Numer. Math. (2007) 106:41–68
DOI 10.1007/s00211-006-0054-x

Numerische Mathematik

On the numerical condition of a generalized Hankel eigenvalue problem

B. Beckermann · G. H. Golub · G. Labahn

Received: 26 March 2005 / Revised: 5 July 2006 /
Published online: 22 December 2006

Abstract The generalized eigenvalue problem $\widetilde{H}y = \lambda Hy$ with H a Hankel matrix and $\widetilde{H}$ the corresponding shifted Hankel matrix occurs in number of applications such as the reconstruction of the shape of a polygon from its moments, the determination of abscissa of quadrature formulas, of poles of Padé approximants, or of the unknown powers of a sparse black box polynomial in computer algebra. In many of these applications, the entries of the Hankel matrix are only known up to a certain precision. We study the sensitivity of the nonlinear application mapping the vector of Hankel entries to its generalized eigenvalues. A basic tool in this study is a result on the condition number of Vandermonde matrices with not necessarily real abscissas which are possibly row-scaled.

Mathematics Subject Classification (2000) 15A18 · 65F35 · 15A12 · 30E10

B. Beckermann was supported in part by INTAS research network NaCCA 03-51-6637.
G. H. Golub was supported in part by DOE grant DE-FC-02-01ER41177.
G. Labahn was supported in part by NSERC and MITACS Canada grants.

B. Beckermann (✉)
Laboratoire Painlevé UMR 8524 (ANO-EDP), UFR Mathématiques – M3,
UST Lille, 59655 Villeneuve d'Ascq Cedex, France
e-mail: bbecker@math.univ-lille1.fr

G. H. Golub
Fletcher Jones Professor of Computer Science, Stanford University,
Stanford, CA, USA
e-mail: golub@sccm.stanford.edu

G. Labahn
David R. Cheriton School of Computer Science, University of Waterloo,
Waterloo, ON, Canada
e-mail: glabahn@uwaterloo.ca

1 Introduction

Given complex numbers $h_0, h_1, \ldots, h_{2n-1}$, which we refer to as moments, we are interested in the generalized eigenvalue problem (abbreviated as GEP)

$$\widetilde{H}_n y^R = \lambda H_n y^R, \quad y^L \widetilde{H}_n = \lambda y^L H_n \tag{1}$$

for the Hankel matrices

$$H_n := \begin{bmatrix} h_0 & h_1 & \cdots & h_{n-1} \\ h_1 & h_2 & \cdots & h_n \\ \vdots & \vdots & & \vdots \\ h_{n-1} & h_n & \cdots & h_{2n-2} \end{bmatrix}, \quad \widetilde{H}_n := \begin{bmatrix} h_1 & h_2 & \cdots & h_n \\ h_2 & h_3 & \cdots & h_{n+1} \\ \vdots & \vdots & & \vdots \\ h_n & h_{n+1} & \cdots & h_{2n-1} \end{bmatrix}.$$

Such eigenvalue problems appear in applications such as shape reconstruction of a polygon from its moments [15], determination of abscissa of quadrature formulas [10,11] and the determination of the unknown powers of a multivariate sparse black box polynomial in computer algebra [12,14]. The shape from moments problem is applicable to a wide variety of inverse problems of uniform density regions related to general elliptical equations appearing for example in statistics and probability, geophysics, and computed tomography (see for instance the references in [15]). In the case of sparse multivariate interpolation the applications include approximate multivariate factorization and decompositions of approximately specified polynomial systems (see the references in [14]).

In many applications, the moments h_k are obtained by measurements and hence are affected by some error. In this paper we are interested in the sensitivity of the generalized eigenvalues with respect to such errors in the moments. In other words, we are interested in the sensitivity of the application

$$\begin{aligned} G_n : \mathbb{C}^{2n} &\mapsto \mathbb{C}^n \\ (h_0, \ldots, h_{2n-1}) &\mapsto (\lambda_1, \ldots, \lambda_n) \end{aligned} \tag{2}$$

which maps the moments to the n generalized eigenvalues defined by (1). Up to first order, this sensitivity is measured by the norm of the Jacobian of this non-linear map. In the case $\lambda_j \in \mathbb{R}$ and $c_j > 0$, which is related to orthogonality on the real line, such a study was done previously by Gautschi [10,11] followed by other authors [5,9]. In our applications, however, we are more interested in the case where λ_j and c_j are complex.

In the applications mentioned earlier the input is a set of given moments h_k satisfying

$$h_k = \sum_{j=1}^{n} c_j \lambda_j^k, \quad k = 0, \ldots, 2n-1, \tag{3}$$

with $c_j \in \mathbb{C} \setminus \{0\}$ and distinct $\lambda_j \in \mathbb{C}$. In such applications, the quantities c_j and λ_j are unknown, and need to be computed from the moments. Notice that the λ_j in formula (3) are just the generalized eigenvalues of problem (1). Conversely, one may in fact show that if (1) has distinct generalized eigenvalues $\lambda_1, \ldots, \lambda_n$ then there exist suitable c_j such that Eq. (3) holds true. Thus, the generalized eigenvalue problem (1) is a convenient way of writing the problem of finding the λ_j from the moments h_k. Additional details for the above points are given later in Sect. 2.1.

This paper has two main contributions. The first is the (rather surprising) result that the sensitivity of (1) with respect to both structured and unstructured perturbations can be measured essentially by the same quantity. This gives a theoretical justification for a numerical observation in [15]. Our second contribution is to give lower bounds for the (relative) sensitivity for the most sensitive eigenvalue of (1). These lower bounds are given in terms of the condition number of the underlying Hankel matrix H_n. For a subclass of such problems (which are relevant to the applications mentioned earlier) there are lower bounds for the most sensitive eigenvalue that is given in terms of an associated Vandermonde matrix formed with the help of the (possibly complex) λ_j multiplied by some diagonal factor. In order to obtain more precise information about the sensitivity in terms of these λ_j, we finally analyze more closely the condition number of such scaled Vandermonde matrices. Our analysis shows that this sensitivity potentially grows exponentially with the size of the matrix.

The remainder of the paper is organized as follows. In Sect. 2 we recall the notion of ill-disposed generalized eigenvalues from [15] and compare the sensitivity of (1) with respect to both structured and unstructured perturbations. In Sect. 3 we study the sensitivity for the most sensitive eigenvalue of (1) and its relation to the condition number of the underlying Hankel matrix. In this section we also give lower bounds for the most sensitive eigenvalue for a subclass of problems relevant for our applications, and study the dependency of this lower bound on the distribution of the λ_j in the complex plane. In Sect. 4 we illustrate the impact of our results by discussing some simple examples coming from the applications of shape detection and the reconstruction of a sparse black box polynomial. The proof of the main Theorem of Sect. 3 is given in Sect. 5. Finally, a conclusion and discussion for further work is given in Sect. 6.

Notations: We let $\|\cdot\|$ be the Euclidean vector norm and the spectral matrix norm and $\|S\|_F = \sqrt{\mathrm{trace}(S^*S)}$ the subordinate Frobenius matrix norm, where S^* denotes the adjoint of the matrix S. The jth canonical vector is denoted by e_j. For a continuous function $f : \mathbb{C} \mapsto \mathbb{C}$, and some compact set $E \subset \mathbb{C}$, we consider the norms

$$\|f\|_{L_2(\partial\mathbb{D})} := \left[\frac{1}{2\pi} \int_0^{2\pi} |f(\mathrm{e}^{it})|^2 \,\mathrm{d}t \right]^{1/2}, \quad \|f\|_{L_\infty(E)} := \max_{z \in E} |f(z)|,$$

where $\mathbb{D}$ denotes the closed unit disk, and $\partial\mathbb{D}$ its boundary. Given a polynomial $p(z) = a_0 + a_1 z + \cdots + a_n z^n$, we consider its vector of coefficients $\vec{p} := (a_0, a_1, \ldots, a_n)^t$, the length of this vector being clear from the context. One easily verifies that, for any polynomial p of degree at most $n-1$, the following bounds hold

$$\deg p < n : \quad \|\vec{p}\| = \|p\|_{L_2(\partial\mathbb{D})}, \quad \frac{1}{\sqrt{n}}\|p\|_{L_\infty(\partial\mathbb{D})} \leq \|p\|_{L_2(\partial\mathbb{D})} \leq \|p\|_{L_\infty(\partial\mathbb{D})}. \tag{4}$$

2 Unstructured and structured perturbations of the GEP

In this section we give some background for the generalized eigenvalue problem (1). In particular, we show that in the generic case it can always be formulated in terms of problem (3) and give a simple well-known formula for the generalized eigenvectors. We also recall the perturbation analysis of this problem for the case of unstructured perturbations from [15] and then give the main result of this section, a perturbation analysis which takes into account the special Hankel structure of our input matrices.

2.1 Preliminaries

Let us first show that the generalized eigenvalue problem (1) is equivalent to finding λ_j in Eq. (3) from the moments h_k in the generic case. Thus, suppose problem (1) has moments defined by (3). Then one has a simple formula for the corresponding generalized eigenvalues and eigenvectors. Indeed in this case one has the factorizations of the two Hankel matrices

$$H_n = V_n^t \operatorname{diag}(c_1, \ldots, c_n) V_n = W_n^t W_n, \quad \widetilde{H}_n = V_n^t \operatorname{diag}(\lambda_1 c_1, \ldots, \lambda_n c_n) V_n, \tag{5}$$

where ^{t} denotes taking of a transpose (without taking conjugates), and

$$V_n = \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{bmatrix}, \quad W_n := \operatorname{diag}(\sqrt{c_1}, \ldots, \sqrt{c_n}) V_n, \tag{6}$$

a Vandermonde matrix and a row-scaled Vandermonde matrix, respectively. It follows from (5) that the generalized eigenvalues of (1) are given by the abscissa λ_j of (3), with corresponding right and left eigenvectors given by $y^R = (y^L)^t = V_n^{-1} e_j$.

Conversely suppose that (1) has distinct generalized eigenvalues $\lambda_1, \ldots, \lambda_n$. Then it is well-known [2] that the polynomial $q(z) := \det(\widetilde{H}_n - z H_n)$ is the denominator of the nth Padé approximant at infinity of the series $h(z) = h_0 z^{-1} + h_1 z^{-2} + \cdots$. Thus the generalized eigenvalues of (1) are just the poles of the nth Padé approximant. In this case we have a partial fraction decomposition for the nth Padé approximant given by

$$\sum_{j=0}^{2n-1} \frac{h_j}{z^{j+1}} = \sum_{j=1}^{n} \frac{c_j}{z - \lambda_j} + \mathcal{O}(z^{-2n-1})_{z\to\infty},$$

with $c_j \in \mathbb{C} \setminus \{0\}$. Equating coefficients leads to the representation (3).

2.2 Analysis of errors

Golub et al. [15, Sect. 3.3] considered perturbations of the generalized eigenvalue problem (1), namely

$$\begin{aligned} [\widetilde{H}_n + \epsilon \widetilde{E}_n] y^R(\epsilon) &= \lambda(\epsilon)[H_n + \epsilon E_n] y^R(\epsilon), \\ y^L(\epsilon)[\widetilde{H}_n + \epsilon \widetilde{E}_n] &= \lambda(\epsilon) y^L(\epsilon)[H_n + \epsilon E_n], \end{aligned} \tag{7}$$

where $\epsilon > 0$ is small, and the perturbation matrices $\widetilde{E}_n, E_n$ normalized such that $\|\tilde{E}_n\| \le 1$, $\|E_n\| \le 1$. In a small neigborhood around a simple eigenvalue $\lambda(0) = \lambda_j$ (and thus $y^R(0) = (y^L(0))^t = V_n^{-1} e_j$), the function $\epsilon \mapsto \lambda(\epsilon) = \lambda_j(\epsilon)$ is differentiable, with its derivative at zero being given by

$$\frac{\mathrm{d}\lambda_j}{\mathrm{d}\epsilon}(0) = \frac{y^L(0)\,[\tilde{E}_n - \lambda_j E_n]\, y^R(0)}{y^L(0)\, H_n\, y^R(0)}.$$

Plugging in explicitly the eigenvectors and using the first formula of (5) leads to the expression

$$\frac{\mathrm{d}\lambda_j}{\mathrm{d}\epsilon}(0) = \frac{(V_n^{-1} e_j)^t\,[\widetilde{E}_n - \lambda_j E_n]\, V_n^{-1} e_j}{(V_n^{-1} e_j)^t\, H_n\, V_n^{-1} e_j} = \frac{(V_n^{-1} e_j)^t\,[\widetilde{E}_n - \lambda_j E_n]\, V_n^{-1} e_j}{c_j}. \tag{8}$$

Following [15, Sect. 3.3], we say that a generalized eigenvalue λ_j is *ill-disposed* if the right-hand side of (8) multiplied with $\|H_n\| + \|\widetilde{H}_n\|$ is "large".

Using the link (6) between V_n and W_n, and exploiting the given information on the norms of E_n and $\widetilde{E}_n$, we obtain from (8) the simpler upper bound

$$\left| \frac{\mathrm{d}\lambda_j}{\mathrm{d}\epsilon}(0) \right| \le (|\lambda_j| + 1)\, \|W_n^{-1} e_j\|^2. \tag{9}$$

This upper bound appears to be quite rough. In addition, the bound is based on unstructured perturbations of Hankel matrices which a priori should lead to a severe over estimation of the actual sensitivity since in applications the perturbations will be structured. For instance, if we are interested in the sensitivity of λ_j with respect to small perturbations of the second moment h_2, we should consider as perturbation matrices $E_n, \widetilde{E}_n$ those matrices obtained from $H_n, \widetilde{H}_n$ by replacing h_2 by 1, and all other moments by 0.

However, it follows from Theorem 2.1 below that the sensitivity with respect to structured perturbations essentially gives the same result as (9). This fact was already observed numerically in [15]. To make a clear statement about first order sensitivity of the generalized eigenvalues λ_j with respect to perturbations on the moments h_k, we have to compute the Jacobian of the non-linear map G_n defined in (2), and to estimate the norms of its rows.

Theorem 2.1 *Suppose that* (1) *has distinct possibly complex eigenvalues* $\lambda_1, \dots, \lambda_n$*, and let* c_j*, and* W_n*, as in* (3)*, and* (6)*, respectively. Then for all* $j = 1, \dots, n$ *we have*

$$\left\| \left(\frac{\partial \lambda_j}{\partial h_k} \right)_{k=0,\dots,2n-1} \right\| = \eta_{j,n} \left(|\lambda_j| + 1 \right) \|W_n^{-1} e_j\|^2,$$

where $1/(2n) \le \eta_{j,n} \le \sqrt{2n}$*.*

Proof We start with the general observation that polynomial language is very useful for expressing the inverse of a Vandermonde matrix V_n with abscissa $\lambda_1, \dots, \lambda_n$. Indeed, with the corresponding Lagrange polynomials

$$\ell_j(z) = \prod_{k=1,\dots,n, k \neq j} \frac{z - \lambda_k}{\lambda_j - \lambda_k} = \frac{\omega(z)}{(z - \lambda_j)\omega'(\lambda_j)}, \quad \omega(z) = \prod_{k=1}^{n} (z - \lambda_k), \tag{10}$$

we get the relation $\ell_j(\lambda_k) = \delta_{j,k}$ or, in other words,

$$j = 1, \dots, n: \quad V_n^{-1} e_j = \vec{\ell}_j, \qquad \|V_n^{-1} e_j\| = \|\ell_j\|_{L_2(\partial\mathbb{D})}. \tag{11}$$

Consider a fixed $j \in \{1, .., n\}$ and a fixed $k \in \{0, 1, \dots, 2n-1\}$. In order to measure the first order sensitivity of the generalized eigenvalue λ_j with respect to perturbations in the moment h_k, we proceed as in (8) and obtain the formula

$$\frac{\partial \lambda_j}{\partial h_k} = \frac{(V_n^{-1} e_j)^t [\tilde{E}_n - \lambda_j E_n] V_n^{-1} e_j}{c_j}.$$

However, now the perturbation matrices E_n and $\tilde{E}_n$ have a special form, namely the matrix E_n (respectively $\tilde{E}_n$) is the Hankel matrix obtained from the zero matrix by replacing the $(k+1)$th anti-diagonal by ones (respectively the kth anti-diagonal). Again it is useful to express this formula in polynomial language. With the corresponding Lagrange polynomials $\ell_j(z) = \sum_\kappa \ell_{j,\kappa} z^\kappa$ as in (10) and the two polynomials

$$p(z) = \ell_j(z)^2 = \sum_\kappa p_\kappa z^\kappa, \quad q(z) = (z-\lambda_j)\ell_j(z)^2 = \sum_\kappa q_\kappa z^\kappa,$$

we get from (11) that

$$(V_n^{-1}e_j)^t E_n V_n^{-1} e_j = \sum_{m=0}^{k} \ell_{m,j}\ell_{k-m,j} = p_k,$$
$$(V_n^{-1}e_j)^t \tilde{E}_n V_n^{-1} e_j = \sum_{m=0}^{k} \ell_{m,j}\ell_{k-1-m,j} = p_{k-1},$$

and thus

$$\frac{\partial \lambda_j}{\partial h_k} = \frac{1}{c_j}[p_{k-1} - \lambda_j p_k] = \frac{q_k}{c_j}.$$

Using different techniques, a similar formula has been found by Gautschi (see [10, Sect. 3.2] and [11]) for the case $\lambda_j \in \mathbb{R}$ and $c_j > 0$.

Since $W_n^{-1}e_j = V_n^{-1}e_j/\sqrt{c_j} = \vec{\ell}_j/\sqrt{c_j}$ by (6) and (11), in order to establish Theorem 2.1 it only remains to show that

$$\frac{|\lambda_j|+1}{2n}\|\vec{\ell}_j\|^2 \le \|\vec{q}\| \le \sqrt{2n}(|\lambda_j|+1)\,\|\vec{\ell}_j\|^2. \tag{12}$$

Notice that $\vec{q} = B\vec{p}$, with B being a matrix of size $(2n)\times(2n-1)$ obtained from the zero matrix by replacing the main diagonal by $-\lambda_j$, and the lower diagonal by 1. The singular values of B are easily calculated: since B^*B is similar to the tridiagonal Toeplitz matrix with main diagonal $|\lambda_j|^2+1$ and lower and upper diagonals $|\lambda_j|$, its eigenvalues are given by $|\lambda_j|^2+1-2|\lambda_j|\cos(\pi m/(2n))$, $m = 1, \dots, 2n-1$, and hence

$$\frac{\|\vec{q}\|}{\|\vec{p}\|} = \frac{\|B\vec{p}\|}{\|\vec{p}\|}\begin{cases} \le \sqrt{|\lambda_j|^2+1-2|\lambda_j|\cos\left(\pi\frac{2n-1}{2n}\right)} \le 1+|\lambda_j|, \\ \ge \sqrt{|\lambda_j|^2+1-2|\lambda_j|\cos\left(\frac{\pi}{2n}\right)} \ge (1+|\lambda_j|)\sin\left(\frac{\pi}{4n}\right) \ge \frac{1+|\lambda_j|}{2n}. \end{cases}$$

In addition, since $p(z) = \ell_j(z)^2$, we get from (4) and the Cauchy–Schwarz inequality that

$$\|\vec{p}\|^2 = \frac{1}{2\pi}\int_{|w|=1} |p(w)|^2\,|\mathrm{d}w| \ge \left[\frac{1}{2\pi}\int_{|w|=1} |p(w)|\,|\mathrm{d}w|\right]^2 = \|\vec{\ell}_j\|^4,$$

and that

$$\begin{aligned}\|\vec{p}\|^2 &= \frac{1}{2n}\sum_{m=0}^{2n-1} |p(\mathrm{e}^{2\pi i m/(2n)})|^2 \\ &\leq \frac{1}{2n}\left(\sum_{m=0}^{2n-1} |p(\mathrm{e}^{2\pi i m/(2n)})|\right)^2 = 2n\left(\frac{1}{2n}\sum_{m=0}^{2n-1} |\ell_j(\mathrm{e}^{2\pi i m/(2n)})|^2\right)^2 = 2n\|\vec{\ell}_j\|^4.\end{aligned}$$

Consequently, the inequalities (12) are valid, which proves Theorem 2.1.

As a consequence of Theorem 2.1 we can say that, at least up to the factor $\eta_{j,n}$, the ill-disposedness of the generalized eigenvalue λ_j of (1) with respect to either structured and to unstructured perturbations can be measured by checking whether

$$\rho_j := (|\lambda_j| + 1)\, \|W_n^{-1} e_j\|^2\, (\|H_n\| + \|\widetilde{H}_n\|) \tag{13}$$

is “large”. The term $\eta_{j,n}$ is close to some modest power of n depending on the choice of the norm and on the techniques chosen to prove Theorem 2.1. Though for particular applications it might be desirable to have sharper results, we believe that this factor $\eta_{j,n}$ should be neglected in the measurement of ill-disposedness. As pointed out by Wilkinson [17], “usually the bound itself is weaker that it might have been because of the necessity of restricting the mass of details to a reasonable level and because of the limitations imposed by expressing the errors in terms of matrix norms.”

3 Sensitivity of the GEP

In formula (13) of the preceding section we introduced the quantities ρ_j for measuring the conditioning of the (nonlinear) generalized eigenvalue problem (1). The aim of this section is to relate our conditioning measures to those used for linear problems, in particular to the condition number of the underlying Hankel matrix H_n and the row-scaled Vandermonde matrix W_n. We are primarily concerned with determining when at least one of the generalized eigenvalues is ill-disposed. Our results will enable us (in the second part of this section) to describe classes of configurations of $\lambda_1, \ldots, \lambda_n$ where at least one of the quantities ρ_j will be large.

3.1 Sensitivity of the GEP in terms of condition numbers

The following result is an interesting direct consequence of Theorem 2.1.

Corollary 3.1 *If the Hankel matrix H_n is ill-conditioned then at least one of the generalized eigenvalues of* (1) *is ill-disposed. More precisely, we have the estimate*

$$\rho_1 + \rho_2 + \cdots + \rho_n \geq \|H_n\| \, \|H_n^{-1}\|.$$

Proof From (13) we have that $\rho_j \geq \|W_n^{-1} e_j\|^2 \, \|H_n\|$ for each j. Thus

$$\begin{aligned} \sum_{j=1}^{n} \rho_j &\geq \|H_n\| \sum_{j=1}^{n} \|W_n^{-1} e_j\|^2 = \|H_n\| \operatorname{trace}((W_n^{-1})^* W_n^{-1}) \\ &\geq \|H_n\| \, \|(W_n^{-1})^* W_n^{-1}\| \geq \|H_n\| \, \|(W_n^t W_n)^{-1}\| = \|H_n\| \, \|H_n^{-1}\|. \end{aligned}$$

Hankel matrices are suspected to be notoriously ill-conditioned (for instance, the famous Hilbert matrix is a Hankel matrix). In the case of positive definite Hankel matrices (or, equivalently, $\lambda_j \in \mathbb{R}$ and $c_j > 0$ for all j) we can be even more precise. Namely, in [4] it is shown that any positive definite Hankel matrix of order n has a spectral condition number bounded below by $3.2 \times 10^{n-1}/(16n)$. Thus Corollary 3.1 confirms a result of Gautschi [10,11] saying that, in case $c_j > 0$ and $\lambda_j \in \mathbb{R}$, the sensitivity of the map G_n is increasing exponentially in n.

Of course there are also Hankel matrices which are well-conditioned (for example the counter identity), and therefore we should say more about the case of non-real data. Here the following special case occurs quite often in applications.

Definition 3.2 *We say that the unit disk case holds if the* (*unperturbed*) *moments h_k are generated by* (3) *with $|c_j| \leq 1$ and distinct $|\lambda_j| \leq 1$.*

For the case of shape detection from moments we always have $|c_j| \leq 1$ by [15, Eq. (2.2)], and the assumption $|\lambda_j| \leq 1$ will be true after scaling of the moments (such a scaling was proposed to be useful in [15, Sect. 4]). For the application of reconstructing a sparse black box polynomial, we have $|\lambda_j| = 1$, and the assumptions on c_j can be seen as reasonable assumptions on the scaling of the unknown coefficients c_j. These applications are discussed later in Sect. 4.

For the unit disk case, we have the following direct consequence of Theorem 2.1.

Corollary 3.3 *Suppose that the unit disk case holds. If the row-scaled Vandermonde matrix W_n defined in* (6) *is ill-conditioned then at least one of the generalized eigenvalues of* (1) *is ill-disposed. More precisely, we have the estimate*

$$\rho_1 + \cdots + \rho_n \geq \frac{\|H_n\|}{n^2} \left(\|W_n\|_F \, \|W_n^{-1}\|_F\right)^2.$$

Proof From (13) we know that

$$\rho_1 + \cdots + \rho_n \geq \|H_n\| \sum_{j=1}^{n} \|W_n^{-1} e_j\|^2 = \|H_n\| \, \|W_n^{-1}\|_F^2.$$

Thus the assertion follows by observing that in the unit disk case we have the rough bound $\|W_n\|_F \leq n$.

The factor n^{-2} in front of the Frobenius condition number of W_n in Corollary 3.3 depends on the choice of the norm and the techniques of proof, and thus does not seem to be significant. Also, in typical examples, the quantity $\|H_n\|$ is of magnitude 1. Thus, roughly speaking, Corollary 3.3 tells us that, within the set of generalized eigenvalues, there is at least one for which the sensitivity is bounded below by the square of the condition number of W_n.

In comparing Corollaries 3.1 and 3.3 we recall that $H_n = W_n^t W_n$. Thus, the square of the Frobenius condition number of W_n might potentially be much larger than the condition number of the Hankel matrix H_n, and so may lead to sharper bounds.

3.2 Sensitivity of the GEP with respect to distribution of eigenvalues

In Corollary 3.3 we related the sensitivity of the generalized eigenvalue problem (1) to the Frobenius condition number of the row-scaled Vandermonde matrix W_n defined in (6). Clearly, a small $|c_j|$ will deteriorate the condition number of W_n. However, even in the case of c_j minimizing the Frobenius condition number, the condition number may still be "large" depending on the distribution of the abscissa λ_j in the unit disk. The aim of the remainder of this section is to make this last point more precise.

Definition 3.4 *For a given compact set E in the complex plane, define*

$$\gamma_n(E) = \inf \left\{ \min_{D \text{ diagonal}} \|DV_n\|_F \, \|(DV_n)^{-1}\|_F : \lambda_j \in E \text{ distinct} \right\}. \tag{14}$$

Besides being of interest in its own right, the quantity $\gamma_n(E) \geq n$ gives an alternate lower bound for ill-disposedness, as shown in the following result.

Lemma 3.5 *Let E be a compact set with $\lambda_1, \ldots, \lambda_n \in E$. Then*

$$\rho_1 + \cdots + \rho_n \;\geq\; \gamma_n(E). \tag{15}$$

Moreover, in the unit disk case we have that

$$\rho_1 + \cdots + \rho_n \;\geq\; \frac{\|H_n\|}{n^2} \gamma_n(E)^2. \tag{16}$$

Proof We start by observing that, for fixed $\lambda_1, \ldots, \lambda_n$, it is not difficult to determine the diagonal factor in (14) which minimizes the Frobenius condition number. Indeed, using the Cauchy–Schwarz inequality we find that

$$\|DV_n\|_F \, \|(DV_n)^{-1}\|_F \ge \sum_{j=1}^{n} \|(DV_n)^t e_j\| \, \|(DV_n)^{-1} e_j\| = \sum_{j=1}^{n} \|V_n^t e_j\| \, \|V_n^{-1} e_j\|, \tag{17}$$

with equality if $D = \mathrm{diag}\big(\sqrt{\|V_n^{-1} e_j\| / \|V_n^t e_j\|}\big)$.

In order to show (15) we first observe that for the row-scaled Vandermonde matrix W_n of (6) we have

$$\|H_n\| \, \|W_n^{-1} e_j\| = \|W_n^t W_n\| \, \|W_n^{-1} e_j\| \ge \|W_n^t e_j\| = \sqrt{|c_j|} \, \|V_n^t e_j\|,$$

and $\|W_n^{-1} e_j\| = \|V_n^{-1} e_j\| / \sqrt{|c_j|}$. Thus, by (13),

$$\rho_1 + \cdots + \rho_n \ \ge \sum_{j=1}^{n} \|H_n\| \, \|W_n^{-1} e_j\|^2 \ge \sum_{j=1}^{n} \|V_n^t e_j\| \, \|V_n^{-1} e_j\| \ge \gamma_n(E),$$

where the last inequality follows from (17). Finally, in the unit disk case, Eq. (16) follows directly from Corollary 3.3 and Definition 3.4.

For the remainder of this section we will give lower bounds for $\gamma_n(E)$ and thus for the measure of ill-disposedness of our generalized eigenvalues. Notice that if $\partial\mathbb{D} \subset E$, then we may choose as λ_j the nth roots of unity. In this case $W_n = V_n/\sqrt{n}$ is unitary, and thus $\gamma_n(E)$ attains its smallest possible value, namely n. We will show in Theorem 3.6 below that, for sufficiently "nice" sets E, $\gamma_n(E)$ can be bounded below and above by a quantity $\gamma(E)^n$ multiplied by a modest power of n. Here $\gamma(E) \ge 1$ measures which part of the unit circle is not part of E. When E contains the unit circle $\partial\mathbb{D}$ then $\gamma(E) = 1$. If however $\partial\mathbb{D} \not\subset E$, then $\gamma(E) > 1$, and thus $\gamma_n(E)$ grows exponentially, with the rate depending on the size of the part of the unit circle not included in E.

In order to give a precise definition of $\gamma(E)$, we require some tools from complex analysis. In order to simplify our work we will assume in the following that E is a simply connected compact set. Denote by ϕ the Riemann map which maps the complement of E conformally onto the complement of the closed unit disk $\mathbb{D}$. Then the Green function of E with pole at y is given by

$$g_E(z,y) = \begin{cases} \log\left(\left|\dfrac{1 - \overline{\phi(y)}\phi(z)}{\phi(z) - \phi(y)}\right|\right) & \text{for } z, y \in \overline{\mathbb{C}} \setminus E, \\ 0 & \text{otherwise,} \end{cases} \tag{18}$$

with g_E being continuous in both variables. Notice that $g_E(z,y) \geq 0$ for $z, y \in \mathbb{C}$, and strictly positive if both z and y lie in the complement of E. Thus the constant

$$\gamma(E) = \exp\left(\max_{x\in\partial\mathbb{D}} \frac{1}{2\pi} \int_0^{2\pi} g_E(x, \mathrm{e}^{it})\, \mathrm{d}t \right) \tag{19}$$

is at least 1, with equality if and only if $\partial\mathbb{D} \subset E$. It will be shown in the proof of Lemma 5.2 that in some cases we have a simplified expression for $\gamma(E)$, namely

$$E \subset \mathbb{D}: \quad \gamma(E) = \exp\left(\max_{x\in\partial\mathbb{D}} g_E(x, \infty) \right). \tag{20}$$

For E a real interval, it follows from [4, Remark 3.4] (where the spectral condition number was considered) that

$$\gamma_{n+1}(E) \geq \frac{1}{2\sqrt{2n+2}} (\gamma(E)^n - 2\gamma(E)^{-n}), \tag{21}$$

with the right-hand side clearly being exponentially increasing. Inequality (21) shows, for example, that positive definite Hankel matrices of order n have a condition number bounded below by some term exponentially increasing in n. In fact, a behavior similar to (21) is true for any sufficiently "nice" set E.

Theorem 3.6 *For any compact set $E \subset \mathbb{C}$ which is regular with respect to the Dirichlet problem we have that*

$$\lim_{n\to\infty} \gamma_n(E)^{1/n} = \gamma(E),$$

and $\gamma_n(E) \leq n^2 \gamma(E)^{n-1}$ for all $n \geq 1$. If, in addition E is simply connected and of bounded variation V, then we have that

$$\gamma_n(E) \geq \frac{1}{\sqrt{n}} \left[\frac{\pi}{2V} \gamma(E)^{n-1} - 1 - \frac{V}{\pi} \right]. \tag{22}$$

Proof See Sect. 5.

With respect to the second part of Theorem 3.6 we notice that any simply connected compact set is known to be regular with respect to the Dirichlet problem. We also recall, for example from [1], that E is of bounded variation V if, given some parametrization $\beta : [0,1] \mapsto \partial E$ of the boundary of E, there exists a tangent at almost every $\beta(s)$, forming an angle $\theta(s)$ with the positive real axis, and if θ has a total variation bounded by V. For instance, convex sets are of bounded variation 2π.

In order to make the statements of Theorem 3.6 more precise, we need to know both V and the Riemann map. The latter can be determined for various domains, for example for sectors $E = \{re^{it} : \alpha' \le t \le \alpha + \alpha', 0 \le r \le 1\}$ (here $V = 2\pi$), for a rectangle E with edges $\pm a \pm ib$ (here $V = 2\pi$), or more generally for a polygon (see any advanced textbook on conformal mappings). Here we will have a look at two particular examples.

Example 3.7 Consider the case of a subarc of the unit circle $E = \{\mathrm{e}^{it} : \alpha' \le t \le \alpha + \alpha'\}$ with $0 < \alpha < 2\pi$. Here one easily verifies that $V = 2\pi + 2\alpha$. The Riemann map ϕ is found as $\phi_1 \circ \phi_2$, where

$$\phi_1(z) = z/\tan(\alpha/4) + \sqrt{(z/\tan(\alpha/4))^2 - 1}$$

which maps $\overline{\mathbb{C}} \setminus [-\tan(\frac{\alpha}{4}), \tan(\frac{\alpha}{4})]$ conformally on $\overline{\mathbb{C}} \setminus \mathbb{D}$ and

$$\phi_2(z) = i\frac{\mathrm{e}^{-i\alpha'-i\alpha/2}z - 1}{\mathrm{e}^{-i\alpha'-i\alpha/2}z + 1}$$

which maps $\overline{\mathbb{C}} \setminus E$ conformally on $\overline{\mathbb{C}} \setminus [-\tan(\frac{\alpha}{4}), \tan(\frac{\alpha}{4})]$. Since $E \subset \mathbb{D}$, we may apply (18) and (20), leading to

$$\gamma(E) = \max_{\alpha+\alpha' \le t \le \alpha'+2\pi} \left| \frac{1 - \phi(\mathrm{e}^{it})\overline{\phi(\infty)}}{\phi(\mathrm{e}^{it}) - \phi(\infty)} \right|, \quad \phi(\infty) = i\frac{1+\cos(\frac{\alpha}{4})}{\sin(\frac{\alpha}{4})},$$
$$\phi_1(\mathrm{e}^{i(t+\alpha'+\frac{\alpha}{2})}) = i\tan\left(\frac{t}{2}\right).$$

Thus

$$\gamma(E) = |\phi(\infty)| = \frac{1+\cos(\alpha/4)}{\sin(\alpha/4)} = \frac{1}{\tan(\alpha/8)}, \quad 0 < \alpha < 2\pi. \tag{23}$$

Example 3.8 For an ellipse $E \subset \mathbb{D}$ with foci at $\pm c$ and half axes $c(R \pm 1/R)/2$, $R \ge 1$ (including the case of the interval $[-c, c]$ for $R = 1$) we have $V = 2\pi$ (since E is convex), and $\phi(z) = \Phi(z/c)/R$, $\phi(\infty) = \infty$ where $\Phi(z) = z + \sqrt{z^2 - 1}$. If $0 < c \le 2/(R + 1/R)$ then $E \subset \mathbb{D}$, and a simple computation shows that

$$\gamma(E) = |\phi(i)| = \frac{1}{R}\left(\frac{1}{c} + \sqrt{\frac{1}{c^2} + 1}\right).$$

In particular, for the special case $c = 2/(R + 1/R)$ of ellipses in $\mathbb{D}$ containing ± 1 we find that $\gamma(E) = (1 + \sqrt{1 + c^2})/(1 - \sqrt{1 - c^2})$. Notice that this is increasing in c (that is as the eccentricity increases), tends to 1 for $c \to 0$ (that is the case of small eccentricity, $E = \mathbb{D}$) and to $1 + \sqrt{2}$ for $c \to 1$ (the case of large eccentricity, $E = [-1, 1]$). If, on the other hand we have $2/(R - 1/R) \le c$, then $\mathbb{D} \subset E$ and hence $\gamma(E) = 1$. Finally, if $2/(R - 1/R) > c > 2/(R + 1/R)$ then we need to apply formula (19), with the maximum being attained for $x = i$ by symmetry. Since the resulting expressions are complicated, we omit details.

We can conclude from (23) and Theorem 3.6 that, if a part of the unit circle is omitted by the generalized eigenvalues of (1), then some of them are necessarily ill-disposed, even if n is not too large. However, the following example shows that for well-disposedness it is not enough to simply require that the generalized eigenvalues are distributed throughout the unit circle and are well separated.

Example 3.9 Let $E_n = \{\lambda_1, \ldots, \lambda_n\}$, with $n = 6m + 1$, be the set

$$\left\{\exp\left(\frac{2\pi ij}{4m+1}\right) : j=0,\ldots,4m\right\} \cup \left\{\exp\left(\frac{2\pi i(2j+1)}{8m+2}\right) : j=-m,-m+1,\ldots,m-1\right\},$$

a subset of the unit circle which has no significant gaps, (see Fig. 1 for $m = 4$.) We claim that

$$\gamma_n(E_n) \geq \frac{\rho^m}{n} \tag{24}$$

for $\rho = \exp\left(\frac{8\text{Catalan}}{\pi}\right) \approx 10.30 > 1$, so that there is exponential growth.

To see this claim, suppose without loss of generality that $\lambda_1 = 1$. We make use of (17) and of the formulas (4) and (11) in order to conclude that

$$\gamma_n(E_n) \geq \|V_n^t e_1\| \, \|V_n^{-1} e_1\| = \sqrt{n} \, \|V_n^{-1} e_1\| \geq \max_{|z|=1} |\ell_1(z)| \geq |\ell_1(-1)|,$$

with ℓ_1 the correspond first Lagrange polynomial, see (10). With our choice of abscissa we have that

$$\prod_{j=1}^{n}(z - \lambda_j) = (z^{4m+1} - 1)\,\widetilde{\omega}(z), \qquad \text{where } \widetilde{\omega}(z) = \prod_{j=-m}^{m-1}\left(z - \exp\left(\frac{2\pi i(2j+1)}{8m+2}\right)\right),$$

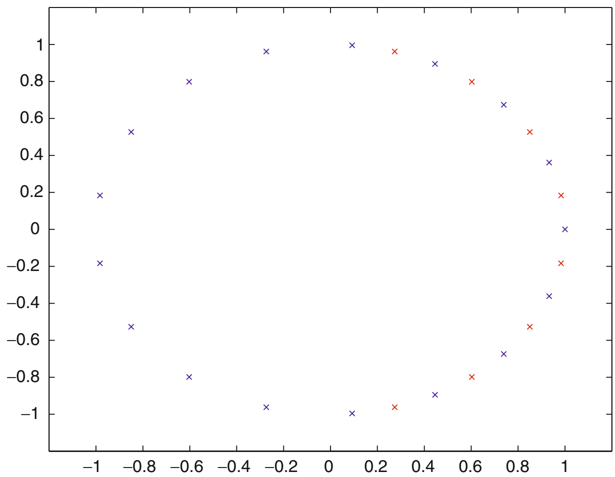

Fig. 1 The eigenvalues of Example 3.9 for $m = 4$

and thus, again by (10),

$$|\ell_1(-1)| = \frac{|\widetilde{\omega}(-1)|}{(4m+1)\,|\widetilde{\omega}(1)|} \geq \frac{|\widetilde{\omega}(-1)|}{n\,|\widetilde{\omega}(1)|}.$$

Notice that in the right-hand expression there are no longer the $(4m+1)$th roots of unity, and that the zeros of $\widetilde{\omega}$ all lie in the right-hand half circle. This enables us to show exponential growth. Indeed, the mapping $x \mapsto \log(1/\tan(x))$ is both decreasing and concave on $[0, \pi/4]$, and hence

$$\begin{aligned}
\log\left(\frac{|\widetilde{\omega}(-1)|}{|\widetilde{\omega}(1)|}\right) &= \sum_{j=-m}^{m-1} \log\left(\left|\frac{-1-\exp\left(\frac{2\pi i(2j+1)}{8m+2}\right)}{1-\exp\left(\frac{2\pi i(2j+1)}{8m+2}\right)}\right|\right) \\
&= 2\sum_{j=0}^{m-1} \log\left(\frac{1}{\left|\tan\left(\frac{\pi(2j+1)}{8m+2}\right)\right|}\right) \\
&\geq 2\sum_{j=0}^{m-1} \log\left(\frac{1}{\left|\tan\left(\frac{\pi}{4}\frac{2j+1}{2m}\right)\right|}\right) \mathrm{d}t \\
&\geq 2m\int_0^1 \log\left(\frac{1}{\left|\tan\left(\frac{\pi t}{4}\right)\right|}\right) \mathrm{d}t = \frac{8m\mathrm{Catalan}}{\pi}.
\end{aligned}$$

In the last estimate we have used the classical fact that the midpoint quadrature rule gives an upper bound for the corresponding integral of some concave function. This proves (24). We remark that the term ρ in (24) can in fact be shown to be optimal.

4 Applications of the generalized Hankel eigenvalue problem

In this section we briefly describe two applications of the generalized Hankel eigenvalue problem, sparse polynomial interpolation and the shape from moments problem.

4.1 Sparse interpolation of black box polynomials

Sparse interpolation is the problem of finding a sparse standard representation of a multivariate polynomial $f(x_1,\ldots,x_d)$ represented as a black-box. That is we want to find constants c_i and powers $m_{i,j}$ such that one can write f in its standard form

$$f(x_1,\ldots,x_d) = \sum_j^t c_j x_1^{m_{1,j}} \cdots x_d^{m_{d,j}}. \tag{25}$$

Here the number of nonzero terms t is typically considered as being known. The sparse interplation problem in a numerical context has been introduced and studied by Giesbrecht et al. [12,14]. Applications of sparse interpolation include its use in approximate polynomial factorization and the decomposition of approximately specified polynomial systems. We refer the reader to the references in [14] for more information on these applications.

Assuming that an upper bound of the degree for each x_j is available (which is typically the case), the initial approach found in [14] evaluates the polynomial at the points

$$h_k = f(v_1^k, \ldots, v_d^k), \quad k = 0, \ldots, 2t-1$$

with each $v_j = \mathrm{e}^{\frac{2\pi i}{p_j}}$. Here p_j are relatively prime integers with p_j larger than the upper bound of the degree of term x_j. If we set

$$\lambda_j = \mathrm{e}^{2\pi i\left(\frac{m_{1,j}}{p_1}+\cdots+\frac{m_{d,j}}{p_d}\right)},$$

then this sparse interpolation problem becomes one of determining the c_j and λ_j satisfying

$$h_k = \sum_{j=1}^{t} c_j {\lambda_j}^k, \quad k = 0, \ldots, 2t-1.$$

Once the λ_j are known, the individual powers $m_{i,j}$ are determined by taking logarithms and applying Chinese remaindering. We refer the reader to [14] for further details. Notice that in this application each λ_j lies on the unit circle.

Example 4.1 Consider the sparse interpolation problem for the simple univariate polynomial

$$f(x) = \sum_{j=1}^{11} c_j x^{j-1} + \sum_{j=1}^{10} c_{11+j} x^{p+j-11}$$

having 21 terms and p as our upper bound. In this example the corresponding abscissas on the unit circle are $\left\{1, \mathrm{e}^{\frac{2\pi i}{p}}, \ldots, \mathrm{e}^{\frac{20\pi i}{p}}, \mathrm{e}^{\frac{-20\pi i}{p}}, \ldots, \mathrm{e}^{\frac{-2\pi i}{p}}\right\}$. In this case (22) together with Example 3.7 for $\alpha = \frac{20\pi}{p}$ gives a lower bound for the conditioning of the corresponding generalized eigenvalues. Indeed for a degree as small as $p = 101$ the resulting condition number is already bounded below by 10^{20}, illustrating the conditioning problem for this example.

To reduce the ill-conditioning in the sparse interpolation problem, [14] adopts a strategy of evaluating points at $v_j = \mathrm{e}^{2\pi i \frac{r_j}{p_j}}$ with r_j a *random* point between

1 and $p_j - 1$, where p_j is a distinct prime and

$$\lambda_j = \mathrm{e}^{2\pi i\left(\frac{m_{1,j}r_1}{p_1}+\cdots+\frac{m_{d,j}r_d}{p_d}\right)}.$$

They show that, for such a random choice of root of unity, the condition number of the associated problem is reasonable with probability at least $\frac{1}{2}$, independent of the input polynomial.

Of course for an unlucky choice of a root of unity, the conditioning may still be high. In the example above having an r between 1 and $\frac{p}{20}$ only ensures a distribution on the half circle. Here (22) together with Example 3.7 for $\alpha = \pi$ give the lower bound

$$\gamma_{21}(E) \geq \frac{(1+\sqrt{2})^{20}}{8\sqrt{21}} - \frac{5}{\sqrt{21}} \approx 1.23 \times 10^6,$$

which together with (15) and (16) implies that at least some of the λ_j are quite ill-disposed. The authors in [14] propose to restart the interpolation if such ill-conditioning is encountered, as another random choice will again be well-conditioned with reasonable probability.

4.2 The shape from moments problem

The shape from moments problem is to find the vertices z_j of a polygonal domain D using complex moments. The problem has a wide variety of uses in both pure and applied mathematics, for example for solving inverse problems for uniform regions related to general elliptical equations. We refer the reader to [15] and the references therein for examples of such applications.

The method takes advantage of a formula originally due to Motzkin and Schoenberg [6, also the references given there] and Davis [6,7] which states that for any simply connected polygonal domain D with t vertices $z_1, \ldots, z_t$, there are weights $c_1, \ldots, c_t$ such that for any function f analytic in the closure of D

$$\int\int_D f''(z)\,\mathrm{d}x\mathrm{d}y = \sum_{j=1}^{t} c_j f(z_j). \tag{26}$$

In fact, there is an explicit formula for the c_j in terms of the z_j (see, e.g., [15, Eq. (2.2)]), and $|c_j| = |\sin(\psi_j)|$, with ψ_j being the outer angle of D at z_j.

If we let

$$h_k = k(k-1)\int\int_D z^{k-2}\mathrm{d}x\,\mathrm{d}y$$

be *weighted complex moments*, then formula (26) becomes

$$h_k = \sum_{j=1}^{t} c_j \, z_j^k$$

so that our shape from moments problems attempts to find $c_1, \ldots, c_t$ and $z_1, \ldots, z_t$ from the knowledge of h_k for $0 \leq k \leq 2t-1$.

In practice one also transforms these moments via shifts in order to center the problem around the center of mass of the polygonal region. In addition, in [15] these transformed moments are scaled in order that the region can be moved inside the unit disk. If the region is already inside the unit disk then the scaling is done in such a way as to enlarge the region so as to take more space inside the disk.

As a consequence of the scaling used in the shape from moments problem the unit disk case holds, and so implies that $\|H_n\| + \|\widetilde{H}_n\| \leq 2n^2$. In addition, $\|H_n\| \geq |h_2|$, which by (26) equals twice the volume of D. Thus, in view of (13), the quantity

$$\|W_n^{-1} e_j\|^2 = \frac{1}{|\sin(\psi_j)|} \|V_n^{-1} e_j\|^2$$

is a good measure for ill-disposedness of the generalized eigenvalue λ_j of problem (1), or, in other words, of the sensitivity of the problem of recovering the vertex λ_j from the moments. The following example shows that, even for domains where $|\sin(\psi_j)|$ is not too small, the quantities $\|V_n^{-1} e_j\|$ are quite often exponentially increasing in n, implying that the corresponding vertices are ill-disposed.

Example 4.2 Let $\{\lambda_1, \ldots, \lambda_n\}$ with $n = 2m$ be the set

$$\lambda_{2j} = R \, \mathrm{e}^{2\pi i \frac{j}{m}} \quad \text{and} \quad \lambda_{2j-1} = r \, \mathrm{e}^{\pi i \frac{(2j-1)}{m}} \quad j = 1, \ldots, m$$

with $1 \geq R \geq r > 0$. Consider the shape from moments problem used to reconstruct the star pattern with m equally spaced vertices λ_{2j-1} on an inside circle of radius r and m equally spaced vertices λ_{2j} on an outside circle of radius R. See Fig. 2 for $n = 18$.

In order to describe the sensitivity of the problem of recovering vertex λ_j from the moments, we determine the asymptotic behavior of $\|V_n^{-1} e_{2j}\|$ and of $\|V_n^{-1} e_{2j-1}\|$ (which by symmetry do not depend on j) for $n \to \infty$. We again make use of the formulas (10) and (11) from Sect. 2 expressing $V_n^{-1} e_j$ in terms of Lagrange polynomials. Here we observe that the polynomial ω of (10) can be factored as $\omega(z) = \omega_1(z) \, \omega_2(z)$ where

$$\omega_1(z) = \prod_{j=1}^{m} (z - \lambda_{2j-1}) = z^m + r^m, \quad \omega_2(z) = \prod_{j=1}^{m} (z - \lambda_{2j}) = z^m - R^m.$$

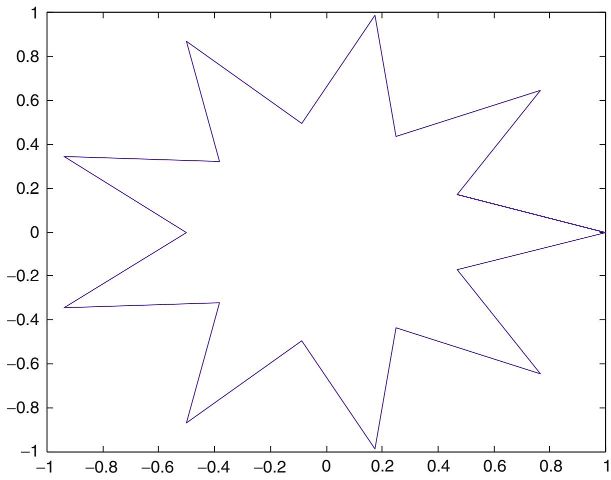

Fig. 2 The polygon of Example 4.2 for $n = 18$ with $R = 1$ and $r = \frac{1}{2}$

Consequently, $|\omega'(\lambda_{2j})| = mR^{m-1}(R^m + r^m)$, and, by (11),

$$\|V_n^{-1}e_{2j}\|^2 = \frac{\|\omega(z)/(z-\lambda_{2j})\|^2_{L_2(\partial\mathbb{D})}}{|\omega'(\lambda_{2j})|^2} = \frac{\|\omega_1(z)\sum_{k=0}^{m-1} z^{m-1-k}\lambda_{2j}^k\|^2_{L_2(\partial\mathbb{D})}}{|\omega'(\lambda_{2j})|^2}$$
$$= \frac{\epsilon_m}{m^2R^{4m-2}},$$

where $\epsilon_m \in [1/4, 2m]$. Similarly one can show that

$$\|V_n^{-1}e_{2j-1}\|^2 = \frac{\|\omega_2(z)\sum_{k=0}^{m-1} z^{m-1-k}\lambda_{2j-1}^k\|^2_{L_2(\partial\mathbb{D})}}{|mr^{m-1}(r^m + R^m)|^2} = \frac{\epsilon'_m}{m^2r^{2m-2}R^{2m}},$$

with $\epsilon'_m \in [1/4, 2m]$. Thus, for inner vertices, our measure for ill-disposedness is exponentially increasing in n unless $r = R = 1$ (i.e., D is a regular polygon with $2n$ vertices). Moreover, if the outer vertices do not lie on the unit circle $R = 1$ but strictly inside the unit disk ($R < 1$), then also our measure for ill-disposedness is exponentially increasing in n. This confirms a previously mentioned claim that one should scale D in a way that it takes as much space as possible in the unit disk.

5 Proof for theorem 3.6

Our proof of Theorem 3.6 is divided in several parts. First, in Lemma 5.1 we relate the quantity $\gamma_n(E)$ defined in (14) to the solution $\tilde{\gamma}_n(E)$ of an extremal problem for weighted polynomials in the complex plane, namely, the problem of finding the quantity

$$\tilde{\gamma}_n(E) := \max\left\{\frac{\|P\|_{L_\infty(\partial\mathbb{D})}}{\|\rho^n P\|_{L_\infty(E)}} : P \text{ is a polynomial of degree} \le n\right\}, \tag{27}$$

where $\rho(z) := 1/\max(|z|, 1)$. In Lemma 5.2 we report about some recent results on weighted polynomials and relate the quantity $\tilde{\gamma}_n(E)$ to $\gamma(E)$ defined in (19) and (20). For the special case of simply connected $E \subset \mathbb{D}$ we give, in Lemma 5.3, sharper estimates for a quantity closely related to $\tilde{\gamma}_n(E)$. Finally, we follow the Lemma with a proof of Theorem 3.6.

Lemma 5.1 *Let $E \subset \mathbb{C}$ be compact and $n \ge 1$ an integer. Then*

$$\frac{1}{\sqrt{n}}\,\tilde{\gamma}_{n-1}(E) \le \gamma_n(E) \le n^2\,\tilde{\gamma}_{n-1}(E).$$

Proof Consider $\lambda_1, \dots, \lambda_n \in E$, and the corresponding Lagrange polynomials ℓ_j as defined in (10). For any polynomial P of degree at most $n-1$ we obtain, using the Lagrange interpolation formula and (4),

$$\begin{aligned}\|P\|_{L_\infty(\partial\mathbb{D})} \le \sqrt{n}\,\|P\|_{L_2(\partial\mathbb{D})} &\le \sqrt{n}\sum_{j=1}^{n}|P(\lambda_j)|\,\|\ell_j\|_{L_2(\partial\mathbb{D})}.\\ &\le \sqrt{n}\,\|\rho^{n-1}P\|_{L_\infty(E)}\sum_{j=1}^{n}\frac{1}{\rho(\lambda_j)^{n-1}}\,\|\ell_j\|_{L_2(\partial\mathbb{D})}.\end{aligned}$$

Since $\|\ell_j\|_{L_2(\partial\mathbb{D})} = \|V_n^{-1}e_j\|$ by (11) and $\|V_n^t e_j\| = [1+|\lambda_j|^2+\dots+|\lambda_j|^{2n-2}]^{1/2} \ge 1/\rho(\lambda_j)^{n-1}$ with the weight function $\rho(z) = 1/\max(1,|z|)$ as defined above, it follows that

$$\|P\|_{L_\infty(\partial\mathbb{D})} \le \sqrt{n}\,\|\rho^{n-1}P\|_{L_\infty(E)}\,\gamma_n(E)$$

for any polynomial P of degree at most $n-1$. Thus $\tilde{\gamma}_{n-1}(E) \le \sqrt{n}\,\gamma_n(E)$, as claimed in the assertion of Lemma 5.1.

In order to show the other estimate, we first have to choose "good" points $\lambda_j \in E$ such that the discrete set $E_n := \{\lambda_1, \dots, \lambda_n\}$ satisfies $\tilde{\gamma}_{n-1}(E_n) \approx \tilde{\gamma}_{n-1}(E)$. Here we will take the set of weighted Fekete points[1] defined as follows (cf. [16, Sect. III.1]). If we explicitly denote $V_n(\lambda_1, \dots, \lambda_n)$ for a Vandermonde matrix with abscissa $\lambda_1, \dots, \lambda_n$, then the weighted Fekete points are those points in E

[1] For special sets E some other points like Fejer points or suitable alternants might lead to sharper estimates. Compare with [4].

maximizing the expression

$$\det\, V_n(\lambda_1,\ldots,\lambda_n) \prod_{k=1}^{n} \rho(\lambda_k)^{n-1}.$$

For the corresponding Lagrange polynomials ℓ_j we have

$$\begin{aligned}\|\rho^{n-1}\ell_j\|_{L_\infty(E)} &= \rho^{n-1}(\lambda_j) \max_{z\in E} \frac{\rho(z)^{n-1}}{\rho(\lambda_j)^{n-1}} \frac{\det\, V_n(\lambda_1,\ldots,\lambda_{j-1},z,\lambda_{j+1},\ldots,\lambda_n)}{\det\, V_n(\lambda_1,\ldots,\lambda_n)} \\ &\le \rho^{n-1}(\lambda_j)\end{aligned}$$

by definition of the Fekete points. Hence, for any polynomial P of degree at most $n-1$, by the Lagrange interpolation formula,

$$1 \le \frac{\|\rho^{n-1}P\|_{L_\infty(E)}}{\|\rho^{n-1}P\|_{L_\infty(E_n)}} \le \sum_{j=1}^{n} \frac{|P(\lambda_j)|}{\|\rho^{n-1}P\|_{L_\infty(E_n)}}\, \|\rho^{n-1}\ell_j\|_{L_\infty(E)} \le n,$$

implying that $\tilde{\gamma}_{n-1}(E) \le \tilde{\gamma}_{n-1}(E_n) \le n\tilde{\gamma}_{n-1}(E)$. Also, by (the sharpness of) the Hölder inequality applied to the Lagrange interpolation formula, we have that, for any $\zeta \in \mathbb{C}$,

$$\max\left\{\frac{|P(\zeta)|}{\|\rho^n P\|_{L_\infty(E_n)}} :\ P \text{ is a polynomial of degree} \le n-1\right\} = \sum_{j=1}^{\infty} |\ell_j(\zeta)|/\rho(\lambda_j)^{n-1}.$$

Thus

$$\begin{aligned} n\tilde{\gamma}_{n-1}(E) \ge \tilde{\gamma}_{n-1}(E_n) &= \max_{\zeta\in\partial\mathbb{D}} \sum_{j=1}^{n} |\ell_j(\zeta)|/\rho(\lambda_j)^{n-1} \\ &\ge \left[\max_{\zeta\in\partial\mathbb{D}} \sum_{j=1}^{n} |\ell_j(\zeta)|^2/\rho(\lambda_j)^{2n-2}\right]^{1/2} \\ &\ge \left[\sum_{j=1}^{n} \frac{1}{\rho(\lambda_j)^{2n-2}} \frac{1}{2\pi} \int_{|\zeta|=1} |\ell_j(\zeta)|^2\, |\mathrm{d}\zeta|\right]^{1/2} = \left[\sum_{j=1}^{n} \frac{\|\vec{\ell}_j\|^2}{\rho(\lambda_j)^{2n-2}}\right]^{1/2} \\ &\ge \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \frac{\|\vec{\ell}_j\|}{\rho(\lambda_j)^{n-1}}. \end{aligned}$$

On the other hand, by (14), (17), and (11),

$$\gamma_n(E) \leq \gamma_n(E_n) = \sum_{j=1}^{n} \|\vec{\ell}_j\| \sqrt{1 + |\lambda_j|^2 + \cdots + |\lambda_j|^{2n-2}} \leq \sqrt{n} \sum_{j=1}^{n} \frac{\|\vec{\ell}_j\|}{\rho(\lambda_j)^{n-1}},$$

and consequently the second inequality of Lemma 5.1 is true.

The following result on the weighted polynomial extremal problem (27) was obtained in a more general setting by Maskhar and Saff in 1985. We refer the reader to [16, Sect. III.2] and the references therein.

Lemma 5.2 *Let $E \subset \mathbb{C}$ be compact, and regular with respect to the Dirichlet problem. Then for any integer $n \geq 1$ we have that $\tilde{\gamma}_n(E) \leq \gamma(E)^n$, and*

$$\lim_{n\to\infty} \tilde{\gamma}_n(E)^{1/n} = \gamma(E).$$

Proof The logarithmic potential of a probability measure μ is defined by

$$U^{\mu}(z) := \int \log\left(\frac{1}{|z-t|}\right) \mathrm{d}\mu(t).$$

According to [16, Theorems I.1.3 and I.4.8], given some function Q continuous on E there exists a unique probability measure μ_Q (called the equilibrium measure in the presence of an external field Q) supported on E and a constant c_Q such that

$$U^{\mu_Q}(z) + Q(z) \begin{cases} = c_Q & \text{for } z \text{ lying in the support of } \mu_Q, \\ \geq c_Q & \text{for } z \in E. \end{cases}$$

In addition, μ_Q has a continuous potential by the assumption on E. The weighted Bernstein–Walsh inequality [16, Theorem III.2.1] tells us that, for any integer $n \geq 0$, for any polynomial P_n of degree at most n, and for any $z \in \mathbb{C}$, we have

$$|P_n(z)| \leq \mathrm{e}^{n(c_Q - U^{\mu_Q}(z))} \, \|\mathrm{e}^{-nQ} P_n\|_{L_\infty(E)}.$$

Moreover, this inequality is sharp since by [16, Theorem III.5.3] and the following remarks we have for all $z \in \mathbb{C}$

$$\mathrm{e}^{c_Q - U^{\mu_Q}(z)} = \lim_{n\to\infty} \sup\left\{\left[\frac{|P_n(z)|}{\|\mathrm{e}^{-nQ} P_n\|_{L_\infty(E)}}\right]^{1/n} : P_n \text{ a polynomial of degree } n\right\}.$$

Thus the assertion of the Lemma 5.2 follows by showing that the constant $\gamma(E)$ of (19) equals

$$\gamma(E) = \max_{z\in\partial\mathbb{D}} \exp(c_Q - U^{\mu_Q}(z)) \tag{28}$$

for our external field $Q = \log(1/\rho)$. Using [16, Example 0.5.7] we have that

$$Q(z) = \log\left(\frac{1}{\rho(z)}\right) = \log(\max(1,|z|)) = -\int_{-\pi}^{\pi} \log\left(\frac{1}{|z-t|}\right)\frac{\mathrm{d}t}{2\pi},$$

that is, the external field coincides with (-1) times the logarithmic potential U^ν of a probability measure ν. Thus we know from [16, Example II.4.8]) that $\mu_Q = \hat{\nu}$, the balayage measure of ν onto E (see, e.g., [16, Theorem II.4.7]), and

$$\begin{aligned} &U^{\hat{\nu}}(z) + Q(z) = U^{\hat{\nu}}(z) - U^{\nu}(z) = F(\infty) - F(z), \quad \text{where} \\ &F(z) := \int g_E(z,t)\,\mathrm{d}\nu(t) = \frac{1}{2\pi}\int_0^{2\pi} g_E(z,\mathrm{e}^{it})\,\mathrm{d}t. \end{aligned}$$

The latter formula follows from integration [16, Eq. (4.32) of Chap. II] with respect to ν. Notice that, since the Green function is continuous, the function F equals zero on E, and hence $c_Q = F(\infty)$. Finally, by Eq. (19) we have that $\gamma(E) = \|F\|_{L_\infty(\partial\mathbb{D})}$, and Q is identically zero on $\partial\mathbb{D}$. Thus (28) holds.

It only remains to establish formula (20). Let $E \subset \mathbb{D}$. Then U^ν is constant on E, and hence $\hat{\nu}$ is the equilibrium measure of E. Using the classical link between Green functions and potentials of equilibrium measures (cf. [16, Eq. (4.8) of Chap. I]), we conclude that $F(z) = g_{\partial\mathbb{D}}(z,\infty) - g_E(z,\infty) + c$ for some constant c. Comparing the values on E yields that $c = 0$ and so implies the alternate representation (20) in the case $E \subset \mathbb{D}$.

Lemma 5.3 *Let $E \subset \mathbb{C}$ be a compact simply connected set of bounded rotation V. As before, let ϕ map $\overline{\mathbb{C}} \setminus E$ conformally to $\overline{\mathbb{C}} \setminus \mathbb{D}$, with $\phi(\infty) = \infty$, and $\phi'(\infty) > 0$, and let*

$$Q(x) = \prod_{k=1}^{M}(x-\omega_k), \quad \omega_k \in \mathbb{C}\setminus E.$$

Then for all $z \notin E$ and $m \geq M$

$$\max\left\{\frac{|P(z)/Q(z)|}{\|P/Q\|_{L_\infty(E)}} : \deg P \leq m\right\} \geq \frac{\pi}{V}|\phi(z)|^{m-M}\prod_{j=1}^{M}\left|\frac{1-\phi(z)\overline{\phi(\omega_k)}}{\phi(z)-\phi(\omega_k)}\right| - 1 - \frac{\pi}{V}.$$

Proof Consider the function

$$\tilde{R}(w) = w^{m-M} \prod_{k=1}^{M} \frac{1 - w\overline{\phi(\omega_k)}}{w - \phi(\omega_k)}$$

which is a Blaschke product, analytic in $\mathbb{D}$, and of modulus 1 on $\partial\mathbb{D}$, and so in particular satisfies

$$\|\tilde{R}\|_{L_\infty(\partial\mathbb{D})} = 1.$$

Following [1,8] we also consider the Faber map T which for a rational function R analytic in $|w| \le \theta$, $\theta > 1$, is defined by

$$r(z) := T(R)(z) = \frac{1}{2\pi i} \int\limits_{|\phi(\zeta)|=\theta} \frac{R(\phi(\zeta))}{\zeta - z}\,\mathrm{d}\zeta, \quad |\phi(z)| < \theta.$$

It is known (see, e.g., [8, Lemma 1.1 and Proposition 1.2]) that, with R, also $T(R)$ is rational, and $z \mapsto T(R)(z) - R(\phi(z))$ has an analytic continuation outside E, including infinity. In particular, $T(R)$ has all its poles outside of E, namely at the images under ψ, the inverse mapping of ϕ, of R. For the operator norm we know from [1] that

$$\frac{\|T(R)\|_{L_\infty(\partial E)}}{\|R\|_{L_\infty(\partial\mathbb{D})}} \le \|T\| \le \frac{V}{\pi}.$$

As a consequence, the function $\tilde{P}$ defined by

$$\tilde{P}/Q = \tilde{r} = T(\tilde{R})$$

is a polynomial of degree at most m, and, with the help of the maximum principle, we get for $z \notin E$

$$\begin{aligned}
\frac{|\tilde{r}(z)|}{\|\tilde{r}\|_{L_\infty(E)}} &\ge \frac{\pi}{V} \frac{|\tilde{R}(\phi(z))| - |\tilde{r}(z) - \tilde{R}(\phi(z))|}{\|\tilde{R}\|_{L_\infty(\partial\mathbb{D})}} = \frac{\pi}{V}\Big[|\tilde{R}(\phi(z))| - |\tilde{r}(z) - \tilde{R}(\phi(z))|\Big] \\
&\ge \frac{\pi}{V}\Big[|\tilde{R}(\phi(z))| - \|\tilde{r} - \tilde{R}\circ\phi\|_{L_\infty(\partial E)}\Big] \\
&\ge \frac{\pi}{V}\Big[|\tilde{R}(\phi(z))| - \|\tilde{r}\|_{L_\infty(\partial E)} - \|\tilde{R}\|_{L_\infty(\partial\mathbb{D})}\Big] \\
&\ge \frac{\pi}{V}|\tilde{R}(\phi(z))| - 1 - \frac{\pi}{V},
\end{aligned}$$

as claimed in Lemma 5.3.

Proof of Theorem 3.6 The first part of the Theorem follows by combining Lemmas 5.1 and 5.2. Suppose now that E is simply connected and of bounded variation. We consider first the case $\partial\mathbb{D} \subset E$. Since $\|V_n^{-1}e_j\|\,\|V_n^t e_j\| \geq |e_j^t V_n V_n^{-1} e_j| = 1$, we see from (17) that $\gamma_n(E) \geq n$. On the other hand, for the abscissa $\lambda_1, \ldots, \lambda_n$ being the nth roots of unity we know that $V_n^{-1} = \frac{1}{n}V_n^*$ and $\|V_n^* e_j\| = \sqrt{n}$, and thus $\gamma_n(E) = n$ by (14) and (17). Thus in this case the last part of Theorem 3.6 is trivially true since $\gamma(E) = 1$.

If $\partial\mathbb{D} \not\subset E$ then $\gamma(E) > 1$ by (19), and there exists a $z' \in \partial\mathbb{D} \setminus E$ with

$$\log(\gamma(E)) = \frac{1}{2\pi}\int_0^{2\pi} g_E(z', \mathrm{e}^{it})\,\mathrm{d}t = \frac{1}{2\pi}\int_0^{2\pi} g(\mathrm{e}^{it})\,\mathrm{d}t,$$

where $g(z) := g_E(z', z) - \log\left(\frac{1}{|z'-z|}\right)$, a continuous function on $\partial\mathbb{D}$. Notice that

$$\log(\gamma(E)^n) = \frac{1}{\int_0^{2\pi/n}\mathrm{d}t}\int_0^{2\pi/n}\sum_{k=1}^{n} g(\mathrm{e}^{2\pi i/k}\mathrm{e}^{it})\,\mathrm{d}t \leq \sum_{k=1}^{n} g(\mathrm{e}^{2\pi i/k}\mathrm{e}^{it'})$$

for some $t' \in [0, 2\pi/n]$. Since g is continuous we may suppose without loss of generality that $z' \notin \{\mathrm{e}^{2\pi i/k}\mathrm{e}^{it'} : k = 1, \ldots, n\}$. Taking into account (18) and writing $\omega_k = \mathrm{e}^{2\pi i/k}\mathrm{e}^{it'}$, we conclude that

$$\gamma(E)^n \leq |(z')^n - \mathrm{e}^{it'n}| \prod_{\omega_k \notin E}\left|\frac{1 - \phi(z')\,\overline{\phi(\omega_k)}}{\phi(z') - \phi(\omega_k)}\right|. \tag{29}$$

Let

$$Q(z) = \prod_{\omega_k \notin E}(z - \omega_k), \quad \tilde{Q}(z) = \prod_{\omega_k \in E}(z - \omega_k), \quad m := \deg Q,$$

and observe that, for $z \in E$,

$$\rho(z)^n\,|Q(z)|\,|\tilde{Q}(z)| = \frac{|z^n - \mathrm{e}^{it'n}|}{\max\{1, |z|^n\}} \leq 2.$$

Hence,

$$\begin{aligned}
\gamma_n(E) &\geq \max\left\{\frac{|P(z')|}{\|\rho^n P\|_{L_\infty(E)}} : \deg P \leq n\right\} \\
&\geq \max\left\{\frac{|\tilde{Q}(z')P(z')|}{\|\rho^n \tilde{Q}P\|_{L_\infty(E)}} : \deg P \leq m\right\} \\
&\geq \frac{|(z')^n - \mathrm{e}^{it'n}|}{2}\max\left\{\frac{|P(z')/Q(z')|}{\|P/Q\|_{L_\infty(E)}} : \deg P \leq m\right\}.
\end{aligned}$$

Replacing the right-hand expression by the lower bound found in Lemma 5.3 for $M = m$ and using (29) together with the trivial estimate $|(z')^n - \mathrm{e}^{it'n}| \leq 2$, we obtain

$$\frac{\pi}{2V}\gamma(E)^n - 1 - \frac{\pi}{V} \leq \gamma_n(E) \leq \gamma(E)^n,$$

the right-hand bound having been already established in Lemma 5.2. Together with Lemma 5.1 we may conclude that the second part of Theorem 3.6 is also true in the case $\partial\mathbb{D} \not\subset E$.

6 Conclusions

In this paper we have shown that structured and unstructured errors of the GEP for Hankel matrices (1) are bounded by essentially the same quantities. We also relate the relative sensitivity of the most sensitive eigenvalue to the condition number of the underlying Hankel matrix. Finally, a new result on the growth of the condition number of special scaled Vandermonde matrices with complex nodes is also given. This result shows that the conditioning of the generalized eigenvalue problem (1) for a subclass of important problems of interest grows exponentially in terms of a quantity that depends on the largest distance between eigenvalues.

In this paper we have analyzed only the sensitivity with respect to ordinary moments. It is natural to ask what can be done with modified moments. Recall that a well-known procedure of improving the sensitivity of the GEP is to consider the equivalent problem

$$T_n^t(\widetilde{H}_n - \lambda H_n)T_n y = 0$$

where T_n is some suitably chosen invertible upper triangular matrix. By considering the $(k+1)$th column of T_n as the coefficient vector of some polynomial p_k of degree k for $k = 0, \ldots, n-1$, we have from (3) that this new GEP can be rewritten as $(\widetilde{\mathbf{H}} - \lambda\mathbf{H})y = 0$, where the entries of $\widetilde{\mathbf{H}} = T_n^t\widetilde{H}_n T_n$, and $\mathbf{H} = T_n^t H_n T_n$, respectively, are given by

$$\widetilde{\mathbf{H}}_{k,\ell} = \sum_{j=1}^{n} c_j\lambda_j p_k(\lambda_j)p_\ell(\lambda_j), \quad \mathbf{H}_{k,\ell} = \sum_{j=1}^{n} c_j p_k(\lambda_j)p_\ell(\lambda_j).$$

Such matrices are usually referred to as modified Gramians. The entries $\sum_{j=1}^{n} c_j p_k(\lambda_j)$ of the (possibly scaled) first column/row of $\mathbf{H}$ are usually referred to as modified moments.

Following the ideas of Sect. 2.2 one may show that an (unstructured) ϵ–perturbation of these two matrices similar to (7) leads to a perturbation of the jth eigenvalue, which in first order can be bounded above by

$$\left|\frac{\mathrm{d}\lambda_j}{\mathrm{d}\epsilon}(0)\right| \leq (|\lambda_j|+1)\|\mathbf{W}^{-1}e_j\|^2, \quad \mathbf{W} = \mathrm{diag}\,(\sqrt{c_j})\,\mathbf{V}, \quad \mathbf{V} = \Big(p_{k-1}(\lambda_j)\Big)_{j,k=1,\dots,n}.$$

That is, we have to replace the Vandermonde matrix in (9) by a generalized Vandermonde matrix. The norm of the inverse and the condition number of such (possibly row-scaled) generalized Vandermonde matrices has been discussed by a number of authors for various choices of bases and particular abscissa (for a summary see for instance [3]). A good choice of a family of polynomials p_k may dramatically improve the sensitivity of the GEP. For instance, we can produce a unitary $\mathbf{W}$ by taking as p_k the orthonormal polynomials with respect to the hermitian scalar product $\ll p, q \gg = \sum_{j=1}^{n} |c_j| \overline{p(\lambda_j)} q(\lambda_j)$. As another example, choosing the corresponding formal orthogonal polynomials leads to $\mathbf{H}$ being the identity matrix (and thus $\mathbf{W}^{-1} = \mathbf{W}^t$) and $\widetilde{\mathbf{H}}$ being a (in general non hermitian) tridiagonal matrix. Here

$$\|\mathbf{W}^{-1}e_j\|^2 = \frac{\sum_{\ell=0}^{n-1} |p_\ell(\lambda_j)|^2}{|\sum_{\ell=0}^{n-1} p_\ell(\lambda_j)^2|}$$

and thus the sensitivity will depend on the growth behaviour of the formal orthogonal polynomials of the support of (formal) orthogonality. This in turn is a subject of research where (up to the simple case of real orthogonality $\lambda_j \in \mathbb{R}$ and $c_j > 0$) very few results are known.

In summary, a suitable choice of the family of polynomials may lead to a sparse or structured eigenvalue problem or to well-conditioned GEPs (or sometimes even both). Unfortunately, in general for our applications the entries of the modified Gramians are not available and need to be computed which is a potential new source of errors. In addition, it seems to us that in problems like shape detection or sparse interpolation in general we do not have sufficient a priori information to design a "good" family of polynomials p_k. As such it remains a topic for future research to alter our GEP in order to arrive at better behaved numerical procedures for effective computation.

Of course there are other GEPs that remain of interest to us for specific applications. As an example, sparse interpolation problems in computer algebra often look for other bases in which sparse representations may be possible. These include bases in terms of Cheyshev polynomials, Pochhammer (factorial) polynomials and others (c.f. [13]). It would be of interest to study the structured perturbations of these problems and the geometric properties of the set of generalized eigenvalues that might effect condition numbers of such problems.

References

1. Anderson, J.M.: The Faber operator. In: Rational Approximation and Interpolation. Proceedings Tampa, Lecture Notes, vol. 1105. Springer, Berlin Heidelberg New York (1983)

2. Baker, G.A., Graves-Morris, P.R.: Padé approximants.In: Encyclopedia of Mathematics, 2nd edn. Cambridge University Press, New York (1995)
3. Beckermann, B.: On the numerical condition of polynomial bases: estimates for the condition number of Vandermonde, Krylov and Hankel matrices. Habilitationsschrift, Universität Hannover (1996)
4. Beckermann, B.: The condition number of real Vandermonde, Krylov and positive definite Hankel matrices. Numer. Math. **85**, 553–577 (2000)
5. Beckermann, B., Bourreau, E.: How to choose modified moments? J. Comput. Appl. Math. **98**, 81–98 (1998)
6. Davis, P.J.: Triangle formulas in the complex plane. Math. Comput. **18**(88), 569–577 (1964)
7. Davis, P.J.: Plane regions determined by complex moments. J. Approx. Theory **19**(2), 5148–5153 (1977)
8. Ellacott, S.W., Saff, E.B.: Computing with the Faber transform. In: Rational Approximation and Interpolation. Proceedings Tampa, Lecture Notes, vol. 1105. Springer, Berlin Heidelberg New York, 412–418 (1983)
9. Fischer, H.-J.: On the condition of orthogonal polynomials via modified moments. Z. Anal. Anwend. **15**(1), 1–18 (1996)
10. Gautschi, W.: On generating orthogonal polynomials. SIAM J. Sci. Stat. Comput. **3**(3), 298–317 (1982)
11. Gautschi, W.: On the sensitivity of orthogonal polynomials to perturbations in the moments. Numer. Math. **48**, 369–382 (1986)
12. Giesbrecht, M., Labahn, G., Lee, W.-S.: Symbolic-numeric sparse interpolation of multivariate polynomials. In: Proceedings of the 9th Rhine Workshop on Computer Algebra (2004)
13. Giesbrecht, M., Labahn, G., Lee, W.-S.: Symbolic-numeric sparse polynomial interpolation in Chebyshev basis and trigonometric interpolation. In: Proceedings of Computer Algebra in Scientific Computing (CASC 2004), St. Petersburg, Russia, 195–206 (2004)
14. Giesbrecht, M., Labahn, G., Lee, W.-S.: Symbolic-numeric sparse interpolation of multivariate polynomials. In: Proceedings of ISSAC'06. ACM, Genoa, 116–123 (2006)
15. Golub, G.H., Milanfar, P., Varah, J.: A stable numerical method for inverting shape from moments. SIAM J. Sci. Comput. **21**, 1222–1243 (1999)
16. Saff, E.B., Totik, V.: Logarithmic potentials with external fields. Springer, Berlin Heidelberg New York (1997)
17. Wilkinson, J.H.: Modern error analysis. SIAM Rev. **13**, 548–568 (1971)