For $t \in \mathbb{T}$, where $\alpha = t^{1/n}$ denotes any one of the nth roots of t, and $D_n := \text{diag}(\alpha^1, \ldots, \alpha^n)$, it holds that

$$D_n^{-1}(V_n^{\text{per},t} - \lambda I_n)D_n = \alpha(V_n^{\text{per},1} - \lambda \bar{\alpha} I_n)$$

so that $V_n^{\text{per},t}$ is also normal and

$$\text{Spec } V_n^{\text{per},t} = t^{1/n} \text{Spec } V_n^{\text{per},1} = t^{1/n} \mathbb{T}_n. \tag{1.54}$$

(Note that the set $t^{1/n} \mathbb{T}_n$ is independent of which nth root of t we select.) Thus, for each $t \in \mathbb{T}$, the spectrum of V, $\text{Spec } V = \mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$, is well approximated as $n \to \infty$ by $\text{Spec } V_n^{\text{per},t}$, but clearly not by $\text{Spec } V_n = \{0\}$.

The τ method and π method inclusion sets for $\text{Spec } V = \text{Spec}_0 V$, given by (1.22) and (1.25), respectively, reduce in this case to the sets

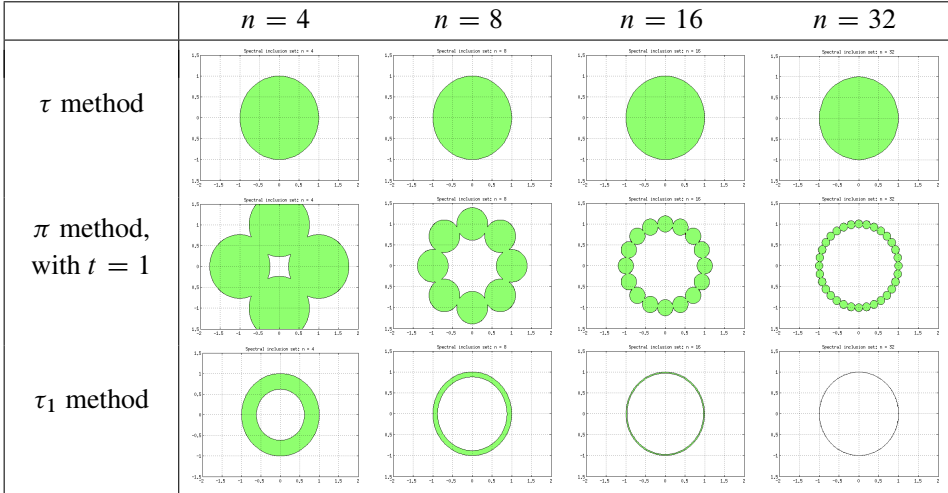$$\Sigma_{\varepsilon_n}^n (V) = \text{Spec}_{\varepsilon_n}(V_n)$$

and

$$\Pi_{\varepsilon_n'}^{n,t}(V) = \text{Spec}_{\varepsilon_n'}(V_n^{\text{per},t}), \quad t \in \mathbb{T}.$$

Since $\|\alpha\|_\infty = 1$ and $\|\gamma\|_\infty = 0$ for this example, we have by Corollary 3.7 and (1.26) that

$$\varepsilon_n = 2\sin\left(\frac{\pi}{4n+2}\right) \quad \text{and} \quad \varepsilon_n' = 2\sin\left(\frac{\pi}{2n}\right).$$

$\text{Spec}_{\varepsilon_n}(V_n)$ and $\text{Spec}_{\varepsilon_n'}(V_n^{\text{per},t})$ are neighbourhoods of $\text{Spec } V_n = \{0\}$ and $\text{Spec } V^{\text{per}_n,t} = t^{1/n}\mathbb{T}_n$, respectively. By (1.22) and (1.25) these neighbourhoods must be large enough

	$n = 4$	$n = 8$	$n = 16$	$n = 32$
τ method				
π method, with $t = 1$				
τ_1 method				

Figure 1.2. The τ, π, and τ_1 inclusion sets, $\Sigma_{\varepsilon_n}^n(V)$, $\Pi_{\varepsilon_n'}^{n,1}(V)$, and $\Gamma_{\varepsilon_n''}^n(V)$, respectively, for $n = 4, 8, 16$, and 32, each an inclusion set for Spec $V = \mathbb{T}$.

to cover Spec $V = \mathbb{T}$. As we will see in Section 8.1, for each $n \in \mathbb{N}$, $\Sigma_{\varepsilon_n}^n(V)$ is precisely the closed unit disc, while (we recall this and other properties of pseudospectra below in Section 2), since $V_n^{\mathrm{per},t}$ is normal, $\mathrm{Spec}_{\varepsilon_n'}(V_n^{\mathrm{per},t}) = \Pi_{\varepsilon_n'}^{n,t}(V)$ is the closed ε_n' neighbourhood of $t^{1/n}\mathbb{T}_n$. Further, as we will see in Section 8.1, $\Gamma_{\varepsilon_n''}^n(V)$, the τ_1 method inclusion set given by (1.25), is the closed annulus with outer radius 1, inner radius $1 - (\varepsilon_n'')^2$, where

$$\varepsilon_n'' = 2\sin\left(\frac{\pi}{2(n+1)}\right).$$

Of course any intersection of these inclusion sets is also an inclusion set. In particular, see Section 8.1,

$$\bigcap_{t \in \mathbb{T}} \Pi_{\varepsilon_n'}^{n,t}(V) = \bigcap_{t \in \mathbb{T}} (t^{1/n}\mathbb{T}_n + \varepsilon_n'\overline{\mathbb{D}})$$

is also a closed annulus, with outer radius 1, inner radius $1 - (\varepsilon_n')^2$. The τ, π, and τ_1 inclusion sets for spec $V = \mathbb{T}$ are shown for a range of values of n, selecting $t = 1$ for the π method, in Figure 1.2.

This simple example illustrates the sharpness of the inclusions (1.22), (1.25), and (1.30). As we will see in Section 8.1, the inclusion sets, which are the closed disc $\overline{\mathbb{D}}$ for the τ method, the union $t^{1/n}\mathbb{T}_n + \varepsilon_n'\overline{\mathbb{D}}$ of closed discs around shifted roots of unity for the π method, and a closed annulus for the τ_1 method, do not cover spec $V = \mathbb{T}$ if ε_n, ε_n', and ε_n'' are any smaller than the values given above.

This example also illustrates Theorem 1.2 on convergence of the τ_1 method, that $\Gamma^n_{\varepsilon''_n}(V) \xrightarrow{H} \operatorname{spec} V$ as $n \to \infty$. Since also, for each $t \in \mathbb{T}$, $\Pi^{n,t}_{\varepsilon_n}(V) \xrightarrow{H} \operatorname{spec} V$ as $n \to \infty$, this example also illustrates that the π method can be convergent. (Generalising this example, we will show in Section 8.4, by application of Theorem 1.4, that the π method is convergent for all tridiagonal $A \in L(E)$ with scalar entries that are pseudoergodic in the sense of Davies [31], this class including all tridiagonal Laurent matrices, where the diagonals α, β, and γ are constant.) Of course, this is also an example illustrating that the τ method need not be convergent, an example where the τ method suffers from spectral pollution (see Definition 1.1), so that the conditions of Theorem 1.3 are not satisfied.

1.6. Related work

Let us give more details about previous work related to this paper, on inclusions sets and approximation algorithms for the spectrum and pseudospectrum of bounded linear operators. See also the recent, overlapping review in [25].

Given a Banach space Y, and $A \in L(Y)$, a trivial but important inclusion set for $\operatorname{Spec} A$ is the ball $\|A\|\overline{\mathbb{D}} = \{z \in \mathbb{C} : |z| \le \|A\|\}$. In the case that Y is a Hilbert space, a well-known, sharper inclusion set is $\overline{\operatorname{Num} A}$, the closure of $\operatorname{Num} A$, the numerical range of A. Important generalisations of these respective inclusion sets are the higher order hull and (in the case that Y is Hilbert space) the higher order numerical range, denoted by $\operatorname{Hull}_n A$ and $\operatorname{Num}_n A$, respectively, and defined by (e.g., [33, Section 9.4])

$$\operatorname{Hull}_n A := \bigcap_{\deg(p) \le n} \{z : |p(z)| \le \|p(A)\|\},$$

$$\operatorname{Num}_n A := \bigcap_{\deg(p) \le n} \{z : p(z) \in \overline{\operatorname{Num}(p(A))}\}$$

for $n \in \mathbb{N}$. Note that $\operatorname{Num}_1 A = \overline{\operatorname{Num} A}$ and, rather surprisingly (see [33, Theorem 9.4.5] and [12]), $\operatorname{Hull}_n A = \operatorname{Num}_n A$, for each n, in the case that Y is a Hilbert space. Clearly, the higher order hulls and numerical ranges form decreasing sequences, i.e., $\operatorname{Hull}_{n+1} A \subset \operatorname{Hull}_n A$ and $\operatorname{Num}_{n+1} A \subset \operatorname{Num}_n A$, for $n \in \mathbb{N}$, so that

$$\operatorname{Hull}_n A \xrightarrow{H} \operatorname{Hull}_\infty A := \bigcap_{n \in \mathbb{N}} \operatorname{Hull}_n A \quad \text{and} \quad \operatorname{Num}_n A \xrightarrow{H} \operatorname{Num}_\infty A := \bigcap_{n \in \mathbb{N}} \operatorname{Num}_n A$$

as $n \to \infty$. (Of course, $\operatorname{Num}_\infty A$ is defined only in the Hilbert space case, in which case $\operatorname{Hull}_\infty A = \operatorname{Num}_\infty A$, since $\operatorname{Hull}_n A = \operatorname{Num}_n A$ for each n.)

These inclusion sets may be sharp or asymptotically sharp as $n \to \infty$. In particular, [33, Lemma 9.4.4], $\operatorname{Hull}_2(A) = \operatorname{Num}_2(A) = \operatorname{Spec}(A)$ in the Hilbert space if A is self-adjoint. Further [71, Theorem 2.10.3] $\operatorname{Hull}_\infty A = \widehat{\operatorname{Spec} A}$, the complement of the

unbounded component of $\mathbb{C} \setminus \operatorname{Spec} A$, so that $\operatorname{Hull}_n A \xrightarrow{H} \operatorname{Spec} A$ as $n \to \infty$ if and only if $\mathbb{C} \setminus \operatorname{Spec} A$ has only one component. But, while these sequences of inclusion sets are of significant theoretical interest and converge in many cases to $\operatorname{Spec} A$, it is unclear, for general operators A and particularly for larger n, how to realise these sets computationally. (However, see the related work of Frommer et al. [43] on computation of inclusion sets for pseudospectra via approximate computation of numerical ranges of resolvents of the operator, and work of Bögli and Marletta [6] on inclusion sets for spectra via numerical ranges of related operator pencils.)

For finite matrices, as we have noted already in Section 1.1, a standard inclusion set for the spectrum (the set of eigenvalues) is provided by Gershgorin's theorem [45, 86]. This result has been generalised, independently in Ostrowski [72], Feingold and Varga [41], and Fiedler and Pták [42] (and see [86, Chapter 6]) to cases where the entries of the $N \times N$ matrix $A = [a_{i,j}]$ are themselves submatrices[8]. A further extension to the case where each entry $a_{i,j}$ is an operator between Banach spaces has been made by Salas [76]. To make clear the connection with this paper, consider the case where each entry $a_{i,j} \in L(X)$, for some Banach space X. The main result of [76] in that case, expressed in the language of pseudospectra, is that

$$\operatorname{Spec} A \subset \bigcup_{i=1}^{N} \operatorname{Spec}_{r_i} a_{i,i}, \quad \text{where } r_i := \sum_{j=1, \, j \neq i}^{N} \|a_{i,j}\|_{L(X)}, \qquad (1.55)$$

which is close to (1.1) in the case $n = 1$, when $\varepsilon_n = r(A)$ and $r(A)$ is an upper bound for $\sup_{i \in \mathbb{Z}} r_i$. Of course, (1.55) is for finite rather than infinite matrices. Several authors have generalised the Gershgorin theorem to infinite matrices with scalar entries, but the focus has been on cases where the infinite matrix A is an unbounded operator with a discrete spectrum; see [37, 38, 83] and the references therein. A simple generalisation of the Gershgorin theorem to the case where A is a tridiagonal matrix of the form (1.15), with scalar entries, that is a bounded operator on $\ell^2(\mathbb{Z})$ is proved as [20, Theorem 2.50]. This has (1.1) for $n = 1$ as an immediate corollary.

Probably the most natural idea to approximate $\operatorname{spec}_\varepsilon A$, where $A \in \operatorname{BO}(E)$ or $\operatorname{BDO}(E)$, with $E = \ell^2(\mathbb{N})$ or $\ell^2(\mathbb{Z})$, is to use the pseudospectra, $\operatorname{spec}_\varepsilon A_n$, of the finite sections of A, i.e., $A_n := [a_{i,j}]_{i,j=1}^{n}$ or $A_n := [a_{i,j}]_{i,j=-n}^{n}$, respectively. In some cases, including Toeplitz operators [10, 75] and semi-infinite random Jacobi operators [19],

---

[8]These papers are arguably the earliest occurrence of pseudospectra in the literature (cf. [85, Section I.6]). The Gershgorin-type enclosure for the spectrum obtained in each of these papers (e.g., [41, equation (3.1)]), which is essentially the right-hand side of (1.55), is thought of in [41] as the union of N so-called Gershgorin sets G_j, $j = 1, \ldots, N$. The jth set G_j (though this language is not used), is precisely a (closed) pseudospectrum of $a_{j,j}$, the jth matrix on the diagonal.

$\mathrm{spec}_\varepsilon A_n \overset{H}{\longrightarrow} \mathrm{spec}_\varepsilon A$ as $n \to \infty$. (In some instances where the finite section method is not effective, the pseudospectra of periodised finite sections (cf. our π method) can converge to $\mathrm{spec}_\varepsilon A$; see [21] and Remark 8.5.) But in general, as noted already in Section 1.1, the sequence $\mathrm{spec}_\varepsilon A_n$ does not converge to $\mathrm{spec}_\varepsilon A$; its cluster points typically contain $\mathrm{spec}_\varepsilon A$ but also contain points outside Spec A (see, e.g., [79]).

Even in the self-adjoint case one observes this effect of so-called spectral pollution; see, e.g., [34]. As we noted in Section 1.1, Davies and Plum [34], building on Davies [29], proposed, as a method to locate eigenvalues of self-adjoint operators in spectral gaps while avoiding spectral pollution, to compute $\nu(P_m(A - \lambda I)|_{E_n})$, with $m > n$, as an approximation to $\nu(A - \lambda I)$ (which coincides with $\mu(A - \lambda I)$, defined by (1.9), in this self-adjoint setting). Here, $(P_k)_{k\in\mathbb{N}}$ is any sequence of orthogonal projections onto finite-dimensional subspaces that is strongly convergent to the identity, and $E_n := P_n(E)$ is the range of P_n. The significance of this proposal is that, in this self-adjoint setting: (a) Spec $A = \{\lambda \in \mathbb{R} : \nu(A - \lambda I) = 0\}$ (see (1.9)); (b) $\nu(P_m(A - \lambda I)|_{E_n}) \to \nu((A - \lambda I)|_{E_n})$ as $m \to \infty$; (c) $\nu((A - \lambda I)|_{E_n}) \to \nu(A - \lambda I)$ as $n \to \infty$, uniformly for $\lambda \in \mathbb{R}$; indeed [34, Lemma 5],

$$\nu(A - \lambda I)|_{E_n} \le \nu(A - \lambda I) \le \nu((A - \lambda I)|_{E_n}) + \eta_n \qquad (1.56)$$

for $\lambda \in \mathbb{R}$, where $\eta_n \to 0$ as $n \to \infty$. To relate this to our results above, suppose that $E = \ell^2(\mathbb{Z}, X)$ and $P_n := P_{2n+1,-n-1}$, for $n \in \mathbb{N}$, where $P_{n,k}$ is defined by (1.18). Then, if $A \in L(E)$ is tridiagonal, and using the notations of Section 1.1,

$$\nu(P_m(A - \lambda I)|_{E_n}) = \nu((A - \lambda I)|_{E_n}) = \nu((A - \lambda I)_n^+)$$

for $m \ge n + 1$, so that the second inequality of (1.56) bears a resemblance to Proposition 5.1 (except that, importantly, (1.56) can give no information on the dependence of η_n on A).

These ideas are moved to the non-self-adjoint case in Hansen [55, 56] (see also [2, 25, 26, 61]), where an approximation $\mathcal{S}_\varepsilon^{n,m}(A)$, for $n, m \in \mathbb{N}$, $m \ge n$, to $\mathrm{Spec}_\varepsilon A$ is proposed. This approximation, written in terms of lower norms using (7.1), to match the work of Davies and Plum [34], is given by

$$\mathcal{S}_\varepsilon^{n,m}(A) := \{\lambda \in \mathbb{C} : \min(\nu(P_m(A - \lambda I)|_{E_n}), \nu(P_m(A - \lambda I)^*|_{E_n})) \le \varepsilon\}, \quad (1.57)$$

where $E_n = P_n(E)$, E is some separable Hilbert space, and $(P_n)_{n\in\mathbb{N}}$ is a sequence of projection operators converging strongly to the identity. Using (7.1), in the case we consider where $E = \ell^2(\mathbb{Z}, X)$, and if we define $P_n := P_{2n+1,-n-1}$, for $n \in \mathbb{N}$, and if $A \in L(E)$ is tridiagonal, then, if $m \ge n + 1$, $\mathcal{S}_\varepsilon^{n,m}(A)$ is the specific approximation (1.5), which we have already contrasted with the new approximations in this paper in Section 1.1. Key features of the approximation (1.57) from the perspectives of the

current work are: (i) that one can determine whether $\lambda \in \mathcal{S}_{\varepsilon}^{n,m}(A)$ in finitely many arithmetic operations (via the characterisations (7.3) and (7.4) in terms of positive definiteness and singular values of rectangular matrices, see [56]); and that: (ii) if A band-dominated and $m = 2n$, then $\mathcal{S}_{\varepsilon}^{n,m}(A) \xrightarrow{H} \operatorname{spec}_{\varepsilon} A$ as $n \to \infty$ (see [2]).

Hansen [55,56] extends (1.57) to provide, in the Hilbert space context, a sequence of approximations, based on computations with finite rectangular matrices, also to the so-called (N, ε)-pseudospectra, $\operatorname{spec}_{N,\varepsilon} A$, defined in [56, Definition 1.2]. These generalise the pseudospectrum (note that $\operatorname{spec}_{0,\varepsilon} A = \operatorname{spec}_{\varepsilon} A$), and are another sequence of inclusions sets: one has that $\operatorname{Spec} A + \varepsilon \mathbb{D} \subset \operatorname{spec}_{N,\varepsilon} A$, for $\varepsilon > 0$ and $N \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$; further, for every $\delta > \varepsilon$, $\operatorname{spec}_{N,\varepsilon} \subset \operatorname{Spec} A + \delta \mathbb{D}$, for all sufficiently large N. These approximation ideas are extended to the general Banach space case in Seidel [78].

The Solvability Complexity Index (SCI), that we discuss in Section 1.4.5, Section 7, and Section 8.4.2, is introduced in a specific context in Hansen [56], and in the general form that we use it in this paper in [2, 3]. Building on the work in [56], Ben-Artzi et al. [2] consider the computation of $\operatorname{Spec}_{\varepsilon} A$, and hence, via (1.8), $\operatorname{Spec} A$, in the case when $A \in L(E)$ and E is a separable Hilbert space. Via an orthonormal basis $(e_i)_{i \in \mathbb{N}}$ for E, we may identify E with $\ell^2(\mathbb{N})$ and A with its matrix representation $A = [a_{i,j}]_{i,j \in \mathbb{N}}$, where $a_{i,j} := (Ae_j, e_i)$. On the assumption that only the values $a_{i,j}$ are available as data, Ben-Artzi et al. [2], building on [56], discuss the existence or otherwise of algorithms for computing $\operatorname{Spec}_{\varepsilon} A$. In particular, the authors construct in [2] for each $\varepsilon > 0$ the sequence of mappings $(\Gamma_{\varepsilon,n})_{n \in \mathbb{N}}$ from $L(E)$ to $\mathbb{C}^C$ given by

$$\Gamma_{\varepsilon,n}(A) := \mathcal{S}_{\varepsilon}^{n,2n}(A) \cap \operatorname{Grid}(n), \quad \text{where } \operatorname{Grid}(n) := \frac{1}{n}(\mathbb{Z} + i\mathbb{Z}). \tag{1.58}$$

It is clear, from the above discussion, that $\Gamma_{\varepsilon,n}(A)$ can be computed for each n from finitely many of the entries $a_{i,j}$ in finitely many operations, and it is not difficult, as a consequence of the properties of $\mathcal{S}_{\varepsilon}^{n,2n}(A)$ noted above, to see (cf. Proposition 1.10) that $\Gamma_{\varepsilon,n}(A) \xrightarrow{H} \operatorname{Spec}_{\varepsilon} A$ as $n \to \infty$ whenever the matrix A is band-dominated. Thus, where $(\varepsilon_m)_{m \in \mathbb{N}}$ is any positive null sequence, $\operatorname{Spec} A$ can be computed as the iterated limit

$$\operatorname{Spec} A = \lim_{m \to \infty} \operatorname{Spec}_{\varepsilon_m} A = \lim_{m \to \infty} \lim_{n \to \infty} \Gamma_{\varepsilon_m,n}(A). \tag{1.59}$$

In the same paper [2] (and see [23]) they show moreover, that this double limit is optimal; it cannot be replaced by a single limit. Precisely, they show that there exists no sequence of algorithms $(\gamma_n)_{n \in \mathbb{N}}$, from $L(E)$ to $\mathbb{C}^C$, such that $\gamma_n(A) \xrightarrow{H} \operatorname{Spec}(A)$ for every band-dominated matrix A and $\gamma_n(A)$ can be computed, for each n, with only finitely many of the matrix entries $a_{i,j}$ as input. In the language introduced in [2], the SCI of computing the spectrum of a band-dominated operator from its matrix entries is two.

2. Pseudospectra and Globevnik's property

We have introduced in Section 1.3 definitions of the open and closed ε-pseudospectra of $B \in L(Y)$, for a Banach space Y. A natural question is: what is the relationship between these two definitions? Clearly, for every $\varepsilon > 0$, $\mathrm{spec}_\varepsilon B \subset \mathrm{Spec}_\varepsilon B$, so that also $\overline{\mathrm{spec}_\varepsilon B} \subset \mathrm{Spec}_\varepsilon B$ (since $\mathrm{Spec}_\varepsilon B$ is closed). It is known [46,47,80] that equality holds, i.e., $\mathrm{Spec}_\varepsilon B = \overline{\mathrm{spec}_\varepsilon B}$, if Y has *Globevnik's property*, that is: Y is finite-dimensional or Y or its dual Y^* is a complex uniformly convex Banach space (see [80, Definition 2.4 (ii)]). Hilbert spaces have Globevnik's property but also $E = \ell^2(\mathbb{Z}, X)$ has Globevnik's property if X has it [36]. On the other hand, Banach spaces Y and operators $B \in L(Y)$ are known [80–82] such that $\overline{\mathrm{spec}_\varepsilon B} \subsetneq \mathrm{Spec}_\varepsilon B$.

Let us note a few further properties of pseudospectra that we will utilise. For a Banach space Y, $B \in L(Y)$, and $\varepsilon > 0$, let

$$\mathrm{spec}_\varepsilon^0(B) := \big\{ \lambda \in \mathbb{C} : \text{ there exists } x \in Y \text{ with } \|(A - \lambda I)x\| < \varepsilon \|x\| \big\}, \qquad (2.1)$$

and let $\mathrm{Spec}_\varepsilon^0 B$ denote the right-hand side of the above equation with the $<$ replaced by $\leq$. Then (this is [85, p. 31] plus (1.9)), we have the following useful characterisations of $\mathrm{spec}_\varepsilon B$:

$$\mathrm{spec}_\varepsilon B = \bigcup_{\|T\| < \varepsilon} \mathrm{Spec}\,(B + T) = \mathrm{Spec}\,B \cup \mathrm{spec}_\varepsilon^0 B = \mathrm{spec}_\varepsilon^0(B) \cup \mathrm{spec}_\varepsilon^0(B^*), \quad \varepsilon > 0. \tag{2.2}$$

For $\mathrm{Spec}_\varepsilon B$ we have (see [81] and (1.9)) the weaker statement that

$$\mathrm{Spec}_\varepsilon B \supset \bigcup_{\|T\| \leq \varepsilon} \mathrm{Spec}\,(B + T) \supset \mathrm{Spec}\,B \cup \mathrm{Spec}_\varepsilon^0 B = \mathrm{Spec}_\varepsilon^0(B) \cup \mathrm{Spec}_\varepsilon^0(B^*), \quad \varepsilon > 0. \tag{2.3}$$

It follows from (2.2) and (2.3) that, for $\varepsilon > 0$, $\mathrm{Spec}\,B + \varepsilon\mathbb{D} \subset \mathrm{spec}_\varepsilon B$ and $\mathrm{Spec}\,B + \varepsilon\overline{\mathbb{D}} \subset \mathrm{Spec}_\varepsilon B$. We note that equality holds in these inclusions (see, e.g., [33, p. 247]) if Y is a Hilbert space and B is normal. Further, it follows from the first equality in (2.2) that, for $B, T \in L(Y)$ and $\varepsilon, \delta > 0$,

$$\mathrm{spec}_\varepsilon(B + T) \subset \mathrm{spec}_{\varepsilon + \delta} B \quad \text{if } \|T\| \leq \delta. \tag{2.4}$$

If Y is finite-dimensional or is a Hilbert space the first $\supset$ in (2.3) can be replaced by $=$ (see [51, Theorem 3.27] and [81]), so that we have also that

$$\mathrm{Spec}_\varepsilon(B + T) \subset \mathrm{Spec}_{\varepsilon + \delta} B \quad \text{if } \|T\| \leq \delta. \tag{2.5}$$

But the first $\supset$ in (2.3) cannot be replaced by $=$ for every Banach space Y and every $B \in L(Y)$, even if Y has Globevnik's property, as shown by examples of Shargorodsky [81]. The second $\supset$ in (2.3) can be replaced by equality if Y is finite-dimensional,

but not, in general, otherwise. As an example relevant to this paper let $A \in L(E)$ have the matrix representation (1.15) with $\alpha = \gamma = 0$ and $\beta_k = \tanh(k)$, $k \in \mathbb{Z}$. Then, $E = \ell^2(\mathbb{Z})$ is a Hilbert space so that the first $\supset$ in (2.3) is equality. However, Spec $A = \overline{\{\beta_k : k \in \mathbb{Z}\}}$, so that $\pm 1 \in$ Spec $A \subset [-1, 1]$, which implies that $\lambda = 1 + \varepsilon \in \operatorname{Spec}_\varepsilon A \setminus \operatorname{Spec} A$. But this $\lambda \notin \operatorname{Spec}_\varepsilon^0 A$, for if $x \in E$ with $x \neq 0$ then

$$\|(A - \lambda I)x\|^2 = \sum_{k \in \mathbb{Z}} (1 + \varepsilon - \beta_k)^2 |x_k|^2 > \varepsilon^2 \|x\|, \tag{2.6}$$

since $\beta_k < 1$ for all $k \in \mathbb{Z}$.

We have noted in Section 1.3 the simple identity (1.8), and the same identity with $\operatorname{Spec}_\varepsilon$ replaced by $\operatorname{spec}_\varepsilon$. Similarly, immediately from the definitions, for every Banach space Y and $B \in L(Y)$, generalising (1.8), we have that

$$\operatorname{Spec}_\varepsilon B = \bigcap_{\varepsilon' > \varepsilon} \operatorname{Spec}_{\varepsilon'} B = \bigcap_{\varepsilon' > \varepsilon} \operatorname{spec}_{\varepsilon'} B, \quad \varepsilon \geq 0. \tag{2.7}$$

Recalling the results on Hausdorff convergence of Section 1.3, this implies, for every positive null sequence (η_n) and every $\varepsilon \geq 0$, that

$$\operatorname{Spec}_{\varepsilon + \eta_n} B \xrightarrow{H} \operatorname{Spec}_\varepsilon B \quad \text{and} \quad \operatorname{spec}_{\varepsilon + \eta_n} B \xrightarrow{H} \operatorname{Spec}_\varepsilon B. \tag{2.8}$$

Recalling the discussion of Hausdorff convergence in Section 1.3, if Y has Globevnik's property, so that $\overline{\operatorname{spec}_\varepsilon B} = \operatorname{Spec}_\varepsilon B$, then we can, if we prefer, write (2.8) with $\operatorname{Spec}_\varepsilon B$ replaced by $\operatorname{spec}_\varepsilon B$.

3. The τ method: Principal submatrices

In this section, we prove the inclusion (1.22) and compute the optimal value of the "penalty" term ε_n. The proof of (1.22) is so central to this paper that we want to sketch its main idea, captured in Proposition 3.3 below, beforehand.

Recalling (2.2), suppose that $\lambda \in \operatorname{spec}_\varepsilon A \setminus \operatorname{Spec} A = \operatorname{spec}_\varepsilon^0 A$ and $x = (x_j)_{j \in \mathbb{Z}} \in E$ is a corresponding pseudomode, i.e.,

$$\|(A - \lambda I)x\| < \varepsilon \|x\|. \tag{3.1}$$

Consider all finite subvectors of x, of some fixed length $n \in \mathbb{N}$, namely,

$$x_{n,k} := (x_{k+1}, \ldots, x_{k+n}) \in X^n, \quad k \in \mathbb{Z}, \tag{3.2}$$

and consider the corresponding principal submatrices $A_{n,k}$ of A introduced in (1.19). We claim that, for at least one $k \in \mathbb{Z}$, it holds, for some $\varepsilon_n > 0$ to be specified, that

$$\|(A_{n,k} - \lambda I_n)x_{n,k}\| < (\varepsilon + \varepsilon_n) \|x_{n,k}\|, \tag{3.3}$$

and hence, $\lambda \in \operatorname{spec}_{\varepsilon + \varepsilon_n} A_{n,k}$.

One way to show this is to suppose that, conversely, for all $k \in \mathbb{Z}$, the opposite of (3.3) holds, and then to take all these (opposite) inequalities and add their squares up over $k \in \mathbb{Z}$, which, after some computation, contradicts (3.1). Although (3.3) need not hold for every $k \in \mathbb{Z}$, it does hold "in the quadratic mean" over all $k \in \mathbb{Z}$, and therefore for at least one $k \in \mathbb{Z}$.

This computation reveals that (3.3) holds with $\varepsilon_n = O(n^{-1/2})$ as $n \to \infty$; see (3.18). However, one can do better: instead of the "sharp cutoff" (3.2), we can introduce weights $w_1, \ldots, w_n > 0$ and put

$$x_{n,k} := (w_1 x_{k+1}, \ldots, w_n x_{k+n}) \in X^n, \quad k \in \mathbb{Z}.$$

With this modification (3.3) holds for some $k \in \mathbb{Z}$ but now with ε_n dependent on the choice of $w_1, \ldots, w_n$. By varying the weight vector $w = (w_1, \ldots, w_n)$ we can minimise ε_n; the minimal

$$\varepsilon_n = O(n^{-1}) \quad \text{as } n \to \infty.$$

We structure these arguments as follows. In Section 3.1 we prove, as Theorem 3.5, a version of (1.22) with a formula for ε_n dependent on the choice of $w_1, \ldots, w_n$. The main idea of the proof, outlined above, is captured in Proposition 3.3. In Section 3.2 we minimise ε_n over the choice of the weights $w_1, \ldots, w_n$; the full proof of (1.22), establishing the minimal formula for ε_n, is given as Theorem 3.6. This formula for ε_n, while amenable to numerical computation, is complex, requiring minimisation of a function of one variable on a finite interval, which function is defined implicitly through the solution of a nonlinear equation. Corollary 3.7 states an explicit formula for ε_n in the case that A is bidiagonal, and Corollaries 3.8 and 3.9 provide more explicit upper bounds for ε_n in the general tridiagonal case. These are of value as discussed in Remark 1.5.

Before continuing with the above plan, let us note properties of the sets $\sigma_\varepsilon^n(A)$ and $\Sigma_\varepsilon^n(A)$, defined by (1.21), that appear in (1.22). Clearly, for all $n \in \mathbb{N}$ (and all tridiagonal $A \in L(E)$), $\sigma_\varepsilon^n(A)$ is open, for $\varepsilon > 0$, and $\Sigma_\varepsilon^n(A)$ is closed, for $\varepsilon \geq 0$. We have further that

$$\overline{\sigma_\varepsilon^n(A)} \subset \Sigma_\varepsilon^n(A), \quad \varepsilon > 0. \tag{3.4}$$

The following proposition is a characterisation of $\sigma_\varepsilon^n(A)$ and $\Sigma_\varepsilon^n(A)$ that is key to the proof of (1.22). This uses the notations

$$\begin{aligned}
\hat{\sigma}_\varepsilon^n(A) &:= \{\lambda \in \mathbb{C} : \mu_n^\dagger(A - \lambda I) < \varepsilon\}, \quad \varepsilon > 0, \\
\hat{\Sigma}_\varepsilon^n(A) &:= \{\lambda \in \mathbb{C} : \mu_n^\dagger(A - \lambda I) \leq \varepsilon\}, \quad \varepsilon \geq 0,
\end{aligned} \tag{3.5}$$

for $n \in \mathbb{N}$, where $\mu_n^\dagger$ is defined by (1.27). Note that (1.12) and (1.27) imply that

$$|\mu_n^\dagger(A) - \mu_n^\dagger(B)| \leq \|A - B\|,$$

so that $\hat{\sigma}_\varepsilon^n(A)$ is open and $\hat{\Sigma}_\varepsilon^n(A)$ is closed. Note also that

$$\hat{\Sigma}_\varepsilon^n(A) = \bigcap_{\varepsilon' > \varepsilon} \hat{\sigma}_{\varepsilon'}^n(A), \quad \varepsilon \geq 0. \tag{3.6}$$

Proposition 3.1. *For every tridiagonal $A \in L(E)$ and $n \in \mathbb{N}$, it holds that $\sigma_\varepsilon^n(A) = \hat{\sigma}_\varepsilon^n(A)$, for $\varepsilon > 0$. If X has the Globevnik property, then also $\overline{\sigma_\varepsilon^n(A)} = \Sigma_\varepsilon^n(A) = \hat{\Sigma}_\varepsilon^n(A)$, for $\varepsilon > 0$.*

Proof. The fact that $\sigma_\varepsilon^n(A) = \hat{\sigma}_\varepsilon^n(A)$, for $\varepsilon > 0$, follows easily from (1.10), recalling that, for every $\varepsilon \in \mathbb{R}$, the infimum of a set of real numbers is $< \varepsilon$ if and only if one of the numbers in the set is $< \varepsilon$.

Suppose that $\lambda \in \Sigma_\varepsilon^n(A)$. Then, there exists a sequence $(\lambda_j) \subset \mathbb{C}$ and a sequence $(k_j) \subset \mathbb{Z}$ such that $\lambda_j \to \lambda$ and $\lambda_j \in \mathrm{Spec}_\varepsilon(A_{n,k_j})$, so that $\mu(A_{n,k_j} - \lambda_j I_n) \leq \varepsilon$. By (1.12) it follows that

$$\mu_n^\dagger(A - \lambda I) = \inf_{k \in \mathbb{Z}} \mu(A_{n,k} - \lambda I_n) \leq \mu(A_{n,k_j} - \lambda I_n) \leq \varepsilon + |\lambda - \lambda_j|,$$

for each j, so that $\mu_n^\dagger(A - \lambda I) \leq \varepsilon$ and $\lambda \in \hat{\Sigma}_\varepsilon^n(A)$. Thus, $\Sigma_\varepsilon^n(A) \subset \hat{\Sigma}_\varepsilon^n(A)$.

Define $D \in L(E)$ by $D := \mathrm{Diag}\{A_{n,k} : k \in \mathbb{Z}\}$, and note that, for all $\lambda \in \mathbb{C}$,

$$\mu(D - \lambda I) = \inf_{k \in \mathbb{Z}} \mu(A_{n,k} - \lambda I_n) = \mu_n^\dagger(A - \lambda I), \tag{3.7}$$

so that, by (1.10), $\hat{\sigma}_\varepsilon^n(A) = \mathrm{spec}_\varepsilon D$ and $\hat{\Sigma}_\varepsilon^n(A) = \mathrm{Spec}_\varepsilon D$. Noting also (3.4), we have shown that $\overline{\mathrm{spec}_\varepsilon D} = \overline{\sigma_\varepsilon^n(A)} \subset \Sigma_\varepsilon^n(A) \subset \hat{\Sigma}_\varepsilon^n(A) = \mathrm{Spec}_\varepsilon D$. But if X has the Globevnik property, then $\overline{\mathrm{spec}_\varepsilon D} = \mathrm{Spec}_\varepsilon D$, completing the proof. ∎

3.1. Proof of the inclusion (1.22)

We start our proof of (1.22) with a very simple lemma that proves helpful throughout.

Lemma 3.2. *For $a, b \in \mathbb{R}$ and $\theta > 0$ we have $(a + b)^2 \leq a^2(1 + \theta) + b^2(1 + \theta^{-1})$, with equality if and only if $a\theta = b$.*

Proof. Clearly, $(a + b)^2 \leq (a + b)^2 + (a\theta^{\frac{1}{2}} - b\theta^{-\frac{1}{2}})^2 = a^2(1 + \theta) + b^2(1 + \theta^{-1})$ with equality if and only if $a\theta^{\frac{1}{2}} - b\theta^{-\frac{1}{2}} = 0$, i.e., $a\theta = b$. ∎

The following results lead to Theorem 3.5 that provides justification of (1.22), initially with a formula for ε_n that depends on general weights $w_1, \ldots, w_n$.

Proposition 3.3. *Let $\varepsilon > 0$ and $n \in \mathbb{N}$, suppose that $w_j \in \mathbb{R}$, for $j = 1, \ldots, n$, with at least one w_j non-zero, and that $A \in L(E)$ is tridiagonal and $\|Ax\| \leq \varepsilon$, for some $x \in E$ with $\|x\| = 1$. Then, for some $k \in \mathbb{Z}$ it holds that*

$$\nu(A_{n,k}) \leq \varepsilon + \varepsilon_n,$$

with

$$\varepsilon_n := \|\alpha\|_\infty \sqrt{\frac{T_n^-}{S_n}} + \|\gamma\|_\infty \sqrt{\frac{T_n^+}{S_n}}, \tag{3.8}$$

where

$$S_n := \sum_{j=1}^n w_j^2, \quad T_n^- := \sum_{j=1}^n (w_{j-1} - w_j)^2, \quad T_n^+ := \sum_{j=1}^n (w_{j+1} - w_j)^2,$$

with $w_0 := 0$ and $w_{n+1} := 0$.

Proof. Let $y = Ax$, so $\|y\| \le \varepsilon$. For $j = 0, \dots, n$, let

$$E_j := |w_{j+1} - w_j|. \tag{3.9}$$

Note that

$$\sum_{j=1}^n E_j^2 = T_n^+ \quad \text{and} \quad \sum_{j=1}^n E_{j-1}^2 = T_n^-. \tag{3.10}$$

For $k \in \mathbb{Z}$, let $a_k := \|A_{n,k}\tilde{x}_{n,k}\|$, where $\tilde{x}_{n,k} = (w_1 x_{k+1}, w_2 x_{k+2}, \dots, w_n x_{k+n})^T$, and let $b_k := \|\tilde{x}_{n,k}\|$. We will prove that $a_k \le (\varepsilon + \varepsilon_n) b_k$ for some $k \in \mathbb{Z}$ with $b_k \ne 0$, which will show that $\nu(A_{n,k}) \le \varepsilon + \varepsilon_n$.

Note first that, using the notation (3.9),

$$a_k^2 = \sum_{j=1}^n \|w_{j-1}\alpha_{j+k-1}x_{j+k-1} + w_j\beta_{j+k}x_{j+k} + w_{j+1}\gamma_{j+k+1}x_{j+k+1}\|_X^2 \tag{3.11}$$

$$= \sum_{j=1}^n \|(w_{j-1} - w_j)\alpha_{j+k-1}x_{j+k-1} + w_j y_{j+k} + (w_{j+1} - w_j)\gamma_{j+k+1}x_{j+k+1}\|_X^2$$

$$\le \sum_{j=1}^n (w_j\|y_{j+k}\|_X + E_{j-1}\|\alpha\|_\infty\|x_{j+k-1}\|_X + E_j\|\gamma\|_\infty\|x_{j+k+1}\|_X)^2. \tag{3.12}$$

So, for all $\theta > 0$ and $\phi > 0$, by Lemma 3.2,

$$a_k^2 \le \sum_{j=1}^n \Big[(1 + \theta)w_j^2\|y_{j+k}\|_X^2$$

$$+ (1 + \theta^{-1})\big(E_{j-1}\|\alpha\|_\infty\|x_{j+k-1}\|_X + E_j\|\gamma\|_\infty\|x_{j+k+1}\|_X\big)^2 \Big]$$

$$\le \sum_{j=1}^n \Big[(1 + \theta)w_j^2\|y_{j+k}\|_X^2 + (1 + \theta^{-1})$$

$$\cdot \big((1 + \phi)E_{j-1}^2\|\alpha\|_\infty^2\|x_{j+k-1}\|_X^2 + (1 + \phi^{-1})E_j^2\|\gamma\|_\infty^2\|x_{j+k+1}\|_X^2\big) \Big].$$

$$\tag{3.13}$$

Thus, and recalling (3.10) and that $\|x\| = 1$,

$$\sum_{k \in \mathbb{Z}} a_k^2 \le \sum_{j=1}^{n} \Big[(1 + \theta) \|y\|^2 w_j^2 + (1 + \theta^{-1})$$

$$\cdot ((1 + \phi) E_{j-1}^2 \|\alpha\|_\infty^2 + (1 + \phi^{-1}) E_j^2 \|\gamma\|_\infty^2) \Big]$$

$$= (1 + \theta) \|y\|^2 S_n + (1 + \theta^{-1})[(1 + \phi) \|\alpha\|_\infty^2 T_n^- + (1 + \phi^{-1}) \|\gamma\|_\infty^2 T_n^+].$$

Similarly, $\sum_{k \in \mathbb{Z}} b_k^2 = S_n \|x\|^2 = S_n$. Now, by Lemma 3.2 and recalling (3.8),

$$\inf_{\phi > 0} [(1 + \phi) \|\alpha\|_\infty^2 T_n^- + (1 + \phi^{-1}) \|\gamma\|_\infty^2 T_n^+] \le \varepsilon_n^2 S_n.$$

Thus,

$$\sum_{k \in \mathbb{Z}} a_k^2 \le \big((1 + \theta) \|y\|^2 + (1 + \theta^{-1}) \varepsilon_n^2 \big) \sum_{k \in \mathbb{Z}} b_k^2, \tag{3.14}$$

for all $\theta > 0$. Applying Lemma 3.2 again, we see that

$$\inf_{\theta > 0} \big((1 + \theta) \|y\|^2 + (1 + \theta^{-1}) \varepsilon_n^2 \big) = (\|y\| + \varepsilon_n)^2 \le (\varepsilon + \varepsilon_n)^2, \tag{3.15}$$

so that

$$\sum_{k \in \mathbb{Z}} a_k^2 \le (\varepsilon + \varepsilon_n)^2 \sum_{k \in \mathbb{Z}} b_k^2.$$

Thus, either $a_k > (\varepsilon + \varepsilon_n) b_k$, for some $k \in \mathbb{Z}$, in which case also $a_k < (\varepsilon + \varepsilon_n) b_k$, for some $k \in \mathbb{Z}$, which implies that $b_k > 0$, or $a_k \le (\varepsilon + \varepsilon_n) b_k$, for all $k \in \mathbb{Z}$, so for some k with $b_k \ne 0$. ∎

Corollary 3.4. *Let $n \in \mathbb{N}$, suppose that $w_j \in \mathbb{R}$, for $j = 1, \ldots, n$, with at least one w_j non-zero, and that $A \in L(E)$ is tridiagonal. Then,*

$$\inf_{k \in \mathbb{Z}} \nu(A_{n,k}) \le \nu(A) + \varepsilon_n, \tag{3.16}$$

where ε_n is given by (3.8).

Proof. Let $\eta > 0$. By definition of $\nu(A)$ there exists $x \in E$ with $\|x\| = 1$ such that $\|Ax\| \le \nu(A) + \eta$. By Proposition 3.3, $\nu(A_{n,k}) \le \nu(A) + \varepsilon_n + \eta$, for some $k \in \mathbb{Z}$. Since this holds for all $\eta > 0$ the result follows. ∎

It is not true, for all tridiagonal $A \in L(E)$, that $\nu(A_{n,k}) \le \nu(A) + \varepsilon_n$, for some $k \in \mathbb{Z}$ (a strengthened version of (3.16)). For consider $B = A - \lambda I$, where A is as defined below (2.5), so that $X = \mathbb{C}$, and choose $\lambda = 1 + \varepsilon$, for some $\varepsilon \ge 0$. For $x = (x_k)_{k \in \mathbb{Z}} \in E = \ell^2(\mathbb{Z})$, $(Bx)_k = (\beta_k - 1 - \varepsilon) x_k$, $k \in \mathbb{Z}$, with $\beta_k = \tanh(k)$, so that $\nu(B) \le 1 + \varepsilon - \beta_k$, for all $k \in \mathbb{Z}$, i.e., $\nu(B) \le \varepsilon$. But (cf. (2.6)), for each $k \in \mathbb{Z}$, $\nu(B_{n,k}) \ge 1 + \varepsilon - \beta_{k+n} > \varepsilon = \varepsilon + \varepsilon_n$, since $\varepsilon_n = 0$ as B is diagonal.

Theorem 3.5. *Suppose that $n \in \mathbb{N}$ and $w_j \in \mathbb{R}$, for $j = 1, \ldots, n$, with at least one w_j non-zero. Then,*

$$\mathrm{Spec}_\varepsilon A \subset \hat{\Sigma}^n_{\varepsilon + \varepsilon_n}(A) \quad \varepsilon \geq 0 \quad and \quad \mathrm{spec}_\varepsilon A \subset \sigma^n_{\varepsilon + \varepsilon_n}(A), \quad \varepsilon > 0, \qquad (3.17)$$

with ε_n given by (3.8), in particular, the second inclusion of (1.22) holds with ε_n given by (3.8). If X has Globevnik's property, then also the first inclusion of (1.22) holds with ε_n given by (3.8).

Proof. To see that the second inclusion in (3.17)/(1.22) holds, suppose $\lambda \in \mathrm{spec}_\varepsilon A$. Then, by (1.10) and the definition (1.9), either $\nu(A - \lambda I) < \varepsilon$ or $\nu(A^* - \lambda I) < \varepsilon$. By Corollary 3.4 it follows that either $\nu(A_{n,k} - \lambda I_n) < \varepsilon + \varepsilon_n$, for some $k \in \mathbb{Z}$, or $\nu(A^*_{n,k} - \lambda I_n) < \varepsilon + \varepsilon_n$, for some $k \in \mathbb{Z}$. Thus, recalling the notation (1.9), $\mu(A_{n,k} - \lambda I_n) < \varepsilon + \varepsilon_n$, for some $k \in \mathbb{Z}$, so that $\lambda \in \mathrm{spec}_{\varepsilon + \varepsilon_n} A_{n,k}$ by (1.10), so that $\lambda \in \sigma^n_{\varepsilon + \varepsilon_n}(A)$. The first inclusion in (3.17) follows from the second inclusion by taking intersections, noting (2.7), the first sentence of Proposition 3.1, and (3.6). This implies, by the second sentence of Proposition 3.1, that the first inclusion of (1.22) holds if X has the Globevnik property. ∎

If we evaluate (3.8) in the case when all weights $w_1, \ldots, w_n$ are equal to 1, we get that

$$S_n = n \quad and \quad T_n^- = T_n^+ = 1,$$

so that

$$\varepsilon_n = \frac{1}{\sqrt{n}}(\|\alpha\|_\infty + \|\gamma\|_\infty), \qquad (3.18)$$

and hence, $\varepsilon_n \to 0$ as $n \to \infty$. In the next subsection we will see that we can achieve much smaller values for ε_n, smaller by a factor $< 2\pi/\sqrt{n}$ for large n (Corollary 3.8), by minimising ε_n as a function of the weight vector w.

3.2. Minimising ε_n via variation of the weight vector

We will now minimise the "penalty term" ε_n as a function, (3.8), of the weight vector $w = (w_1, \ldots, w_n)^T \in \mathbb{R}^n$. That is, we will compute

$$\inf_{w \in \mathbb{R}^n, \, w \neq 0} \varepsilon_n = \inf_{w \in \mathbb{R}^n, \, \|w\| = 1} \varepsilon_n,$$

the equality of these two expressions following easily from (3.8). It follows from the compactness of the set $\{w \in \mathbb{R}^n : \|w\| = 1\}$ that this infimum is achieved by some weight vector $w = (w_1, \ldots, w_n)^T \in \mathbb{R}^n$, so that the infimum is a minimum. This is important because it implies that Theorem 3.5 applies with this minimal weight vector and ε_n taking its minimal value.

Let us abbreviate $\|\alpha\|_\infty =: r$ and $\|\gamma\|_\infty =: s$, so that

$$\varepsilon_n^2 = \left(r\sqrt{\frac{T_n^-}{S_n}} + s\sqrt{\frac{T_n^+}{S_n}} \right)^2 = \inf_{\tau>0} \left(r^2(1+\tau)\frac{T_n^-}{S_n} + s^2(1+\tau^{-1})\frac{T_n^+}{S_n} \right), \quad (3.19)$$

by (3.8) and Lemma 3.2. (This step (3.19) may appear to be introducing additional complexity, but we shall see that it helpfully reduces a large part of the minimisation to an eigenvalue computation.) Now, put

$$B_n := \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

Then,

$$T_n^- = \|B_n w\|^2 = (B_n w)^T (B_n w) = w^T B_n^T B_n w,$$

while

$$T_n^+ = \|B_n^T w\|^2 = (B_n^T w)^T (B_n^T w) = w^T B_n B_n^T w$$

and $S_n = \|w\|^2 = w^T w$. So,

$$r^2(1+\tau)\frac{T_n^-}{S_n} + s^2(1+\tau^{-1})\frac{T_n^+}{S_n} = \frac{w^T D_n w}{w^T w} \qquad (3.20)$$

with $D_n := r^2(1+\tau)B_n^T B_n + s^2(1+\tau^{-1})B_n B_n^T$. It is easy to see that D_n is symmetric, real and positive definite. In fact, $w^T D_n w \geq [r^2(1+\tau) + s^2(1+\tau^{-1})]\lambda_1$, where $\lambda_1 > 0$ is the smallest eigenvalue of $B_n^T B_n$ and $B_n B_n^T$. Our aim is to compute the sharpest lower bound on (3.20), i.e., the smallest eigenvalue of D_n, since

$$\inf_{w\in\mathbb{R}^n, w\neq 0} \varepsilon_n^2 = \inf_{\tau>0} \inf_{w\in\mathbb{R}^n, w\neq 0} \frac{w^T D_n w}{w^T w}. \qquad (3.21)$$

To compute the infimum (3.21), define $E_n(\phi)$, for $0 \leq \phi \leq 1$, to be the real, symmetric, positive definite matrix $E_n(\phi) := \phi B_n^T B_n + (1-\phi) B_n B_n^T$, so that $E_1(\phi) = 1$ while, for $n = 2, 3, \ldots$,

$$E_n(\phi) = \begin{pmatrix} 1+\phi & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2-\phi \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

Noting that

$$D_n = (r^2(1 + \tau) + s^2(1 + \tau^{-1}))E_n\left(\frac{r^2(1 + \tau)}{r^2(1 + \tau) + s^2(1 + \tau^{-1})}\right),$$

and where $\varrho_n(\phi) := \lambda_{\min}(E_n(\phi)) > 0$ denotes the smallest eigenvalue of $E_n(\phi)$, it follows from (3.21) that

$$\inf_{w \in \mathbb{R}^n, w \neq 0} \varepsilon_n^2 = \inf_{\tau > 0} \left((r^2(1 + \tau) + s^2(1 + \tau^{-1}))\varrho_n\left(\frac{r^2(1 + \tau)}{r^2(1 + \tau) + s^2(1 + \tau^{-1})}\right)\right). \tag{3.22}$$

Further, if $\phi := \frac{r^2(1+\tau)}{r^2(1+\tau)+s^2(1+\tau^{-1})}$, it follows that $\tau = \frac{s^2\phi}{r^2(1-\phi)}$ and that $\{\phi : \tau > 0\} = (0, 1)$. Thus, (3.22) can be rewritten as

$$\inf_{w \in \mathbb{R}^n, w \neq 0} \varepsilon_n^2 = \inf_{0 < \phi < 1} \left[\left(\frac{r^2}{\phi} + \frac{s^2}{1 - \phi}\right)\varrho_n(\phi)\right]. \tag{3.23}$$

The right-hand side of equation (3.23) already solves our minimisation problem. But we calculate now a more explicit expression for $\varrho_n(\phi)$, which will make (3.23) more easily computable.

Clearly, $\varrho_1(\phi) = 1$. If λ is an eigenvalue of $E_n(\phi)$ for $n \geq 2$ then, since $E_n(\phi)$ is positive definite, and applying the Gershgorin circle theorem, we see that $0 < \lambda \leq 4$. It is rather straightforward to show that 4 is not an eigenvalue so that $0 < \lambda < 4$. But, in any case, for $n \geq 2$, $\lambda < 4$ is an eigenvalue of $E_n(\phi)$, with eigenvector $v = (v_n, \ldots, v_1)^T \in \mathbb{R}^n$, if and only if

$$\lambda = 2 - 2\cos\theta = 4\sin^2\frac{\theta}{2} \quad \text{for some } \theta \in (0, \pi), \tag{3.24}$$

and

$$[(1 + \phi) - \lambda]v_n - v_{n-1} = 0, \tag{3.25}$$
$$-v_{j+1} + (2 - \lambda)v_j - v_{j-1} = 0 \quad \text{for } j = 2, \ldots, n - 1, \tag{3.26}$$
$$-v_2 + [(2 - \phi) - \lambda]v_1 = 0. \tag{3.27}$$

Equation (3.26) holds if and only if v_j is a linear combination of $\cos((j - 1)\theta)$ and $\sin((j - 1)\theta)$. To within multiplication by a constant, the linear combination that also satisfies (3.27) is

$$v_j = \sin\theta\cos[(j - 1)\theta] + (\cos\theta - \phi)\sin[(j - 1)\theta]$$
$$= \sin(j\theta) - \phi\sin((j - 1)\theta), \quad j = 1, \ldots, n. \tag{3.28}$$

Substituting in (3.25), we see that $\lambda < 4$ is an eigenvalue of $E_n(\phi)$ for $n \geq 2$ if and only if (3.24) holds and $F(\theta) = 0$, where

$$F(t) := (2\cos t - 1)\sin(nt) - \sin((n - 1)t) + \phi(1 - \phi)\sin((n - 1)t), \quad t \geq 0.$$

Now, for $n \in \mathbb{N}$,

$$F(t) = -\sin(nt) + \sin((n+1)t) + \phi(1-\phi)\sin((n-1)t) \qquad (3.29)$$

$$= 2\sin\frac{t}{2}\cos\left(\left(n+\frac{1}{2}\right)t\right) + \phi(1-\phi)\sin((n-1)t). \qquad (3.30)$$

Thus, noting that $\phi(1-\phi) \le 1/4$ and setting $\sigma := \pi/(n+2) \le \pi/3$, so that $1/2 \le \cos\sigma < 1$, we see from (3.29) that

$$F(\sigma) \le -\sin 2\sigma + \sin\sigma + \frac{1}{4}\sin 3\sigma = \frac{1}{4}\sin\sigma(3 + 4\cos^2\sigma - 8\cos\sigma) \le 0.$$

Moreover, from (3.30), while $F(0) = 0$, $F(t) > \phi(1-\phi)\sin((n-1)t) \ge 0$, for $0 < t < \frac{\pi}{2n+1}$. Let $\theta_n(\phi)$ denote the smallest positive solution of the equation $F(t) = 0$. Then, we have shown that $\theta_n(\phi) \in [\frac{\pi}{2n+1}, \frac{\pi}{n+2}] \subset (0, \pi)$, and hence that the smallest eigenvalue of $E_n(\phi)$ is in $(0, 4)$; precisely,

$$\varrho_n(\phi) = \lambda_{\min}(E_n(\phi)) = 4\sin^2\left(\frac{\theta_n(\phi)}{2}\right). \qquad (3.31)$$

(We have shown this equation for $n \ge 2$, and it holds also for $n = 1$ since $\varrho_1(\phi) = 1$ and $\theta_1(\phi) = \pi/3$.) Further, $\theta_n(\phi)$ is the unique solution of $F(t) = 0$ in $[\frac{\pi}{2n+1}, \frac{\pi}{n+2}]$. To see this in the case $n \ge 2$, note that, defining $\eta := \phi(1-\phi) \in [0, 1/4]$, it follows from (3.30) that

$$F'(t) = \cos\frac{t}{2}\cos\left(\left(n+\frac{1}{2}\right)t\right) - (2n+1)\sin\frac{t}{2}\sin\left(\left(n+\frac{1}{2}\right)t\right)$$
$$+ \eta(n-1)\cos((n-1)t)$$
$$= \cos\left(\left(n+\frac{1}{2}\right)t\right)\mathcal{A}_n(t) - \sin\left(\left(n+\frac{1}{2}\right)t\right)\mathcal{B}_n(t),$$

where $\mathcal{A}_n(t) := \cos\frac{t}{2} + \eta(n-1)\cos\frac{3t}{2}$ and $\mathcal{B}_n(t) := (2n+1)\sin\frac{t}{2} - \eta(n-1)\sin\frac{3t}{2}$. Now, for $n \ge 2$ and $t \in (\frac{\pi}{2n+1}, \frac{\pi}{n+2})$, it holds that $\cos((n+\frac{1}{2})t) < 0$, $\sin((n+\frac{1}{2})t) > 0$, $\mathcal{A}_n(t) > 0$, and $\mathcal{B}_n(t) > 0$ (the last inequality holding since $\eta(n-1)\sin\frac{3t}{2} \le 3\eta(n-1)\sin\frac{t}{2} \le \frac{3}{4}(n-1)\sin\frac{t}{2}$). Thus, $F'(t) < 0$ for $t \in (\frac{\pi}{2n+1}, \frac{\pi}{n+2})$, which implies that F has at most one zero in $[\frac{\pi}{2n+1}, \frac{\pi}{n+2}]$.

So, we have shown that $\theta_n(\phi)$ is the unique solution of $F(t) = 0$ in $[\frac{\pi}{2n+1}, \frac{\pi}{n+2}]$. Combining this with (3.23) and (3.31), we see that we have completed our aim of minimising ε_n^2, proving the following corollary of Theorem 3.5.

Theorem 3.6. *For all $n \in \mathbb{N}$ the second inclusion in (1.22) holds with*

$$\varepsilon_n := 2 \inf_{0 < \phi < 1}\left(\sqrt{\frac{\|\alpha\|_\infty^2}{\phi} + \frac{\|\gamma\|_\infty^2}{1-\phi}}\sin\frac{\theta_n(\phi)}{2}\right), \qquad (3.32)$$

where $\theta_n(\phi)$ is the unique solution in the range $[\frac{\pi}{2n+1}, \frac{\pi}{n+2}]$ of the equation $F(t) = 0$, i.e., of

$$2 \sin \frac{t}{2} \cos\left(\left(n + \frac{1}{2}\right)t\right) + \phi(1 - \phi) \sin((n-1)t) = 0. \tag{3.33}$$

If X has the Globevnik property, then also the first inclusion in (1.22) holds with ε_n given by (3.32).

The following properties of θ_n in Theorem 3.6 are worth noting:

$$\theta_1(\phi) = \frac{\pi}{3}, \quad 0 \le \phi \le 1; \tag{3.34}$$

$$\theta_n(\phi) = \theta_n(1 - \phi), \quad 0 \le \phi \le 1, \, n \in \mathbb{N}; \tag{3.35}$$

$$\frac{\pi}{2n+1} = \theta_n(0) = \theta_n(1) \le \theta_n(\phi) \le \theta_n\left(\frac{1}{2}\right), \quad 0 \le \phi \le 1, \, n \in \mathbb{N}. \tag{3.36}$$

These properties are evident from the definition of θ_n, except for the bound $\theta_n(\phi) \le \theta_n(\frac{1}{2})$. One way to see this bound is via (3.31), as it follows from $E_n(\frac{1}{2}) = \frac{1}{2}(E_n(\phi) + E_n(1-\phi))$ that

$$\varrho_n\left(\frac{1}{2}\right) = \frac{1}{2} \min_{\|w\|=1} (\phi w^T E_n(\phi)w + w^T E_n(1 - \phi)w)$$

$$\ge \frac{1}{2}[\varrho_n(\phi) + \varrho_n((1-\phi))] = \varrho_n(\phi),$$

since $\varrho_n(\phi) = \varrho_n(1 - \phi)$ as a consequence of (3.35) and (3.31).

We remark also that $\theta_n(\frac{1}{2})$ is the unique solution of the equation

$$G(t) := 2 \cos\left(\left(\frac{n+1}{2}\right)t\right) - \cos\left(\left(\frac{n-1}{2}\right)t\right) = 0 \tag{3.37}$$

in the interval $(\frac{\pi}{n+3}, \frac{\pi}{n+2}]$. To see this claim it is enough to check that (3.37) and (3.33) have the same solutions in $(0, \frac{\pi}{n+2}]$ when $\phi = \frac{1}{2}$, and to show that $\theta_n(\frac{1}{2}) > \frac{\pi}{n+3}$. To see the first statement multiply both sides of (3.37) by the positive term $2 \sin \frac{(n+1)t}{2} - \sin \frac{(n-1)t}{2}$, giving

$$2 \sin((n+1)t) - 2 \sin(nt) + \frac{1}{2} \sin((n-1)t) = 0.$$

This is equivalent to (3.37) with $\phi = \frac{1}{2}$, so that $\theta_n(\frac{1}{2})$ is the unique solution of $G(t) = 0$ in $(0, \frac{\pi}{n+2}]$. To see the second statement, observe that $G(\frac{\pi}{n+3}) = 2 \sin \frac{\pi}{n+3} - \sin \frac{2\pi}{n+3} > 0$, whereas, where $\sigma = \frac{\pi}{n+2}$, $G(\sigma) = 2 \sin \frac{\sigma}{2} - \sin \frac{3\sigma}{2} = \sin \frac{\sigma}{2}(4 \sin^2 \frac{\sigma}{2} - 1) \le 0$, since $0 < \frac{\sigma}{2} \le \frac{\pi}{6}$. Thus, $\theta_n(\frac{1}{2}) > \frac{\pi}{n+3}$.

In general, it appears not possible to explicitly compute the value of ϕ for which the bracket in (3.32) is minimised. But it follows easily from the bound (3.36) that, in the case $r = \|\alpha\|_\infty = 0$ the infimum is attained by setting $\phi = 0$, and this infimum

is $\varepsilon_n = 2\|\gamma\|_\infty \sin\frac{\pi}{4n+2}$. Likewise, when $s = \|\gamma\|_\infty = 0$ we see that the infimum is attained by setting $\phi = 1$, and this infimum is $\varepsilon_n = 2\|\alpha\|_\infty \sin\frac{\pi}{4n+2}$. Thus, we have computed the mimimal value of ε_n completely explicitly in the case when A is bidiagonal, giving the following corollary.

Corollary 3.7. *If $n \in \mathbb{N}$ and $\gamma = 0$ or $\alpha = 0$, then* (3.32) *reduces to*

$$\varepsilon_n = 2(\|\alpha\|_\infty + \|\gamma\|_\infty) \sin\left(\frac{\pi}{4n+2}\right). \tag{3.38}$$

The example of the shift operator (see Example 1.11 and Section 8.1 below) shows that the value of ε_n in Corollary 3.7 is the best possible, in the sense that, for every $n \in \mathbb{N}$, there exists a bidiagonal $A \in L(E)$ for which (1.22) is not true for any smaller value for ε_n.

In the case that $r = \|\alpha\|_\infty \neq 0$ and $s = \|\gamma\|_\infty \neq 0$ it is not clear what the infimum in (3.32) is explicitly when $n \geq 2$. However, since $\theta_n(\phi)$ depends continuously on ϕ and is bounded as a function of ϕ on $(0, 1)$, it is clear that the infimum is attained at some $\phi \in (0, 1)$. Further, as a consequence of (3.35), it is easy to see that the infimum is attained for some ϕ in the range $(0, \frac{1}{2})$ if $r < s$, for some ϕ in the range $(\frac{1}{2}, 1)$ if $r > s$. A simple choice of ϕ which has these properties and which reduces to the optimal values of ϕ, $\phi = 0$ and $\phi = 1$, respectively, when $r = 0$ and $s = 0$, is $\phi = \frac{r}{r+s} = \frac{\|\alpha\|_\infty}{\|\alpha\|_\infty + \|\gamma\|_\infty}$. This choice is further motivated by the fact that this is the unique ϕ that attains the infimum in (3.32) in the case $n = 1$ (which is an easy calculation in view of (3.34)). If we evaluate the bracket in (3.32) for this choice we obtain the following corollary which includes Corollary 3.7 as a special case.

Corollary 3.8. *For all $n \in \mathbb{N}$, ε_n given by* (3.32) *satisfies*

$$\varepsilon_n \leq \eta_n := 2(\|\alpha\|_\infty + \|\gamma\|_\infty) \sin\frac{\theta}{2},$$

where θ is the unique solution in the range $[\frac{\pi}{2n+1}, \frac{\pi}{n+2}]$ of the equation

$$2 \sin\left(\frac{t}{2}\right) \cos\left(\left(n + \frac{1}{2}\right)t\right) + \frac{\|\alpha\|_\infty\|\gamma\|_\infty}{(\|\alpha\|_\infty + \|\gamma\|_\infty)^2} \sin((n-1)t) = 0.$$

In particular, $\eta_1 = \varepsilon_1 = \|\alpha\|_\infty + \|\gamma_\infty\|_\infty$.

The following even more explicit result is obtained if we evaluate the bracket in (3.32) for $\phi = \frac{1}{2}$, noting the equivalent characterisation of $\theta_n(\frac{1}{2})$ above, that it is the unique solution in $(\frac{\pi}{n+3}, \frac{\pi}{n+2}]$ of (3.37).

Corollary 3.9. *For all $n \in \mathbb{N}$, ε_n given by* (3.32) *satisfies*

$$\varepsilon_n \leq \eta'_n := 2\sqrt{2}\sqrt{\|\alpha\|_\infty^2 + \|\gamma\|_\infty^2} \sin\frac{\theta}{2},$$

where θ is the unique solution in the range $(\frac{\pi}{n+3}, \frac{\pi}{n+2}]$ of (3.37).

Note that Corollaries 3.8 and 3.9 give the same upper bound for ε_n when $\|\alpha\|_\infty = \|\gamma\|_\infty$, but when $\|\alpha\|_\infty \neq \|\gamma\|_\infty$, as a consequence of (3.36), Corollary 3.8 is sharper ($\eta_n < \eta'_n$).

4. The π method: Periodised principal submatrices

In this section, we prove the inclusion (1.25) that is the equivalent of (1.22) but with periodised submatrices (1.20) instead of (1.19). We start with Proposition 4.1, the equivalent of Proposition 3.3 for the π method. The matrix referenced in this proposition is a generalisation of the π-method matrix $A_{n,k}^{\mathrm{per},t}$ of (1.20), defined, for any $\alpha, c \in \ell^\infty(\mathbb{Z}, L(X))$, by

$$A_{n,k}^{\alpha,c} := A_{n,k} + B_{n,k}^{\alpha,c}, \quad k \in \mathbb{Z}, \ n \in \mathbb{N},$$

where $B_{n,k}^{\alpha,c}$ is the $n \times n$ matrix whose entry in row i and column j is $\delta_{i,1}\delta_{j,n}\alpha_k + \delta_{i,n}\delta_{j,1}c_k$. Thus, for $n \geq 3$, $A_{n,k}^{\alpha,c}$ takes the form (1.20), but with the top right entry $t\alpha_k$ replaced by α_k, and the bottom left entry $\bar{t}\gamma_{k+n+1}$ replaced by c_k. Importantly,

$$A_{n,k}^{\mathrm{per},t} = A_{n,k}^{\alpha,c} \quad \text{if } \alpha_k := t\alpha_k \quad \text{and} \quad c_k := \bar{t}\gamma_{k+n+1}, \quad k \in \mathbb{Z}. \tag{4.1}$$

(Of course, also $A_{n,k}^{\alpha,c} = A_{n,k}$, if $\alpha = c = 0$.)

Proposition 4.1. *Let $\varepsilon > 0$ and $n \in \mathbb{N}$, suppose that $w_j \in \mathbb{R}$, for $j = 1, \ldots, n$, with at least one w_j non-zero, that $\alpha, c \in \ell^\infty(\mathbb{Z}, L(X))$, with $\|\alpha\|_\infty \leq \|\alpha\|_\infty$ and $\|c\|_\infty \leq \|\gamma\|_\infty$, and that $A \in L(E)$ is tridiagonal and $\|Ax\| \leq \varepsilon$, for some $x \in E$ with $\|x\| = 1$. Then, for some $k \in \mathbb{Z}$ it holds that*

$$\nu(A_{n,k}^{\alpha,c}) \leq \varepsilon + \varepsilon'_n,$$

with

$$\varepsilon'_n := (\|\alpha\|_\infty + \|\gamma\|_\infty)\sqrt{\frac{T_n}{S_n}}, \tag{4.2}$$

where

$$S_n = \sum_{j=1}^n w_j^2 \quad \text{and} \quad T_n = (w_1 + w_n)^2 + \sum_{j=1}^{n-1}(w_{j+1} - w_j)^2,$$

with

$$w_0 := 0 \quad \text{and} \quad w_{n+1} := 0.$$

Proof. Let $y = Ax$, so $\|y\| \leq \varepsilon$. We use the notation E_j, for $j = 0, \ldots, n$, defined in (3.9) and where $\tilde{x}_{n,k} := (w_1 x_{k+1}, w_2 x_{k+2}, \ldots, w_n x_{k+n})^T$, put $a_k := \|A_{n,k}^{\alpha,c}\tilde{x}_{n,k}\|$

and $b_k := \|\tilde{x}_{n,k}\|$, for $k \in \mathbb{Z}$. As in the proof of Proposition 3.3, we have that

$$\sum_{k\in\mathbb{Z}} b_k^2 = S_n.$$

Further, for $n \geq 2$ and $k \in \mathbb{Z}$, using (3.9) and the Kronecker delta notation,

$$a_k^2 = \|w_1 y_{k+1} + w_n \alpha_k x_{k+n} - w_1 \alpha_k x_k + (w_2 - w_1)\gamma_{k+2} x_{k+2}\|_X^2$$

$$+ \sum_{j=2}^{n-1} \|w_j y_{k+j} + (w_{j-1} - w_j)\alpha_{k+j-1} x_{k+j-1} + (w_{j+1} - w_j)\gamma_{k+j+1} x_{k+j+1}\|_X^2$$

$$+ \|w_n y_{k+n} + (w_{n-1} - w_n)\alpha_{k+n-1} x_{k+n-1} + w_1 c_k x_{k+1} - w_n \gamma_{k+n+1} x_{k+n+1}\|_X^2$$

$$\leq \sum_{j=1}^{n} (w_j \|y_{k+j}\|_X + \|\alpha\|_\infty c_{j,k} + \|\gamma\|_\infty d_{j,k})^2,$$

(cf. (3.12)) where, for $j = 1, \dots, n$ and $k \in \mathbb{Z}$,

$$c_{j,k} := \delta_{j,1} w_n \|x_{k+n}\|_X + E_{j-1}\|x_{k+j-1}\|_X,$$
$$d_{j,k} := \delta_{j,n} w_1 \|x_{k+1}\|_X + E_j \|x_{k+j+1}\|_X. \tag{4.3}$$

After applying Lemma 3.2 twice, we get (cf. (3.13)) that, for all $\theta > 0$ and $\phi > 0$,

$$a_k^2 \leq \sum_{j=1}^{n} \Big[(1+\theta)w_j^2 \|y_{k+j}\|_X^2 + (1+\theta^{-1})$$

$$\cdot \big((1+\phi)\|\alpha\|_\infty^2 c_{j,k}^2 + (1+\phi^{-1})\|\gamma\|_\infty^2 d_{j,k}^2\big)\Big].$$

Now, recalling that $\|x\| = 1$,

$$\sum_{k\in\mathbb{Z}} c_{1,k}^2 = \sum_{k\in\mathbb{Z}} (w_n \|x_{k+n}\|_X + w_1 \|x_k\|_X)^2$$

$$\leq w_n^2 + 2w_1 w_n \sum_{k\in\mathbb{Z}} \|x_{k+n} x_k\|_X + w_1^2 \leq (w_1 + w_n)^2,$$

so that

$$\sum_{j=1}^{n}\sum_{k\in\mathbb{Z}} c_{j,k}^2 \leq (w_1 + w_n)^2 + \sum_{j=2}^{n}\sum_{k\in\mathbb{Z}} E_{j-1}^2 |x_{k+j-1}|^2 = T_n.$$

Similarly,

$$\sum_{k\in\mathbb{Z}} d_{n,k}^2 \leq (w_1 + w_n)^2 \quad \text{so that} \quad \sum_{j=1}^{n}\sum_{k\in\mathbb{Z}} d_{j,k}^2 \leq T_n.$$

Thus, applying Lemma 3.2 (cf. (3.14) and (3.15)),

$$\sum_{k \in \mathbb{Z}} a_k^2 \le \inf_{\theta, \phi > 0} \left[(1 + \theta) S_n \|y\|^2 + (1 + \theta^{-1}) T_n \right.$$

$$\left. \cdot \left[(1 + \phi) \|\alpha\|_\infty^2 + (1 + \phi^{-1}) \|\gamma\|_\infty^2 \right] \right] \qquad (4.4)$$

$$= \inf_{\theta > 0} \left[(1 + \theta) \|y\|^2 S_n + (1 + \theta^{-1}) T_n (\|\alpha\|_\infty + \|\gamma\|_\infty)^2 \right]$$

$$\le \inf_{\theta > 0} \left[(1 + \theta) \varepsilon^2 + (1 + \theta^{-1}) \left((\|\alpha\|_\infty + \|\gamma\|_\infty) \sqrt{\frac{T_n}{S_n}} \right)^2 \right] S_n$$

$$= \inf_{\theta > 0} \left[\left((1 + \theta) \varepsilon^2 + (1 + \theta^{-1}) \varepsilon_n'^2 \right) \right] S_n$$

$$= (\varepsilon + \varepsilon_n')^2 \sum_{k \in \mathbb{Z}} b_k^2. \qquad (4.5)$$

Arguing as at the end of Proposition 3.3 we deduce that $a_k \le (\varepsilon + \varepsilon_n') b_k$, for some $k \in \mathbb{Z}$ with $b_k \ne 0$. We reach the same conclusion, via the same bounds (4.4) and (4.5), also in the case $n = 1$ when $S_n = w_1^2$, $T_n = 4w_1^2$, and

$$a_k^2 = w_1^2 \|y_{k+1} + \alpha_k x_{k+1} - \alpha_k x_k + c_k x_{k+1} - \gamma_{k+2} x_{k+2}\|_X^2, \quad k \in \mathbb{Z}.$$

Thus,

$$\nu(A_{n,k}^{\alpha,c}) \le \varepsilon + \varepsilon_n'. \qquad \blacksquare$$

This is the π version of Corollary 3.4, which has an identical proof.

Corollary 4.2. *Let $n \in \mathbb{N}$, suppose that $w_j \in \mathbb{R}$, for $j = 1, \ldots, n$, with at least one w_j non-zero, that $\alpha, c \in \ell^\infty(\mathbb{Z}, L(X))$, with $\|\alpha\|_\infty \le \|\alpha\|_\infty$ and $\|c\|_\infty \le \|\gamma\|_\infty$, and that $A \in L(E)$ is tridiagonal. Then,*

$$\inf_{k \in \mathbb{Z}} \nu(A_{n,k}^{\alpha,c}) \le \nu(A) + \varepsilon_n',$$

where ε_n' is given by (4.2).

To state the main theorem of this section we introduce a generalisation of the notation (1.24). For $n \in \mathbb{N}$ and $\alpha, c \in \ell^\infty(\mathbb{Z}, L(X))$, let

$$\pi_\varepsilon^{n,\alpha,c}(A) := \bigcup_{k \in \mathbb{Z}} \operatorname{spec}_\varepsilon A_{n,k}^{\alpha,c}, \quad \varepsilon > 0 \quad \text{and} \quad \Pi_\varepsilon^{n,\alpha,c}(A) := \overline{\bigcup_{k \in \mathbb{Z}} \operatorname{Spec}_\varepsilon A_{n,k}^{\alpha,c}}, \quad \varepsilon \ge 0.$$
$$(4.6)$$

Clearly, for $n \in \mathbb{N}$ and $t \in \mathbb{T}$,

$$\pi_\varepsilon^{n,t}(A) = \pi_\varepsilon^{n,\alpha,c}(A), \quad \varepsilon > 0 \quad \text{and} \quad \Pi_\varepsilon^{n,t}(A) = \Pi_\varepsilon^{n,\alpha,c}(A), \quad \varepsilon \ge 0, \qquad (4.7)$$

if α and c are given by (4.1). Clearly, for all $n \in \mathbb{N}$ and $\alpha, c \in \ell^\infty(\mathbb{Z}, L(X))$, $\pi_\varepsilon^{n,\alpha,c}(A)$ is open and $\Pi_\varepsilon^{n,\alpha,c}(A)$ is closed. Further, defining

$$\mu_n^{\alpha,c}(B) := \inf_{k \in \mathbb{Z}} \mu(A_{n,k}^{\alpha,c}) = \inf_{k \in \mathbb{Z}} \min \left(\nu(A_{n,k}^{\alpha,c}), \, \nu((A_{n,k}^{\alpha,c})^*) \right), \tag{4.8}$$

it holds, for $n \in \mathbb{N}$ and $\alpha, c \in \ell^\infty(\mathbb{Z}, L(X))$, by arguing as in the proof of Proposition 3.1, that

$$\pi_\varepsilon^{n,\alpha,c}(A) = \left\{ \lambda \in \mathbb{C} : \mu_n^{\alpha,c}(A - \lambda I) < \varepsilon \right\}, \quad \varepsilon > 0, \tag{4.9}$$

and that, if X has the Globevnik property, also

$$\overline{\pi_\varepsilon^{n,\alpha,c}(A)} = \Pi_\varepsilon^{n,\alpha,c}(A) = \left\{ \lambda \in \mathbb{C} : \mu_n^{\alpha,c}(A - \lambda I) \leq \varepsilon \right\}, \quad \varepsilon > 0. \tag{4.10}$$

Thus, if X has the Globevnik property, then

$$\Pi_\varepsilon^{n,\alpha,c}(A) = \bigcap_{\varepsilon' > \varepsilon} \pi_{\varepsilon'}^{n,\alpha,c}(A), \quad \varepsilon > 0. \tag{4.11}$$

In relation to the definition (4.8), note that

$$(A_{n,k}^{\alpha,c})^* = (A^*)_{n,k}^{c,\alpha}. \tag{4.12}$$

Theorem 4.3. *Let $n \in \mathbb{N}$, and suppose that $w_j \geq 0$, for $j = 1, \ldots, n$, with at least one w_j non-zero, and that $\alpha, c \in \ell^\infty(\mathbb{Z}, L(X))$, with $\|\alpha\|_\infty \leq \|\alpha\|_\infty$ and $\|c\|_\infty \leq \|\gamma\|_\infty$, Then, where ε'_n is given by (4.2),*

$$\mathrm{spec}_\varepsilon A \subset \pi_{\varepsilon+\varepsilon'_n}^{n,\alpha,c}(A), \quad \varepsilon > 0. \quad \textit{Moreover,} \quad \mathrm{Spec}_\varepsilon A \subset \Pi_{\varepsilon+\varepsilon'_n}^{n,\alpha,c}(A), \quad \varepsilon \geq 0, \tag{4.13}$$

provided X has the Globevnik property.

Proof. To see that the first inclusion in (4.13) holds, suppose that $\lambda \in \mathrm{spec}_\varepsilon A$. Then, by (2.2), either $\nu(A - \lambda I) < \varepsilon$ or $\nu(A^* - \lambda I) < \varepsilon$. By applications of Corollary 4.2, it follows, in the case $\nu(A - \lambda I) < \varepsilon$, that $\nu(A_{n,k}^{\alpha,c} - \lambda I_n) < \varepsilon + \varepsilon_n$, for some $k \in \mathbb{Z}$, and, in the case $\nu(A^* - \lambda I) < \varepsilon$, recalling (4.12), that

$$\nu((A_{n,k}^{\alpha,c})^* - \lambda I_n) = \nu((A^*)_{n,k}^{c,\alpha} - \lambda I_n) < \varepsilon + \varepsilon_n$$

for some $k \in \mathbb{Z}$. Thus, recalling the notation (1.9), $\mu(A_{n,k}^{\alpha,c} - \lambda I_n) < \varepsilon + \varepsilon_n$, for some $k \in \mathbb{Z}$, so that $\lambda \in \mathrm{spec}_{\varepsilon+\varepsilon_n} A_{n,k}^{\alpha,c}$ by (1.10), so that $\lambda \in \pi_{\varepsilon+\varepsilon_n}^{n,\alpha,c}(A)$. The second inclusion in (4.13) follows from the first inclusion by taking intersections, noting (2.7) and (4.11). $\blacksquare$

In particular, Theorem 4.3 holds with α and c given by (4.1), giving the following corollary.

Corollary 4.4. *Let $n \in \mathbb{N}$, and suppose that $w_j \geq 0$, for $j = 1, \ldots, n$, with at least one w_j non-zero, and that $t \in \mathbb{T}$. Then, the second inclusion of (1.25) holds with ε'_n given by (4.2). If X has the Globevnik property, then also the first inclusion of (1.25) holds with ε'_n given by (4.2).*

As we did above for the τ method, we will now minimise the penalty term ε'_n as a function, (4.2), of the weight vector $w = (w_1, \ldots, w_n)^T$. In the cases $n = 1$ and 2 it is easy to see that ε_n is minimised by taking $w_1 = \cdots = w_n = 1$, giving

$$\varepsilon_1 = 2(\|\alpha\|_\infty + \|\gamma\|_\infty)$$

and

$$\varepsilon_2 = \sqrt{2}(\|\alpha\|_\infty + \|\gamma\|_\infty).$$

For $n \geq 3$, from the definitions of S_n and T_n in Theorem 4.4, we know that $S_n = \|w\|_2^2$ and $T_n = \|B_n w\|_2^2$, where

$$B_n = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ 1 & & & & 1 \end{pmatrix}$$

so that

$$B_n^T B_n = \begin{pmatrix} 2 & -1 & & & 1 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 1 & & & -1 & 2 \end{pmatrix}.$$

Let us seek

$$\inf_{w \in \mathbb{R}^n, w \neq 0} \sqrt{\frac{T_n}{S_n}} = \inf_{w \in \mathbb{R}^n, w \neq 0} \sqrt{\frac{\|B_n w\|_2^2}{\|w\|_2^2}}$$

$$= \min_{w \in \mathbb{R}^n, \|w\|=1} \sqrt{\frac{w^T B_n^T B_n w}{w^T w}} = \sqrt{\lambda_{\min}(B_n^T B_n)}, \qquad (4.14)$$

noting that a vector w achieves this minimum if and only if it is an eigenvector of $B_n^T B_n$ corresponding to the eigenvalue $\lambda_{\min}(B_n^T B_n)$. It is clear that this minimum is strictly positive and that (taking $w_1 = w_2 = \cdots = w_n$) it is no larger than $2/\sqrt{n}$. Further, $\lambda \in (0, 4)$ is an eigenvalue of $B_n^T B_n$ with eigenvector $v = (v_1, \ldots, v_n)^T$ if and only if

$$\lambda = 2(1 - \cos \theta) = 4 \sin^2 \frac{\theta}{2} \quad \text{for some } \theta \in (0, \pi), \qquad (4.15)$$

and

$$v_n + (2 - \lambda)v_1 - v_2 = 0, \tag{4.16}$$

$$-v_{i+1} + (2 - \lambda)v_i - v_{i-1} = 0 \quad \text{for } i = 2, \dots, n-1, \tag{4.17}$$

$$-v_{n-1} + (2 - \lambda)v_n + v_1 = 0. \tag{4.18}$$

Note that if (λ, v) satisfies (4.16)–(4.18) then so does $(\lambda, \tilde{v})$, where $\tilde{v} := (v_n, \dots, v_1)$. Further, $(\lambda, v^\dagger)$, where $v^\dagger := v + \tilde{v}$ satisfies $v_j^\dagger = v_{n+1-j}^\dagger$, $j = 1, \dots, n$. Thus, to compute λ and a corresponding eigenvector v from (4.16)–(4.18), we may assume that

$$v_j = v_{n+1-j}, \quad j = 1, \dots, n. \tag{4.19}$$

As discussed around (3.26), (4.17) holds if and only if v_j is a linear combination of $\cos((j-1)\theta)$ and $\sin((j-1)\theta)$. To within multiplication by a constant, the linear combination that satisfies (4.19) is

$$v_j = \sin((j-1)\theta) + \sin((n-j)\theta), \quad j = 1, 2, \dots, n, \tag{4.20}$$

and this satisfies (4.16) and (4.18) if and only if $F_\pi(\theta) = 0$ where, for $t \in \mathbb{R}$,

$$F_\pi(t) := \sin((n-1)t) + 2\cos(t)\sin((n-1)t) - \sin(t) - \sin((n-2)t)$$
$$= \sin((n-1)t) + \sin(nt) - \sin(t) = 2\sin((n-2)t/2)\cos(nt/2) + \sin(nt).$$

From this last expression for $F_\pi(t)$ it is clear that $F_\pi(\pi/n) = 0$ and that $F_\pi(t) > 0$ for $0 < t < \pi/n$. Thus, $\lambda_{\min}(B_n^T B_n)$ is given by (4.15) with $\theta = \frac{\pi}{n}$. Since the corresponding eigenvector given by (4.20) satisfies $v_j > 0$, $j = 1, \dots, n$, the following corollary follows from Theorem 4.4 and (4.14) and the observations above about the cases $n = 1$ and 2.

Corollary 4.5. *The conclusions of Theorem 4.3 hold with*

$$\varepsilon_n' := 2(\|\alpha\|_\infty + \|\gamma\|_\infty)\sin\frac{\pi}{2n}. \tag{4.21}$$

In particular, for $n \in \mathbb{N}$, the second inclusion of (1.25) holds with ε_n' given by (4.21), as does the first inclusion of (1.25) if X has Globevnik's property.

Again, as in Corollary 3.7, the example of the shift operator (see Example 1.11 and Section 8.1 below) shows that this value of ε_n' is the best possible.

5. The τ_1 method: One-sided truncations

5.1. Proof of the inclusions (1.30)

In this section, which follows the patterns of Section 3 and Section 4, we establish the inclusions (1.30).

Proposition 5.1. *Let $\varepsilon > 0$ and $n \in \mathbb{N}$, suppose that $w_j \in \mathbb{R}$, for $j = 1, \ldots, n$, with at least one w_j non-zero, and that $A \in L(E)$ is tridiagonal and $\|Ax\| \leq \varepsilon$, for some $x \in E$ with $\|x\| = 1$. Then, for some $k \in \mathbb{Z}$, where $A_{n,k}^+$ is defined by (1.32), it holds that*

$$\nu(A_{n,k}^+) \leq \varepsilon + \varepsilon_n'', \tag{5.1}$$

where

$$\varepsilon_n'' := (\|\alpha\|_\infty + \|\gamma\|_\infty)\sqrt{\frac{T_n}{S_n}} \tag{5.2}$$

with

$$S_n = \sum_{j=1}^{n} w_j^2 \quad and \quad T_n = w_1^2 + w_n^2 + \sum_{j=1}^{n-1}(w_{i+1} - w_i)^2.$$

Proof. Let $y = Ax$, so $\|y\| \leq \varepsilon$. We argue as in the proof of Proposition 3.3, using the notation E_j, defined in (3.9), for $j = -1, 0, \ldots, n, n+1$, with $w_{-1} := w_0 := w_{n+1} := w_{n+2} := 0$. Where $\tilde{x}_{n,k} := (w_1 x_{k+1}, w_2 x_{k+2}, \ldots, w_n x_{k+n})^T$, put $a_k := \|A_{n,k}^+ \tilde{x}_{n,k}\|$ and $b_k := \|\tilde{x}_{n,k}\|$, for $k \in \mathbb{Z}$. As in the proof of Proposition 3.3, we have that $\sum_{k \in \mathbb{Z}} b_k^2 = S_n$. Further, cf. (3.11) and (3.12),

$$a_k^2 = \|w_1 \gamma_{k+1} x_{k+1}\|_X^2 + \|w_n \alpha_{k+n} x_{k+n}\|_X^2$$
$$+ \sum_{j=1}^{n} \|\alpha_{j+k-1} w_{j-1} x_{j+k-1} + \beta_{j+k} w_j x_{j+k} + \gamma_{j+k+1} w_{j+1} x_{j+k+1}\|_X^2$$
$$\leq \sum_{j=0}^{n+1}(w_j \|y_{j+k}\|_X + E_{j-1}\|\alpha\|_\infty \|x_{j+k-1}\|_X + E_j\|\gamma\|_\infty \|x_{j+k+1}\|_X)^2.$$

So, for all $\theta > 0$ and $\phi > 0$, using Lemma 3.2 twice,

$$a_k^2 \leq \sum_{j=0}^{n+1} \Big[(1+\theta)\big(w_j \|y_{j+k}\|_X\big)^2 + (1+\theta^{-1})$$
$$\cdot \big((1+\phi)|\alpha_{i-1}|^2\big(E_{j-1}\|\alpha\|_\infty \|x_{j+k-1}\|_X\big)^2$$
$$+ (1+\phi^{-1})\big(E_j\|\gamma\|_\infty \|x_{j+k+1}\|_X\big)^2\big)\Big].$$

Thus, for all $\theta, \phi > 0$,

$$\sum_{k \in \mathbb{Z}} a_k^2 \leq (1+\theta)S_n\|y\|^2 + (1+\theta^{-1})[(1+\phi)T_n\|\alpha\|_\infty^2 + (1+\phi^{-1})T_n\|\gamma\|_\infty^2].$$

Arguing as in the proof of Proposition 3.3 (cf. (3.14) and (3.15)) it follows that

$$\sum_{k \in \mathbb{Z}} a_k^2 \leq (\varepsilon + \varepsilon_n'')^2 \sum_{k \in \mathbb{Z}} b_k^2,$$

so that $a_k \leq (\varepsilon + \varepsilon_n'')b_k$, for some $k \in \mathbb{Z}$ with $b_k \neq 0$, which implies that $\nu(A_{n,k}^+) \leq \varepsilon + \varepsilon_n''$. ∎

The above result has the following straightforward corollary (cf. Corollary 3.4).

Corollary 5.2. *Let $n \in \mathbb{N}$, suppose that $w_j \in \mathbb{R}$, for $j = 1, \ldots, n$, with at least one w_j non-zero, and that $A \in L(E)$ is tridiagonal. Then,*

$$\inf_{k \in \mathbb{Z}} \nu(A_{n,k}^+) \leq \nu(A) + \varepsilon_n'', \tag{5.3}$$

where ε_n'' is given by (5.2).

The proof of the following main theorem is very similar to those of Theorems 3.5 and 4.3.

Theorem 5.3. *Suppose that $n \in \mathbb{N}$ and $w_j \in \mathbb{R}$, for $j = 1, \ldots, n$, with at least one w_j non-zero. Then, (1.30) holds with ε_n'' given by (5.2).*

Proof. We have justified already, below (1.31), the inclusions from the left in (1.30). To see that the second inclusion from the right in (1.30) holds, let $\lambda \in \mathrm{spec}_\varepsilon(A)$. Then, either

$$\nu(A - \lambda I) < \varepsilon \quad \text{or} \quad \nu(A^* - \lambda I) < \varepsilon.$$

By Corollary 5.2 it follows that either $\nu(A_{n,k}^+ - \lambda I_n^+) < \varepsilon + \varepsilon_n''$, for some $k \in \mathbb{Z}$, or $\nu((A^*)_{n,k}^+ - \lambda I_n^+) < \varepsilon + \varepsilon_n''$, for some $k \in \mathbb{Z}$. Thus, by (1.33), $\mu_n(A - \lambda I) < \varepsilon + \varepsilon_n''$, so that $\lambda \in \gamma_{\varepsilon+\varepsilon_n''}^n(A)$. The first inclusion from the right in (1.30) follows by taking intersections, recalling (2.7). ∎

Following the pattern of Section 3.2 and Section 4, we now minimise ε_n'' as a function of the weight vector $w = (w_1, \ldots, w_n)^T \in \mathbb{R}^n$ over $\mathbb{R}^n \setminus \{0\}$, or equivalently over $\{w \in \mathbb{R}^n : \|w\| = 1\}$. In order to minimise ε_n'', given by (5.2), we need to minimise T_n/S_n, where

$$S_n = \|w\|^2$$

and

$$T_n = w_1^2 + (w_2 - w_1)^2 + \cdots + (w_n - w_{n-1})^2 + w_n^2 = \|Bw\|^2,$$

with

$$B = \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & -1 & 1 & \\ & & & -1 & 1 \\ & & & & -1 \end{pmatrix}_{(n+1) \times n}$$

so that

$$B^T B = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}_{n \times n}.$$

Clearly,

$$\inf_{\|w\| \neq 0} \frac{T_n}{S_n} = \inf_{\|w\| \neq 0} \frac{\|Bw\|^2}{\|w\|^2} = \inf_{\|w\| \neq 0} \frac{w^T B^T B w}{w^T w} = \lambda_{\min}(B^T B),$$

the smallest eigenvalue of $B^T B$. One could now compute the eigenvalues of $B^T B$ as in the previous sections Section 3.2 and Section 4. In this simple case ($-B^T B$ is the discrete Laplacian), this is a standard result (e.g., [7]). We have that

$$\mathrm{spec}\,(B^T B) = \left\{ \lambda_j := 2 - 2 \cos \frac{j\pi}{(n+1)} = 4 \sin^2 \frac{j\pi}{2(n+1)} : j \in \{1, \dots, n\} \right\}, \quad (5.4)$$

so that $\lambda_{\min}(B^T B) = \lambda_1$. Hence, the minimal value for ε_n'' is

$$\varepsilon_n'' := (\|\alpha\|_\infty + \|\gamma\|_\infty) \sqrt{\frac{T_n}{S_n}} = (\|\alpha\|_\infty + \|\gamma\|_\infty) \sqrt{\lambda_1} = 2(\|\alpha\|_\infty + \|\gamma\|_\infty) \sin \frac{\pi}{2(n+1)}.$$

This minimum is realised [7] by the choice $w = (\sin \frac{j\pi}{n+1})_{j=1}^n$ for the weight vector in Theorem 5.3.

Corollary 5.4. *For all $n \in \mathbb{N}$, the inclusions* (1.30) *hold with*

$$\varepsilon_n'' := 2(\|\alpha\|_\infty + \|\gamma\|_\infty) \sin \frac{\pi}{2(n+1)}. \quad (5.5)$$

The example of the shift operator (Example 1.11 and Section 8.1) shows that the above formula is sharp; (1.30) does not hold for all tridiagonal $A \in L(E)$ if ε_n'' is any smaller than the above value.

5.2. Does an analogue of Proposition 3.1 hold for the τ_1 method?

For operators $B \in L(X)$ (in particular, when $X = \mathbb{C}^n$ and B is a finite square matrix), the resolvent norm

$$\lambda \mapsto \|(B - \lambda I)^{-1}\|$$

on $\rho(B) := \mathbb{C} \setminus \mathrm{Spec}\, B$ is a subharmonic function, subject to a maximum principle; it cannot have a local maximum in $\rho(B)$. Likewise, its reciprocal, which (recall (1.9) and (1.12)) is the Lipschitz continuous function

$$f_B : \lambda \mapsto \mu(B - \lambda I),$$

cannot have a local minimum on $\rho(B)$, indeed cannot have a local minimum on $\mathbb{C}$, except that it takes the minimum value zero on Spec B.

The proof of the identity

$$\overline{\sigma_\varepsilon^n(A)} = \Sigma_\varepsilon^n(A) = \hat{\Sigma}_\varepsilon^n(A)$$

for the τ method in Proposition 3.1 rests indirectly on this property, and the stronger result that f_B cannot be locally constant on any open subset of $\rho(B)$, which holds if X has the Globevnik property. The same is true for the proof of the corresponding property (4.10) for the π method. But the corresponding relationship for the τ_1 method, relating the two τ_1 sets defined in (1.29), that $\overline{\gamma_\varepsilon^n(A)} = \Gamma_\varepsilon^n(A)$, does not hold for every $n \in \mathbb{N}$, $\varepsilon > 0$, and every tridiagonal A, as the example we give below shows.

The issue is that the appropriate version of the mapping f_B can have local (non-zero) minima, in the case that B is a finite rectangular matrix, as discussed in [85, Section X.46] and in [77]. An example that illustrates this is the 4×2 matrix

$$B := \begin{pmatrix} \delta & 0 \\ 0 & 0 \\ 1 & 1 \\ 0 & \delta \end{pmatrix} \quad \text{with } \delta \in \left(0, \frac{1}{2}\right).$$

The relevant version of the mapping f_B, namely, $f_B(\lambda) := \nu(B - \lambda I_2^+)$, where I_2^+ is as defined above (1.33) (with $X = \mathbb{C}$), has global minima (over all $\lambda \in \mathbb{C}$) at 0 and 1 (see the calculations in Lemma 5.5 below), with the values

$$f(0) = f(1) = \nu(B) = \delta > 0.$$

To obtain our counterexample to $\overline{\gamma_\varepsilon^n(A)} = \Gamma_\varepsilon^n(A)$ we build a tridiagonal bi-infinite matrix around B, given by[9]

$$A := \begin{pmatrix} \ddots & \ddots & & & & & & \\ ,\ddots & 0 & \delta & & & & & \\ & 1 & 0 & \delta & & & & \\ & & 1 & 0 & & & & \\ & & & 1 & 1 & & & \\ & & & & \delta & 1 & & \\ & & & & & \delta & 1 & \\ & & & & & & \ddots & \ddots \end{pmatrix}. \tag{5.6}$$

<hr>

[9]This is an example of a so-called *paired Laurent operator*, a class for which the spectrum can be computed explicitly. See, e.g., [51, Section 4.4.1], [64, Section 3.7.3].

The sets $\gamma_\varepsilon^n(A)$ and $\Gamma_\varepsilon^n(A)$ are defined by (1.29) in terms of $\mu_n(A - \lambda I)$, which in turn is expressed in terms of lower norms of rectangular matrices in (1.33). For the case $n = 2$ there are only three distinct matrices $A_{n,k}^+$ as k ranges over $\mathbb{Z}$, namely,

$$
B_1 := \begin{pmatrix} \delta & \\ 0 & \delta \\ 1 & 0 \\ & 1 \end{pmatrix}, \quad
B_2 := B = \begin{pmatrix} \delta & \\ 0 & 0 \\ 1 & 1 \\ & \delta \end{pmatrix}, \quad
B_3 := \begin{pmatrix} 0 & \\ 1 & 0 \\ \delta & 1 \\ & \delta \end{pmatrix},
$$

and only five different matrices $(A^*)_{n,k}^+$ with $n = 2$ and $k \in \mathbb{Z}$, namely,

$$
C_1 = \begin{pmatrix} 1 & \\ 0 & 1 \\ \delta & 0 \\ & \delta \end{pmatrix}, \quad
C_2 = \begin{pmatrix} 1 & \\ 0 & 1 \\ \delta & 0 \\ & 0 \end{pmatrix}, \quad
C_3 = \begin{pmatrix} 1 & \\ 0 & 1 \\ 0 & 1 \\ & 0 \end{pmatrix},
$$

$$
C_4 = \begin{pmatrix} 1 & \\ 1 & \delta \\ 0 & 1 \\ & 0 \end{pmatrix}, \quad
C_5 = \begin{pmatrix} \delta & \\ 1 & \delta \\ 0 & 1 \\ & 0 \end{pmatrix}.
$$

Putting, for $j \in \{1, 2, 3\}$ and $k \in \{1, \dots, 5\}$,

$$
f_j(\lambda) := \nu(B_j - \lambda I_2^+) \quad \text{and} \quad g_k(\lambda) := \nu(C_k - \lambda I_2^+), \quad \lambda \in \mathbb{C},
$$

we have

$$
\mu_n(A - \lambda I) = \min \{ f_j(\lambda), g_k(\lambda) : j \in \{1, 2, 3\}, k \in \{1, \dots, 5\} \} =: h(\lambda). \quad (5.7)
$$

Lemma 5.5. *Let $\delta \in (0, \frac{1}{2})$. Then, the following statements hold.*

(a) $f_2(\lambda) \geq \delta$ *for all $\lambda \in \mathbb{C}$.*

(b) $0 < \delta = f_2(0) < \frac{1}{2} < f_3(0) = g_5(0) < g_2(0) = g_3(0) = g_4(0) = 1 < f_1(0) = g_1(0)$.

As a consequence, the function f_2, and so, also the function $h(\lambda) = \mu_n(A - \lambda I)$, have positive local minima at zero, namely, $\mu_n(A) = h(0) = f_2(0) = \delta$.

Proof. (a) For all $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{C}^2$, we have

$$
\|(B_2 - \lambda I)x\| = \left\| \begin{pmatrix} \delta x_1 \\ -\lambda x_1 \\ x_1 + (1 - \lambda)x_2 \\ \delta x_2 \end{pmatrix} \right\| \geq \left\| \begin{pmatrix} \delta x_1 \\ 0 \\ 0 \\ \delta x_2 \end{pmatrix} \right\| = \delta \|x\|,
$$

so that $f_2(\lambda) = \nu(B_2 - \lambda I) \geq \delta$ for all $\lambda \in \mathbb{C}$.

(b) We start with the computation of $f_j(0)$ for $j \in \{1, 2, 3\}$. Firstly, $f_1(0) = \nu(B_1) = \sqrt{1 + \delta^2} > 1$, since $\|B_1 x\|^2 = (1 + \delta^2)\|x\|^2$ for all $x \in \mathbb{C}^2$. Secondly, $f_2(0) = \nu(B_2) \geq \delta$ holds by (a). To see equality, observe that $\|B_2 x\| = \delta \|x\|$ for $x = \binom{1}{-1}$. Thirdly, $f_3(0) = \nu(B_3) \in [1 - \delta, 1)$, since $\|B_3 x - \delta(0, 0, x_1, x_2)^\top\| = \|x\|$, so that $\|B_3 x\| \geq \|x\| - \delta \|x\|$ for all $x = \binom{x_1}{x_2} \in \mathbb{C}^2$, whence $\nu(B_3) \geq 1 - \delta$. Computation of $\|B_3 x\|$ with $x = \binom{1}{-\delta}$ shows that $\nu(B_3) < 1$.

Carrying on to the $g_k(0)$ with $k \in \{1, \ldots, 5\}$, again writing $x \in \mathbb{C}^2$ as $x = \binom{x_1}{x_2}$, we get $g_2(0) = \nu(C_2) = 1$, since $\|C_2 x\| \geq \|(x_1, x_2, 0, 0)^\top\| = \|x\|$ for all $x \in \mathbb{C}^2$, with equality for $x = \binom{0}{1}$; $g_3(0) = \nu(C_3) = 1$, since $\|C_3 x\| \geq \|(x_1, x_2, 0, 0)^\top\| = \|x\|$ for all $x \in \mathbb{C}^2$, with equality for $x = \binom{1}{0}$; $g_4(0) = \nu(C_4) = 1$, since $\|C_4 x\| \geq \|(x_1, 0, x_2, 0)^\top\| = \|x\|$ for all $x \in \mathbb{C}^2$, with equality for $x = \binom{\delta}{-1}$. Finally, denoting the flip isometry $(x_1, \ldots, x_m) \mapsto (x_m, \ldots, x_1)$ on $\mathbb{C}^m$ by J_m, we have $C_1 = J_4 B_1 J_2$ and $C_5 = J_4 B_3 J_2$, showing that $g_1(0) = f_1(0)$ and $g_5(0) = f_3(0)$. ∎

By Lemma 5.5, the tridiagonal matrix (5.6) has the properties that, for $n = 2$ and $0 < \delta < 1/2$, $0 \in \Gamma_\delta^n(A) = \{\lambda \in \mathbb{C} : \mu_n(A - \lambda I) \leq \delta\}$ but $\gamma_\delta^n(A) = \{\lambda \in \mathbb{C} : \mu_n(A - \lambda I) < \delta\}$ does not intersect $\varepsilon \mathbb{D}$, for some $\varepsilon > 0$. Thus, $0 \notin \overline{\gamma_\delta^n(A)}$, and

$$\overline{\gamma_\delta^n(A)} \subsetneq \Gamma_\delta^n(A). \tag{5.8}$$

6. Proof of Theorem 1.9 and a band-dominated generalisation

The following result is a generalisation of Theorem 1.9 to the band-dominated case. It reduces to Theorem 1.9 if A is tridiagonal (in which case $A^{(n)} = A$, for each n, so that $w_n = 1$ and also $A_n = A$ and $\delta_n = 0$), provided also X is finite-dimensional or a Hilbert space (in which case $\eta_n = 0$). Note that the requirement that $\{A_{ij} : i, j \in \mathbb{Z}\} \subset L(X)$ be relatively compact is satisfied automatically if X is finite-dimensional.

Theorem 6.1. *Suppose that $A \in \mathrm{BDO}(E)$ and that $\{a_{i,j} : i, j \in \mathbb{Z}\} \subset L(X)$ is relatively compact. Let $A^{(n)} \in \mathrm{BO}(E)$ be defined, for $n \in \mathbb{N}$, by $A^{(n)} := A$ if A if tridiagonal, otherwise by (1.43) for some $\mathfrak{p} \geq 0$ such that $\delta_n := \|A - A^{(n)}\| \to 0$ as $n \to \infty$, and let $A_n := \mathcal{I}_{w_n} A^{(n)} \mathcal{I}_{w_n}^{-1}$, where w_n is the band-width of $A^{(n)}$ and $\mathcal{I}_b$ is defined as above (1.14). In the case that X is finite-dimensional or a Hilbert space set $\eta_n := 0$, for $n \in \mathbb{N}$, otherwise let $(\eta_n)_{n \in \mathbb{N}}$ denote any positive null sequence. Then, for every $n \in \mathbb{N}$ there exists a finite set $K_n \subset \mathbb{Z}$ such that*

$$\forall k \in \mathbb{Z} \ \exists j \in K_n \ \text{such that}$$
$$\|(A_n)_{n,k}^+ - (A_n)_{n,j}^+\| \leq 1/n \quad \text{and} \quad \|((A_n)^*)_{n,k}^+ - ((A_n)^*)_{n,j}^+\| \leq 1/n. \tag{6.1}$$

Further, for $\varepsilon \geq 0$ and $n \in \mathbb{N}$, using the notations (1.47) and (1.42),

$$\Gamma_\varepsilon^{n,\mathrm{fin}}(A_n) \subset \mathrm{Spec}_{\varepsilon + \delta_n + \eta_n} A \quad \text{and} \quad \mathrm{Spec}_\varepsilon A \subset \Gamma_{\varepsilon + \delta_n + \eta_n + \varepsilon_n''(A_n) + 1/n}^{n,\mathrm{fin}}(A_n), \tag{6.2}$$

where

$$\Gamma_\varepsilon^{n,\mathrm{fin}}(A_n) = \left\{\lambda \in \mathbb{C} : \mu_n^{\mathrm{fin}}(A_n - \lambda I) \leq \varepsilon\right\} \tag{6.3}$$

with

$$\mu_n^{\mathrm{fin}}(A_n - \lambda I) = \min_{j \in K_n} \min\left(\nu((A_n)_{n,j}^+ - \lambda I_n^+), \nu(((A_n)^*)_{n,j}^+ - \lambda I_n^+)\right), \quad \textit{for } \lambda \in \mathbb{C}. \tag{6.4}$$

Moreover, for $\varepsilon \geq 0$, $\Gamma_{\varepsilon+2\delta_n+2\eta_n+\varepsilon_n''(A_n)+1/n}^{n,\mathrm{fin}}(A_n) \xrightarrow{H} \mathrm{Spec}_\varepsilon A$ as $n \to \infty$, in particular, $\Gamma_{\varepsilon_n''(A_n)+2\delta_n+2\eta_n+1/n}^{n,\mathrm{fin}}(A_n) \xrightarrow{H} \mathrm{Spec}\, A$.

Proof. Fix $n \in \mathbb{N}$. To see (6.1) note first that the relative compactness of the matrix entries of A implies, by the definition of A_n, relative compactness also of the entries of A_n. Further, since the set of matrix entries of A_n is relatively compact, so that every sequence has a convergent subsequence, the same is true for

$$\mathcal{S}_n := \left\{((A_n)_{n,k}^+, ((A_n)^*)_{n,k}^+) : k \in \mathbb{Z}\right\} \subset \left((X^{w_n})^{(n+2)\times n}\right)^2. \tag{6.5}$$

Thus, there exists $K_n \subset \mathbb{Z}$ such that (6.1) holds since relatively compact sets are totally bounded.

To see (6.2), note that it follows from (1.45), (1.30), and (2.4) (if $\eta_n > 0$) or (2.5) (if $\eta_n = 0$) that $\Gamma_\varepsilon^{n,\mathrm{fin}}(A_n) \subset \mathrm{Spec}_\varepsilon A_n \subset \mathrm{Spec}_{\varepsilon+\delta_n+\eta_n} A$. Again, applying (2.4)/(2.5),

$$\mathrm{Spec}_\varepsilon A \subset \mathrm{Spec}_{\varepsilon+\delta_n+\eta_n} A_n \subset \Gamma_{\varepsilon+\delta_n+\eta_n+\varepsilon_n''(A_n)}^n(A_n) \subset \Gamma_{\varepsilon+\delta_n+\eta_n+\varepsilon_n''(A_n)+1/n}^{n,\mathrm{fin}}(A_n),$$

by (1.30) and (1.45). Thus, for $n \in \mathbb{N}$ and $\varepsilon \geq 0$,

$$\mathrm{Spec}_\varepsilon A \subset \Gamma_{\varepsilon+\delta_n+\eta_n+\varepsilon_n''(A_n)+1/n}^{n,\mathrm{fin}}(A_n) \subset \mathrm{Spec}_{\varepsilon+2\delta_n+2\eta_n+\varepsilon_n''(A_n)+1/n} A.$$

Arguing as above Theorem 1.6, we have that $\varepsilon_n''(A_n) \to 0$ as $n \to \infty$. This implies that $\Gamma_{\varepsilon+2\delta_n+2\eta_n+\varepsilon_n''(A_n)+1/n}^{n,\mathrm{fin}}(A_n) \xrightarrow{H} \mathrm{Spec}_\varepsilon A$ as $n \to \infty$, by (2.8). ∎

Remark 6.2. In the theorem above we note that there exists a finite set $K_n \subset \mathbb{Z}$ such that (6.1) holds. An equivalent statement is to say that there exists a finite set $S_n \subset \mathcal{S}_n$, where $\mathcal{S}_n$ is defined by (6.5), such that

$$\forall k \in \mathbb{Z} \; \exists (B_1, B_2) \in S_n \text{ such that } \|(A_n)_{n,k}^+ - B_1\| \leq 1/n \text{ and } \|((A_n)^*)_{n,k}^+ - B_2\| \leq 1/n. \tag{6.6}$$

The above theorem thus remains true if we replace (6.1) with (6.6) and replace (6.4) with the formula

$$\mu_n^{\mathrm{fin}}(A_n - \lambda I) = \min_{(B_1,B_2) \in S_n} \min\left(\nu(B_1 - \lambda I_n^+), \nu(B_2 - \lambda I_n^+)\right), \quad \text{for } \lambda \in \mathbb{C}. \tag{6.7}$$

Importantly, with these changes, the proof still holds if, instead of requiring $S_n \subset \mathcal{S}_n$, we make the weaker requirement that $S_n \subset \overline{\mathcal{S}_n}$, the closure of $\mathcal{S}_n$ in $((X^{w_n})^{(n+2)\times n})^2$:

the elements of S_n need not be in S_n only in $\overline{S_n}$. This observation is helpful because there are instances where it is much easier to identify $\overline{S_n}$ than S_n.

Theorem 1.9 and the above result are both based on the τ_1 method. Versions of these results hold also for the τ and π methods, at least in the case that A is tridiagonal, provided (this is a substantial assumption) these methods do not suffer from spectral pollution for the particular operator A. Here is a version (cf. Theorem 1.9) for the case that A is tridiagonal (recall from Section 1.4 that every banded A can be written in tridiagonal form), written in the way suggested by the above remark. We use in this theorem the notation that, for tridiagonal $B \in L(E)$,

$$B_{n,k}^{(M)} := \begin{cases} B_{n,k} & \text{if } M = \tau, \\ B_{n,k}^{\text{per},t} & \text{if } M = \pi, \end{cases} \tag{6.8}$$

where t in (6.8) is some fixed value with $t \in \mathbb{T}$, and $B_{n,k}$ and $B_{n,k}^{\text{per},t}$ are as defined in (1.19) and (1.20).

Theorem 6.3. *Suppose that X satisfies Globevnik's property, that $A \in L(E)$ is tridiagonal, and that $\{a_{i,j} : i, j \in \mathbb{Z}\} \subset L(X)$ is relatively compact, and, where $M = \tau$ or π, suppose that the M method does not suffer from spectral pollution for A in the sense of Definition 1.1. Let $(\eta_n)_{n \in \mathbb{N}}$ be the null sequence from that definition, and $(\eta_n')_{n \in \mathbb{N}}$ be any other positive null sequence. Then, for every $n \in \mathbb{N}$ there exists a finite set $S_n \subset \overline{S_n^{(M)}}$, where*

$$S_n^{(M)} := \{A_{n,k}^{(M)} : k \in \mathbb{Z}\} \subset X^{n \times n}, \tag{6.9}$$

such that

$$\forall k \in \mathbb{Z} \; \exists B \in S_n \text{ such that } \|A_{n,k}^{(M)} - B\| \leq 1/n. \tag{6.10}$$

Further, for $\varepsilon \geq 0$ and $n \in \mathbb{N}$,

$$\Delta_\varepsilon^{n,\text{fin},M}(A) \subset \text{spec}_{\varepsilon + \eta_n + \eta_n'} A \quad \text{and} \quad \text{Spec}_\varepsilon A \subset \Delta_{\varepsilon + \varepsilon_n^{(M)} + 1/n}^{n,\text{fin},M}(A), \tag{6.11}$$

where $\varepsilon_n^{(M)} := \varepsilon_n$ for $M = \tau$, $:= \varepsilon_n'$ for $M = \pi$, with ε_n and ε_n' defined as in (3.6) and (1.26), and

$$\Delta_\varepsilon^{n,\text{fin},M}(A) := \{\lambda \in \mathbb{C} : \min_{B \in S_n} \mu(B - \lambda I_n) \leq \varepsilon\}. \tag{6.12}$$

Moreover, for $\varepsilon \geq 0$, $\Delta_{\varepsilon + \varepsilon_n^{(M)} + 1/n}^{n,\text{fin},M}(A) \xrightarrow{H} \text{Spec}_\varepsilon A$ as $n \to \infty$, in particular,

$$\Delta_{\varepsilon_n^{(M)} + 1/n}^{n,\text{fin},M}(A) \xrightarrow{H} \text{Spec } A.$$

Proof. The fact that (6.10) holds follows from the relative compactness of $\{a_{i,j} : i, j \in \mathbb{Z}\} \subset L(X)$, arguing as in the proof of Theorem 1.9 (and see Remark 6.2). Now, by (1.21) and (1.10),

$$\Sigma_\varepsilon^n(A) = \left\{\lambda \in \mathbb{C} : \inf_{B \in S^{(\tau)}} \mu(B - \lambda I_n) \le \varepsilon\right\} = \left\{\lambda \in \mathbb{C} : \inf_{B \in \overline{S^{(\tau)}}} \mu(B - \lambda I_n) \le \varepsilon\right\},$$

by (1.12), so that $\Delta_\varepsilon^{n,\mathrm{fin},\tau}(A) \subset \Sigma_\varepsilon^n(A)$. Similarly, $\Delta_\varepsilon^{n,\mathrm{fin},\pi}(A) \subset \Pi_\varepsilon^{n,t}(A)$ by (1.24), (1.10), and (1.12). Since $\Sigma_\varepsilon^n(A) \subset \sigma_{\varepsilon+\eta_n'}^n(A)$ and $\Pi_\varepsilon^n(A) \subset \pi_{\varepsilon+\eta_n'}^n(A)$, the first inclusion in (6.11) follows from the assumption of the absence of spectral pollution. Note that (cf. (1.45)), if (6.10) holds, then, for $\varepsilon \ge 0$,

$$\Sigma_\varepsilon^n(A) \subset \Delta_{\varepsilon+1/n}^{n,\mathrm{fin},\tau}(A) \quad \text{and} \quad \Pi_\varepsilon^{n,t}(A) \subset \Delta_{\varepsilon+1/n}^{n,\mathrm{fin},\pi}(A),$$

as a consequence of (1.12) and since

$$\sigma_\varepsilon^n(A) = \left\{\lambda \in \mathbb{C} : \inf_{B \in S^{(\tau)}} \mu(B - \lambda I_n) < \varepsilon\right\},$$

and the same representation holds for $\pi_\varepsilon^{n,t}(A)$, with $S^{(\tau)}$ replaced by $S^{(\pi)}$. These inclusions and (1.22) and (1.25) imply the second inclusion in (6.11). Thus, for $n \in \mathbb{N}$ and $\varepsilon \ge 0$,

$$\mathrm{Spec}_\varepsilon A \subset \Delta_{\varepsilon+\varepsilon_n^{(M)}+1/n}^{n,\mathrm{fin},M}(A) \subset \mathrm{spec}_{\varepsilon+\varepsilon_n^{(M)}+\eta_n+\eta_n'+1/n} A,$$

so that

$$\Delta_{\varepsilon+\varepsilon_n^{(M)}+1/n}^{n,\mathrm{fin},M}(A) \xrightarrow{H} \mathrm{Spec}_\varepsilon A \quad \text{as } n \to \infty,$$

by (2.8). ∎

7. The scalar case: Computational aspects and the solvability complexity index

In this section we focus on the case that, for some $p \in \mathbb{N}$, $X = \mathbb{C}^p$, which we equip with the Euclidean norm (the 2-norm), so that X and $E = \ell^2(\mathbb{Z}, X) = \ell^2(\mathbb{Z}, \mathbb{C}^p)$ are Hilbert spaces, and the entries of the matrix representation $[a_{i,j}]$ of $A \in L(E)$ are $p \times p$ complex matrices, i.e., $L(X) = \mathbb{C}^{p \times p}$. This includes the special case that $p = 1$, so that $E = \ell^2(\mathbb{Z})$ and the entries $a_{i,j}$ are just complex numbers[10].

---

[10] This is a special case, but recall, as noted at the top of the paper, that, given a separable Hilbert space Y, our results for this special case apply to any $A \in L(Y)$ that, with respect to some orthonormal basis $(e_i)_{i \in \mathbb{Z}} \subset Y$, has a matrix representation $a_{i,j} = (Ae_j, e_i) \in \mathbb{C}$ that is banded or band-dominated.

Our goals in this section are to provide computational details regarding the determination of membership of our τ, π, and τ_1 inclusion sets in this case, to prove Proposition 1.10 and an extension of this result to the band-dominated case, and to note implications regarding the solvability complex index. In particular, a key conclusion in this section is that the computational problem of determining the spectrum of a band-dominated operator, given the inputs that we assume, has solvability complexity index $\mathrm{SCI}_A = 1$, indeed is in the class Π_1^A (our notations are those of [2, 25]).

In the above set up, in which X and E are Hilbert spaces, standard formulae are available for computing the lower norms that our methods require. For any Hilbert spaces Y and Z and any $B \in L(Y, Z)$ we have (see, e.g., [63, Section 1.3] or [64, Section 2.4])

$$(\nu(B))^2 = \nu(B'B) = \min \mathrm{Spec}\,(B'B) = (s_{\min}(B))^2, \qquad (7.1)$$

where $B' \in L(Z, Y)$ denotes the Hilbert space adjoint of B and its smallest singular value is denoted by $s_{\min}(B) := (\min \mathrm{Spec}\,(B'B))^{1/2}$, so that (recall our notation (1.9))

$$\mu(B) = (\min(\nu(B'B), \nu(BB')))^{1/2} = \min(s_{\min}(B), s_{\min}(B')). \qquad (7.2)$$

In particular, the above holds when $Y = \mathbb{C}^M$ and $Z = \mathbb{C}^N$, for some $M, N \in \mathbb{N}$, in which case B is an $N \times M$ matrix and B' is the conjugate transpose of B. In particular (see the discussion below (1.9)), if $M = N$, so that $Y = Z$ and B is a square matrix, then

$$\mu(B) = \nu(B) = s_{\min}(B) = (\min \mathrm{Spec}\,(B'B))^{1/2} = (\nu(B'B))^{1/2}. \qquad (7.3)$$

To decide whether a particular $\lambda \in \mathbb{C}$ is contained in one of our inclusion sets it is not necessary to actually compute any lower norms, it is enough to decide whether or not $\nu(B) \geq \varepsilon$ or is $> \varepsilon$, for some non-negative threshold ε, and for $B \in \mathcal{B}$, where $\mathcal{B}$ is some set of $N \times M$ matrices, depending on which inclusion set method we choose (τ, π, or τ_1). By (7.1), this is equivalent to deciding whether the $M \times M$ Hermitian matrix $B'B - \varepsilon^2 I_M$ is positive semi-definite or positive definite, in particular,

$$\nu(B) \geq \varepsilon \iff B'B - \varepsilon^2 I_M \text{ is positive semi-definite.} \qquad (7.4)$$

This can be tested (cf., [26, Supplementary materials]) by Gaussian elimination or by attempting an LDL^T version of the Cholesky decomposition [2, Proposition 10.1]. These algorithms only require finitely many arithmetic operations. Precisely, $O(M^3)$ operations are sufficient, indeed $O(w^2 M)$ operations suffice if $B'B$ has band-width w (see, e.g., [50, Algorithm 4.3.5]).

Let us emphasise in this case we are considering where X is a finite-dimensional Hilbert space, X (and so, also E) satisfy Globevnik's property (see Section 2), so that

all the inclusions (1.22), (1.25), and (1.30) apply. In this finite-dimensional case the expressions for the key inclusion sets for the τ and π methods simplify somewhat: we have by Proposition 3.1 that $\sigma^n_\varepsilon(A)$ and $\Sigma^n_\varepsilon(A)$ are given by (1.28) and, noting (7.3), the expression (1.27) for $\mu^\dagger(A - \lambda I)$ simplifies to

$$\mu^\dagger_n(A - \lambda I) = \inf_{k \in \mathbb{Z}} \nu(A_{n,k} - \lambda I_n), \quad n \in \mathbb{N}, \ \lambda \in \mathbb{C}. \tag{7.5}$$

Similarly, recalling from (4.7) that $\pi^{n,t}_\varepsilon(A) = \pi^{n,\mathfrak{a},\mathfrak{c}}_\varepsilon(A)$ and $\Pi^{n,t}_\varepsilon(A) = \Pi^{n,\mathfrak{a},\mathfrak{c}}_\varepsilon(A)$, when $\mathfrak{a}$ and $\mathfrak{c}$ are given by (4.1), and letting $\mu^{\pi,t}_n$ denote the function $\mu^{\mathfrak{a},\mathfrak{c}}_n$ defined by (4.8) in the case that $\mathfrak{a}$ and $\mathfrak{c}$ are given by (4.1), we have by (4.9), (4.10), and (7.3) that, for all $n \in \mathbb{N}$ and $t \in \mathbb{T}$,

$$\begin{aligned}
\pi^{n,t}_\varepsilon(A) &= \{\lambda \in \mathbb{C} : \mu^{\pi,t}_n(A - \lambda I) < \varepsilon\}, \quad \varepsilon > 0, \\
\Pi^{n,t}_\varepsilon(A) &= \{\lambda \in \mathbb{C} : \mu^{\pi,t}_n(A - \lambda I) \leq \varepsilon\}, \quad \varepsilon \geq 0,
\end{aligned} \tag{7.6}$$

where

$$\mu^{\pi,t}_n(A - \lambda I) = \inf_{k \in \mathbb{Z}} \nu(A^{\mathrm{per},t}_{n,k} - \lambda I_n), \quad n \in \mathbb{N}, \ \lambda \in \mathbb{C}, \ t \in \mathbb{T}. \tag{7.7}$$

The corresponding representations for the τ_1 method inclusion sets are (1.29) and (1.33).

The above representations, and the discussion around (7.4), make clear that, in the case that $\{A_{n,k} : k \in \mathbb{Z}\}$ is finite, which, as we have noted in Section 1.4.4, holds if and only if the matrix representation of A has only finitely many distinct entries, we can determine whether a given $\lambda \in \mathbb{C}$ is in one of the τ-method-related sets, $\sigma^n_\varepsilon(A)$ and $\Sigma^n_\varepsilon(A)$, in only finitely many arithmetic operations (assuming that we have already determined the finite set $\{A_{n,k} : k \in \mathbb{Z}\}$, or at least its cardinality). Similar remarks apply to the π and τ_1 methods. In the remainder of this section we consider questions of computation in finitely many operations in cases where A may have infinitely many distinct entries, proving Proposition 1.10 and an extension to band-dominated case.

Proof of Proposition 1.10. Suppose $n \in \mathbb{N}$. Given the inputs specified in the theorem, to compute the finite set $\Gamma^n_{\mathrm{fin}}(A)$, given by (1.50), one only needs to calculate $R := \alpha_{\max} + \beta_{\max} + \gamma_{\max}$ and check whether or not each of the finitely many points in $\mathrm{Grid}(n, R)$, given by (1.49), are in $\Gamma^{n,\mathrm{fin}}_{\varepsilon^*_n + 2/n}(A)$, where $\Gamma^{n,\mathrm{fin}}_\varepsilon(A)$ is given by (1.47) and (1.48). But this is just a question of deciding whether or not $\nu(B) \leq \varepsilon^*_n + 2/n$ for finitely many matrices B, where ε^*_n is given by (1.51), and, as noted above, this can be done in finitely many arithmetic operations.

Recall that $K_n = \mathcal{A}_p(A, n)$ so that (1.46) holds. To see that $\mathrm{Spec}\, A \subset \hat{\Gamma}^n_{\mathrm{fin}}(A)$ it is enough, by Theorem 1.9, and since $\mathrm{Spec}\, A \subset \|A\|\bar{\mathbb{D}}$, to show that $\Gamma^{n,\mathrm{fin}}_{\varepsilon''_n + 1/n}(A) \cap \|A\|\bar{D} \subset \hat{\Gamma}^n_{\mathrm{fin}}(A)$. So, suppose that $\lambda' \in \Gamma^{n,\mathrm{fin}}_{\varepsilon''_n + 1/n}(A)$ and $|\lambda'| \leq \|A\|$. Then, since

$R \geq \|A\|$, it is easy to see that there exists $\lambda \in \mathrm{Grid}(n, R)$ with $|\lambda - \lambda'| \leq \sqrt{2}/n$. Since $\lambda' \in \Gamma^{n,\mathrm{fin}}_{\varepsilon''_n + 1/n}(A)$ and recalling (1.11), we see that there exists $k \in K_n$ such that either $\nu(A^+_{n,k} - \lambda I^+_n) \leq \varepsilon''_n + (1 + \sqrt{2})/n$ or $\nu((A^*)^+_{n,k} - \lambda I^+_n) \leq \varepsilon''_n + (1 + \sqrt{2})/n < \varepsilon''_n + 3/n$, so that $\lambda \in \Gamma^{n,\mathrm{fin}}_{\varepsilon''_n + 3/n}(A) \subset \Gamma^{n,\mathrm{fin}}_{\varepsilon^*_n + 3/n}(A)$, i.e., $\lambda \in \Gamma^n_{\mathrm{fin}}(A)$. This implies, since $|\lambda - \lambda'| \leq \sqrt{2}/n < 2/n$, that $\lambda' \in \hat{\Gamma}^n_{\mathrm{fin}}(A)$. Thus, $\mathrm{Spec}\, A \subset \hat{\Gamma}^n_{\mathrm{fin}}(A)$, but also, by the definition (1.50) and Theorem 1.9, $\Gamma^n_{\mathrm{fin}}(A) \subset \Gamma^{n,\mathrm{fin}}_{\varepsilon^*_n + 3/n}(A) \subset \mathrm{Spec}_{\varepsilon^*_n + 3/n}(A)$, so that

$$\mathrm{Spec}\, A \subset \hat{\Gamma}^n_{\mathrm{fin}}(A) \subset \mathrm{Spec}_{\varepsilon^*_n + 3/n}(A) + \frac{2}{n}\overline{\mathbb{D}}.$$

It follows from (2.8) that $\mathrm{Spec}_{\varepsilon^*_n + 3/n}(A) + \frac{2}{n}\overline{\mathbb{D}} \xrightarrow{H} \mathrm{Spec}\, A$ as $n \to \infty$, so that also $\hat{\Gamma}^n_{\mathrm{fin}}(A) \to \mathrm{Spec}\, A$. Since $d_H(\Gamma^n_{\mathrm{fin}}(A), \hat{\Gamma}^n_{\mathrm{fin}}(A)) \leq 2/n$, we have also that $\Gamma^n_{\mathrm{fin}}(A) \to \mathrm{Spec}\, A$. ∎

Let us, as promised, extend Proposition 1.10 and the discussion of Section 1.4.5 to the band-dominated case. Let $\Omega = \mathrm{BDO}(E)$, with $X = \mathbb{C}$, so that $E = \ell^2(\mathbb{Z})$ and the matrix entries of $A \in \Omega$ are complex numbers. The mappings we will need, to provide us with the inputs needed to compute a sequence of approximations to $\mathrm{Spec}\, A$, for $A \in \Omega$, are $\mathcal{A}_p, \mathcal{B}_p, \mathcal{C}_p$, for $p \in \mathbb{N}$, as defined in Section 1.4.5, plus the mapping $\mathcal{D} = (\mathcal{D}_1, \mathcal{D}_2) : \Omega \times \mathbb{N} \to \Omega^n_T \times \mathbb{R}$, $(\mathcal{A}, n) \mapsto (A_n, \eta_n)$, where $A_n := \mathcal{I}_n A^{(n)} \mathcal{I}_n^{-1}$, $\mathcal{I}_n$ is defined below (1.14), $A^{(n)}$ is defined by (1.43) with $\mathfrak{p} = 1$, and $\eta_n \geq 0$ is such that $\eta_n \geq \|A - A^{(n)}\|$ and $\eta_n \to 0$ as $n \to \infty$.

The sequence of approximations to $\mathrm{Spec}\, A$ is $(\Gamma^n_{\mathrm{fin}}(A))_{n \in \mathbb{N}}$, given by

$$\Gamma^n_{\mathrm{fin}}(A) := \Gamma^{n,\mathrm{fin}}_{\varepsilon^*_n(A_n) + 3/n + \eta_n}(A_n) \cap \mathrm{Grid}(n, R_n), \quad n \in \mathbb{N}, \tag{7.8}$$

where $(A_n, \eta_n) := \mathcal{D}(A, n)$, $(\alpha_{\max}, \beta_{\max}, \gamma_{\max}) := \mathcal{B}_n(A_n)$, $R_n := \alpha_{\max} + \beta_{\max} + \gamma_{\max} + \eta_n$ (so that R_n is an upper bound for $\|A\|$), $K_n \subset \mathbb{Z}$, in the definition (1.47) and (1.48), is given by $K_n := \mathcal{A}_n(A_n, n)$, and (cf. (1.42) and (1.51))

$$\varepsilon^*_n(A_n) := (\alpha_{\max} + \gamma_{\max})\frac{22}{7(n+1)} \geq 2(\alpha_{\max} + \gamma_{\max}) \sin \frac{\pi}{2(n+1)} \geq \varepsilon''_n(A_n), \tag{7.9}$$

where $\varepsilon''_n(A_n)$ is defined by (1.42). As we note in the following theorem, each element of this sequence can be computed in finitely many arithmetical operations, given finitely many evaluations of the input maps $\mathcal{A}_p, \mathcal{B}_p, \mathcal{C}_p$, and $\mathcal{D}$.

Theorem 7.1. *For $A \in \Omega$ and $n \in \mathbb{N}$, $\Gamma^n_{\mathrm{fin}}(A)$, as defined in (7.8), can be computed in finitely many arithmetic operations, given $\eta_n := \mathcal{D}_2(A, n)$, $K_n := \mathcal{A}_n(A_n, n)$, where $A_n := \mathcal{D}_1(A, n)$, $(\alpha_{\max}, \beta_{\max}, \gamma_{\max}) := \mathcal{B}_n(A_n)$ and for $k \in K_n$,*

$$((A_n)^+_{n,k}, ((A_n)A^*)^+_{n,k}) := \mathcal{C}_n(A_n, k, n).$$

Further, $\Gamma_{\text{fin}}^n(A) \xrightarrow{H} \operatorname{Spec} A$ *as* $n \to \infty$, *and also*

$$\widehat{\Gamma}_{\text{fin}}^n(A) := \Gamma_{\text{fin}}^n(A) + \frac{2}{n}\overline{\mathbb{D}} \xrightarrow{H} \operatorname{Spec} A$$

as $n \to \infty$, *with* $\operatorname{Spec} A \subset \widehat{\Gamma}_{\text{fin}}^n(A)$ *for each* $n \in \mathbb{N}$.

We will prove Theorem 7.1 shortly, but let us interpret this theorem as a result relating to the solvability complexity index (SCI) of [2, 3, 25], as we did for Proposition 1.10 in Section 1.4.5. Inspecting the definition of $\Gamma_{\text{fin}}^n(A)$, note that it is not necessary to explicitly compute the action of both components of $\mathcal{D}$, it is enough to evaluate $\mathcal{D}_2$, the second component of $\mathcal{D}$, to obtain $\eta_n = \mathcal{D}_2(A, n)$, and then to evaluate compositions of $\mathcal{D}_1$ with $\mathcal{A}_n$, $\mathcal{B}_n$, and $\mathcal{C}_n$, to obtain $K_n := (\mathcal{A}_n(\cdot, n) \circ \mathcal{D}_1)(A)$, $(\alpha_{\max}, \beta_{\max}, \gamma_{\max}) := (\mathcal{B}_n \circ \mathcal{D}_1)(A)$, $((A_n)_{n,k}^+, ((A_n)^*)_{n,k}^+) := (\mathcal{C}_n(\cdot, k, n) \circ \mathcal{D}_1)(A)$. Equipping $\mathbb{C}^C$ with the Hausdorff metric as in Section 1.3, the mappings

$$\Omega \to \mathbb{C}^C, \quad A \mapsto \Gamma_{\text{fin}}^n(A), \quad A \mapsto \widehat{\Gamma}_{\text{fin}}^n(A)$$

are, for each $n \in \mathbb{N}$, general algorithms in the sense of [2, 25], with evaluation set (in the sense of [2, 25])

$$\Lambda := \big\{ \mathcal{A}_n(\cdot, n) \circ \mathcal{D}_1, \mathcal{B}_n \circ \mathcal{D}_1, \mathcal{C}_n(\cdot, k, n) \circ \mathcal{D}_1, \mathcal{D}_2 : k \in \mathbb{Z}, n \in \mathbb{N} \big\};$$

each function in Λ has domain Ω, and each can be expressed in terms of finitely many complex-valued functions; see the footnote below (1.52). Further, $\widehat{\Gamma}_{\text{fin}}^n(A)$ can be computed in finitely many arithmetic operations and specified using finitely many complex numbers (the elements of $\Gamma_{\text{fin}}^n(A)$ and the value of n). Thus, where $\Xi : \Omega \to \mathbb{C}^C$ is the mapping given by $\Xi(A) := \operatorname{Spec} A$, for $A \in \Omega$, the computational problem $\{\Xi, \Omega, \mathbb{C}^C, \Lambda\}$ has arithmetic SCI, in the sense of [2,25], equal to one; more precisely, since also $\operatorname{Spec} A \subset \widehat{\Gamma}_{\text{fin}}^n(A)$, for each $n \in \mathbb{N}$ and $A \in \Omega$, this computational problem is in the class Π_1^A, as defined in [2,25].

Proof of Theorem 7.1. Suppose $n \in \mathbb{N}$. Note first that $R := \sup\{R_n : n \in \mathbb{N}\} < \infty$ as (cf. (1.17)) $R_n \leq 3\|A_n\| + \eta_n$, (η_n) is a null sequence, and $\|A_n\| = \|A^{(n)}\| \leq \|A\| + \eta_n$. Thus, arguing as in the proof of Proposition 1.10, $\Gamma_{\text{fin}}^n(A)$ can be computed in finitely many arithmetic operations.

Since $K_n = \mathcal{A}_n(A_n, n)$, (1.46) holds with A replaced by A_n. To see that $\operatorname{Spec} A \subset \widehat{\Gamma}_{\text{fin}}^n(A)$, note that, by the first inclusion in (2.3), $\operatorname{Spec} A \subset \operatorname{Spec}_{\eta_n} A^{(n)} = \operatorname{Spec}_{\eta_n} A_n$. Further, by Theorem 1.9, $\operatorname{Spec}_{\eta_n} A_n \subset \Gamma_{\eta_n + \varepsilon_n''(A_n) + 1/n}^{n, \text{fin}}$. Since also $\operatorname{Spec} A \subset \|A\|\overline{\mathbb{D}}$, it follows that $\operatorname{Spec} A \subset \Gamma_{\eta_n + \varepsilon_n''(A_n) + 1/n}^{n, \text{fin}} \cap \|A\|\overline{D}$. Since $R_n \geq \|A\|$, it follows by arguing as in the proof of Proposition 1.10 that $\operatorname{Spec} A \subset \widehat{\Gamma}_{\text{fin}}^n(A_n)$. Also, by definition

and Theorem 1.9,

$$\Gamma_{\text{fin}}^n(A_n) \subset \Gamma_{\varepsilon_n^*(A_n)+3/n+\eta_n}^{n,\text{fin}}(A_n) \subset \text{Spec}_{\varepsilon_n^*(A_n)+3/n+\eta_n}(A_n),$$

and $\text{Spec}_{\varepsilon_n^*(A_n)+3/n+\eta_n}(A_n) \subset \text{Spec}_{\varepsilon_n^*(A_n)+3/n+2\eta_n}(A)$, by (2.5). Thus,

$$\text{Spec } A \subset \widehat{\Gamma}_{\text{fin}}^n(A) \subset \text{Spec}_{\varepsilon_n^*(A_n)+3/n+2\eta_n}(A) + \frac{2}{n}\overline{\mathbb{D}},$$

so that, by (2.8), $\widehat{\Gamma}_{\text{fin}}^n(A) \to \text{Spec } A$. Since $d_H(\Gamma_{\text{fin}}^n(A), \widehat{\Gamma}_{\text{fin}}^n(A)) \le 2/n$, we have also that $\Gamma_{\text{fin}}^n(A) \to \text{Spec } A$. ∎

8. Pseudoergodic operators and other examples

In this section, as in the previous section, our focus is the case that, for some $p \in \mathbb{N}$, $X = \mathbb{C}^p$, which we equip with the Euclidean norm, so that X and $E = \ell^2(\mathbb{Z}, X) = \ell^2(\mathbb{Z}, \mathbb{C}^p)$ are Hilbert spaces, and the matrix entries $a_{i,j}$ are $p \times p$ complex matrices. This includes the special case $p = 1$, so that $E = \ell^2(\mathbb{Z})$ and the entries $a_{i,j} \in \mathbb{C}$. Our aim is to illustrate various of the results in earlier sections in the case that A is the tridiagonal matrix (1.15). We focus particularly on the study and computation of $\Sigma_{\varepsilon_n}^n(V)$, $\Pi_{\varepsilon_n'}^{n,t}(V)$, and $\Gamma_{\varepsilon_n''}^n(V)$, defined in (1.21), (1.24), and (1.29), respectively, that, by (1.22), (1.25), and (1.30), are inclusion sets for Spec A, and the convergence of these sequences of inclusion sets as $n \to \infty$.

8.1. The shift operator

As a first example, we consider the shift operator V from Example 1.11 in more detail. As we noted in Example 1.11, Spec V is the unit circle $\mathbb{T}$ for this operator. Moreover, we can obtain exact analytical descriptions for the inclusion sets for Spec V provided by the τ, π, and τ_1 methods in (1.22), (1.25), and (1.30). These are $\Sigma_{\varepsilon_n}^n(V)$, $\Pi_{\varepsilon_n'}^{n,t}(V)$, and $\Gamma_{\varepsilon_n''}^n(V)$, respectively, defined in (1.21), (1.24), and (1.29). In this case, as we noted in Example 1.11, $\|\alpha\|_\infty = 1$ and $\|\gamma\|_\infty = 0$ so that, by Corollary 3.7, (1.26), and (1.31), $\varepsilon_n = 2\sin\frac{\pi}{4n+2}$, $\varepsilon_n' = 2\sin\frac{\pi}{2n}$, and $\varepsilon_n'' = 2\sin\frac{\pi}{2n+2}$.

In general, computation of these inclusion sets for an operator A requires consideration of infinitely many submatrices of A, indexed by $k \in \mathbb{Z}$. But, as we have noted already in Example 1.11, in the case $A = V$ the relevant $n \times n$ submatrices for the τ and π methods are all the same: we have (see (1.53)) that $A_{n,k} = V_n$ and $A_{n,k}^{\text{per},t} = V_n^{\text{per},t}$, for $k \in \mathbb{Z}$, where V_n and $V_n^{\text{per},t}$ are defined in (1.53), so that, for $n \in \mathbb{N}$,

$$\Sigma_{\varepsilon_n}^n(V) = \text{Spec}_{\varepsilon_n}(V_n) \quad \text{and} \quad \Pi_{\varepsilon_n'}^{n,t}(V) = \text{Spec}_{\varepsilon_n'}(V_n^{\text{per},t}), \quad t \in \mathbb{T}. \tag{8.1}$$

Similarly, recalling the definitions (1.21) and (1.24),

$$\sigma_{\varepsilon_n}^n(V) = \mathrm{spec}_{\varepsilon_n}(V_n) \quad \text{and} \quad \pi_{\varepsilon_n'}^{n,t}(V) = \mathrm{spec}_{\varepsilon_n'}(V_n^{\mathrm{per},t}), \quad t \in \mathbb{T}.$$

For the τ_1 method the relevant submatrices for computing $\gamma_{\varepsilon_n''}^n(A)$ and $\Gamma_{\varepsilon_n''}^n(A)$, given by (1.29), are $A_{n,k}^+$ and $(A^*)_{n,k}^+$, for $k \in \mathbb{Z}$, this notation is defined in (1.32). In the case $A = V$ we have, for $n \in \mathbb{N}$ and $k \in \mathbb{Z}$, that

$$A_{n,k}^+ = V_n^+ := \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ \hline & & V_n & & \\ \hline 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

and

$$(A^*)_{n,k}^+ = (V^*)_n^+ := \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ \hline & & V_n^T & & \\ \hline 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Thus, by (1.29), (1.33), and (7.1), we have

$$\gamma_{\varepsilon_n''}^n(V) = \{\lambda \in \mathbb{C} : v(\lambda) < \varepsilon_n''\} \quad \text{and} \quad \Gamma_{\varepsilon_n''}^n(V) = \{\lambda \in \mathbb{C} : v(\lambda) \le \varepsilon_n''\}, \quad (8.2)$$

where

$$v(\lambda) := \min(v(V_n^+ - \lambda I_n^+), v((V^*)_n^+ - \lambda I_n^+))$$

$$= \min(s_{\min}(V_n^+ - \lambda I_n^+), s_{\min}((V^*)_n^+ - \lambda I_n^+)).$$

With $\mathbb{D}$ and $\mathbb{T}_n$ as defined in Section 1.3, it follows from (8.1) and (8.2) that

$$\Sigma_{\varepsilon_n}^n(V) = \overline{\mathbb{D}}, \tag{8.3}$$

$$\Pi_{\varepsilon_n'}^{n,t}(V) = t^{1/n}\mathbb{T}_n + \varepsilon_n'\overline{\mathbb{D}} \quad \text{for } t \in \mathbb{T}, \tag{8.4}$$

and that

$$\Gamma_{\varepsilon_n''}^n(V) = \begin{cases} \overline{\mathbb{D}} \setminus (1 - (\varepsilon_n'')^2)\mathbb{D}, & n \ge 3, \\ \overline{\mathbb{D}}, & n = 1, 2. \end{cases} \tag{8.5}$$

We have shown (8.4) already in Example 1.11, and we will show the other identities above in a moment. Clearly, as predicted by (1.22), (1.25) and (1.30), these identities show that

$$\mathrm{Spec}\, V \subset \Sigma_{\varepsilon_n}^n(V) \cap \Pi_{\varepsilon_n'}^{n,t}(V) \cap \Gamma_{\varepsilon_n''}^n(V), \quad t \in \mathbb{T},$$

but, as we will see in the calculations below, for $n \in \mathbb{N}$ and $t \in \mathbb{T}$,

$$\mathrm{Spec}\, V \not\subset \sigma_{\varepsilon_n}^n(V), \quad \mathrm{Spec}\, V \not\subset \pi_{\varepsilon_n'}^{n,t}(V), \quad \mathrm{Spec}\, V \not\subset \gamma_{\varepsilon_n''}^n(V), \tag{8.6}$$

since

$$\sigma_{\varepsilon_n}^n(V) = \mathbb{D}, \quad \pi_{\varepsilon_n'}^{n,t}(V) = t^{1/n}\mathbb{T}_n + \varepsilon_n'\mathbb{D},$$

$$\gamma_{\varepsilon_n''}^n(V) = \begin{cases} \mathbb{D} \setminus (1 - (\varepsilon_n'')^2)\overline{\mathbb{D}}, & n \geq 2, \\ \mathbb{D}, & n = 1. \end{cases} \tag{8.7}$$

This makes clear that ε_n, ε_n' and ε_n'' are just as large as necessary for (1.22), (1.25), and (1.30) to hold. Observe also that the above expressions imply that

$$d_H(\operatorname{Spec} V, \Gamma_{\varepsilon_n''}^n(V)) = 4\sin^2\frac{\pi}{2(n+1)} \quad \text{and} \quad d_H(\operatorname{Spec} V, \Pi_{\varepsilon_n'}^n(V)) = 2\sin\frac{\pi}{2n},$$

so that

$$\Gamma_{\varepsilon_n''}^n(V) \xrightarrow{H} \operatorname{Spec} V \quad \text{as } n \to \infty,$$

as predicted by Theorem 1.2, and also $\Pi_{\varepsilon_n'}^n(V) \xrightarrow{H} \operatorname{Spec} V$ (this is a special case of Theorem 8.1 below). However, $d_H(\operatorname{Spec} V, \Gamma_{\varepsilon_n'}^n(V)) = 1$ for all $n \in \mathbb{N}$.

Proof of formulas (8.3)–(8.6). Fix $n \in \mathbb{N}$ and put $R_n(\lambda) := \|(V_n - \lambda I_n)^{-1}\|$ for $\lambda \in \mathbb{C}$ (with $R_n(0) := \infty$), where V_n is as in (1.53). Firstly, R_n is invariant under rotation around 0. Indeed, let $\varphi \in [0, 2\pi)$ and put $D_n := \operatorname{diag}(e^{i\varphi}, e^{i2\varphi}, \ldots, e^{in\varphi})$. Then, D_n is an isometry and $V_n D_n = e^{-i\varphi} D_n V_n$, so that

$$R_n(e^{i\varphi}\lambda) = \|(V_n - e^{i\varphi}\lambda I_n)^{-1}\| = \|e^{-i\varphi}(e^{-i\varphi}V_n - \lambda I_n)^{-1}\|$$
$$= \|(D_n^{-1}V_n D_n - \lambda I_n)^{-1}\| = \|D_n^{-1}(V_n - \lambda I_n)^{-1}D_n\| = R_n(\lambda).$$

Secondly, R_n is strictly monotonously decreasing on the half-axis $(0, +\infty)$. Indeed, suppose $0 < \lambda_1 < \lambda_2 < \infty$ and $R_n(\lambda_1) \leq R_n(\lambda_2)$. Since $\lim_{\lambda\to\infty} R_n(\lambda) = 0$, there is a $\lambda_3 \geq \lambda_2$ with $R_n(\lambda_3) = R_n(\lambda_1)$. Thus, R_n attains its supremum

$$M := \max_{\lambda \in [\lambda_1, \lambda_3]} R_n(\lambda) \geq R_n(\lambda_2) \geq R_n(\lambda_1) = R_n(\lambda_3)$$

in the open disc

$$U := \frac{\lambda_1 + \lambda_3}{2} + \frac{\lambda_3 - \lambda_1}{2}\mathbb{D},$$

which contradicts the maximum principle of [46], noted in [80, Theorem 2.2]. Finally, recalling (7.1) and the notations B_n, $E_n(\phi)$, and $\varrho_n(\phi)$ of Section 3.2, $R_n(1) = \|B_n^{-1}\|$ and $1/\|B_n^{-1}\|^2 = (\nu(B_n))^2 = \min \operatorname{Spec}(B_n^T B_n) = \min \operatorname{Spec}(E_n(1)) = \varrho_n(1)$. Further, by (3.31) and (3.36), $\varrho_n(1) = \varepsilon_n^2$. Thus, $R_n(1) = 1/\varepsilon_n$. Taking all of this together, we get that

$$\Sigma_{\varepsilon_n}^n(V) = \{\lambda \in \mathbb{C} : R_n(\lambda) > 1/\varepsilon_n\} = \overline{\mathbb{D}}$$

and

$$\sigma_{\varepsilon_n}^n(V) = \mathbb{D}.$$

We have shown in Example 1.11 that $V_n^{\mathrm{per},t}$ is normal and $\mathrm{Spec}\, V_n^{\mathrm{per},t} = t^{1/n}\mathbb{T}_n$, so that

$$\Pi_\varepsilon^{n,t}(V) = \mathrm{Spec}_\varepsilon V_n^{\mathrm{per},t} = \mathrm{Spec}\, V_n^{\mathrm{per},t} + \varepsilon\overline{\mathbb{D}} = t^{1/n}\mathbb{T}_n + \varepsilon\overline{\mathbb{D}}, \quad \varepsilon > 0. \qquad (8.8)$$

Clearly, for each $t \in \mathbb{T}$, $\Pi_\varepsilon^{n,t}(V)$ covers $\mathrm{Spec}\, V = \mathbb{T}$ if and only if it covers $t^{1/n}\mathbb{T}_{2n}$. The latter is the case if and only if ε is greater than or equal to the Euclidean distance of two adjacent points of $\mathbb{T}_{2n}$, which is $2\sin\frac{\pi}{2n} = \varepsilon_n'$. Thus, $\mathrm{Spec}\, V \subset \Pi_{\varepsilon_n'}^{n,t}(V)$, but $\mathrm{Spec}\, V \not\subset \pi_{\varepsilon_n'}^{n,t}(V) = t^{1/n}\mathbb{T}_n + \varepsilon_n'\mathbb{D}$, since the points $t^{1/n}\mathfrak{S}_n$, where $\mathfrak{S}_n := \mathbb{T}_{2n} \setminus \mathbb{T}_n$, which are points of intersection of the discs of radius ε_n' centred on $\mathbb{T}_n$, are not contained in $\pi_{\varepsilon_n'}^{n,t}(V)$. Nor does $\pi_{\varepsilon_n'}^{n,t}(V)$ contain the other points of intersection of these discs, the points $rt^{1/n}\mathfrak{S}_n$, with $r = 1 - (\varepsilon_n')^2$, or any of the points $rt^{1/n}\mathfrak{S}_n$, with $0 \le r \le 1 - (\varepsilon_n')^2$ or with $r \ge 1$. On the other hand, by elementary geometry, $\pi_{\varepsilon_n'}^{n,t}(V)$ contains the annulus $\mathbb{D} \setminus (1 - (\varepsilon_n')^2)\overline{\mathbb{D}}$. Thus,

$$\bigcap_{t\in\mathbb{T}} \pi_{\varepsilon_n'}^{n,t}(V) = \bigcap_{t\in\mathbb{T}}(t^{1/n}\mathbb{T}_n + \varepsilon_n'\mathbb{D}) = \mathbb{D} \setminus (1 - (\varepsilon_n')^2)\overline{\mathbb{D}},$$

and, by (4.10), also $\bigcap_{t\in\mathbb{T}} \Pi_{\varepsilon_n'}^{n,t}(V) = \overline{\mathbb{D}} \setminus (1 - (\varepsilon_n')^2)\mathbb{D}$.

To see that (8.5) holds, note that, for $n \in \mathbb{N}$, recalling (7.1), and where $D_n := I_n$ if $\lambda = 0$, $D_n := \mathrm{diag}(w^1, \ldots, w^n)$ with $w := |\lambda|/\overline{\lambda}$ if $\lambda \ne 0$,

$$\begin{aligned}
(\nu(V_n^+ - \lambda I_n^+))^2 &= (\nu((V - \lambda I)|_{E_{n,0}}))^2 = \min \mathrm{Spec}(P_{n,0}(V - \lambda I)'(V - \lambda I)|_{E_{n,0}}) \\
&= \min \mathrm{Spec}((1 + |\lambda|^2)I_n - \lambda V_n^T - \overline{\lambda} V_n) \\
&= \min \mathrm{Spec}(D_n^{-1}((1 + |\lambda|^2)I_n - |\lambda|(V_n^T + V_n))D_n) \\
&= \min \mathrm{Spec}((1 + |\lambda|^2)I_n - |\lambda|(V_n^T + V_n)) \\
&= 1 + |\lambda|^2 - 2|\lambda|c_n,
\end{aligned}$$

by (5.4), where $c_n := \cos\frac{\pi}{n+1} = 1 - \frac{1}{2}(\varepsilon_n'')^2$. Similarly,

$$\begin{aligned}
(\nu((V^*)_n^+ - \lambda I_n^+))^2 &= \min \mathrm{Spec}(P_{n,0}(V' - \lambda I)'(V' - \lambda I)|_{E_{n,0}}) \\
&= \min \mathrm{Spec}((1 + |\lambda|^2)I_n - \lambda V_n - \overline{\lambda} V_n^T) \\
&= \min \mathrm{Spec}((1 + |\lambda|^2)I_n - |\lambda|(V_n^T + V_n)) \\
&= 1 + |\lambda|^2 - 2|\lambda|c_n,
\end{aligned}$$

so that

$$v(\lambda) = (1 + |\lambda|^2 - 2|\lambda|c_n)^{1/2}, \quad \lambda \in \mathbb{C}.$$

Now, $\lambda \in \Gamma_{\varepsilon_n''}^n(V)$ if and only if $v(\lambda) \le \varepsilon_n''$, so if and only if

$$(1 - |\lambda|)^2 \le (\varepsilon_n'')^2(1 - |\lambda|). \qquad (8.9)$$

If $\varepsilon_n'' \geq 1$, which is the case for $n = 1$ and 2, then (8.9) holds if and only if $|\lambda| \leq 1$. If $\varepsilon_n'' < 1$, which is the case for $n \geq 3$, then (8.9) holds if and only if $1 - (\varepsilon_n'')^2 \leq |\lambda| \leq 1$. Similarly, $\lambda \in \gamma_{\varepsilon_n''}^n(V)$ if and only if $\upsilon(\lambda) < \varepsilon_n''$, which holds if and only if $|\lambda| < 1$ for $n = 1$, if and only if $1 - (\varepsilon_n'')^2 < |\lambda| < 1$, for $n \geq 2$ (note that $\varepsilon_2'' = 1$).   ∎

8.2. Block-Laurent operators

Suppose now, generalising the shift operator example of Section 8.1, that $X = \mathbb{C}^p$, for some $p \in \mathbb{N}$, and that A is given by (1.15) with each diagonal constant, so that, for some $\hat{\alpha}, \hat{\beta}, \hat{\gamma} \in L(X) = \mathbb{C}^{p \times p}$, $\alpha_j = \hat{\alpha}$, $\beta_j = \hat{\beta}$, and $\gamma_j = \hat{\gamma}$, for $j \in \mathbb{Z}$, i.e., A is a *block-Laurent* operator (a *Laurent* operator in the case $p = 1$). Generalising the notation of Section 8.1, let $V \in L(E)$ denote the forward shift operator, so that $(Vx)_j = x_{j+1}$, for $x \in E = \ell^2(\mathbb{Z}, X)$ and $j \in \mathbb{Z}$. Then[11],

$$A = \hat{\alpha} V^1 + \hat{\beta} V^0 + \hat{\gamma} V^{-1} = a(V),$$

where a is the *Laurent polynomial*

$$a(z) := \hat{\alpha} z^1 + \hat{\beta} z^0 + \hat{\gamma} z^{-1}, \quad z \in \mathbb{T}.$$

It is well known that A is unitarily equivalent to multiplication by a, the *symbol* of A. Precisely (see, e.g., Theorem 4.4.9 and its proof in [33]), the Fourier operator $\mathcal{F} : E = \ell^2(\mathbb{Z}, X) \to L^2(\mathbb{T}, X)$, defined by

$$(\mathcal{F} x)(z) = \frac{1}{\sqrt{2\pi}} \sum_{j \in \mathbb{Z}} x_j z^{-j}, \quad z \in \mathbb{T},$$

is unitary, and

$$A = \mathcal{F}^{-1} M_a \mathcal{F},$$

where $M_a : L^2(\mathbb{T}, X) \to L^2(\mathbb{T}, X)$ is operation of multiplication by $a \in C(\mathbb{T}, L(X))$, i.e., $(M_a \phi)(z) = a(z)\phi(z)$, for $\phi \in L^2(\mathbb{T}, X)$, $z \in \mathbb{T}$. Thus, [33, Theorem 4.4.9]

$$\text{Spec } A = \text{Spec } M_a = \bigcup_{z \in \mathbb{T}} \text{Spec } a(z), \tag{8.10}$$

and, similarly,

$$\text{spec}_\varepsilon A = \text{spec}_\varepsilon M_a = \bigcup_{z \in \mathbb{T}} \text{spec}_\varepsilon a(z),$$

$$\text{Spec}_\varepsilon A = \text{Spec}_\varepsilon M_a = \bigcup_{z \in \mathbb{T}} \text{Spec}_\varepsilon a(z), \quad \varepsilon > 0. \tag{8.11}$$

---

[11]In the following and similar expressions $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are to be understood as operating pointwise on the elements of a given $x = (x_i)_{i \in \mathbb{Z}} \in E$, so that, e.g., $(\hat{\alpha} x)_i := \hat{\alpha} x_i$, for $i \in \mathbb{Z}$.

In this block-Laurent case similar spectral computations hold for the matrices $A_{n,k}^{\mathrm{per},t}$, given by (1.20), that arise in the π method. For $n \in \mathbb{N}$, $t \in \mathbb{T}$, and all $k \in \mathbb{Z}$,

$$A_{n,k}^{\mathrm{per},t} = \hat{\alpha} V_{n,t}^1 + \hat{\beta} V_{n,t}^0 + \hat{\gamma} V_{n,t}^{-1} = a(V_{n,t}),$$

where

$$V_{n,t} := \begin{pmatrix} 0 & & & tI_X \\ I_X & 0 & & \\ & \ddots & \ddots & \\ & & I_X & 0 \end{pmatrix}_{n \times n}.$$

A simple computation shows that, for any $s \in \mathbb{T}$ with $s^n = t$ (cf. the calculation in Example 1.11),

$$V_{n,t} = sD_n^{-1} V_{n,1} D_n \quad \text{with} \quad D_n := \mathrm{diag}(s^1 I_X, s^2 I_X, \ldots, s^n I_X),$$

so that

$$A_{n,k}^{\mathrm{per},t} = a(V_{n,t}) = a(sD_n^{-1} V_{n,1} D_n) = D_n^{-1} a(sV_{n,1}) D_n, \qquad (8.12)$$

so that $A_{n,k}^{\mathrm{per},t}$ is unitarily equivalent to $a(sV_{n,1})$. Moreover (see, e.g., [35] and [85, Section II.7]), $V_{n,1}$ can be diagonalised as $V_{n,1} = \mathscr{F}_n^{-1} M_n \mathscr{F}_n$, where $\mathscr{F}_n : X^n \to X^n$ is the discrete Fourier transform

$$\mathscr{F}_n = \frac{1}{\sqrt{n}} (\omega^{jk} I_X)_{j,k=1}^n \quad \text{and} \quad M_n := \mathrm{diag}(\omega^1 I_X, \omega^2 I_X, \ldots, \omega^n I_X)$$

with $\omega := \exp(2\pi i/n)$. Combining this observation with (8.12), we see that

$$A_{n,k}^{\mathrm{per},t} = D_n^{-1} \mathscr{F}_n^{-1} a(sM_n) \mathscr{F}_n D_n \quad \text{with} \quad a(sM_n) = \mathrm{diag}(a(s\,\omega^1), \ldots, a(s\,\omega^n)),$$
$$(8.13)$$

in particular that $A_{n,k}^{\mathrm{per},t}$ is unitarily equivalent to $a(sM_n)$. Since $\{\omega^1, \omega^2, \ldots, \omega^n\} = \{z \in \mathbb{C} : z^n = 1\} = \mathbb{T}_n$, we see, similarly to (8.10) and (8.11), that

$$\mathrm{Spec}\, A_{n,k}^{\mathrm{per},t} = \mathrm{Spec}\, a(sM_n) = \bigcup_{z \in s\mathbb{T}_n} \mathrm{Spec}\, a(z), \qquad (8.14)$$

and that for $\varepsilon > 0$,

$$\mathrm{spec}_\varepsilon A_{n,k}^{\mathrm{per},t} = \bigcup_{z \in s\mathbb{T}_n} \mathrm{spec}_\varepsilon a(z) \quad \text{and} \quad \mathrm{Spec}_\varepsilon A_{n,k}^{\mathrm{per},t} = \bigcup_{z \in s\mathbb{T}_n} \mathrm{Spec}_\varepsilon a(z). \qquad (8.15)$$

Note that the set $s\mathbb{T}_n$ is independent of which root s we choose as the solution of $s^n = t$.

The above observations lead, applying Theorem 1.4, to the following result about convergence of the π method in the case that A is block-Laurent. Of course, we

already have, as a consequence of Theorem 1.2, that the τ_1 method is convergent (recall that X has Globevnik's property as it is finite-dimensional), and we know, from the example in Section 8.1, that the τ method is not convergent for tridiagonal block–Laurent operators in general, even in the special case that $X = \mathbb{C}$.

Theorem 8.1. *If $X = \mathbb{C}^p$, equipped with the Euclidean norm, and $A \in L(E)$ is a tridiagonal block-Laurent operator, then, for every $k \in \mathbb{Z}$, $n \in \mathbb{N}$, and $t \in \mathbb{T}$, $A_{n,k}^{\mathrm{per},t} = A_{n,0}^{\mathrm{per},t}$,*

$$\Pi_\varepsilon^{n,t}(A) = \mathrm{Spec}_\varepsilon A_{n,0}^{\mathrm{per},t} \quad for\ \varepsilon \geq 0 \quad and \quad \pi_\varepsilon^{n,t}(A) = \mathrm{spec}_\varepsilon A_{n,0}^{\mathrm{per},t} \quad for\ \varepsilon > 0. \quad (8.16)$$

Further, the π method does not suffer from spectral pollution for A for all $t \in \mathbb{T}$; indeed, for all $n \in \mathbb{N}$ and $\varepsilon > 0$,

$$\mathrm{Spec}\, A = \bigcup_{t \in \mathbb{T}} \mathrm{Spec}\, A_{n,0}^{\mathrm{per},t}, \quad \mathrm{spec}_\varepsilon A = \bigcup_{t \in \mathbb{T}} \mathrm{spec}_\varepsilon A_{n,0}^{\mathrm{per},t},$$
$$\mathrm{Spec}_\varepsilon A = \bigcup_{t \in \mathbb{T}} \mathrm{Spec}_\varepsilon A_{n,0}^{\mathrm{per},t}, \quad (8.17)$$

so that, as $n \to \infty$,

$$\Pi_{\varepsilon+\varepsilon_n'}^{n,t}(A) = \mathrm{Spec}_{\varepsilon+\varepsilon_n'} A_{n,0}^{\mathrm{per},t} \xrightarrow{H} \mathrm{Spec}_\varepsilon A \quad for\ \varepsilon \geq 0,$$

in particular, $\Pi_{\varepsilon_n'}^{n,t}(A) = \mathrm{Spec}_{\varepsilon_n'} A_{n,0}^{\mathrm{per},t} \xrightarrow{H} \mathrm{Spec}\, A$, and

$$\pi_{\varepsilon+\varepsilon_n'}^{n,t}(A) = \mathrm{spec}_{\varepsilon+\varepsilon_n'} A_{n,0}^{\mathrm{per},t} \xrightarrow{H} \mathrm{spec}_\varepsilon A \quad for\ \varepsilon > 0.$$

Proof. The fact that $A_{n,k}^{\mathrm{per},t} = A_{n,0}^{\mathrm{per},t}$, for $k \in \mathbb{Z}$, and the representations (8.16), are clear from the definitions (1.20) and (1.24) of $A_{n,k}^{\mathrm{per},t}$, $\Pi_\varepsilon^{n,t}(A)$, and $\pi_\varepsilon^{n,t}(A)$. The fact that (8.17) holds follows by comparing (8.10) and (8.11) with (8.14) and (8.15), noting that $\cup_{t \in \mathbb{T}}(s\mathbb{T}_n) = \mathbb{T}$. The second of the equalities in (8.17) implies, by Definition 1.1, that, for every $t \in \mathbb{T}$, the π method does not suffer from spectral pollution for A. The remaining results follow from (8.16) and Theorem 1.4, noting that X satisfies Globevnik's property (see Section 2) since X is finite-dimensional. ∎

8.3. Periodic tridiagonal matrices

Suppose A is given by (1.15) with $X = \mathbb{C}$ so that $E = \ell^2(\mathbb{Z})$, with all three diagonals p-periodic for some $p \in \mathbb{N}$. Then, A can be understood as a block-Laurent operator, so as to make use of the results from the previous section. Precisely, as in our discussion around (1.14), define $E_p := \ell^2(\mathbb{Z}, X^p)$, equipping $X^p = \mathbb{C}^p$ with the Euclidean norm (1.14), and let $A_p := \mathcal{I}_p^{-1} V^k A V^{-k} \mathcal{I}_p$, for some $k \in \mathbb{Z}$, where V is the shift operator

as defined in Section 8.1. Then, $A_p \in L(E_p)$ is a tridiagonal block-Laurent operator with constant entries on the diagonals, $\hat{\alpha}, \hat{\beta}, \hat{\gamma} \in L(X) = \mathbb{C}^{p \times p}$, where

$$(\hat{\alpha} \mid \hat{\beta} \mid \hat{\gamma}) = \begin{pmatrix} \alpha_{p+k} & \beta_{1+k} & \gamma_{2+k} & & & \\ & \alpha_{1+k} & \beta_{2+k} & \ddots & & \\ & & \ddots & \ddots & \gamma_{p+k} & \\ & & & \alpha_{p+k-1} & \beta_{p+k} & \gamma_{1+k} \end{pmatrix}_{p \times 3p}.$$

Noting that the $n \times n$ matrix $(A_p)_{n,0}^{\mathrm{per},t}$ with entries in $\mathbb{C}^{p \times p}$ is identical to the $(np) \times (np)$ matrix $A_{np,k}^{\mathrm{per},t}$, we have the following corollary of Theorem 8.1. The first equation of (8.18) is well known (see [10, 11, 33]). In the case $n = 1$ it expresses Spec A as the union over t of the eigenvalues of the $p \times p$ matrix Spec $A_{p,k}^{\mathrm{per},t}$.

Corollary 8.2. *Suppose that* $X = \mathbb{C}$ *and* $A \in L(E)$ *is tridiagonal with, for some* $p \in \mathbb{N}$, $\alpha_{j+p} = \alpha_j$, $\beta_{j+p} = \beta_j$, $\gamma_{j+p} = \gamma_j$, *for* $j \in \mathbb{Z}$. *Then, for every* $k \in \mathbb{Z}$, $n \in \mathbb{N}$, *and* $\varepsilon > 0$,

$$\mathrm{Spec}\, A = \bigcup_{t \in \mathbb{T}} \mathrm{Spec}\, A_{pn,k}^{\mathrm{per},t}, \quad \mathrm{spec}_\varepsilon A = \bigcup_{t \in \mathbb{T}} \mathrm{spec}_\varepsilon A_{pn,k}^{\mathrm{per},t},$$
$$\mathrm{Spec}_\varepsilon A = \bigcup_{t \in \mathbb{T}} \mathrm{Spec}_\varepsilon A_{pn,k}^{\mathrm{per},t}. \tag{8.18}$$

Further, as $n \to \infty$, *if* $k \in \mathbb{Z}$ *and* $t \in \mathbb{T}$, *then*

$$\mathrm{Spec}_{\varepsilon+\varepsilon_n'} A_{pn,k}^{\mathrm{per},t} \xrightarrow{H} \mathrm{Spec}_\varepsilon A \quad \text{for } \varepsilon \geq 0, \quad \text{in particular} \quad \mathrm{Spec}_{\varepsilon_n'} A_{pn,k}^{\mathrm{per},t} \xrightarrow{H} \mathrm{Spec}\, A,$$

and

$$\mathrm{spec}_{\varepsilon+\varepsilon_n'} A_{pn,k}^{\mathrm{per},t} \xrightarrow{H} \mathrm{spec}_\varepsilon A \quad \text{for } \varepsilon > 0.$$

The above result implies that, for $\varepsilon \geq 0$ and $t \in \mathbb{T}$,

$$\Pi_{\varepsilon+\varepsilon_n'}^{pn,t}(A) = \bigcup_{k=1}^{p} \mathrm{Spec}_{\varepsilon+\varepsilon_n'} A_{pn,k}^{\mathrm{per},t} \xrightarrow{H} \mathrm{Spec}_\varepsilon A,$$

as $n \to \infty$, and similarly that $\pi_{\varepsilon+\varepsilon_n'}^{pn,t}(A) \xrightarrow{H} \mathrm{spec}_\varepsilon A$, for $\varepsilon > 0$. A natural question is: does $\Pi_{\varepsilon+\varepsilon_n'}^{n,t}(A) \xrightarrow{H} \mathrm{Spec}_\varepsilon A$, as $n \to \infty$, for $\varepsilon \geq 0$? I.e., does the whole sequence converge and not just the subsequence $\Pi_{\varepsilon+\varepsilon_n'}^{pn,t}(A)$? The answer is not necessarily!

As an example where the whole sequence does not converge, suppose that A has p-periodic diagonals with $p = 2$, and that the entries of A are chosen so that the block-Laurent operator $A_p = \mathcal{I}_p^{-1} A \mathcal{I}_p$ has $\hat{\alpha} = \hat{\gamma} = 0$ and

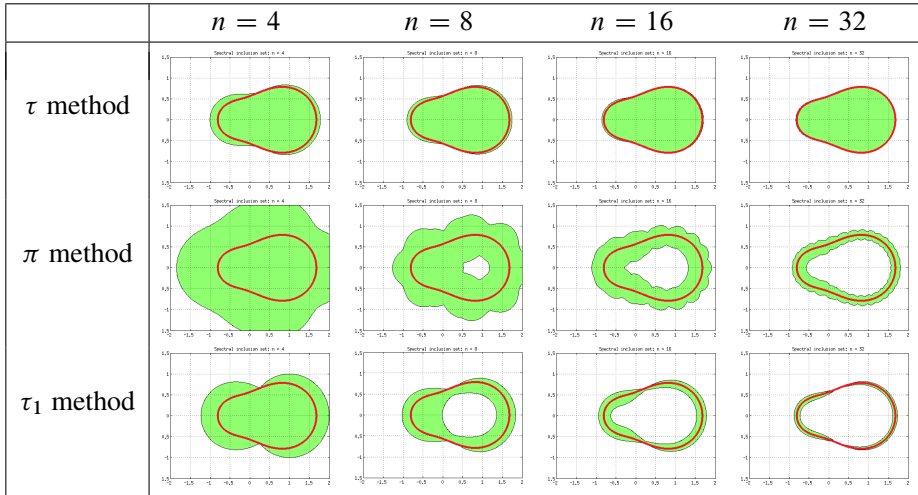$$\hat{\beta} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$
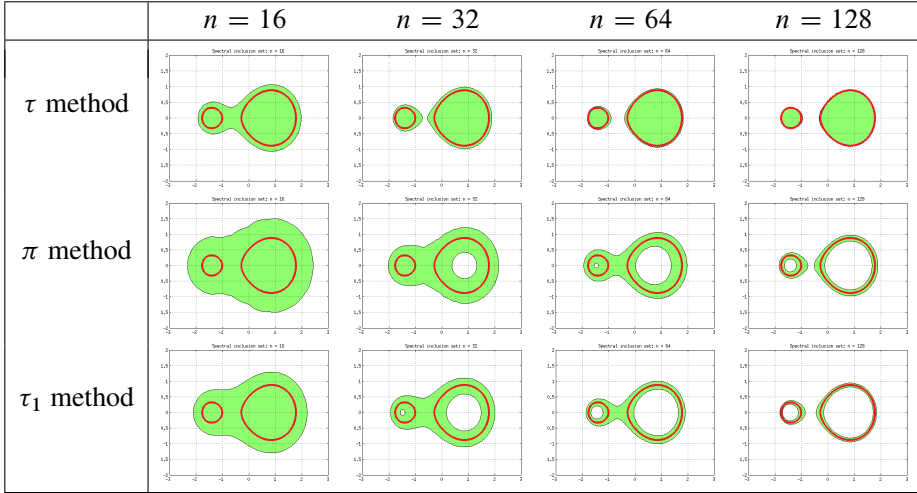
Figure 8.1. The τ, π, and τ_1 inclusion sets, $\Sigma_{\varepsilon_n}^n(A)$, $\Pi_{\varepsilon_n'}^{n,1}(A)$, and $\Gamma_{\varepsilon_n''}^n(A)$, respectively, for $n = 16, 32, 64$, and 128, in the case that $\alpha = (\cdots, 0, 0, 0, \cdots)$, $\beta = (\cdots, -\frac{3}{2}, 1, 1, \cdots)$ and $\gamma = (\cdots, 1, 2, 1, \cdots)$ are 3-periodic. These are each inclusion sets for Spec A, given by the first equation of (8.18) with $n = 1$, and shown in each figure as the red curve.

so that A_2 is block diagonal and A_2 and A are self-adjoint. Then, Spec $A =$ Spec $A_2 =$ Spec $\hat{\beta} = \{-1, 1\}$; indeed, since A is normal, we have further that $\mathrm{Spec}_\varepsilon A = \mathrm{Spec}\, A + \varepsilon\overline{\mathbb{D}} = \{-1, 1\} + \varepsilon\overline{\mathbb{D}}$, for $\varepsilon \geq 0$. But, if n is odd, then, for $t \in \mathbb{T}$, the last column of $A_{n,0}^{\mathrm{per},t}$ is zero, so that $0 \in \mathrm{Spec}\, A_{n,0}^{\mathrm{per},t}$, so that (see Section 2) $\varepsilon\overline{\mathbb{D}} \subset \mathrm{Spec}_\varepsilon A_{n,0}^{\mathrm{per},t}$, for $\varepsilon \geq 0$. Thus, $(\varepsilon + \varepsilon_n')\overline{\mathbb{D}} \subset \Pi_{\varepsilon+\varepsilon_n'}^{n,t}(A)$, for $\varepsilon \geq 0$, in particular, $(\varepsilon + \varepsilon_n')\mathrm{i} \in \Pi_{\varepsilon+\varepsilon_n'}^{n,t}(A)$. Thus, for $\varepsilon \geq 0$,

$$d_H(\Pi_{\varepsilon+\varepsilon_n'}^{n,t}(A), \mathrm{Spec}_\varepsilon A) \geq \mathrm{dist}((\varepsilon + \varepsilon_n')\mathrm{i}, \mathrm{Spec}_\varepsilon A)$$
$$= \sqrt{1 + (\varepsilon + \varepsilon_n')^2} - \varepsilon \geq \sqrt{1 + \varepsilon^2} - \varepsilon,$$

so $\Pi_{\varepsilon+\varepsilon_n'}^{n,t}(A)$ does not converge to $\mathrm{Spec}_\varepsilon A$ as $n \to \infty$.

On the other hand, the numerical results in Figures 8.1 and 8.2, both examples with $p = 3$, suggest that there are periodic operators for which the whole sequence does converge. It appears in both figures that $\Pi_{\varepsilon_n'}^{n,t}(A) \xrightarrow{H} \mathrm{Spec}\, A$ as $n \to \infty$ (note that the values of n we choose are not multiples of 3). Note also, as predicted by Theorem 1.2, that $\Gamma_{\varepsilon_n''}^n \xrightarrow{H} \mathrm{Spec}\, A$ as $n \to \infty$. Further, it appears in these two examples (and also in Example 1.11) that $\Sigma_{\varepsilon_n}^n(A)$, the τ method inclusion set, converges to $\widehat{\mathrm{Spec}}\, A$, as $n \to \infty$, where $\widehat{\mathrm{Spec}}\, A$ denotes the complement of the unbounded component of $\mathbb{C} \setminus \mathrm{Spec}\, A$.

Figure 8.2. The τ, π, and τ_1 inclusion sets, $\Sigma^n_{\varepsilon_n}(A)$, $\Pi^{n,1}_{\varepsilon'_n}(A)$, and $\Gamma^n_{\varepsilon''_n}(A)$, respectively, for $n = 4, 8, 16$, and 32, in the case that $\alpha = (\cdots, 0, 0, 0, \cdots)$, $\beta = (\cdots, -\frac{1}{2}, 1, 1, \cdots)$ and $\gamma = (\cdots, 1, 1, 1, \cdots)$ are 3-periodic. These are each inclusion sets for Spec A, given by the first equation of (8.18) with $n = 1$, and shown in each figure as the red curve.

Figures 8.1 and 8.2 are plotted in Matlab, using Matlab's inbuilt contouring routine. In particular, the plots of $\Sigma^n_{\varepsilon_n}(A)$ are produced by noting that the boundary of $\Sigma^n_{\varepsilon_n}(A)$ is, by Proposition 3.1 and (3.5), and recalling (7.5), the contour on which

$$\mu^\dagger_n(A - \lambda I) = \min_{k=1,\ldots,p} \nu(A_{n,k} - \lambda I_n)$$
$$= \left(\min_{k=1,\ldots,p} \min \text{Spec}((A_{n,k} - \lambda I_n)'(A_{n,k} - \lambda I_n)) \right)^{1/2} \quad (8.19)$$

has the value ε_n. This contour is determined by Matlab's contouring routine, interpolating values of the right-hand side of (8.19) plotted on a fine grid of values of λ covering the respective regions shown in Figures 8.1 and 8.2. $\Pi^{n,1}_{\varepsilon'_n}(A)$ and $\Gamma^n_{\varepsilon''_n}(A)$ are plotted similarly, starting from the representations (7.6) and (7.7) for $\Pi^{n,1}_{\varepsilon'_n}(A)$, and (1.29) and (1.33) for $\Gamma^n_{\varepsilon''_n}(A)$.

8.4. Pseudoergodic and random operators

Let Σ_α, Σ_β and Σ_γ be nonempty compact subsets of $\mathbb{C}$ and let A be of the form (1.15) with $\alpha_j \in \Sigma_\alpha$, $\beta_j \in \Sigma_\beta$ and $\gamma_j \in \Sigma_\gamma$ for all $j \in \mathbb{Z}$. Following Davies [31] (and see [19] or [21, Definition 1.1]) we say that A is *pseudoergodic with respect to* Σ_α, Σ_β, *and* Σ_γ if, for all $\varepsilon > 0$, $n \in \mathbb{N}$ and all tridiagonal matrices $B_n \in \mathbb{C}^{n \times n}$ with

subdiagonal entries in Σ_α, main diagonal entries in Σ_β, and superdiagonal entries in Σ_γ, there exists a $k \in \mathbb{Z}$ such that $\|A_{n,k} - B_n\| < \varepsilon$. We say that A is *pseudoergodic* if, for some non-empty compact Σ_α, Σ_β, and $\Sigma_\gamma \subset \mathbb{C}$, A is pseudoergodic with respect to Σ_α, Σ_β, and Σ_γ. Pseudoergodicity was introduced by Davies [31] to study spectral properties of random operators while eliminating probabilistic arguments. Indeed (see [31, Section 5]), if the matrix entries in (1.15) are independent (or at least not fully correlated) random variables with probability measures whose supports are Σ_α, Σ_β, and Σ_γ, then by the usual probabilistic arguments (sometimes termed the *infinite monkey theorem*) the operator A is pseudoergodic almost surely.

The study of random and pseudoergodic operators, even restricting attention to the tridiagonal case that we consider here, has a large literature, both in the self-adjoint and, more recently, in the non-self-adjoint case. A readable, short introduction is [85, Sections 36 and 37], and see [19] and the references therein; see also [64, Section 3.4.10] or [14] for pseudoergodicity interpreted in the language of limit operators. Falling within the scope of the results of this section are the following.

(1) Random (self-adjoint, bi-infinite) Jacobi operators, including discrete versions of Schrödinger operators with random bounded potentials (so $\Sigma_\alpha = \Sigma_\gamma = \{1\}$ and $\Sigma_\beta \subset \mathbb{R}$), including cases exhibiting Anderson localisation (see, e.g., [73] and the references therein).

(2) The (bi-infinite) non-self-adjoint version of the Anderson model, with applications to superconductors and biological growth problems [1,57–59,70], introduced by Hatano and Nelson [57–59]. In this case Σ_α and Σ_γ are real singleton sets (the α and γ diagonals are constant) and $\Sigma_\beta \subset \mathbb{R}$. For this model the behaviour of $\operatorname{Spec} A_{n,k}^{\mathrm{per},1}$ as $n \to \infty$ has been described by Goldsheid and Khoruzhenko [48, 49], $\operatorname{Spec} A$ in this case has been studied by a variety of methods in Davies [30–32] and Martínez [68, 69]. Moreover, approximations to $\operatorname{Spec}_\varepsilon A$ have been computed by Colbrook [21], via the representation he establishes that $\operatorname{Spec}_\varepsilon A = \lim_{n \to \infty} \operatorname{Spec}_\varepsilon A_{2n+1,-n-1}^{\mathrm{per},1}$.

(3) The so-called Feinberg–Zee random hopping matrix, the case $\Sigma_\alpha = \{-1, 1\}$, $\Sigma_\beta = \{0\}$, $\Sigma_\gamma = \{1\}$, which has been studied in [1,13,14,16,17,24,39,40,52, 53,62], and which we will return to below; also, its generalisation to the case $\Sigma_\alpha = \{\pm\sigma\}$, for some $\sigma \in (0, 1]$, in Chandler-Wilde and Davies [16] (and see [1, 24]).

(4) The special case when A is bidiagonal, meaning that $\Sigma_\alpha = \{0\}$ or $\Sigma_\gamma = \{0\}$; see [65, 84].

Given non-empty compact sets Σ_α, Σ_β, $\Sigma_\gamma \subset \mathbb{C}$, let $M(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma)$ denote the set of operators A of the form (1.15) such that $\alpha_j \in \Sigma_\alpha$, $\beta_j \in \Sigma_\beta$, $\gamma_j \in \Sigma_\gamma$, for $j \in \mathbb{Z}$. The key result in the theory of pseudoergodic operators is the following: if A

is pseudoergodic with respect to Σ_α, Σ_β, and Σ_γ, then

$$
\begin{aligned}
\operatorname{Spec}_\varepsilon A &= \bigcup_{B \in M(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma)} \operatorname{Spec}_\varepsilon B, \quad \varepsilon \geq 0, \\
\operatorname{spec}_\varepsilon A &= \bigcup_{B \in M(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma)} \operatorname{spec}_\varepsilon B, \quad \varepsilon > 0.
\end{aligned}
\tag{8.20}
$$

The first of these equalities in the case $\varepsilon = 0$ dates back to [31]; the second is shown analogously as [19, Corollary 4.12]; the first for $\varepsilon > 0$ then follows by taking closures in the second identity[12].

Clearly, (8.20) implies that if $B \in M(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma)$ then

$$
\operatorname{Spec}_\varepsilon B \subset \operatorname{Spec}_\varepsilon A,
$$

for $\varepsilon \geq 0$, and $\operatorname{spec}_\varepsilon B \subset \operatorname{spec}_\varepsilon A$, for $\varepsilon > 0$. In particular, this is true if B is a Laurent operator, i.e., if, for some $\hat\alpha \in \Sigma_\alpha$, $\hat\beta \in \Sigma_\beta$, and $\hat\gamma \in \Sigma_\gamma$, $\alpha_j = \hat\alpha$, $\beta_j = \hat\beta$, $\gamma_j = \hat\gamma$, for $j \in \mathbb{Z}$, in which case, by (8.17) with $n = 1$,

$$
\operatorname{Spec} B = \mathcal{E}(\hat\alpha, \hat\beta, \hat\gamma) := \{z\hat\alpha + \hat\beta + z^{-1}\hat\gamma : z \in \mathbb{T}\},
$$

which is an ellipse in the complex plane (degenerating to a line if $|\hat\alpha| = |\hat\gamma|$). It will be key to the proof that the π method does not suffer from spectral pollution, in Theorem 8.3 below, that also $\operatorname{spec}_\varepsilon B \subset \operatorname{spec}_\varepsilon A$, if $B \in M(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma)$ and the coefficients α, β, and γ of B are periodic, so that Corollary 8.2 applies.

Let $\mathcal{E}_\cap$ denote the set of $z \in \mathbb{C}$ which are in the interior of $\mathcal{E}(\hat\alpha, \hat\beta, \hat\gamma)$, for every $\hat\alpha \in \Sigma_\alpha$, $\hat\beta \in \Sigma_\beta$, and $\hat\gamma \in \Sigma_\gamma$, and let

$$
\alpha_* := \min_{z \in \Sigma_\alpha} |z|, \quad \alpha^* := \max_{z \in \Sigma_\alpha} |z|, \quad \gamma_* := \min_{z \in \Sigma_\gamma} |z|, \quad \gamma^* := \max_{z \in \Sigma_\gamma} |z|. \tag{8.21}
$$

The main theoretical result of this section is the following. Recall, as is the case for any tridiagonal $A \in L(E)$ (irrespective of pseudoergodicity), that, in addition to the results shown in this theorem, the inclusions (1.22), (1.25), and (1.30) apply to A.

Theorem 8.3. *Suppose that $X = \mathbb{C}$, so that $E = \ell^2(\mathbb{Z})$, that Σ_α, Σ_β, $\Sigma_\gamma \subset \mathbb{C}$ are non-empty and compact, and that A is pseudoergodic with respect to Σ_α, Σ_β, and*

---

[12]In more detail, taking closures in the second identity, since $\overline{\operatorname{spec}_\varepsilon B} = \operatorname{Spec}_\varepsilon B$, for any tridiagonal B with scalar entries, we see that

$$
\bigcup_{B \in M(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma)} \operatorname{Spec}_\varepsilon B \subset \overline{\bigcup_{B \in M(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma)} \operatorname{spec}_\varepsilon B} = \operatorname{Spec}_\varepsilon A.
$$

Since $A \in M(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma)$, also $\operatorname{Spec}_\varepsilon A \subset \bigcup_{B \in M(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma)} \operatorname{Spec}_\varepsilon B$.

Σ_γ. Then, the π method does not suffer from spectral pollution for A for any $t \in \mathbb{T}$. Further, the τ method does not suffer from spectral pollution for A if either: (a) $\gamma_* \leq \alpha^*$ and $\alpha_* \leq \gamma^*$; or (b) $E_\cap$ is empty. Moreover, as $n \to \infty$,

$$\Gamma^n_{\varepsilon+\varepsilon''_n}(A) \xrightarrow{H} \mathrm{Spec}_\varepsilon A, \quad \varepsilon \geq 0 \quad and \quad \gamma^n_{\varepsilon+\varepsilon''_n}(A) \xrightarrow{H} \mathrm{spec}_\varepsilon A, \quad \varepsilon > 0; \quad (8.22)$$

if $t \in \mathbb{T}$, also

$$\Pi^{n,t}_{\varepsilon+\varepsilon'_n}(A) \xrightarrow{H} \mathrm{Spec}_\varepsilon A, \quad \varepsilon \geq 0 \quad and \quad \pi^{n,t}_{\varepsilon+\varepsilon'_n}(A) \xrightarrow{H} \mathrm{spec}_\varepsilon A, \quad \varepsilon > 0. \quad (8.23)$$

Further, if (a) or (b) holds, also

$$\Sigma^n_{\varepsilon+\varepsilon_n}(A) \xrightarrow{H} \mathrm{Spec}_\varepsilon A, \quad \varepsilon \geq 0 \quad and \quad \sigma^n_{\varepsilon+\varepsilon_n}(A) \xrightarrow{H} \mathrm{spec}_\varepsilon A, \quad \varepsilon > 0. \quad (8.24)$$

Proof. For $k \in \mathbb{Z}$, $n \in \mathbb{N}$, $t \in \mathbb{T}$, and $\varepsilon > 0$, applying Corollary 8.2, we see that $\mathrm{spec}_\varepsilon A^{\mathrm{per},t}_{n,k} \subset \mathrm{spec}_\varepsilon B$ for the unique $B \in M(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma)$ with n-periodic coefficients that has $B^{\mathrm{per},t}_{n,k} = A^{\mathrm{per},t}_{n,k}$. But $\mathrm{spec}_\varepsilon B \subset \mathrm{spec}_\varepsilon A$ by (8.20). Thus, the π method does not suffer from spectral pollution for A for any $t \in \mathbb{T}$. Let $E_+ := \ell_2(\mathbb{N}, \mathbb{C})$ and suppose that $A_+ \in L(E_+)$ is tridiagonal and pseudoergodic, i.e., $A_+ \in \Psi E_+(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma)$, in the notation of [19, Section 1]. Then, for $k \in \mathbb{Z}$, $n \in \mathbb{N}$, and $\varepsilon > 0$, $\mathrm{spec}_\varepsilon A_{n,k} \subset \mathrm{spec}_\varepsilon A_+$, by [19, Corollary 4.14 (b)]. Moreover, $\mathrm{spec}_\varepsilon A_+ = \mathrm{spec}_\varepsilon A$, by [19, Proposition 4.13 (c)], if (a) or (b) hold. Thus, if (a) or (b) hold, the τ method does not suffer from spectral pollution for A.

Noting that $X = \mathbb{C}$ trivially has the Globevnik property, (8.22) follows from Theorem 1.2, (8.23) follows from Theorem 1.4, and, provided that (a) or (b) holds, (8.24) follows from Theorem 1.3. ∎

Remark 8.4. In the case of the Feinberg–Zee random hopping matrix, i.e., the case that $\Sigma_\alpha = \{-1, 1\}$, $\Sigma_\beta = \{0\}$, $\Sigma_\gamma = \{1\}$, we have that $\alpha_* = \alpha^* = \gamma_* = \gamma^* = 1$, so that (a) holds in the above theorem and $\Sigma^n_{\varepsilon+\varepsilon_n}(A) \xrightarrow{H} \mathrm{Spec}_\varepsilon A$, for $\varepsilon \geq 0$. This result, for the specific Feinberg–Zee random hopping case, was shown previously, starting from (1.22), in [14, Theorem 4.9].

Remark 8.5. We note that Colbrook in [21] has proved results similar to the results in the above theorem for the π method, by similar arguments. He shows in [21], for very general classes of pseudoergodic operators A, including tridiagonal pseudoergodic operators as defined here, that the pseudospectra of periodised finite section approximations converge to those of A, precisely that

$$\mathrm{Spec}_\varepsilon A^{\mathrm{per},1}_{2n+1,-n-1} \xrightarrow{H} \mathrm{Spec}_\varepsilon A \quad \text{as } n \to \infty, \quad (8.25)$$

for every $\varepsilon > 0$.

8.4.1. The case that Σ_α, Σ_β, and Σ_γ are finite sets. When Σ_α, Σ_β, and Σ_γ are finite sets, as holds, for example, in the Feinberg–Zee case, then the infinite collections of finite matrices, indexed by $k \in \mathbb{Z}$, in the definitions (1.21), (1.24), and (1.29)/(1.33) of the inclusion sets

$$\Sigma^n_{\varepsilon+\varepsilon_n}(A), \quad \sigma^n_{\varepsilon+\varepsilon_n}(A), \quad \Pi^{n,t}_{\varepsilon+\varepsilon'_n}(A), \quad \pi^n_{\varepsilon+\varepsilon'_n}(A), \quad \Gamma^n_{\varepsilon+\varepsilon''_n}(A), \quad \gamma^n_{\varepsilon+\varepsilon''_n}(A),$$

are in fact (large) finite sets of matrices. Let $M_n(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma)$ denote the set of all $n \times n$ tridiagonal matrices with subdiagonal entries in Σ_α, diagonal entries in Σ_β, and superdiagonal entries in Σ_γ. This set has cardinality $N_\alpha^{n-1} N_\beta^n N_\gamma^{n-1}$, where $N_\alpha :=$ $|\Sigma_\alpha|$, $N_\beta := |\Sigma_\beta|$, and $N_\gamma := |\Sigma_\gamma|$. Further, recalling the notation (6.9), it is easy to see that if A is pseudoergodic with respect to Σ_α, Σ_β, and Σ_γ, then

$$S_n^{(\tau)} := \{A_{n,k} : k \in \mathbb{Z}\} = M_n(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma). \tag{8.26}$$

Similarly, where $B_n^t(\Sigma_\alpha, \Sigma_\gamma)$, for some fixed $t \in \mathbb{T}$, denotes the set of all $n \times n$ matrices B with entries $b_{i,j} = \delta_{i,1}\delta_{j,n}t\hat{\alpha} + \delta_{i,n}\delta_{j,1}\bar{t}\hat{\gamma}$, $i, j = 1, \ldots, n$, for some $\hat{\alpha} \in \Sigma_\alpha$ and $\hat{\gamma} \in \Sigma_\gamma$, it holds that

$$S_n^{(\pi)} := \{A_{n,k}^{\mathrm{per},t} : k \in \mathbb{Z}\} = M_n(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma) + B_n^t(\Sigma_\alpha, \Sigma_\gamma), \tag{8.27}$$

which set has cardinality no larger than $N_\alpha^n N_\beta^n N_\gamma^n$, indeed exactly this cardinality for $n \geq 3$.

In the Feinberg–Zee case these sets have cardinalities $|S_n^{(\tau)}| = 2^{n-1}$ and $|S_n^{(\pi)}| = 2^n$, respectively. Thus, given $\lambda \in \mathbb{C}$, determining whether $\lambda \in \Sigma^n_{\varepsilon+\varepsilon_n}(A)$ requires half the computation of determining whether $\lambda \in \Pi^n_{\varepsilon+\varepsilon_n}(A)$. Since both these inclusion sets converge to $\mathrm{Spec}_\varepsilon A$, for $\varepsilon \geq 0$, by Theorem 8.3 and Remark 8.4, this suggests computing $\Sigma^n_{\varepsilon+\varepsilon_n}(A)$ rather than $\Pi^n_{\varepsilon+\varepsilon_n}(A)$, as an approximation to $\mathrm{Spec}_\varepsilon A$.

In Figure 8.3, taken from [14], we plot $\Sigma^n_{\varepsilon_n}(A) \supset \mathrm{Spec}\, A$ for $n = 6$, 12, and 18. Concretely, by (1.28), (7.5), and (8.26), we have

$$\Sigma^n_{\varepsilon_n}(A) = \left\{\lambda \in \mathbb{C} : \min_{B \in M_n(\{\pm 1\}, \{0\}, \{1\})} \nu(B - \lambda I_n) \leq \varepsilon_n\right\}$$

and, cf. (7.4),

$$\nu(B - \lambda I_n) > \varepsilon_n \iff (B - \lambda I_n)'(B - \lambda I_n) - \varepsilon_n^2 I_n \text{ is positive definite.}$$

Thus, to decide whether $\lambda \in \Sigma^n_{\varepsilon_n}(A)$ one has to decide whether or not 2^{n-1} Hermitian matrices of size $n \times n$ are positive definite. For $n = 18$ this requires consideration of the positive definiteness of 131,072 matrices of size 18×18 for each λ on some grid covering the domain of interest.

In Figure 8.3 we also plot $\overline{\mathrm{Num}\, A}$, the closure of the numerical range of A, computed in [14] to be the square $S = \mathrm{conv}\{\pm 2, \pm 2\mathrm{i}\}$. Disappointingly, $\overline{\mathrm{Num}\, A}$
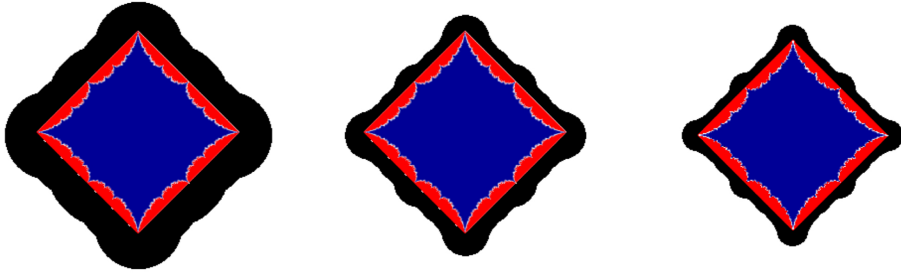
Figure 8.3. Plots in black of $\Sigma^n_{\varepsilon_n}(A)$, for $n = 6, 12$, and 18, in the Feinberg–Zee case that A is pseudoergodic with respect to $\Sigma_\alpha = \{-1, 1\}$, $\Sigma_\beta = \{0\}$, and $\Sigma_\gamma = \{1\}$. $\Sigma^n_{\varepsilon_n}(A)$ is an inclusion set for Spec A for each n, by (1.22), and $\Sigma^n_{\varepsilon_n}(A) \xrightarrow{H}$ Spec A as $n \to \infty$, by Theorem 8.3 and Remark 8.4. Shown in red in each plot is the closure of the numerical range of A, which (see [14, Lemma 3.1]) is the square with corners at ± 2 and $\pm 2i$. This is also an inclusion set for Spec A (see Section 1.6). Overlaid in blue in each plot is the set $\pi_{30} \cup \mathbb{D}$, where π_n (which can be computed by Corollary 8.2) is the union of Spec B over all $B \in M(\{-1, 1\}, \{0\}, \{1\})$ that are periodic with period $p \leq n$. For each $n \in \mathbb{N}$, $\pi_n \subset$ Spec A by (8.20), and $\mathbb{D} \subset$ Spec A by [13, Theorem 2.3].

is a sharper inclusion set than $\Sigma^n_{\varepsilon_n}(A)$, at least for these modest values of n, i.e., $\Sigma^n_{\varepsilon_n}(A) \supset S = \overline{\text{Num } A} \supset$ Spec A, for $n = 6, 12, 18$. If the inclusion $\Sigma^n_{\varepsilon_n}(A) \supset \overline{\text{Num } A}$ held for all $n \in \mathbb{N}$ it would follow, since Spec $A \subset \overline{\text{Num } A}$ and $\Sigma^n_{\varepsilon_n}(A) \xrightarrow{H}$ Spec A as $n \to \infty$, that Spec $A = \overline{\text{Num } A} = S$. To show that this is not the case we carried out calculations in [14] to demonstrate that $\lambda \notin \Sigma^n_{\varepsilon_n}(A)$ for $n = 34$ and $\lambda = \frac{3}{2} + \frac{1}{2}i$, a point on the boundary of the square S. This calculation is a matter of checking that $2^{33} \approx 8.6 \times 10^9$ matrices of size 34×34 are positive definite. This is a large calculation, but one suited to a massively parallel computer architecture, with each processor assigned a subset of the nearly 9 billion matrices to check their positive definiteness. We note also that this calculation is an example of using our inclusion sets to show that a particular operator is invertible (see Remark 1.8). These calculations, showing that, for this particular choice of λ, $\lambda \notin \Sigma^n_{\varepsilon_n}(A)$ for $n = 34$, prove that $\lambda \notin$ Spec A, i.e., $\lambda I - A$ is invertible. For more information about these calculations and those in Figure 8.3, see [14]. For a review of what is known about Spec A for this Feinberg–Zee case, and a list of challenging open problems, see [17].

8.4.2. The case that Σ_α, Σ_β, and Σ_γ are infinite sets, and the SCI. Consider now the case that the sets Σ_α, Σ_β, and Σ_γ are not necessarily finite. (We have in mind applications where one or more of these sets is infinite, or all the sets are finite but the cardinality of $M_n(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma)$ is unfeasibly large for computation as in the previous subsection.) It is easy to see that, if A is pseudoergodic with respect to Σ_α, Σ_β, and

Σ_γ, then, where $S^{(\tau)}$ and $S^{(\pi)}$ are defined by (8.26) and (8.27),

$$\overline{S_n^{(\tau)}} = \overline{\{A_{n,k} : k \in \mathbb{Z}\}} = M_n(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma) \tag{8.28}$$

and, for the chosen fixed $t \in \mathbb{T}$,

$$\overline{S_n^{(\pi)}} = \overline{\{A_{n,k}^{\mathrm{per},t} : k \in \mathbb{Z}\}} = M_n(\Sigma_\alpha, \Sigma_\beta, \Sigma_\gamma) + B_n^t(\Sigma_\alpha, \Sigma_\gamma). \tag{8.29}$$

(These equations are equivalent to (8.26) and (8.27) in the case that Σ_α, Σ_β, and Σ_γ are all finite.)

This is a case where it is relatively clear how to select a finite set $S_n \subset \overline{S_n^{(M)}}$, for $M = \tau$ or π, such that (6.10) holds, so that the approximations to $\mathrm{Spec}_\varepsilon A$ proposed in Theorem 6.3 can be computed. Indeed, suppose that we can find sequences of finite sets $(\Sigma_\alpha^n)_{n\in\mathbb{N}} \subset \Sigma_\alpha$, $(\Sigma_\beta^n)_{n\in\mathbb{N}} \subset \Sigma_\beta$, and $(\Sigma_\gamma^n)_{n\in\mathbb{N}} \subset \Sigma_\gamma$ such that

$$d_H(\Sigma_G, \Sigma_G^n) \le \frac{1}{3n} \quad \text{for } n \in \mathbb{N} \text{ and } G \in \{\alpha, \beta, \gamma\}. \tag{8.30}$$

Then, $S_n = S_n^{(M)}$, as defined in the following lemma, is finite and satisfies (6.10).

Lemma 8.6. *Suppose that A is pseudoergodic with respect to Σ_α, Σ_β, and Σ_γ, and set, for $n \in \mathbb{N}$,*

$$S_n^{(M)} := \begin{cases} M_n(\Sigma_\alpha^n, \Sigma_\beta^n, \Sigma_\gamma^n) & \text{if } M = \tau, \\ M_n(\Sigma_\alpha^n, \Sigma_\beta^n, \Sigma_\gamma^n) + B_n^t(\Sigma_\alpha^n, \Sigma_\gamma^n) & \text{if } M = \pi. \end{cases}$$

Then, $S_n^{(M)}$ is a finite subset of $\overline{S_n^{(M)}}$ and (6.10) holds with $S_n = S_n^{(M)}$.

Proof. Let $M = \tau$ or π and let $S_n := S_n^{(M)}$. It is clear that S_n is finite, since Σ_α^n, Σ_β^n, and Σ_γ^n are finite. Further, for $G = \alpha$, β, or γ, and every $n \in \mathbb{N}$ and $z \in \Sigma_G$, there exists $z_n \in \Sigma_G^n$ such that $|z - z_n| \le 1/(3n)$. But this implies, for every $B \in S_n^{(M)} = \{A_{n,k}^{(M)} : k \in \mathbb{Z}\}$, where $A_{n,k}^{(M)}$ is defined by (6.8), that there exists $B^n \in S_n$ such that $|b_{i,j} - b_{i,j}^n| \le 1/(3n)$, for $i, j = 1, \ldots, n$, where $b_{i,j}$ and $b_{i,j}^n$ denote the elements in row i and column j of B and B^n, respectively. But this implies that

$$\sum_{j=1}^n |b_{i,j} - b_{i,j}^n| \le 1/n$$

and

$$\sum_{i=1}^n |b_{i,j} - b_{i,j}^n| \le 1/n,$$

which imply that $\|B - B^n\| \le 1/n$. ∎

Combining the above lemma with Theorem 6.3 and with the results on absence of spectral pollution in Theorem 8.3, we have immediately the following result, noting that the definition (8.31) coincides with (6.12) by (7.3).

Corollary 8.7. *Suppose that A is pseudoergodic with respect to Σ_α, Σ_β, and Σ_γ, and that $S_n^{(M)}$, for $n \in \mathbb{N}$ and $M = \tau$ or π, is defined as in Lemma 8.6, where $(\Sigma_\alpha^n)_{n \in \mathbb{N}} \subset \Sigma_\alpha$, $(\Sigma_\beta^n)_{n \in \mathbb{N}} \subset \Sigma_\beta$, and $(\Sigma_\gamma^n)_{n \in \mathbb{N}} \subset \Sigma_\gamma$ are finite sets satisfying (8.30). Further, let*

$$\Delta_\varepsilon^{n,\mathrm{fin},M}(A) := \{\lambda \in \mathbb{C} : \min_{B \in S_n^{(M)}} \nu(B - \lambda I_n) \leq \varepsilon\}, \tag{8.31}$$

for $n \in \mathbb{N}$, $M = \tau$ or π, and $\varepsilon \geq 0$. Then, for $\varepsilon \geq 0$ and $n \in \mathbb{N}$,

$$\mathrm{Spec}_\varepsilon A \subset \Delta_{\varepsilon+\varepsilon_n'+1/n}^{n,\mathrm{fin},\pi}(A), \tag{8.32}$$

and $\Delta_{\varepsilon+\varepsilon_n'+1/n}^{n,\mathrm{fin},\pi}(A) \xrightarrow{H} \mathrm{Spec}_\varepsilon A$ as $n \to \infty$, in particular, $\Delta_{\varepsilon_n'+1/n}^{n,\mathrm{fin},\pi}(A) \xrightarrow{H} \mathrm{Spec}\, A$. If, moreover, conditions (a) or (b) in Theorem 8.3 are satisfied, than also

$$\mathrm{Spec}_\varepsilon A \subset \Delta_{\varepsilon+\varepsilon_n+1/n}^{n,\mathrm{fin},\tau}(A) \tag{8.33}$$

for $\varepsilon \geq 0$ and $n \in \mathbb{N}$, and $\Delta_{\varepsilon+\varepsilon_n+1/n}^{n,\mathrm{fin},\tau}(A) \xrightarrow{H} \mathrm{Spec}_\varepsilon A$ as $n \to \infty$, in particular, $\Delta_{\varepsilon_n+1/n}^{n,\mathrm{fin},\tau}(A) \xrightarrow{H} \mathrm{Spec}\, A$.

As we did in Section 1.4.5 and Section 7, we can go beyond the above corollary and define a sequence of approximations $(\Pi_{\mathrm{fin}}^n(A))_{n \in \mathbb{N}}$ to $\mathrm{Spec}\, A$, when A is pseudo-ergodic, such that each $\Pi_{\mathrm{fin}}^n(A)$ is a finite set that can be computed in finitely many arithmetic operations, and $\Pi_{\mathrm{fin}}^n(A) \xrightarrow{H} \mathrm{Spec}\, A$ as $n \to \infty$.

Let Ω_Ψ denote the set of all tridiagonal matrices (1.15), with $\alpha, \beta, \gamma \in \ell^\infty(\mathbb{Z}, \mathbb{C})$, that are pseudoergodic. For a given $A \in \Omega_\Psi$, with coefficients $\alpha, \beta, \gamma \in \ell^\infty(\mathbb{Z}, \mathbb{C})$, let

$$\Sigma_\alpha := \overline{\{\alpha_j : j \in \mathbb{Z}\}}, \quad \Sigma_\beta := \overline{\{\beta_j : j \in \mathbb{Z}\}}, \quad \Sigma_\gamma := \overline{\{\gamma_j : j \in \mathbb{Z}\}}.$$

Note that, with these definitions, A is pseudoergodic with respect to Σ_α, Σ_β, and Σ_γ. As in Section 1.4.5, let $P(S)$ denote the power set of a set S. The inputs we need to compute $\Pi_{\mathrm{fin}}^n(A)$, for $n \in \mathbb{N}$, are provided by the following mappings (cf. Section 1.4.5).

(1) A mapping $\mathcal{B} : \Omega_\Psi \to \mathbb{R}^3$, $A \mapsto (\alpha_{\max}, \beta_{\max}, \gamma_{\max})$, such that $\alpha_{\max} \geq \|\alpha\|_\infty$, $\beta_{\max} \geq \|\beta\|_\infty, \gamma_{\max} \geq \|\gamma\|_\infty$.

(2) A mapping $\mathcal{E} : \Omega_\Psi \times \mathbb{N} \to (P(\mathbb{C}))^3$, $(A, n) \mapsto (\Sigma_\alpha^n, \Sigma_\beta^n, \Sigma_\gamma^n)$, such that $\Sigma_\alpha^n \subset \Sigma_\alpha$, $\Sigma_\beta^n \subset \Sigma_\beta$, and $\Sigma_\gamma^n \subset \Sigma_\gamma$ are finite sets satisfying (8.30).

Note that $\mathcal{B}$ is the mapping $\mathcal{B}_1$ of Section 1.4.5, restricted to $\Omega_\Psi \subset \Omega_T^1$.

Given $A \in \Omega_\Psi$, for $n \in \mathbb{N}$ let (cf. (4.21))

$$\varepsilon_n^\dagger := (\alpha_{\max} + \gamma_{\max}) \frac{22}{7n} \geq 2(\alpha_{\max} + \gamma_{\max}) \sin \frac{\pi}{2n} \geq \varepsilon_n', \qquad (8.34)$$

where $(\alpha_{\max}, \beta_{\max}, \gamma_{\max}) := \mathcal{B}(A)$. Recalling the definition (1.49), define the sequence $(\Pi_{\text{fin}}^n(A))_{n \in \mathbb{N}}$ by

$$\Pi_{\text{fin}}^n(A) := \Delta_{\varepsilon_n^\dagger + 3/n}^{n, \text{fin}, \pi}(A) \cap \text{Grid}(n, R), \quad n \in \mathbb{N}, \qquad (8.35)$$

where $R := \alpha_{\max} + \beta_{\max} + \gamma_{\max}$ (so that R is an upper bound for $\|A\|$), and where $\Delta_\varepsilon^{n, \text{fin}, \pi}(A)$, for $n \in \mathbb{N}$ and $\varepsilon \geq 0$, is defined by equation (8.31), with $S_n^{(\pi)}$ defined as in Lemma 8.6, for some $t \in \mathbb{T}$, with $(\Sigma_\alpha^n, \Sigma_\beta^n, \Sigma_\gamma^n) := \mathcal{E}(A, n)$.

Proposition 8.8. *For $A \in \Omega_\Psi$ and $n \in \mathbb{N}$, $\Pi_{\text{fin}}^n(A)$ can be computed in finitely many arithmetic operations, given $(\alpha_{\max}, \beta_{\max}, \gamma_{\max}) := \mathcal{B}(A)$ and $(\Sigma_\alpha^n, \Sigma_\beta^n, \Sigma_\gamma^n) := \mathcal{E}(A, n)$. Further, $\Pi_{\text{fin}}^n(A) \xrightarrow{H} \text{Spec } A$ as $n \to \infty$, and also*

$$\hat{\Pi}_{\text{fin}}^n(A) := \Pi_{\text{fin}}^n(A) + \frac{2}{n}\overline{\mathbb{D}} \xrightarrow{H} \text{Spec } A$$

as $n \to \infty$, with $\text{Spec } A \subset \hat{\Pi}_{\text{fin}}^n(A) \subset \text{Spec}_{\varepsilon_n^\dagger + 3/n}(A) + \frac{2}{n}\overline{\mathbb{D}}$ for each $n \in \mathbb{N}$.

We omit the proof of the above result which is a minor variant of the proof of Proposition 1.10 in Section 7. But let us interpret this proposition as a result relating to the solvability complexity index (SCI) of [2, 25], as we did for Proposition 1.10 in Section 1.4.5. Equipping $\mathbb{C}^C$ with the Hausdorff metric as in Section 1.3, the mappings

$$\Omega_\Psi \to \mathbb{C}^C, \quad A \mapsto \Pi_{\text{fin}}^n(A), \quad A \mapsto \hat{\Pi}_{\text{fin}}^n(A),$$

are, for each $n \in \mathbb{N}$, general algorithms in the sense of [2, 25], with evaluation set (in the sense of [2, 25])

$$\Lambda := \{\mathcal{B}, \mathcal{E}(\cdot, n) : n \in \mathbb{N}\}; \qquad (8.36)$$

each function in Λ has domain Ω, and each can be expressed in terms of finitely many complex-valued functions; cf. the footnote below (1.52). Further, $\hat{\Pi}_{\text{fin}}^n(A)$ can be computed in finitely many arithmetic operations and specified using finitely many complex numbers (the elements of $\Pi_{\text{fin}}^n(A)$ and the value of n). Thus, where $\Xi : \Omega_\Psi \to \mathbb{C}^C$ is the mapping given by

$$\Xi(A) := \text{Spec } A$$

for $A \in \Omega_\Psi$, the computational problem $\{\Xi, \Omega_\Psi, \mathbb{C}^C, \Lambda\}$ has arithmetic SCI, in the sense of [2,25], equal to one; more precisely, since $\text{Spec } A \subset \hat{\Pi}_{\text{fin}}^n(A)$, for each $n \in \mathbb{N}$ and $A \in \Omega_\Psi$, this computational problem is in the class Π_1^A, as defined in [2, 25].

Remark 8.9 (Rate of convergence). It follows from the last inclusions in Proposition 8.8 that, if $A \in \Omega_\Psi$, i.e., if A is pseudoergodic with respect to some non-empty compact sets Σ_α, Σ_β, and Σ_γ, then

$$d_H(\operatorname{Spec} A, \widehat{\Pi}_{\mathrm{fin}}^n(A)) \leq \delta_n := d_H(\operatorname{Spec} A, \operatorname{Spec}_{\varepsilon_n^\dagger + 3/n}(A)) + 2/n \to 0,$$

as $n \to \infty$ by (2.8). Further note, by (8.20), that δ_n depends only on n and the sets Σ_α, Σ_β, and Σ_γ. Thus, the convergence

$$\widehat{\Pi}_{\mathrm{fin}}^n(A) \xrightarrow{H} \operatorname{Spec} A$$

is uniform in A, for all A that are pseudoergodic with respect to Σ_α, Σ_β, and Σ_γ. By contrast, the rate of convergence of the approximation (8.25) to $\operatorname{Spec}_\varepsilon A$ (note that $\operatorname{Spec}_\varepsilon A$ is given in terms of Σ_α, Σ_β, and Σ_γ by (8.20)) depends on the particular realisation of A.

To make this concrete (see the related discussions in [14, Section 4.3] and [44]), consider the Feinberg–Zee case that $\Sigma_\alpha = \{-1, 1\}$, $\Sigma_\beta = \{0\}$, and $\Sigma_\gamma = \{1\}$, and suppose A, taking the form (1.15), is a random Feinberg–Zee matrix, with $\beta_j = 0$ and $\gamma_j = 1$, for $j \in \mathbb{Z}$, and with the entries $\alpha_j \in \{-1, 1\}$ i.i.d. random variables, with $p := \operatorname{Pr}(\alpha_j = -1) \in (0, 1)$. Then, almost surely, A is pseudoergodic with respect to Σ_α, Σ_β, and Σ_γ, and, by the above result, the convergence $\Pi_{\mathrm{fin}}^n(A) \xrightarrow{H} \operatorname{Spec} A$ is (almost surely) independent of the particular realisation of A and independent of p. But the rate of convergence of (8.25) depends on p. In particular, given any $n \in \mathbb{N}$, it holds, with arbitrarily high probability if p is sufficiently small, that $\alpha_j = 1$, for $|j| \leq n$, so that $A_{2n+1,-n-1}^{\mathrm{per},1}$ is symmetric, $\operatorname{Spec} A_{2n+1,-n-1}^{\mathrm{per},1} \subset \mathbb{R}$, and $\operatorname{Spec} A_{2n+1,-n-1}^{\mathrm{per},1} \subset \mathbb{R} + \varepsilon \bar{D}$. But this implies that

$$d_H(\operatorname{Spec}_\varepsilon A, \operatorname{Spec} A_{2n+1,-n-1}^{\mathrm{per},1}) \geq 2,$$

since $2\mathrm{i} \in \operatorname{Spec} A$ by [17, Theorem 2.5 and Lemma 2.6], so that $(2 + \varepsilon)\mathrm{i} \in \operatorname{Spec}_\varepsilon A$.

9. Open problems and directions for further work

We see a number of open problems and possible directions for future work. Firstly, in our own work in preparation we are considering two extensions to the work reported above, both of which are announced, in part, in the recent conference paper [15]. The first is that our inclusion results (1.22), (1.25), and (1.30), for tridiagonal bi-infinite matrices of the form (1.15), have natural extensions to semi-infinite matrices and, indeed, finite matrices. (The inclusion sets for finite matrices can be viewed as a generalisation of the Gershgorin theorem, which (see Section 1.1) was a key inspiration

for the work we report here.) The second extension, announced in part in [15], is that it is possible, building on the results in this paper for the case of the spectrum, to write down a sequence of approximations, defined in terms of spectral properties of finite matrices, that is convergent to the *essential spectrum*, with each approximation also an inclusion set for the essential spectrum.

In addition to the above, we see the following open questions/possible directions for future work.

(1) *How can the algorithms, implicit in the definitions of our sequences of approximations, be implemented so as to optimise convergence as a function of computation time?* Progress towards efficient implementation of something close to the approximation on the right-hand side of (6.2), building on the work in (a draft of) this paper, has been made in the work by Lindner and Schmidt [66, 77].

(2) *Can the algorithms in this paper be used or adapted to resolve open questions regarding the spectra of bounded linear operators arising in applications?* We have indicated one application of this sort in Section 8.4.1, an example which is a pseudoergodic operator where the set of τ-method matrices,

$$ S_n^{(\tau)} = \{A_{n,k} : k \in \mathbb{Z}\}, $$

has cardinality $|S_n^{(\tau)}|$ that is finite but grows exponentially with n. There are interesting potential applications, notably the Fibonacci Hamiltonian and non-self-adjoint variants [27, 28, 44], where $|S_n^{(\tau)}|$ grows only linearly with n, allowing exploration of spectral properties with much larger values of n. At the other extreme, where S_n (and the corresponding π and τ_1 method sets) are uncountable and so must be approximated as in Theorems 6.3 and 7.1, one might seek to use or adapt our methods to study the spectra of (bounded) integral operators $\mathcal{A}$ on $L^2(\mathbb{R})$, for example those that arise in problems of wave scattering by unbounded rough surfaces (e.g., [87]). Via a standard discretisation (see, e.g., [64, Section 1.2.3] or [18, Chapter 8]) the operator $\mathcal{A}$ on $L^2(\mathbb{R})$ is unitarily equivalent to an operator A on $E = \ell^2(\mathbb{Z}, X)$ with $X = L^2[0, 1]$. Thus, concretely, one might use the inclusion sets in this paper, in the manner indicated in Remark 1.8, to study invertibility of $\lambda I - \mathcal{A}$ on $L^2(\mathbb{R})$, for some specific $\lambda \in \mathbb{C}$ of relevance to applications.

(3) *Can our inclusion set analysis for $E = \ell^2(\mathbb{Z}, X)$ be extended to $E^p := \ell^p(\mathbb{Z}, X)$, for $1 \leq p \leq \infty$?* Although the spectrum of an operator of the form (1.15) does not change if we regard A as acting on E^p, with $p \in [1, \infty]$, rather than on E [74, Corollary 2.5.4], the pseudospectrum does depend on p. We note that many results on approximation of pseudospectra of operators on E

extend to operators on E^p (see, e.g., [8, 14, 19, 21]), and that Proposition 6 in Lindner and Seidel [67], derived for other purposes, can be seen as a key step in this direction.

(4) *Can our inclusion sets and convergence results for $E = \ell^2(\mathbb{Z}, X)$ be extended to $\ell^2(\mathbb{Z}^N, X)$ for arbitrary $N \in \mathbb{N}$?* We see no reason why this extension should not be possible. Indeed, Proposition 6 in Lindner and Seidel [67] can be seen as a first version (a version without the weights w_j) of an extension of our key Proposition 5.1 to this more general case, indeed to the case $E = \ell^p(\mathbb{Z}^N, X)$, for $p \in [1, \infty]$. (Note that the proof of Proposition 6 provided in [67, pp. 909–910] takes inspiration from the proof of Proposition 5.1 in this paper and in [20].) This extension would be attractive so as to tackle a wider range of applications (see, e.g., the multidimensional applications tackled in [21]).

(5) *Can our results, which are all for bounded linear operators on $E = \ell^2(\mathbb{Z}, X)$, be extended to (unbounded) closed, densely defined operators?* Such an extension would be attractive for many applications. We note that other recent works on approximation of spectra and/or pseudospectra (e.g., [4,5,22,25,34]) extend to this setting.

References

[1] A. Amir, N. Hatano, and D. R. Nelson, Non-Hermitian localization in biological networks. *Phys. Rev. E* **93** (2016), no. 4, article no. 042310 MR 3704458

[2] 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. [v1] 2015, [v5] 2020, arXiv:1508.03280v5

[3] J. Ben-Artzi, A. C. Hansen, O. Nevanlinna, and M. Seidel, New barriers in complexity theory: On the solvability complexity index and the towers of algorithms. *C. R. Math. Acad. Sci. Paris* **353** (2015), no. 10, 931–936 Zbl 1343.68078 MR 3411224

[4] S. Bögli, Convergence of sequences of linear operators and their spectra. *Integral Equations Operator Theory* **88** (2017), no. 4, 559–599 Zbl 1466.47012 MR 3694623

[5] S. Bögli, Local convergence of spectra and pseudospectra. *J. Spectr. Theory* **8** (2018), no. 3, 1051–1098 Zbl 1500.47011 MR 3831156

[6] S. Bögli and M. Marletta, Essential numerical ranges for linear operator pencils. *IMA J. Numer. Anal.* **40** (2020), no. 4, 2256–2308 Zbl 1466.15021 MR 4167046

[7] A. Böttcher and S. M. Grudsky, *Spectral properties of banded Toeplitz matrices*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005 Zbl 1089.47001 MR 2179973

[8] A. Böttcher, S. M. Grudsky, and B. Silbermann, Norms of inverses, spectra, and pseudo-spectra of large truncated Wiener–Hopf operators and Toeplitz matrices. *New York J. Math.* **3** (1997), 1–31 Zbl 0887.47025 MR 1443134

[9] A. Böttcher and M. Lindner, Pseudospectrum. *Scholarpedia* **3** (2008), no. 3, article no. 2680

[10] A. Böttcher and B. Silbermann, *Introduction to large truncated Toeplitz matrices*. Universitext, Springer, New York, 1999 Zbl 0916.15012 MR 1724795

[11] A. Böttcher and B. Silbermann, *Analysis of Toeplitz operators*. 2nd edn., Springer Monogr. Math., Springer, Berlin, 2006 Zbl 1098.47002 MR 2223704

[12] J. V. Burke and A. Greenbaum, Characterizations of the polynomial numerical hull of degree k. *Linear Algebra Appl.* **419** (2006), no. 1, 37–47 Zbl 1106.15018 MR 2263108

[13] S. N. Chandler-Wilde, R. Chonchaiya, and M. Lindner, Eigenvalue problem meets Sierpinski triangle: computing the spectrum of a non-self-adjoint random operator. *Oper. Matrices* **5** (2011), no. 4, 633–648 Zbl 1301.47050 MR 2906853

[14] S. N. Chandler-Wilde, R. Chonchaiya, and M. Lindner, On the spectra and pseudospectra of a class of non-self-adjoint random matrices and operators. *Oper. Matrices* **7** (2013), no. 4, 739–775 Zbl 1303.47054 MR 3154570

[15] S. N. Chandler-Wilde, R. Chonchaiya, and M. Lindner, Convergent spectral inclusion sets for banded matrices. *Proc. Appl. Math. Mech.* **23** (2023), no. 4, article no. e202300016

[16] S. N. Chandler-Wilde and E. B. Davies, Spectrum of a Feinberg–Zee random hopping matrix. *J. Spectr. Theory* **2** (2012), no. 2, 147–179 Zbl 1262.15007 MR 2913876

[17] S. N. Chandler-Wilde and R. Hagger, On symmetries of the Feinberg–Zee random hopping matrix. In *Recent trends in operator theory and partial differential equations*, pp. 51–78, Oper. Theory Adv. Appl. 258, Birkhäuser/Springer, Cham, 2017 Zbl 1367.47048 MR 3643922

[18] S. N. Chandler-Wilde and M. Lindner, Limit operators, collective compactness, and the spectral theory of infinite matrices. *Mem. Amer. Math. Soc.* **210** (2011), no. 989, viii+111 Zbl 1219.47001 MR 2790742

[19] S. N. Chandler-Wilde and M. Lindner, Coburn's lemma and the finite section method for random Jacobi operators. *J. Funct. Anal.* **270** (2016), no. 2, 802–841 Zbl 1328.65128 MR 3425904

[20] R. Chonchaiya, *Computing the spectra and pseudospectra of non-self-adjoint random operators arising in mathematical physics*. Ph.D. thesis, University of Reading, UK, 2010

[21] M. J. Colbrook, Pseudoergodic operators and periodic boundary conditions. *Math. Comp.* **89** (2020), no. 322, 737–766 Zbl 1481.47004 MR 4044449

[22] M. J. Colbrook, On the computation of geometric features of spectra of linear operators on Hilbert spaces. *Found. Comput. Math.* (2022)

[23] M. J. Colbrook, *Infinite-dimensional spectral computations: Foundations, algorithms, and modern applications*. Ph.D. thesis, University of Cambridge, UK, 2024

[24] M. J. Colbrook and A. C. Hansen, On the infinite-dimensional QR algorithm. *Numer. Math.* **143** (2019), no. 1, 17–83 Zbl 1530.65055 MR 3987167

[25] M. J. Colbrook and A. C. Hansen, The foundations of spectral computations via the solvability complexity index hierarchy. *J. Eur. Math. Soc. (JEMS)* **25** (2023), no. 12, 4639–4718 Zbl 07774915 MR 4662300

[26] M. J. Colbrook, B. Roman, and A. C. Hansen, How to compute spectra with error control. *Phys. Rev. Lett.* **122** (2019), no. 25, article no. 250201 MR 3980052

[27] D. Damanik, M. Embree, and A. Gorodetski, Spectral properties of Schrödinger operators arising in the study of quasicrystals. In *Mathematics of aperiodic order*, pp. 307–370, Progr. Math. 309, Birkhäuser/Springer, Basel, 2015 Zbl 1378.81031 MR 3381485

[28] D. Damanik, A. Gorodetski, and W. Yessen, The Fibonacci Hamiltonian. *Invent. Math.* **206** (2016), no. 3, 629–692 Zbl 1359.81108 MR 3573970

[29] E. B. Davies, Spectral enclosures and complex resonances for general self-adjoint operators. *LMS J. Comput. Math.* **1** (1998), 42–74 Zbl 0931.47021 MR 1635727

[30] E. B. Davies, Spectral properties of random non-self-adjoint matrices and operators. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.* **457** (2001), no. 2005, 191–206 Zbl 1014.47002 MR 1843941

[31] E. B. Davies, Spectral theory of pseudo-ergodic operators. *Comm. Math. Phys.* **216** (2001), no. 3, 687–704 Zbl 1044.47031 MR 1815722

[32] E. B. Davies, Spectral bounds using higher-order numerical ranges. *LMS J. Comput. Math.* **8** (2005), 17–45 Zbl 1119.47004 MR 2123129

[33] E. B. Davies, *Linear operators and their spectra*. Cambridge Stud. Adv. Math. 106, Cambridge University Press, Cambridge, 2007 Zbl 1138.47001 MR 2359869

[34] E. B. Davies and M. Plum, Spectral pollution. *IMA J. Numer. Anal.* **24** (2004), no. 3, 417–438 Zbl 1062.65056 MR 2068830

[35] P. J. Davis, *Circulant matrices. pure and applied mathematics*. Wiley-Intersci. Publ., John Wiley, New York-Chichester-Brisbane, 1979 Zbl 0418.15017 MR 0543191

[36] M. M. Day, Some more uniformly convex spaces. *Bull. Amer. Math. Soc.* **47** (1941), 504–507 Zbl 0027.11003 MR 0004068

[37] F. O. Farid and P. Lancaster, Spectral properties of diagonally dominant infinite matrices. I. *Proc. Roy. Soc. Edinburgh Sect. A* **111** (1989), no. 3-4, 301–314 Zbl 0681.47011 MR 1007527

[38] F. O. Farid and P. Lancaster, Spectral properties of diagonally dominant infinite matrices. II. *Linear Algebra Appl.* **143** (1991), 7–17 Zbl 0745.47017 MR 1077721

[39] J. Feinberg and A. Zee, Non-Hermitean localization and de-localization. *Phys. Rev. E* **59** (1999), no. 6, 6433–6443

[40] J. Feinberg and A. Zee, Spectral curves of non-Hermitian Hamiltonians. *Nuclear Phys. B* **552** (1999), no. 3, 599–623 Zbl 0944.82017 MR 1704620

[41] D. G. Feingold and R. S. Varga, Block diagonally dominant matrices and generalizations of the Gerschgorin circle theorem. *Pacific J. Math.* **12** (1962), 1241–1250 Zbl 0109.24802 MR 0151473

[42] M. Fiedler and V. Pták, Generalized norms of matrices and the location of the spectrum. *Czechoslovak Math. J.* **12** (1962), no. 87, 558–571 Zbl 0116.25302 MR 0146667

[43] A. Frommer, B. Jacob, L. Vorberg, C. Wyss, and I. Zwaan, Pseudospectrum enclosures by discretization. *Integral Equations Operator Theory* **93** (2021), no. 1, article no. 9 Zbl 1496.47009 MR 4208116

[44] F. Gabel, D. Gallaun, J. Grossmann, M. Lindner, and R. Ukena, Spectral approximation of generalized Schrödinger operators via approximation of subwords. *Complex Anal. Oper. Theory* **18** (2024), no. 1, article no. 7 Zbl 07785407 MR 4673709

[45] S. Gershgorin, über die Abgrenzung der Eigenwerte einer Matrix. *Izv. Akad. Nauk. USSR Otd. Fiz.-Mat. Nauk* **7** (1931), no. 6, 749–754 Zbl 0003.00102

[46] J. Globevnik, Norm-constant analytic functions and equivalent norms. *Illinois J. Math.* **20** (1976), no. 3, 503–506 Zbl 0322.30040 MR 0448044

[47] J. Globevnik and I. Vidav, On operator-valued analytic functions with constant norm. *J. Functional Analysis* **15** (1974), 394–403 Zbl 0291.47003 MR 0346566

[48] I. Y. Goldsheid and B. A. Khoruzhenko, Distribution of eigenvalues in non-Hermitian Anderson models. *Phys. Rev. Lett.* **80** (1998), no. 13, 2897–2900

[49] I. Y. Goldsheid and B. A. Khoruzhenko, Eigenvalue curves of asymmetric tridiagonal random matrices. *Electron. J. Probab.* **5** (2000), article no. 16 Zbl 0983.82006 MR 1800072

[50] G. H. Golub and C. F. Van Loan, *Matrix computations*. 3rd edn., Johns Hopkins Stud. Math. Sci., Johns Hopkins University Press, Baltimore, MD, 1996 Zbl 0865.65009 MR 1417720

[51] R. Hagen, S. Roch, and B. Silbermann, C^*-*algebras and numerical analysis*. Monogr. Textbooks Pure Appl. Math. 236, Marcel Dekker, New York, 2001 Zbl 0964.65055 MR 1792428

[52] R. Hagger, Symmetries of the Feinberg–Zee random hopping matrix. *Random Matrices Theory Appl.* **4** (2015), no. 4, article no. 1550016 Zbl 1408.47010 MR 3418845

[53] R. Hagger, On the spectrum and numerical range of tridiagonal random operators. *J. Spectr. Theory* **6** (2016), no. 2, 215–266 Zbl 1443.47043 MR 3485943

[54] R. Hagger, M. Lindner, and M. Seidel, Essential pseudospectra and essential norms of band-dominated operators. *J. Math. Anal. Appl.* **437** (2016), no. 1, 255–291 Zbl 1336.47036 MR 3451966

[55] A. C. Hansen, On the approximation of spectra of linear operators on Hilbert spaces. *J. Funct. Anal.* **254** (2008), no. 8, 2092–2126 Zbl 1138.47002 MR 2402104

[56] A. C. Hansen, On the solvability complexity index, the *n*-pseudospectrum and approximations of spectra of operators. *J. Amer. Math. Soc.* **24** (2011), no. 1, 81–124 Zbl 1210.47013 MR 2726600

[57] N. Hatano and D. R. Nelson, Localization transitions in non-Hermitian quantum mechanics. *Phys. Rev. Lett.* **77** (1996), no. 3, 570–573

[58] N. Hatano and D. R. Nelson, Vortex pinning and non-Hermitian quantum mechanics. *Phys. Rev. B* **56** (1997), no. 14, 8651–8673

[59] N. Hatano and D. R. Nelson, Non-Hermitian delocalization and eigenfunctions. *Phys. Rev. B* **58** (1998), no. 13, 8384–8390

[60] F. Hausdorff, *Set theory*. 2nd edn., Chelsea, New York, 1962 Zbl 0488.04001 MR 0141601

[61] E. Heinemeyer, M. Lindner, and R. Potthast, Convergence and numerics of a multisection method for scattering by three-dimensional rough surfaces. *SIAM J. Numer. Anal.* **46** (2008), no. 4, 1780–1798 Zbl 1173.65070 MR 2399395

[62] D. E. Holz, H. Orland, and A. Zee, On the remarkable spectrum of a non-Hermitian random matrix model. *J. Phys. A* **36** (2003), no. 12, 3385–3400 Zbl 1038.15016 MR 1986425

[63] V. G. Kurbatov, *Functional-differential operators and equations*. Math. Appl. 473, Kluwer, Dordrecht, 1999 Zbl 0926.34053 MR 1702280

[64] M. Lindner, *Infinite matrices and their finite sections. An introduction to the limit operator method*. Front. Math., Birkhäuser, Basel, 2006 Zbl 1107.47001 MR 2253942

[65] M. Lindner, A note on the spectrum of bi-infinite bi-diagonal random matrices. *Anal. Appl. (Singap.)* **7** (2009), no. 3, 269–278 Zbl 1181.47042 MR 2542739

[66] M. Lindner and T. Schmidt, Recycling Givens rotations for the efficient approximation of pseudospectra of band-dominated operators. *Oper. Matrices* **11** (2017), no. 4, 1171–1196 Zbl 1378.65112 MR 3711439

[67] M. Lindner and M. Seidel, An affirmative answer to a core issue on limit operators. *J. Funct. Anal.* **267** (2014), no. 3, 901–917 Zbl 1292.47020 MR 3212726

[68] C. Martínez, *Spectral properties of tridiagonal operators*. Ph.D. thesis, Kings College London, UK, 2005

[69] C. Martinez, Spectral estimates for the one-dimensional non-self-adjoint Anderson model. *J. Operator Theory* **56** (2006), no. 1, 59–88 Zbl 1134.47304 MR 2261612

[70] D. R. Nelson and N. M. Shnerb, Non-Hermitian localization and population biology. *Phys. Rev. E* **58** (1998), no. 2, 1383–1403 MR 1637117

[71] O. Nevanlinna, *Convergence of iterations for linear equations*. Lectures Math. ETH Zürich, Birkhäuser, Basel, 1993 Zbl 0846.47008 MR 1217705

[72] A. M. Ostrowski, On some metrical properties of operator matrices and matrices partitioned into blocks. *J. Math. Anal. Appl.* **2** (1961), 161–209 Zbl 0101.25504 MR 0130561

[73] L. Pastur and A. Figotin, *Spectra of random and almost-periodic operators*. Grundlehren Math. Wiss. 297, Springer, Berlin, 1992 Zbl 0752.47002 MR 1223779

[74] V. Rabinovich, S. Roch, and B. Silbermann, *Limit operators and their applications in operator theory*. Oper. Theory Adv. Appl. 150, Birkhäuser, Basel, 2004 Zbl 1077.47002 MR 2075882

[75] L. Reichel and L. N. Trefethen, Eigenvalues and pseudo-eigenvalues of Toeplitz matrices. *Linear Algebra Appl.* **162/164** (1992), 153–185 Zbl 0748.15010 MR 1148398

[76] H. N. Salas, Gershgorin's theorem for matrices of operators. *Linear Algebra Appl.* **291** (1999), no. 1-3, 15–36 Zbl 1018.47004 MR 1685637

[77] T. Schmidt, *Approximation of spectra and pseudospectra on a Hilbert space*. Ph.D. thesis, TU Hamburg, Germany, 2017 Zbl 1381.46001

[78] M. Seidel, On (N, ϵ)-pseudospectra of operators on Banach spaces. *J. Funct. Anal.* **262** (2012), no. 11, 4916–4927 Zbl 1307.47005 MR 2913691

[79] M. Seidel and B. Silbermann, Finite sections of band-dominated operators—norms, condition numbers and pseudospectra. In *Operator theory, pseudo-differential equations, and mathematical physics*, pp. 375–390, Oper. Theory Adv. Appl. 228, Birkhäuser/Springer, Basel, 2013 Zbl 1280.47022 MR 3025505

[80] E. Shargorodsky, On the level sets of the resolvent norm of a linear operator. *Bull. Lond. Math. Soc.* **40** (2008), no. 3, 493–504 Zbl 1147.47007 MR 2418805

[81] E. Shargorodsky, On the definition of pseudospectra. *Bull. Lond. Math. Soc.* **41** (2009), no. 3, 524–534 Zbl 1207.47005 MR 2506836

[82] E. Shargorodsky and S. Shkarin, The level sets of the resolvent norm and convexity properties of Banach spaces. *Arch. Math. (Basel)* **93** (2009), no. 1, 59–66 Zbl 1179.47002 MR 2520644

[83] P. N. Shivakumar, J. J. Williams, and N. Rudraiah, Eigenvalues for infinite matrices. *Linear Algebra Appl.* **96** (1987), 35–63 Zbl 0627.15006 MR 0910985

[84] L. N. Trefethen, M. Contedini, and M. Embree, Spectra, pseudospectra, and localization for random bidiagonal matrices. *Comm. Pure Appl. Math.* **54** (2001), no. 5, 595–623 Zbl 1025.15034 MR 1811128

[85] L. N. Trefethen and M. Embree, *Spectra and pseudospectra. The behavior of nonnormal matrices and operators*. Princeton University Press, Princeton, NJ, 2005 Zbl 1085.15009 MR 2155029

[86] R. S. Varga, *Geršgorin and his circles*. Springer Ser. Comput. Math. 36, Springer, Berlin, 2004 Zbl 1057.15023 MR 2093409

[87] B. Zhang and S. N. Chandler-Wilde, Integral equation methods for scattering by infinite rough surfaces. *Math. Methods Appl. Sci.* **26** (2003), no. 6, 463–488 Zbl 1016.78006 MR 1965717

Simon Chandler-Wilde
Department of Mathematics and Statistics, University of Reading, Reading RG6 6AX, UK;
s.n.chandler-wilde@reading.ac.uk

Ratchanikorn Chonchaiya
Faculty of Science, King Mongkut's University of Technology Thonburi, 10140 Bangkok,
Thailand; hengmath@hotmail.com

Marko Lindner
Institut für Mathematik, TU Hamburg (TUHH), 21073 Hamburg, Germany; lindner@tuhh.de