Noname manuscript No.
(will be inserted by the editor)

The Gauss quadrature for general linear functionals, Lanczos algorithm, and minimal partial realization

Stefano Pozza · **Miroslav S. Pranić**

March 27, 2019

Abstract The concept of Gauss quadrature can be generalized to approximate linear functionals with complex moments. Following the existing literature, this survey will describe such generalization. It is well known that the (classical) Gauss quadrature for positive definite linear functionals is connected with orthogonal polynomials, and with the (Hermitian) Lanczos algorithm. Analogously, the Gauss quadrature for linear functionals is connected with formal orthogonal polynomials, and with the non-Hermitian Lanczos algorithm with look-ahead strategy; moreover, it is related to the minimal partial realization problem. We will review these connections pointing out the relationships between several results established independently in related contexts.

Keywords Linear functionals, matching moments, Gauss quadrature, formal orthogonal polynomials, minimal realization, look-ahead Lanczos algorithm, Mismatch Theorem.

Mathematics Subject Classification (2010) 65D15, 65D32, 65F10, 93B15.

1 Introduction

Let A be a Hermitian positive definite matrix and $\mathbf{v}$ a vector so that $\mathbf{v}^*\mathbf{v} = 1$, where $\mathbf{v}^*$ is the conjugate transpose of $\mathbf{v}$. Consider the specific linear functional

This work has been supported by Charles University Research program No. UNCE/SCI/023.

Stefano Pozza
Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Praha 8, Czech Republic.
E-mail: pozza@karlin.mff.cuni.cz

Miroslav S. Pranić
Department of Mathematics and Informatics, University of Banja Luka, Faculty of Science, M. Stojanovića 2, 51000 Banja Luka, Bosnia and Herzegovina.
E-mail: pranic77m@yahoo.com

$\mathcal{L}$ on the space of polynomials defined by

$$\mathcal{L}(\lambda^j) := \mathbf{v}^* A^j \mathbf{v} = m_j, \quad j = 0, 1, \dots, \tag{1.1}$$

where $m_0, m_1, \dots$ are real numbers known as the *moments* of $\mathcal{L}$. The functional $\mathcal{L}$ can be expressed as the Riemann-Stieltjes integral with a non-decreasing positive distribution function $\mu(\lambda)$ supported on the real axis having finitely many points of increase; see, e.g, [60, Section 3.5],[37, Section 7.1], and [15, Chapter II, Section 3]. The n-node (classical) Gauss quadrature approximating $\mathcal{L}$ is given by the unique n-node quadrature formula which matches the first $2n$ moments, i.e.,

$$\mathcal{L}(\lambda^j) = \int_{\mathbb{R}} \lambda^j \, \mathrm{d}\mu(\lambda) = \sum_{i=1}^{n} \omega_i \, (\lambda_i)^j, \quad j = 0, \dots, 2n-1,$$

with ω_i positive weights and λ_i positive distinct nodes. Classical results of the Gauss quadrature can be found, e.g., in [79, Chapters III and XV], [15, Chapter I, Section 6], [33], [34, Section 1.4], [35, Chapter 3.2], [60, Section 3.2]. The linear functional $\mathcal{L}$ can be associated with a *Jacobi matrix* J_n which is an $n \times n$ real symmetric tridiagonal matrix. For every function f defined on the spectrum of A and J_n, the matrix J_n gives an algebraic expression for the Gauss quadrature, i.e.,

$$\mathbf{v}^* f(A) \, \mathbf{v} = \int_{\mathbb{R}} f(\lambda) \, \mathrm{d}\mu(\lambda) \approx \sum_{i=1}^{n} \omega_i \, f(\lambda_i) = \mathbf{e}_1^T f(J_n) \, \mathbf{e}_1, \tag{1.2}$$

where $f(A)$ and $f(J_n)$ are matrix functions, and $\mathbf{e}_1$ is the first vector of the Euclidean basis ($\mathbf{e}_1^T$ is the relative transpose). The matrix J_n can be obtained by n iterations of the Hermitian Lanczos algorithm with inputs A and $\mathbf{v}$. Indeed, $J_n = V_n^* A V_n$, where V_n is the matrix given by Lanczos whose columns are an orthonormal basis of the Krylov subspace $\{\mathbf{v}, A\mathbf{v}, \dots, A^{n-1}\mathbf{v}\}$. Hence the Hermitian Lanczos algorithm with input $A, \mathbf{v}$, and the Conjugate Gradient method (CG) for the solution of the linear system $A\,\mathbf{x} = \mathbf{v}$ (with zero initial approximation) give a matrix formulation of the Gauss quadrature for $\mathcal{L}$. Figure 1 (see [60, Figure 3.2]) represents the connections described above. Such connections can be derived by the properties of *orthogonal polynomials*; a detailed explanation can be found, e.g., in [60, Chapter 3] and [37] (note that the relationships between CG, Lanczos algorithm, and orthogonal polynomials were already pointed out by Hestenes and Stiefel in their seminal paper published in 1952 [46, Sections 14–17]).

This survey deals with the extension of the connections summarized in Figure 1 to the case of a general linear functional defined on the space $\mathcal{P}$ of the polynomials with generally complex coefficients, $\mathcal{L} : \mathcal{P} \to \mathbb{C}$. We point out that, if not specified otherwise, we will consider linear functionals without the underlying assumption that they are determined by a matrix bilinear form analogous to (1.1). The survey will describe the Gauss quadrature for linear functionals, its matrix formulation, its connection with the non-Hermitian

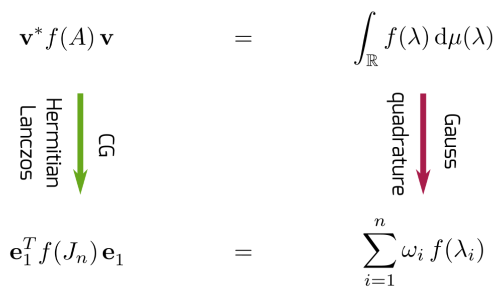


Fig. 1 Visualization of the connections between the (classical) Gauss quadrature, the Hermitian Lanczos algorithm, and the Conjugate Gradient method.

Lanczos algorithm with look-ahead strategy, and its relationship with the minimal partial realization problem. Furthermore, the connections between the incurable breakdown, the exactness of the Gauss quadrature, and the minimal realization problem, will be examined with giving an original proof of the Mismatch Theorem (first proved in [80, Theorem 4.2]). Information about the topics mentioned above and their mutual relationships are scattered in the literature. The survey aims to describe such topics and their connections from the point o view of *formal orthogonal polynomials*. We hope that such a presentation will be of interest for readers working in related different areas.

Regarding the formal orthogonal polynomials and the Gauss quadrature generalization, we will mainly follow the book [22] by Draux where the Gauss quadrature is extended for the approximation of real-valued linear functionals. More precisely, a straightforward extension of Draux's definition to the case of complex-valued linear functionals will be presented. The more recent Gauss quadrature definitions in [64] and in [71, 72], obtained independently of [22], can be seen as a generalization to the complex quasi-definite case. Indeed, for a real quasi-definite linear functional the quadratures in [22, 64, 71, 72] are equivalent. However, some results in [71, 72] do not have a counterpart in the real setting of [22] (for instance, formal orthonormal polynomials may have complex coefficients). The case of a quasi-definite linear functional is simpler to treat; see, e.g., [15, 71, 72]. The survey will first recall the primary results associated with quasi-definite functionals for then deal with the case of a general linear functional.

The paper is organized as follows. Section 2 summarizes basic results of quasi-definite linear functionals. Section 3 concerns technical results of Hankel matrices. Section 4 describes properties of formal orthogonal polynomials and of quasi-orthogonal polynomials with respect to a linear functional $\mathcal{L} : \mathcal{P} \to \mathbb{C}$. The concept of Gauss quadrature for linear functionals and its matrix interpretation can be found respectively in Section 5 and Section 6. The Gauss quadrature connections with the minimal partial realization problem and with the look-ahead Lanczos algorithm are described respectively in Section 7 and Section 8. Section 9 concludes the survey summarizing the links between the Gauss quadrature, minimal partial realization, and look-ahead Lanczos algorithm.

This survey approaches the Lanczos algorithm in a finite dimensional setting. Hence we will not treat infinite dimensional problems. For infinite dimensional problems related to positive definite linear functionals refer, e.g., to [15, Chapter II, Section 3, in particular Theorem 3.1]. For the relationship with infinite dimensional Krylov subspace methods refer, e.g., to [82], [41], and [62, Chapter 5] where many references to original works can be found.

Throughout the survey, we will consider only computations in exact arithmetic. Since rounding errors substantially affect computations with short recurrences, the results described in this survey cannot be applied to finite precision computations without a thorough analysis. Such analysis is out of the scope of this survey. The interested reader can refer to [2] and [20,21] for analysis of the non-Hermitian Lanczos algorithm in finite precision (assuming no breakdown); see also the related works [3,81,67]. As pointed out in [60, Sections 2.5.6 and 5.11], in finite precision arithmetic the short recurrences cannot preserve the biorthogonality or even the linear independence of the computed Krylov subspace basis. Therefore look-ahead techniques for the non-Hermitian Lanczos have a limited impact in computing sufficiently well-conditioned basis. The interplay of look-ahead techniques and rounding errors in practical computations is still an open issue.

2 Quasi-definite linear functionals

Let $\mathcal{L} : \mathcal{P} \to \mathbb{C}$ be a linear functional with complex moments,

$$\mathcal{L}(\lambda^k) = m_k, \quad k = 0, 1, \dots . \tag{2.1}$$

Consider the infinite Hankel matrix composed by the sequence of moments $m_0, m_1, \dots$

$$H = \begin{bmatrix} m_0 & m_1 & m_2 & \cdots \\ m_1 & m_2 & \cdots & \\ m_2 & \vdots & & \\ \vdots & & & \end{bmatrix},$$

and its k-dimensional Hankel submatrix

$$H_{k-1} = H(1:k; 1:k) = \begin{bmatrix} m_0 & m_1 & \dots & m_{k-1} \\ m_1 & m_2 & \dots & m_k \\ \vdots & \vdots & & \vdots \\ m_{k-1} & m_k & \dots & m_{2k-2} \end{bmatrix}, \quad k = 1, 2, \dots , \tag{2.2}$$

with the corresponding determinant Δ_{k-1} (the notation $A(i:j;\ell:n)$ stands for the submatrix of A composed of the elements in the rows from i to j and in the columns from ℓ to n). Hankel matrices have been used in related contexts for more than a century; see, e.g, the seminal paper by Stieltjes [77, Sections 8–11, pp. 624–630] (please notice that we refer to the English translation published by Springer in 1993), the monographs [15, Chapter I] and [22, Chapter 1], and the paper [40, Section 2]. In particular, the zero-nonzero pattern of the sequence $\Delta_0, \Delta_1, \ldots$ characterizes the linear functional $\mathcal{L}$. Indeed, denoting with $\mathcal{P}_k \subset \mathcal{P}$ the subspace of polynomials of degree at most k, the following classes of linear functionals can be defined (see, e.g., [15, Chapter I, Definition 3.1, Definition 3.2 and Theorem 3.4]).

Definition 2.1 *A linear functional $\mathcal{L}$ for which the first $k+1$ Hankel determinants are nonzero, i.e., $\Delta_j \neq 0$ for $j = 0, 1, \ldots, k$, is called* quasi-definite *on $\mathcal{P}_k$. In particular, when $\mathcal{L}$ has real moments $m_0, \ldots, m_{2k}$ and $\Delta_j > 0$ for $j = 0, 1, \ldots, k$ the linear functional is said to be* positive definite *on $\mathcal{P}_k$.*

An n-degree polynomial $p_n(\lambda) \in \mathcal{P}$ is called *formal orthogonal polynomial (FOP)* when it satisfies the *orthogonality conditions* with respect to $\mathcal{L}$

$$\mathcal{L}(p_n\, \lambda^j) = 0, \qquad \text{for } j = 0, \ldots, n-1;$$

refer, e.g., to [22, Introduction and Section 1.1] and [8, Chapter 2]. Notice that in [7] $p_n(\lambda)$ is referred as *general orthogonal polynomial*; cf. the concept of *weak orthogonal polynomial* in [55, definition on p. 137] and [57, Section 2]. The subindex n in the polynomial notation $p_n(\lambda)$ will always stand for the degree of the polynomial and we will not emphasize it further on. Moreover, whenever appropriate the argument λ will be skipped for simplicity of notation.

If $\mathcal{L}$ is quasi-definite on $\mathcal{P}_k$, then there exist unique FOPs $p_0(\lambda), \ldots, p_k(\lambda)$ (we always use the term “unique” for a polynomial in the sense of unique up to multiplication by a nonzero scalar) satisfying the conditions

$$\mathcal{L}(p_j p_n) = 0, \text{ for } j \neq n, \text{ and } \mathcal{L}(p_n^2) \neq 0;$$

see, e.g. [15, Chapter I, Theorem 3.1], [61, Chapter VII, Theorem 1]. In the case in which $\mathcal{L}(p_n^2) = 1$, the polynomials are known as formal *orthonormal* polynomial. A beautiful summary about FOPs in the quasi-definite case can be found in the book by Chihara [15] (notice that Chihara used the simplified term orthogonal polynomials instead of formal orthogonal polynomials). A sequence of formal orthonormal polynomials $p_0(\lambda), \ldots, p_k(\lambda)$ satisfy the three-term recurrence

$$\beta_n p_n(\lambda) = (\lambda - \alpha_{n-1}) p_{n-1}(\lambda) - \beta_{n-1} p_{n-2}(\lambda), \qquad n = 1, \ldots, k, \qquad (2.3)$$

where $\beta_0 = 0$, $p_{-1}(\lambda) = 0$, $p_0(\lambda) = 1/\sqrt{m_0}$ and the coefficients α_{n-1}, β_n are given by

$$\alpha_{n-1} = \mathcal{L}(\lambda p_{n-1} p_{n-1}), \;\; \beta_n = \mathcal{L}(\lambda p_{n-1} p_n).$$

see, e.g., [15, Chapter I, Section 4], [7, Theorem 2.4]. Notice that in order to avoid ambiguity, we always take the principal value of the complex square root, i.e., we consider $arg(\sqrt{c}) \in (-\pi/2, \pi/2]$. The recurrences (2.3) can be written in the compact form as

$$\lambda\, \mathbf{p}(\lambda) = J_n \mathbf{p}(\lambda) + \beta_n p_n(\lambda) \mathbf{e}_n, \quad n = 1, \dots, k,$$

where $\mathbf{p}(\lambda) = [p_0(\lambda), p_1(\lambda), \dots, p_{n-1}(\lambda)]^T$, $\mathbf{e}_n$ is the nth vector of the Euclidean basis, and J_n is the nth complex Jacobi matrix

$$J_n = \begin{bmatrix} \alpha_0 & \beta_1 & & \\ \beta_1 & \alpha_1 & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ & & \beta_{n-1} & \alpha_{n-1} \end{bmatrix}, \quad n = 1, \dots, k; \tag{2.4}$$

more information about complex Jacobi matrices and their properties can be found, e.g., in [4] and in [71, in particular Section 4].

Given a smooth enough function $f(\lambda)$, the Gauss quadrature for quasi-definite linear functionals considered in [71, 72] has the form

$$\mathcal{G}_n(f) := \sum_{i=1}^{\ell} \sum_{j=0}^{s_i - 1} \omega_{i,j}\, f^{(j)}(\lambda_i), \quad n = s_1 + \cdots + s_\ell,$$

and satisfies the following properties.

- G1: the quadrature $\mathcal{G}_n(f)$ has maximal degree of exactness $2n-1$, i.e., it is exact for all polynomials of degree at most $2n-1$;
- G2: the quadrature $\mathcal{G}_n(f)$ is well-defined and it is unique. Moreover, Gauss quadratures with a smaller number of weights also exist and they are unique;
- G3: the quadrature $\mathcal{G}_n(f)$ can be written as the matrix form $m_0\, \mathbf{e}_1^T f(J_n) \mathbf{e}_1$, where J_n is the complex Jacobi matrix associated with $\mathcal{L}$.

A quadrature having properties G1, G2 and G3 exists if and only if the linear functional $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$; see [71, Section 7, in particular Corollaries 7.4 and 7.5] and [72, Theorem 3.1].

Property G3 corresponds to the so called *matching moment property* of the complex Jacobi matrix, i.e., if the complex numbers $m_0, \dots, m_{2n-1}$ define a quasi-definite linear functional (2.1) with associated Jacobi matrix J_n (here and in the following the simplified term *quasi-definite linear functional* and *positive definite linear functional* will stand for linear functionals that are quasi-definite and positive definite on the space of polynomials of sufficiently large degree), then

$$m_0\, \mathbf{e}_1^T\, (J_n)^j \mathbf{e}_1 = m_j, \quad j = 0, \dots, 2n-1; \tag{2.5}$$

see [71, Section 5]. In [29, Theorem 2] the matching moment property was proved for a quasi-definite linear functional given by

$$\mathcal{L}(f) = \mathbf{w}^* f(A) \mathbf{v},$$

where A is a complex matrix and $\mathbf{w}, \mathbf{v}$ are vectors (compare also with [19, Theorem 1]). In [78] it was derived by the Vorobyev method of moments (see in particular Chapter III of [82]).

3 Hankel linear systems

Let $\mathcal{L} : \mathcal{P} \to \mathbb{C}$ be a linear functional with moments $m_0, m_1, \dots$, let $H_0, H_1, \dots$ be the Hankel submatrices (2.2), and let $\mathbf{m}_{k-1}$ be the vector

$$\mathbf{m}_{k-1} = H(1:k; k+1) = [m_k, \dots, m_{2k-1}]^T.$$

Studying the linear system

$$H_{k-1}\,\mathbf{c} = -\mathbf{m}_{k-1} \tag{3.1}$$

will be fundamental in the following sections. In particular, in the case of real moments an analysis of the system (3.1) can be found in Section 1.2 of [22]; its straightforward generalization to the complex case is equivalent to the results given in this section.

Theorem 3.1 *Assume that $\Delta_{k-1} \neq 0$, then $\Delta_k = 0$ if and only if*

$$-\,m_{2k} = c_0 m_k + c_1 m_{k+1} + \dots + c_{k-1} m_{2k-1}, \tag{3.2}$$

where $\mathbf{c} = [c_0, \dots, c_{k-1}]^T$ is the unique solution of the linear system (3.1). *Moreover, if $\Delta_{k-1} \neq 0$ and $\Delta_k = \Delta_{k+1} = \dots = \Delta_{k+j-1} = 0$ for $j \geq 1$, then $\Delta_{k+j} = 0$ if and only if*

$$-\,m_{2k+j} = c_0 m_{k+j} + c_1 m_{k+j+1} + \dots + c_{k-1} m_{2k+j-1}. \tag{3.3}$$

Proof The last row of H_k is a linear combination of the other rows of H_k if and only if (3.2) is satisfied. This proves the first part of the theorem.

For $j \geq 1$ proceed by induction. Consider the submatrix $H(1 : k+1; 1 : k+j+1)$, i.e.,

$$\left[\begin{array}{cc|ccc|c} \multicolumn{2}{c|}{\multirow{2}{*}{H_{k-1}}} & m_k & \dots\ \dots & m_{k+j-1} & m_{k+j} \\ & & \vdots & & \vdots & \vdots \\ m_k & \dots & m_{2k} & \dots\ \dots & m_{2k+j-1} & m_{2k+j} \end{array}\right]. \tag{3.4}$$

By inductive assumption

$$-m_{2k+i} = c_0 m_{k+i} + c_1 m_{k+i+1} + \dots + c_{k-1} m_{2k+i-1},$$

for $i = 0, \dots, j-1$. If

$$-m_{2k+j} = c_0 m_{k+j} + c_1 m_{k+j+1} + \dots + c_{k-1} m_{2k+j-1},$$

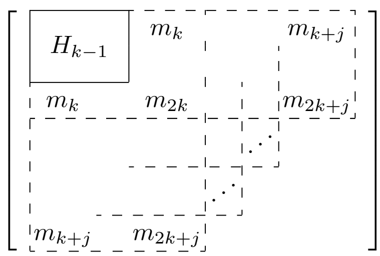
then $\Delta_{k+j} = 0$ since the first $k+1$ rows of H_{k+j} are linearly dependent. On the other hand, if

$$-m_{2k+j} \neq c_0 m_{k+j} + c_1 m_{k+j+1} + \ldots + c_{k-1} m_{2k+j-1},$$

then $\Delta_{k+j} \neq 0$. Indeed, consider the submatrices of H_{k+j}

$$H(1:k+1; 1:k+j+1), H(1:k+2; 1:k+j), \ldots, H(1:k+j+1; 1:k+1),$$

which we represent as follows (dashed lines) in order to help the reader



.

The mentioned submatrices without the last row have rank k. However, with the last row their rank is $k+1$. It means that, for $i = 1, \ldots, j+1$, the row $k+i$ of H_{k+j} is not a linear combination of the rows $1, 2, \ldots, k+i-1$, i.e., H_{k+j} has full rank. $\square$

As a consequence, the following theorem can be derived; see [22, Property 1.6].

Theorem 3.2 *Assume that* $\Delta_{k-1} \neq 0$ *and* $\Delta_k = \Delta_{k+1} = \ldots = \Delta_{k+j-1} = 0$. *Then the system*

$$H_{k+j-1}\,\mathbf{b} = -\mathbf{m}_{k+j-1} \tag{3.5}$$

has (infinitely many) solutions if and only if $\Delta_{k+j} = \Delta_{k+j+1} = \ldots \Delta_{k+2j-1} = 0$.

Proof By Theorem 3.1 the rank of the matrix $H(1:k+1; 1:k+j)$ is k. The linear system composed of the first $k+1$ equations of the system $H_{k+j-1}\mathbf{b} = -\mathbf{m}_{k+j-1}$, i.e., the system

$$H(1:k+1; 1:k+j)\mathbf{b} = -H(1:k+1; k+j+1),$$

has a solution if and only if the moment m_{2k+j} satisfies (3.3); otherwise the rank of the matrix $H(1:k+1; 1:k+j+1)$ is $k+1$. Repeating the argument for $m_{2k+j+1}, \ldots, m_{2k+2j-1}$ concludes the proof. $\square$

4 Polynomials and orthogonality

Let $\mathcal{L} : \mathcal{P} \to \mathbb{C}$ be a linear functional and $H_0, H_1, \dots$ be the sequence of the Hankel submatrices (2.2) composed of the moments of $\mathcal{L}$. Necessary and sufficient conditions for the existence (and uniqueness) of a FOP $p_n(\lambda)$ of degree n are given in the following theorem; see [22, Property 1.14].

Theorem 4.1 *Let $\mathcal{L} : \mathcal{P} \to \mathbb{C}$ be a linear functional. An n-degree monic FOP exists if and only if one of the following conditions is satisfied.*

- $\Delta_{n-1} \neq 0$ *(unique monic FOP);*
- $\Delta_{k-1} \neq 0$ *and* $\Delta_k = \Delta_{k+1} = \ldots = \Delta_{n-1} = \ldots = \Delta_{2n-k-1} = 0$ *(infinitely many monic FOPs);*

where $\Delta_0, \Delta_1, \dots$ are the determinants of the Hankel submatrices $H_0, H_1, \dots$ composed of the moments of $\mathcal{L}$.

Proof A monic FOP of degree n

$$\pi_n(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \ldots + c_1\lambda + c_0$$

exists if and only if $\mathcal{L}(\lambda^j \pi_n) = 0$, for $j = 0, \dots, n-1$, which gives the linear system (3.1) with $k = n$. Therefore if $\Delta_{n-1} \neq 0$, then the polynomial $\pi_n(\lambda)$ exists and is unique. If $\Delta_{n-1} = 0$, then necessary and sufficient conditions for the existence of $\pi_n(\lambda)$ are given by Theorem 3.2: for $\Delta_{k-1} \neq 0$ and $\Delta_k = \Delta_{k+1} = \ldots = \Delta_{n-1} = 0$, there exist infinitely many $\pi_n(\lambda)$ if and only if $\Delta_n = \Delta_{n+1} = \ldots \Delta_{2n-k-1} = 0$. □

Notice that if $\mathcal{L}$ is not quasi-definite, then $p_n(\lambda)$ may not exist or may not be unique. The second item of Theorem 4.1 can be interpreted in the following way: consider the sequence $\Delta_0, \Delta_1, \Delta_2, \dots$. Let $R = R(n-1)$ be the number of zeros in the sequence between Δ_{n-1} and the first nonzero element in the sequence after Δ_{n-1}, i.e., $\Delta_{n-1+j} = 0$ for $j = 1, \dots, R$ and $\Delta_{n+R} \neq 0$. Let $L = L(n-1)$ be the number of zeros in the sequence between Δ_{n-1} and the last nonzero element in the sequence before Δ_{n-1}, i.e., $\Delta_{n-1-j} = 0$ for $j = 1, \dots, L$ and $\Delta_{n-L-2} \neq 0$. A FOP of degree n exists if and only if $R(n-1) > L(n-1)$. Roughly said, there are "more consecutive zeros to the right than to the left".

Among the formal orthogonal polynomials the following cases can be distinguished; see Definition on p. 47 of [22].

Definition 4.2 *A formal orthogonal polynomial (FOP) $p_n(\lambda)$ is called* regular *when $\Delta_{n-1} \neq 0$ (i.e., when it is unique), while it is called* singular *when $\Delta_{n-1} = 0$ (i.e., when it is not unique).*

We remark that by Theorem 3.1 every regular FOP $p_n(\lambda)$ is orthogonal to $\mathcal{P}_{n+R(n-1)-1}$. Knowing all the integers k such that $\Delta_k = 0$ allows determining all the integers n for which a FOP $p_n(\lambda)$ exists.

Example 1. If the zero-nonzero pattern of the sequence of Hankel determinants Δ_k is

$$\begin{array}{lcccccccccccccc} \Delta_k = & * & * & 0 & * & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & * \\ k\ = & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \end{array},$$

then the FOPs of degree $3, 8, 9, 12$ and 13 do not exist. There exist regular FOPs of degree $1, 2, 4, 5, 10$ and 14 and singular FOPs of degree $6, 7$ and 11.

In order to fill the gaps in FOP sequences consider the following class of polynomials.

Definition 4.3 *The polynomial $p_n(\lambda)$ is called* quasi-orthogonal of order k *(or k-*quasi-orthogonal*), with $k < n$, when*

$$\mathcal{L}(p_n \lambda^j) = 0, \quad j = 0, \ldots, n-k-1.$$

Quasi-orthogonal polynomials of order 1 were introduced by Riesz in [74] and then generalized to any order by Chihara in [14]; see also [22, Definition 1.1, p. 51], [23], [25], and compare the definition with the concept of *inner formal orthogonal polynomials* given in [49, Definition 5.2] and of *left and right quasi-formally biorthogonal polynomials* in [27, Definition 3.3].

If $\Delta_{k-1} \neq 0$, then an $(n-k)$-quasi orthogonal polynomial of degree n exists for every n larger than k; see, e.g., [27, Lemma 3.4]. The following theorem will prove it together with the characterization of such polynomials; see discussion on pp. 47–51 of [22].

Theorem 4.4 *Let $\Delta_0, \Delta_1, \ldots$ be the Hankel determinants associated with the linear functional $\mathcal{L}$. Let $\Delta_{k-1} \neq 0$, and $\Delta_{k-1+i} = 0$ for $i = 1, \ldots, j$, and let $\pi_k(\lambda)$ be the regular monic FOP with respect to $\mathcal{L}$. Then all the monic i-quasi-orthogonal polynomials $\pi_{k+i}(\lambda)$ for $i = 1, \ldots, j$ are given by*

$$\pi_{k+i}(\lambda) = \pi_k(\lambda) \prod_{t=1}^{i} (\lambda - \eta_t), \quad \eta_t \in \mathbb{C}. \tag{4.1}$$

Proof The proof is by induction on i. Let $i = 1$. By Theorem 3.1, $\Delta_k = 0$ if and only if $\pi_k(\lambda)$ is orthogonal to all polynomials of degree k. Therefore $\lambda\pi_k(\lambda)$ is a monic polynomial of degree $k+1$ that is orthogonal to $\mathcal{P}_{k-1}$:

$$\mathcal{L}(\lambda\, \pi_k\, q) = \mathcal{L}(\pi_k\, (\lambda\, q)) = 0, \quad \text{for } q(\lambda) \in \mathcal{P}_{k-1}.$$

Moreover, any polynomial of the form $(\lambda - \alpha)\pi_k(\lambda)$, $\alpha \in \mathbb{C}$, is a monic 1-quasi-orthogonal polynomial. On the other side, assume that $p_{k+1}(\lambda)$ is an arbitrary monic polynomial of degree $k+1$ that is orthogonal to $\mathcal{P}_{k-1}$. Then the polynomial $\lambda\pi_k(\lambda) - p_{k+1}(\lambda)$ has the following two properties:

– it is of degree k,
– it is orthogonal to $\mathcal{P}_{k-1}$.

Hence the uniqueness of $\pi_k(\lambda)$ gives $\lambda\pi_k(\lambda) - p_{k+1}(\lambda) = \beta\pi_k(\lambda)$ for a certain complex number β, i.e.,

$$p_{k+1}(\lambda) = (\lambda - \beta)\, \pi_k(\lambda), \quad \text{for some } \beta \in \mathbb{C}.$$

Fix i between 2 and $j-1$, and assume that all the monic i-quasi-orthogonal polynomials of degree $k+i$ are of the form (4.1). By Theorem 3.1, $\Delta_k = \Delta_{k+1} = \ldots = \Delta_{k+i} = 0$ if and only if $\pi_k(\lambda)$ is orthogonal to all polynomials of degree

$k+i$. Therefore $\lambda\pi_{k+i}(\lambda)$ is a monic polynomial of degree $k+i+1$ that is orthogonal to $\mathcal{P}_{k-1}$:

$$\mathcal{L}(\lambda\,\pi_{k+i}\,q) = \mathcal{L}(\pi_k\,(\lambda\,(\lambda-\eta_1)\cdots(\lambda-\eta_i)\,q)) = 0, \qquad \text{for } q(\lambda)\in\mathcal{P}_{k-1}.$$

Clearly, $(\lambda-\alpha)\pi_{k+i}(\lambda)$ is a monic $(i+1)$-quasi-orthogonal polynomial of degree $k+i+1$, for any complex number α. It remains to prove that an arbitrary monic polynomial of degree $k+i+1$ that is orthogonal to $\mathcal{P}_{k-1}$ is of the form $(\lambda-\beta)\pi_{k+i}(\lambda)$, where β is a certain complex number, and $\pi_{k+i}(\lambda)$ is a polynomial of the form (4.1). It can be done similarly to the case $i=1$. □

The proof used the fact that $\pi_k(\lambda)$ is orthogonal to $\mathcal{P}_{k+i}$ as long as $\Delta_{k-1}\neq 0$ and $\Delta_k = \cdots = \Delta_{k+i} = 0$. This property will also be useful in the following part.

Proposition 4.5 *Let $\Delta_0, \Delta_1, \ldots$ be the Hankel determinants associated with the linear functional $\mathcal{L}$ such that $\Delta_{k-1}\neq 0$ and $\Delta_{k+i}=0$ for $i=0,\ldots,2j-1$. Then for $i=0,\ldots,j$, $p_{k+i}(\lambda)$ is a FOP if and only if it is i-quasi-orthogonal.*

Proof Clearly any FOP of degree $k+i$ is i-quasi-orthogonal. Vice versa if $p_{k+i}(\lambda)$ is i-quasi-orthogonal, then it satisfies (4.1). Since $p_k(\lambda)$ is orthogonal to $\mathcal{P}_{k+2j-1}$, if $q(\lambda)\in\mathcal{P}_{k+i-1}$, then $\mathcal{L}(p_{k+i}q) = \mathcal{L}(p_k(\lambda-\eta_1)\cdots(\lambda-\eta_i)q) = 0$ for $i=0,\ldots,j$. □

Consider the sequence of polynomials

$$p_0(\lambda), p_1(\lambda), p_2(\lambda), \ldots \tag{4.2}$$

constructed in the following way: $p_n(\lambda)$ is a regular FOP (when possible) or $p_n(\lambda)$ is a $(n-k)$-quasi-orthogonal polynomial, where $p_k(\lambda)$ is the last regular FOP before $p_n(\lambda)$. For later convenience, we consider every nonzero choice for $p_0(\lambda)$ as a regular FOP. Let us denote by $\nu(0),\nu(1),\nu(2),\ldots$ all the indexes for which $p_{\nu(j)}(\lambda)$ is a regular FOP, i.e., $\Delta_{\nu(j)-1}\neq 0$ (setting $p_{\nu(0)}(\lambda) = p_0(\lambda)\neq 0$, and $\nu(j+1)=\infty$ when $p_{\nu(j)}(\lambda)$ is the last of the regular FOPs). By Theorem 4.4, the quasi-orthogonal polynomials between two consecutive regular FOPs $p_{\nu(j)}(\lambda), p_{\nu(j+1)}(\lambda)$ satisfy the recurrences

$$\beta_n p_n(\lambda) = \lambda p_{n-1}(\lambda) - \sum_{i=\nu(j)}^{n-1}\alpha_{n,i}p_i(\lambda), \quad n=\nu(j)+1,\ldots,\nu(j+1)-1, \tag{4.3}$$

for some coefficients $\alpha_{n,i}\in\mathbb{C}$ and $\beta_n\neq 0$; see [22, Theorem 1.5 and Remark 1.2]. Notice that any choice for $\alpha_{n,i}$ and $\beta_n\neq 0$ defines a $(n-\nu(j))$-quasi-orthogonal polynomials. In particular, there exist families of such polynomials satisfying the two-term recurrences

$$\beta_n p_n(\lambda) = (\lambda-\alpha_{n,n-1})p_{n-1}(\lambda), \quad n=\nu(j)+1,\ldots,\nu(j+1)-1; \tag{4.4}$$

fixing $\alpha_{n,n-1}=0$ gives even simpler recurrences.

Fixing $n = \nu(j+1)$ for some $j \geq 0$, the regular FOP $p_n(\lambda)$ satisfies (see [22, Theorem 1.5 and Remark 1.2] and [40, Theorem 2])

$$\beta_n p_n(\lambda) = \lambda p_{n-1}(\lambda) - \sum_{i=\nu(j)}^{n-1} \alpha_{n,i} p_i(\lambda) - \gamma_n p_{\nu(j-1)}(\lambda), \tag{4.5}$$

with $p_{\nu(-1)}(\lambda) = p_{-1}(\lambda) = 0$, β_n a nonzero coefficient, $\gamma_{\nu(1)} = 0$,

$$\gamma_n = \frac{\mathcal{L}(\lambda p_{n-1} p_{\nu(j)-1})}{\mathcal{L}(p_{\nu(j)-1} p_{\nu(j-1)})} \neq 0, \quad j \geq 1,$$

and $\alpha_{n,i}$ given by

$$\begin{bmatrix} \mathcal{L}(p_{\nu(j)} p_{\nu(j)}) & \dots & \mathcal{L}(p_{\nu(j)} p_{n-1}) \\ \vdots & & \vdots \\ \mathcal{L}(p_{n-1} p_{\nu(j)}) & \dots & \mathcal{L}(p_{n-1} p_{n-1}) \end{bmatrix} \begin{bmatrix} \alpha_{n,\nu(j)} \\ \vdots \\ \alpha_{n,n-1} \end{bmatrix} = \begin{bmatrix} \mathcal{L}(\lambda p_{\nu(j)} p_{n-1}) \\ \vdots \\ \mathcal{L}(\lambda p_{n-1} p_{n-1}) \end{bmatrix}, \tag{4.6}$$

where the matrix of the system is nonsingular; see, e.g., [30, Theorem 2.3]. Notice that for a quasi-definite linear functional the related (regular) formal orthonormal polynomials satisfy the three term recurrences (2.3).

Given $n = 1, 2, \dots$, the recurrences (4.3) and (4.5) can be expressed in the matrix form (see [22, Section 1.7], [70, Section 3]; c.f., [39, pp. 221–222], [40, Figure 2 and Theorem 3], and [28, Equalities (3.4) and (3.5)])

$$\lambda\, \mathbf{p}(\lambda) = T_n\, \mathbf{p}(\lambda) + \beta_n p_n(\lambda) \mathbf{e}_n, \tag{4.7}$$

with $\mathbf{p}(\lambda) = [p_0(\lambda), p_1(\lambda), \dots, p_{n-1}(\lambda)]^T$ and where T_n is the block matrix

$$T_n = \begin{bmatrix} A_0 & \beta_{\nu(1)} & & & \\ \gamma_{\nu(2)} & A_1 & \ddots & & \\ & \ddots & \ddots & \beta_{\nu(\ell)} & \\ & & \gamma_n & A_\ell^n \end{bmatrix}, \tag{4.8}$$

with the coefficients $\beta_{\nu(j)}$ on the first upper diagonal, the coefficients $\gamma_{\nu(j)}$ in the position $(\nu(j), \nu(j-2)+1)$, $\gamma_n = 0$ when $p_n(\lambda)$ is not regular, and

$$A_j^n = \begin{bmatrix} \alpha_{\nu(j)+1,\nu(j)} & \beta_{\nu(j)+1} & 0 & \dots & 0 \\ \alpha_{\nu(j)+2,\nu(j)} & \alpha_{\nu(j)+2,\nu(j)+1} & \beta_{\nu(j)+2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \alpha_{n-1,\nu(j)} & \alpha_{n-1,\nu(j)+1} & \dots & \ddots & \beta_{n-1} \\ \alpha_{n,\nu(j)} & \alpha_{n,\nu(j)+1} & \dots & \dots & \alpha_{n,n-1} \end{bmatrix},$$

for $n = \nu(j)+1, \dots, \nu(j+1)$; $A_j = A_j^{\nu(j+1)}$ for simplicity. Notice that using the recurrences (4.4) with $\alpha_{n,n-1} = 0$ gives the sparse matrix

$$A_j = \begin{bmatrix} 0 & \beta_{\nu(j)+1} & 0 & \dots & 0 \\ 0 & 0 & \beta_{\nu(j)+2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & \beta_{\nu(j+1)-1} \\ \alpha_{\nu(j+1),\nu(j)} & \alpha_{\nu(j+1),\nu(j)+1} & \cdots & \cdots & \alpha_{\nu(j+1),\nu(j+1)-1} \end{bmatrix},$$

with $\alpha_{\nu(j+1),\nu(j)}, \dots, \alpha_{\nu(j+1),\nu(j+1)-1}$ obtained by (4.6).

When the polynomials $p_0(\lambda), \dots, p_n(\lambda)$ are regular FOPs (the linear functional is quasi-definite on $\mathcal{P}_{n-1}$) the blocks A_j are scalars. Therefore T_n is an irreducible tridiagonal matrix since β_j and γ_{j+1} are nonzero for $j = 1, \dots, n-1$. In particular, there exists a sequence of formal orthonormal polynomials so that the matrix T_n is the complex Jacobi matrix (2.4).

5 The Gauss quadrature for linear functionals

Given a linear functional $\mathcal{L}$ and a smooth enough function $f(\lambda)$, consider a quadrature approximating $\mathcal{L}(f)$ of the form (see [22, Chapter 5], [64, Section 2], and [71, Section 7])

$$\mathcal{G}_n(f) := \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \omega_{i,j}\, f^{(j)}(\lambda_i), \quad n = s_1 + \cdots + s_\ell, \tag{5.1}$$

with $\omega_{i,j}$ the weights, λ_i the *distinct* nodes, and s_i the multiplicity of the node λ_i. Notice that the number of nodes ℓ can be less than n. The quadrature (5.1) will be referred as *n-node quadrature* when $\omega_{i,s_i-1} \neq 0$ for $i = 1, \dots, \ell$. Otherwise, the sum of the multiplicities would be smaller than n. For any choice of (distinct) nodes $\lambda_1, \dots, \lambda_\ell$ and their multiplicities s_i, such that $s_1 + \cdots + s_\ell = n$, it is possible to achieve that the quadrature (5.1) is exact for any $f(\lambda) \in \mathcal{P}_{n-1}$. It is necessary and sufficient to set the weights as

$$\omega_{i,j} = \mathcal{L}(h_{i,j}), \tag{5.2}$$

where $h_{i,j}(\lambda)$ are polynomials from $\mathcal{P}_{n-1}$ such that

$$\begin{aligned} h_{i,j}^{(t)}(\lambda_k) &= 1 \quad \text{for } \lambda_k = \lambda_i \text{ and } t = j, \\ h_{i,j}^{(t)}(\lambda_k) &= 0 \quad \text{for } \lambda_k \neq \lambda_i \text{ or } t \neq j, \end{aligned} \tag{5.3}$$

with $k = 1, 2, \dots, \ell$, and $t = 0, 1, \dots, s_i - 1$; see [22, Theorem 5.1] or the proof of Theorem 7.1 in [71]. In this case (5.1) is known as *interpolatory* quadrature, since it can be given by applying $\mathcal{L}$ to the generalized (Hermite) interpolating polynomial for the function $f(\lambda)$ at the nodes λ_i of the multiplicities s_i. An interpolatory quadrature is completely determined by its nodes and multiplicities. Therefore in the following a quadrature $\mathcal{G}_n$ will be said to be determined by a polynomial $p_n(\lambda)$ when it is an interpolatory quadrature (5.1) with λ_i being the roots of p_n, and s_i the corresponding multiplicities of the roots.

The following definition is a straightforward extension to the complex case of the Gauss quadrature introduced by Draux in [22, Chapter 5].

Definition 5.1 *The quadrature (5.1) is called the* n-node Gauss quadrature *when it is exact on the space* $\mathcal{P}_{2n-1}$ *and* $\omega_{i,s_i-1} \neq 0$ *for* $i = 1, \dots, \ell$ *(the number of nodes counting the multiplicities is* n*).*

We point out the following remarks:

- the algebraic degree of exactness of the n-node Gauss quadrature is allowed to be larger than $2n-1$;
- a Gauss quadratures with smaller number of nodes may not exist even when the n-node Gauss quadrature exists.

Hence the n-node Gauss quadrature generally does not satisfy properties G1–G3 in Section 2. However, when $\mathcal{L}$ is a quasi-definite linear functional, then the n-node Gauss quadrature for $\mathcal{L}$ satisfies properties G1–G3, i.e., in this case Definition 5.1 is equivalent to the one in [71, 72].

In order to give conditions for the existence of an n-node Gauss quadrature for a linear functional the following result is needed; see [22, Theorem 5.2], see also [34, Theorem 1.45] for positive definite linear functionals and [71, Theorem 7.1] for quasi-definite linear functionals.

Theorem 5.2 *A quadrature* $\mathcal{G}_n$ *determined by a polynomial* $p_n(\lambda)$ *is exact for all the polynomials in* $\mathcal{P}_{n+k-1}$ *if and only if* $p_n(\lambda)$ *is* $(n-k)$*-quasi-orthogonal.*

Proof Assume $\mathcal{G}_n$ to be exact for every polynomial in $\mathcal{P}_{n+k-1}$. Then $p_n(\lambda)$ is $(n-k)$-quasi-orthogonal. Indeed,

$$\mathcal{L}(p_n q) = \mathcal{G}_n(p_n q) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \omega_{i,j} \, (p_n q)^{(j)}(\lambda_i) = 0, \quad q(\lambda) \in \mathcal{P}_{k-1},$$

since $p_n^{(j)}(\lambda_i) = 0$ for $j = 0, \dots, s_i - 1$, $i = 1, \dots, \ell$. Inversely, let $p_n(\lambda)$ be $(n-k)$-quasi-orthogonal. Any $f(\lambda) \in \mathcal{P}_{n+k-1}$ can be written as $f(\lambda) = p_n(\lambda)q(\lambda) + r(\lambda)$ for some $q(\lambda) \in \mathcal{P}_{k-1}$ and $r(\lambda) \in \mathcal{P}_{n-1}$, giving $\mathcal{L}(f) = \mathcal{L}(r)$. Since $\mathcal{G}_n$ is interpolatory it is exact on $\mathcal{P}_{n-1}$ and thus

$$\mathcal{L}(f) = \mathcal{L}(r) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \omega_{i,j} \, r^{(j)}(\lambda_i).$$

The proof is concluded since $f^{(j)}(\lambda_i) = r^{(j)}(\lambda_i)$ for $j = 0, \dots, s_i - 1$, $i = 1, \dots, \ell$. $\square$

As discussed in Section 4, for every linear functional $\mathcal{L}$ there exists a sequence of polynomials $p_0(\lambda), p_1(\lambda), \dots$ (4.2) so that $p_n(\lambda)$ is a regular FOP (when possible), or $p_n(\lambda)$ is $(n-k)$-quasi-orthogonal, where $p_k(\lambda)$ is the last regular FOP before $p_n(\lambda)$ ($p_0(\lambda) \neq 0$ is assumed to be regular). We denote by $\nu(0) = 0, \nu(1), \dots$ the indexes of the regular FOPs (with $\nu(t+1) = +\infty$ when $\nu(t)$ is the last of the regular FOPs). Theorem 5.2 implies the following corollary (see [22, Theorem 5.2]).

Corollary 5.3 *Let $p_n(\lambda)$ a polynomial in the sequence described above.*

- *If $p_n(\lambda)$ is a regular FOP, then it determines a quadrature $\mathcal{G}_n$ exact for every polynomials in $\mathcal{P}_{2n-1}$.*
- *if $p_n(\lambda)$ is a $(n-k)$-quasi-orthogonal polynomial, then it determines a quadrature $\mathcal{G}_n$ exact for every polynomials in $\mathcal{P}_{n+k-1}$.*

Notice that if $\nu(1) > 1$, then $m_j = 0$ for $j = 0, \dots, \nu(1) - 2$ (see, e.g., [22, Property 1.15]). Hence $\mathcal{G}_n(f) \equiv 0$ for $n = 1, \dots, \nu(1) - 1$.

If $\omega_{i,s_i-1} = 0$ for some i, then the quadrature (5.1) has a smaller number of nodes (counting the multiplicities). The following lemmas deal with this issue; see [22, Theorem 5.3].

Lemma 5.4 *Consider the quadratures $\mathcal{G}_n$ determined by the polynomial $p_n(\lambda)$ in the sequence described above. Given two consecutive regular FOPs $p_{\nu(t)}(\lambda)$ and $p_{\nu(t+1)}(\lambda)$, with $t \geq 1$, then*

$$\mathcal{G}_n = \mathcal{G}_{\nu(t)}, \qquad \textit{for } n = \nu(t) + 1, \dots, \nu(t+1) - 1.$$

Proof Theorem 4.4 gives

$$p_n(\lambda) = p_{\nu(t)}(\lambda) q_{n-\nu(t)}(\lambda),$$

for some polynomial $q_{n-\nu(\lambda)}(\lambda)$. Let $\lambda_1, \dots, \lambda_\ell$ the roots of $p_n(\lambda)$ with multiplicities $s_1, \dots, s_\ell$. The weights of the quadrature $\mathcal{G}_n$ are given by (5.2). Consider the pair i, j so that $(\lambda - \lambda_i)^j$ is not a factor of $p_{\nu(t)}(\lambda)$, i.e., the root λ_i is not a root of $p_{\nu(t)}(\lambda)$ or it is a root of $p_{\nu(t)}(\lambda)$ but with j greater than the multiplicity of λ_i as a root of $p_{\nu(t)}(\lambda)$. Then the $(n-1)$-degree interpolatory polynomial $h_{i,j}(\lambda)$ defined in (5.3) is a multiple of $p_{\nu(t)}(\lambda)$, i.e.,

$$h_{i,j}(\lambda) = p_{\nu(t)}(\lambda) r_{n-\nu(t)-1}(\lambda),$$

for some polynomial $r_{n-\nu(t)-1}(\lambda)$. By Theorem 3.1 $p_{\nu(t)}(\lambda)$ is orthogonal to $\mathcal{P}_{\nu(t+1)-2}$, giving

$$\omega_{i,j} = \mathcal{L}(h_{i,j}) = \mathcal{L}(p_{\nu(t)} r_{n-\nu(t)-1}) = 0.$$

Therefore $\mathcal{G}_n$ has at most $\nu(t)$ nodes. Moreover, each node of $\mathcal{G}_n$ is a node of $\mathcal{G}_{\nu(t)}$ and has multiplicity smaller than or equal to the one of the corresponding node of $\mathcal{G}_{\nu(t)}$.

If λ_i is a root of $p_{\nu(t)}(\lambda)$ with multiplicity j, then there exists a polynomial $\widetilde{h}_{i,j}(\lambda)$ of the kind of (5.3) so that $\widetilde{\omega}_{i,j} = \mathcal{L}(\widetilde{h}_{i,j})$ is the corresponding weight of $\mathcal{G}_{\nu(t)}$. Since $\widetilde{h}_{i,j}(\lambda)$ has degree $\nu(t)-1$ the weight $\widetilde{\omega}_{i,j}$ is given by

$$\widetilde{\omega}_{i,j} = \mathcal{G}_n(\widetilde{h}_{i,j}).$$

Noticing that $\mathcal{G}_n(\widetilde{h}_{i,j}) = \omega_{i,j}$ concludes the proof. □

Lemma 5.5 *If $p_n(\lambda)$ is a regular FOP, with $n \geq 1$, then it determines a quadrature* (5.1) *such that $\omega_{i,s_i-1} \neq 0$, for $i = 1, \dots, \ell$.*

Proof Let t such that $p_{\nu(t)}(\lambda) = p_n(\lambda)$ and $h_{i,j}(\lambda)$ as in (5.3), then

$$\begin{aligned}
\mathcal{L}(h_{i,s_i-1}p_{\nu(t-1)}) &= \sum_{r=1}^{\ell}\sum_{s=0}^{s_i-1}\omega_{r,s}(h_{i,s_i-1}p_{\nu(t-1)})^{(s)}(\lambda_r)\\
&= \sum_{r=1}^{\ell}\sum_{s=0}^{s_i-1}\omega_{r,s}\sum_{u=0}^{s}\binom{s}{u}h_{i,s_i-1}^{(u)}(\lambda_r)p_{\nu(t-1)}^{(s-u)}(\lambda_r)\\
&= \sum_{s=0}^{s_i-1}\omega_{i,s}\sum_{u=0}^{s}\binom{s}{u}h_{i,s_i-1}^{(u)}(\lambda_i)p_{\nu(t-1)}^{(s-u)}(\lambda_i)\\
&= \omega_{i,s_i-1}p_{\nu(t-1)}(\lambda_i).
\end{aligned}$$

Theorem 3.1 gives $\mathcal{L}(h_{i,s_i-1}p_{\nu(t-1)}) \neq 0$, concluding the proof. □

The following theorem summarizes the previous discussion; see [22, Theorems 5.2 and 5.3].

Theorem 5.6 *The n-node Gauss quadrature $\mathcal{G}_n$ exists and is unique if and only if $\Delta_{n-1} \neq 0$. Moreover, if $\Delta_n = \Delta_{n+1} = \cdots = \Delta_{n+j} = 0$, then $\mathcal{G}_n$ has degree of exactness at least $2n+j$. In particular, if $n = \nu(t)$, then $\mathcal{G}_n$ has (maximal) degree of exactness $\nu(t)+\nu(t+1)-2$, with $\nu(t+1) = +\infty$ when n is the last of the regular FOPs.*

Proof By Theorem 5.2, $\mathcal{G}_n$ is exact on $\mathcal{P}_{2n-1}$ if and only if it is determined by a FOP with degree n, i.e., a polynomial $p_n(\lambda)$ orthogonal to $\mathcal{P}_{n-1}$. By Lemma 5.4 if $p_n(\lambda)$ is a singular FOP, then $\mathcal{G}_n$ has not n nodes. Therefore it is not a n-node Gauss quadrature. Considering Lemma 5.5 and noticing that regular FOPs are unique, $\mathcal{G}_n$ exists and is unique if and only if $\Delta_{n-1} \neq 0$. The proof is conclude noticing that Theorem 5.2 and Lemma 5.4 imply that $\mathcal{G}_n$ is exact on $\mathcal{P}_{2n+j}$. □

6 Matrix formulation of the Gauss quadrature

If $\mathcal{L}$ is a quasi-definite linear functional, then the associated complex Jacobi matrix (2.4) satisfies the matching moment property (2.5). The section will prove an extension of the matching moment property for a general sequence of moments using the properties of the formal orthogonal polynomials and of the Gauss quadrature for the linear functionals. The case of a linear functional of the kind $\mathcal{L}(f) = \mathbf{w}^* f(A)\mathbf{v}$ was treated in [42, Theorem 2.10]. We remark that assuming real moments (with a straightforward extension to the complex case), the matching moment property here presented, as well as the ones in [29, 42, 71], can be derived by Theorem 5 of the 1983 paper by Gragg and Lindquist [40], where such property is related to the minimal partial realization problem.

Let $\mathcal{L} : \mathcal{P} \to \mathbb{C}$ be a linear functional and let T_n be the corresponding block tridiagonal matrix (4.8) associated with the sequence of polynomials $p_0(\lambda), \ldots, p_n(\lambda)$. Denote by $p_{\nu(t)}(\lambda)$ the subsequence of the regular FOPs and recall that for $\nu(t) < n < \nu(t+1)$ the polynomials $p_n(\lambda)$ are $(n - \nu(t))$-quasi-orthogonal. Also recall that if $\nu(1) \geq 2$, then $m_j = 0$ for $j = 0, \ldots, \nu(1) - 2$. Since the elements in the superdiagonal of T_n are nonzero the block tridiagonal matrix T_n is nonderogatory, i.e., its eigenvalues have geometric multiplicity 1. Indeed, if λ is an eigenvalue, then deleting the first column and the last row of $T_n - \lambda I$ gives a lower triangular nonsingular matrix (with $I = [\mathbf{e}_1, \ldots, \mathbf{e}_n]$ the identity matrix). Thus the null space of $T_n - \lambda I$ has dimension 1. Proving the matching moment property will need the following lemmas.

Lemma 6.1 *Let T_n and $p_n(\lambda)$ as in* (4.7). *Then $p_n(\lambda)$ is the characteristic polynomial of T_n (up to a nonzero rescaling).*

Proof Rewrite the equality (4.7) as

$$(T_n - \lambda I)\mathbf{p}(\lambda) = -\beta_n p_n(\lambda)\mathbf{e}_n, \tag{6.1}$$

with $\mathbf{p}(\lambda) = [p_0(\lambda), p_1(\lambda), \ldots, p_{n-1}(\lambda)]^T$. Since $p_0(\lambda) \neq 0$ if λ_* is a root of $p_n(\lambda)$, then it is an eigenvalue of T_n with the corresponding eigenvector $\mathbf{p}(\lambda_*)$. Deriving (6.1) gives

$$\mathbf{p}(\lambda) = (T_n - \lambda I)\mathbf{p}'(\lambda) + \beta_n p_n'(\lambda)\mathbf{e}_n,$$

Hence deriving for j times gives

$$j\mathbf{p}^{(j-1)}(\lambda) = (T_n - \lambda I)\mathbf{p}^{(j)}(\lambda) + \beta_n p_n^{(j)}(\lambda)\mathbf{e}_n. \tag{6.2}$$

Since $p_j(\lambda)$ has degree j, then $\mathbf{p}^{(j)}(\lambda) \neq 0$ for $\lambda \in \mathbb{C}$ and for $j = 0, \ldots, n-1$. Therefore if λ_* is a root of $p_n(\lambda)$ with multiplicity s, then (6.2) shows that λ_* is an eigenvalue with algebraic multiplicity s. Since $p_n(\lambda)$ has degree n and T_n has dimension n the proof is concluded. □

We point out that the Lemma 6.1 is a consequence of Lemma 2 in [54]; see also [22, Theorem 1.11].

Lemma 6.2 *Let $T_1, T_2, \ldots$ be a sequence of block tridiagonal matrices* (4.8). *For $n \geq \nu(1) + 1$ the matrices T_{n-1} and T_n satisfy*

$$\mathbf{e}_1^T \,(T_{n-1})^k \,\mathbf{e}_{\nu(1)} = \mathbf{e}_1^T \,(T_n)^k \,\mathbf{e}_{\nu(1)}, \qquad \textit{for } k = 0, \ldots, n-1,$$

where the vectors $\mathbf{e}_1, \mathbf{e}_{\nu(1)}$ have dimension $n-1$ on the left-hand side and n on the right-hand side (we use the same notation for the sake of simplicity).

Proof Consider the n-dimensional vectors

$$\mathbf{u}_k = (T_n)^k \, \mathbf{e}_{\nu(1)}, \quad k = 0, 1, \ldots .$$

If the last element of $\mathbf{u}_k$ is zero for $k = 0, \ldots, n-1$, then

$$\mathbf{e}_1^T \mathbf{u}_k = \mathbf{e}_1^T \, (T_{n-1})^k \, \mathbf{e}_{\nu(1)}, \quad k = 0, \ldots, n-1,$$

proving the lemma. In the following, when the elements from the position i to the position j of a vector are possibly nonzero, we denote them by $*_{i:j}$ ($*_i = *_{i:i}$). Similarly, when the elements from the position i to the position j are null, we denote them by $0_{i:j}$. Direct computations show that

$$\mathbf{u}_1 = \begin{bmatrix} 0_{1:\nu(1)-2} \\ *_{\nu(1)-1:\nu(1)} \\ 0_{\nu(1)+1:n} \end{bmatrix}, \, \mathbf{u}_2 = \begin{bmatrix} 0_{1:\nu(1)-3} \\ *_{\nu(1)-2:\nu(1)} \\ 0_{\nu(1)+1:n} \end{bmatrix}, \ldots, \, \mathbf{u}_{\nu(1)-1} = \begin{bmatrix} *_{1:\nu(1)} \\ 0_{\nu(1)+1:n} \end{bmatrix}.$$

Moreover,

$$\mathbf{u}_{\nu(1)} = \begin{bmatrix} *_{1:\nu(1)} \\ 0_{\nu(1)+1:\nu(2)-1} \\ *_{\nu(2)} \\ 0_{\nu(2)+1:n} \end{bmatrix},$$

and

$$\mathbf{u}_{\nu(1)+1} = \begin{bmatrix} *_{1:\nu(1)} \\ 0_{\nu(1)+1:\nu(2)-2} \\ *_{\nu(2)-1:\nu(2)} \\ 0_{\nu(2)+1:n} \end{bmatrix}, \ldots, \mathbf{u}_{\nu(1)+\nu(2)-1} = \begin{bmatrix} *_{1:\nu(2)} \\ 0_{\nu(2)+1:n} \end{bmatrix}.$$

Repeating the argument gives

$$\mathbf{u}_{n-1} = \begin{bmatrix} *_{1:n-1} \\ 0 \end{bmatrix},$$

concluding the proof. □

Theorem 6.3 (Matching moment property) *Let $\mathcal{L}$ be a linear functional with complex moments $m_0, m_1, \ldots,$ and let T_n be the associated block tridiagonal matrix (4.8) with the corresponding polynomials $p_0(\lambda), \ldots, p_n(\lambda)$. Denote the indexes of the regular FOPs by $\nu(0) = 0, \nu(1), \nu(2), \ldots$. For every $n \geq \nu(1)$ let t be so that $\nu(t) \leq n < \nu(t+1)$, the matrix T_n satisfies*

$$\mu \, m_{\nu(1)-1} \, \mathbf{e}_1^T (T_n)^k \, \mathbf{e}_{\nu(1)} = m_k, \;\; k = 0, \ldots, \nu(t) + \nu(t+1) - 2,$$

with $\mu = (\beta_1 \cdots \beta_{\nu(1)-1})^{-1}$ for $\nu(1) > 1$, $\mu = 1$ for $\nu(1) = 1$, and $\nu(t+1) = +\infty$ when $p_{\nu(t)}$ is the last regular FOP.

Proof Consider the linear functional

$$\mathcal{L}^{(n)}(f) = \mu\, m_{\nu(1)-1}\, \mathbf{e}_1^T f(T_n)\, \mathbf{e}_{\nu(1)}, \quad f(\lambda) \in \mathcal{P}.$$

If the linear functionals $\mathcal{L}$ and $\mathcal{L}^{(n)}$ are identical on the space $\mathcal{P}_{\nu(t)+\nu(t+1)-2}$, then the proof is given. By Lemma 6.1 and the Cayley–Hamilton Theorem, the polynomial $p_n(\lambda)$ satisfies the orthogonality conditions

$$\mathcal{L}^{(n)}(\lambda^k p_n) = \mu\, m_{\nu(1)-1}\, \mathbf{e}_1^T (T_n)^k p_n(T_n)\, \mathbf{e}_{\nu(1)} = 0, \quad k = 0, 1, \ldots . \tag{6.3}$$

Proceeding by induction on n, first consider the case $n = \nu(1) > 1$. Since $T_{\nu(1)}$ is a Hessenberg matrix it satisfies

$$\mathbf{e}_1^T (T_{\nu(1)})^k\, \mathbf{e}_{\nu(1)} = 0 = m_k, \quad k = 0, \ldots, \nu(1) - 2.$$

Direct computations gives $\mathbf{e}_1^T (T_{\nu(1)})^{\nu(1)-1}\, \mathbf{e}_{\nu(1)} = \beta_1 \cdots \beta_{\nu(1)-1} \neq 0$. Therefore

$$\mu\, m_{\nu(1)-1}\, \mathbf{e}_1^T (T_{\nu(1)})^k\, \mathbf{e}_{\nu(1)} = m_k, \quad k = 0, \ldots, \nu(1) - 1, \tag{6.4}$$

which also trivially stands for $n = \nu(1) = 1$. Using property (6.3) and Theorem 5.6, $p_{\nu(1)}(\lambda)$ determines the quadrature $\mathcal{G}_{\nu(1)}^{(\nu(1))}$ for $\mathcal{L}^{(\nu(1))}$ so that $\mathcal{G}_{\nu(1)}^{(\nu(1))}(f) = \mathcal{L}^{(\nu(1))}(f)$ for every $f(\lambda) \in \mathcal{P}$. Moreover, $p_{\nu(1)}(\lambda)$ determines the Gauss quadrature $\mathcal{G}_{\nu(1)}$ for $\mathcal{L}$, exact for polynomials of degree at most $\nu(1) + \nu(2) - 2$. The two quadratures $\mathcal{G}_{\nu(1)}^{(\nu(1))}$ and $\mathcal{G}_{\nu(1)}$ coincide since they have the same weights. Indeed, if $h_{i,j}(\lambda)$ is the interpolatory polynomial (5.3) for $n = \nu(1)$, then the weights of $\mathcal{G}_{\nu(1)}^{(\nu(1))}$ and $\mathcal{G}_{\nu(1)}$ are respectively given by

$$\omega_{i,j}^{(\nu(1))} = \mathcal{L}^{(\nu(1))}(h_{i,j}) \quad \text{and} \quad \omega_{i,j} = \mathcal{L}(h_{i,j}).$$

Since $h_{i,j}(\lambda)$ has degree $\nu(1) - 1$, equality (6.4) gives

$$\omega_{i,j}^{(\nu(1))} = \mathcal{L}^{(\nu(1))}(h_{i,j}) = \mathcal{L}(h_{i,j}) = \omega_{i,j},$$

proving the theorem for $n = \nu(1)$.

Assume $n > \nu(1)$, with t so that $\nu(t) \leq n < \nu(t+1)$, and define the quadrature $\mathcal{G}_n^{(n)}$ for $\mathcal{L}^{(n)}$, determined by the polynomial $p_n(\lambda)$. By (6.3) and Theorem 5.6, $\mathcal{G}_n^{(n)}(f) = \mathcal{L}^{(n)}(f)$ for every $f(\lambda) \in \mathcal{P}$. Furthermore, $p_n(\lambda)$ determines the quadrature $\mathcal{G}_n = \mathcal{G}_{\nu(t)}$ for $\mathcal{L}$, exact for every polynomials of degree at most $\nu(t)+\nu(t+1)-2$. As noticed above, $\mathcal{G}_n^{(n)}$ ad $\mathcal{G}_n$ coincide if and only if the respective weights $\omega_{i,j}^{(n)}$ and $\omega_{i,j}$ coincide. Let $h_{i,j}(\lambda)$ be the interpolatory polynomials (5.3). Since $h_{i,j}(\lambda)$ has degree $n-1$ the weight $\omega_{i,j}^{(n)}$ satisfies

$$\omega_{i,j}^{(n)} = \mathcal{L}^{(n)}(h_{i,j}) = \mu\, m_{\nu(1)-1}\, \mathbf{e}_1^T\, h_{i,j}(T_{n-1})\, \mathbf{e}_{\nu(1)} = \mathcal{L}(h_{i,j}) = \omega_{i,j},$$

where Lemma 6.2 and the inductive assumption were used. □

We recall the definition of *matrix function*. A function $f(\lambda)$ is defined on the spectrum of the given matrix A when for every eigenvalue λ_i of A there exist $f^{(j)}(\lambda_i)$ for $j = 0, 1, \dots, s_i - 1$, with s_i the order of the largest Jordan block of A in which λ_i appears. Consider the Jordan block Λ of the size s corresponding to the eigenvalue λ, then the matrix function $f(\Lambda)$ is defined as

$$f(\Lambda) = \begin{bmatrix} f(\lambda) & \frac{f'(\lambda)}{1!} & \frac{f^{(2)}(\lambda)}{2!} & \dots & \frac{f^{(s-1)}(\lambda)}{(s-1)!} \\ 0 & f(\lambda) & \frac{f'(\lambda)}{1!} & \dots & \frac{f^{(s-2)}(\lambda)}{(s-2)!} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \frac{f'(\lambda)}{1!} \\ 0 & \dots & \dots & 0 & f(\lambda) \end{bmatrix}.$$

Denoting

$$A = W \mathrm{diag}(\Lambda_1, \dots, \Lambda_\nu) W^{-1},$$

the Jordan decomposition of A, the matrix function $f(A)$ is defined as

$$f(A) = W \mathrm{diag}(f(\Lambda_1), \dots, f(\Lambda_\nu)) W^{-1}. \tag{6.5}$$

We refer to [47] for further information and for the equivalence to the other definitions of matrix function.

Consider the block tridiagonal matrix T_n of Theorem 6.3 and its Jordan decomposition $T_n = W \mathrm{diag}(\Lambda_1, \dots, \Lambda_\ell) W^{-1}$. Since T_n is nonderogatory, there are $\lambda_1, \dots, \lambda_\ell$ distinct eigenvalues corresponding to the Jordan blocks $\Lambda_1, \dots, \Lambda_\ell$ of the sizes respectively $s_1, \dots, s_\ell$. If $f(\lambda)$ is a smooth enough function so that $f(T_n)$ is well defined, then the Jordan decomposition of T_n and some algebraic manipulations give

$$\mu \, m_{\nu(1)-1} \, \mathbf{e}_1^T f(T_n) \, \mathbf{e}_{\nu(1)} = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \omega_{i,j} f^{(j)}(\lambda_i), \tag{6.6}$$

with $\omega_{i,j}$ complex weights, μ and $m_{\nu(1)-1}$ as in Theorem 6.3; see [54] and [70, Section 3] for algebraic expressions of the weights. This observation together with the proof of Theorem 6.3 shows that when $n = \nu(t)$ the bilinear form $\mu \, m_{\nu(1)-1} \, \mathbf{e}_1^T f(T_n) \, \mathbf{e}_{\nu(1)}$ is a matrix formulation of the n-node Gauss quadrature $\mathcal{G}_n(f)$ for the linear functional $\mathcal{L}$. Moreover, if $\nu(t) < n < \nu(t+1)$, then Lemma 5.4 gives $\mathcal{G}_n = \mathcal{G}_{\nu(t)}$; hence T_n and $T_{\nu(t)}$ correspond to the same Gauss quadrature $\mathcal{G}_{\nu(t)}$, despite being different.

7 The minimal partial realization and Gauss quadrature

Any triplet $(\mathbf{w}, A, \mathbf{v})$ composed of a matrix A and vectors $\mathbf{v}, \mathbf{w}$, can be associated with a dynamical system

$$\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}t} = A \, \mathbf{z}(t) + \mathbf{v} u(t)$$

$$y(t) = \mathbf{w}^* \mathbf{z}(t),$$

with $\mathbf{z}(t)$ the state vector, $u(t)$ the scalar input (control), and $y(t)$ the scalar output. The transfer function

$$\Gamma(\tau) := \mathbf{w}^*(\tau I - A)^{-1}\,\mathbf{v} = \sum_{j=0}^{\infty} \frac{\mathbf{w}^* A^j\,\mathbf{v}}{\tau^{j+1}}$$

connects $u(t)$ with $y(t)$ and it is obtained applying the Laplace transform; refer, e.g., to [48, Section 2], [68, Section 4], [1, Section 4.1, 4.2 and 11.1]. The series representation holds only for $|\tau|$ large enough, and the coefficients $\{\mathbf{w}^* A^j\,\mathbf{v}\}_{j=0}^{\infty}$ are usually known as *Markov parameters*. The triplet $(\mathbf{w}, A, \mathbf{v})$ is called a realization of Γ. One of the questions in systems theory is to determine all the realizations $(\mathbf{w}, A, \mathbf{v})$ that yield a given (rational) function Γ, or equivalently, its Markov parameters. When the realization matches a finite number of Markov parameters it is said to be a *partial realization*. A partial realization in which A has minimal dimension is called a *minimal partial realization*. Among the extensive literature about the realization problem we refer the reader to the papers by Kalman [52,53], Gilbert [36], Ho and Kalman [48], Gragg [39], Gragg and Lindquist [40], Parlett [68] (which offers an algebraic point of view), and to the monographs by Antoulas [1, Section 4.4] and by Liesen and Strakoš [60, Section 3.9]; see also [65]. In the papers by Chebyshev from 1855–1859 [12,13] and Christoffel from 1858 [16] the concept equivalent to the minimal partial realization is present (without using the name) for a sequence of moments defining a positive definite linear functional; cf. the comment in [11, p. 23]. The seminal paper by Stieltjes on continued fractions published in 1894 [77, Sections 7–8, pp. 623–625, and Section 51, pp. 688–690] provides an instructive description; see also [60, Section 3.9.1] and [73]. The results about the Gauss quadrature for real linear functionals and about the minimal partial realization of a sequence of real numbers appeared in the same year (1983) respectively in the monograph by Draux [22, Chapter 5] and in the paper by Gragg and Lindquist [40]. Section 5 has presented the results by Draux extending them to the complex case. Here the minimal partial realization of a sequence of complex numbers will be described together with the relationships between results in [22] and [40] (with extension to the complex case).

In the following we offer a non-standard formulation of the realization problem in systems theory.

Problem 1: For a given finite sequence of complex numbers

$$m_0, m_1, \dots, m_k, \tag{7.1}$$

find all the triplets $(\mathbf{w}, A, \mathbf{v})$ such that

$$\mathbf{w}^* A^j\,\mathbf{v} = m_j, \quad j = 0, \dots, k.$$

Notice that usually the Markov parameters are defined as $\eta_j = m_{j-1}$.

There always exists a solution of dimension $k+1$ of Problem 1. For instance, take $A \in \mathbb{C}^{k+1 \times k+1}$ and $\mathbf{v}, \mathbf{w} \in \mathbb{C}^{k+1}$ as

$$A = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} m_0 \\ m_1 \\ \vdots \\ m_k \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{7.2}$$

The sequence (7.1) defines the linear functional $\mathcal{L}$ on $\mathcal{P}_k$ with moments

$$\mathcal{L}(\lambda^j) = m_j, \quad j = 0, \dots, k. \tag{7.3}$$

For any solution $(\mathbf{w}, A, \mathbf{v})$ of dimension n, let $\lambda_1, \dots, \lambda_\ell$ be the distinct eigenvalues of A and s_i be the maximal geometric multiplicity of λ_i (the size of the largest Jordan block corresponding to λ_i). Then the definition of matrix function (6.5) and algebraic manipulations give

$$\mathbf{w}^* f(A)\, \mathbf{v} = \sum_{i=1}^{\ell} \sum_{s=0}^{s_i - 1} \omega_{i,s} f^{(s)}(\lambda_i), \quad s_1 + \dots + s_\ell \le n,$$

with $\omega_{i,s}$ complex weights. Therefore every realization of the sequence (7.1) defines a quadrature rule for the linear functional (7.3).

Problem 2: Among all the realizations for (7.1) find those of smallest dimension.

Let n be the smallest index so that the unique n-node Gauss quadrature determined by the regular FOP $p_n(\lambda)$ is exact for every polynomial of degree smaller than or equal to k, i.e.,

$$\mathcal{L}(q) = \mathcal{G}_n(q) = \sum_{i=1}^{\ell} \sum_{s=0}^{s_i - 1} \omega_{i,s} q^{(s)}(\lambda_i), \quad s_1 + \dots + s_\ell = n, \quad q \in \mathcal{P}_k.$$

If T_n is the block tridiagonal matrix (4.8) corresponding to $p_n(\lambda)$, then Theorem 6.3 shows that the triplet $(\mathbf{e}_1, T_n, \mu\, m_{\nu(1)-1} \mathbf{e}_{\nu(1)})$ is a minimal partial realization for (7.1). All the other minimal partial realizations can be expressed as

$$\left(B^* \mathbf{e}_1,\, B^{-1} T_n B,\, \mu\, m_{\nu(1)-1} B^{-1} \mathbf{e}_{\nu(1)} \right), \tag{7.4}$$

with B any $n \times n$ invertible matrix (notice that this is a straightforward extension of the result given in [40, Theorem 5] to complex Markov parameters). Hence any minimal partial realization of a sequence of complex number $m_0, m_1, \dots$ corresponds to a Gauss quadrature for the linear functional having $m_0, m_1, \dots$ as moments.

Finally, we recall the following well-known spectral result about minimal realizations, giving a proof based on the previous developments.

Theorem 7.1 *Consider the matrix A and the vectors $\mathbf{v}, \mathbf{w}$. If the triplet $(\mathbf{c}, S, \mathbf{b})$ is a minimal realization of the sequence of Markov parameters given by*

$$m_j = \mathbf{w}^* A^j \mathbf{v}, \quad j = 0, 1, \ldots,$$

then the spectrum of S is a subset of the spectrum of A.

Proof Let $p_k(\lambda)$ be the characteristic polynomial of the matrix A and consider the linear functional $\mathcal{L}$ defined by

$$\mathcal{L}(q) = \mathbf{w}^* q(A)\, \mathbf{v}, \quad q(\lambda) \in \mathcal{P}.$$

By Lemma 6.1 and the Cayley–Hamilton Theorem the k-degree polynomial $p_k(\lambda)$ is formally orthogonal to every polynomial, i.e., $\mathcal{L}(p_k q) = 0$ for every $q(\lambda) \in \mathcal{P}$. Consider the last regular FOP $p_n(\lambda)$ in the sequence of the FOPs with respect to $\mathcal{L}$. The polynomial $p_k(\lambda)$ is $(k-n)$-quasi-orthogonal (note that $n \leq k$). Hence the roots of $p_n(\lambda)$ are roots of $p_k(\lambda)$ by Theorem 4.4. As discussed above, every minimal realization can be expressed as $\big(B^*\mathbf{e}_1,\, B^{-1}T_n B,\, \mu\, m_{\nu(1)-1} B^{-1}\mathbf{e}_{\nu(1)}\big)$, with T_n the block tridiagonal matrix (4.8) corresponding to $p_n(\lambda)$ and B an invertible matrix. Thus Lemma 6.1 concludes the proof. □

We remark that the previous theorem is a consequence of the Canonical Structure Theorem of the linear system theory; see, e.g., [51], [52, Theorem 5], [36] and the description in [68, Section 7].

8 The look-ahead Lanczos algorithm and Gauss quadrature

Consider a complex matrix A and a complex vector $\mathbf{v}$ of the corresponding dimension. The nth *Krylov subspace* generated by A and $\mathbf{v}$ is the subspace

$$\mathcal{K}_n(A, \mathbf{v}) = \operatorname{span}\{\mathbf{v}, A\,\mathbf{v}, \ldots, A^{n-1}\,\mathbf{v}\},$$

which can be equivalently expressed as

$$\mathcal{K}_n(A, \mathbf{v}) = \{p(A)\,\mathbf{v} : p(\lambda) \in \mathcal{P}_{n-1}\}.$$

The basic facts about Krylov subspaces had been given by Gantmacher in [31]; other results can be found, e.g., in [60, Section 2.2].

Let A be a complex matrix, $\mathbf{v}, \mathbf{w}$ be complex vectors, and $\mathcal{L} : \mathcal{P} \to \mathbb{C}$ be the linear functional defined by

$$\mathcal{L}(p) = \mathbf{w}^* p(A)\, \mathbf{v}, \quad p(\lambda) \in \mathcal{P}. \tag{8.1}$$

Denoting with $\bar{p}(\lambda)$ the polynomial whose coefficients are the conjugates of the coefficients of $p(\lambda)$ and noticing that

$$p(A)^* = \bar{p}(A^*),$$

for $p(\lambda), q(\lambda) \in \mathcal{P}_{n-1}$, give

$$\mathcal{L}(qp) = \mathbf{w}^* q(A) p(A)\, \mathbf{v} = \widehat{\mathbf{w}}^* \widehat{\mathbf{v}},$$

with $\widehat{\mathbf{v}} = p(A)\,\mathbf{v} \in \mathcal{K}_n(A, \mathbf{v})$ and $\widehat{\mathbf{w}} = \bar{q}(A^*)\,\mathbf{w} \in \mathcal{K}_n(A^*, \mathbf{w})$.

The non-Hermitian Lanczos algorithm (formulated by Lanczos in [58] and [59]) gives, when possible, the vectors

$$\mathbf{v}_0, \ldots, \mathbf{v}_{n-1} \quad \text{and} \quad \mathbf{w}_0, \ldots, \mathbf{w}_{n-1},$$

which are respectively basis of $\mathcal{K}_n(A, \mathbf{v})$ and $\mathcal{K}_n(A^*, \mathbf{w})$ satisfying the biorthogonality conditions

$$\mathbf{w}_i^* \mathbf{v}_j = 0, \quad i \neq j, \quad \text{and} \quad \mathbf{w}_i^* \mathbf{v}_i \neq 0, \quad i, j = 0, \ldots, n-1. \tag{8.2}$$

In this case, there exist regular FOPs $p_0(\lambda), \ldots, p_{n-1}(\lambda)$ with respect to the linear functional (8.1) so that

$$\mathbf{v}_j = p_j(A)\,\mathbf{v} \quad \text{and} \quad \mathbf{w}_j = \bar{p}_j(A^*)\,\mathbf{w}, \quad j = 0, \ldots, n-1.$$

Hence bases satisfying (8.2) exist if and only if $\mathcal{L}$ is quasi-definite on $\mathcal{P}_{n-1}$; see, e.g., [72, Theorem 2.1].

In the non-Hermitian Lanczos method, the vectors $\mathbf{v}_j, \mathbf{w}_j$, $j = 0, \ldots, n-1$, are obtained by the three-term recurrences satisfied by the regular FOPs $p_0, \ldots, p_{n-1}$; for details refer to [7, Section 2.7.2], [43,44,45], [76, Chapter 7], [37, Chapter 4], [60, Section 2.4], also refer to the survey [72] where the connection with the Gauss quadrature for quasi-definite linear functionals is described. Considering biorthonormal vectors, i.e., $\mathbf{w}_i^* \mathbf{v}_i = 1$, the non-Hermitian Lanczos algorithm corresponds to the three-term recurrences (2.3) and can be given as Algorithm 8.1; see, e.g., [18,17]. The outputs of the first $n-1$ iterations of Algorithm 8.1 define the matrices

$$V_n = [\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}] \quad \text{and} \quad W_n = [\mathbf{w}_0, \ldots, \mathbf{w}_{n-1}]$$

which satisfy $W_n^* V_n = I$, with I the identity matrix of dimension n. Moreover,

$$AV_n = V_n J_n + \widehat{\mathbf{v}}_n \mathbf{e}_n^T,$$
$$A^* W_n = W_n \bar{J}_n + \widehat{\mathbf{w}}_n \mathbf{e}_n^T,$$

with J_n the complex Jacobi matrix (2.4) associated with the linear functional (8.1), and $\bar{J}_n$ the Jacobi matrix with conjugate elements ($\widehat{\mathbf{v}}_n$ and $\widehat{\mathbf{w}}_n$ are defined in Algorithm 8.1). Therefore the non-Hermitian Lanczos algorithm can be seen as a way to compute J_n and hence the Gauss quadrature for the functional (8.1); see [29, Theorem 2] and also [72] (for the block Lanczos algorithm see, e.g., [26, Section 3]).

If the nth iteration of Algorithm 8.1 gives $\beta_n = 0$, then the algorithm has a *breakdown*. Since $\beta_n = \mathcal{L}(\lambda p_{n-1} p_n)$, a breakdown arises if and only if $\mathcal{L}$ is not quasi-definite on $\mathcal{P}_n$. In this case, the FOP $p_n(\lambda)$ is orthogonal to itself. Therefore there do not exist biorthonormal bases of the Krylov subspaces $\mathcal{K}_{n+1}(A, \mathbf{v})$ and $\mathcal{K}_{n+1}(A^*, \mathbf{w})$. Moreover, there does not exist a regular FOP $p_{n+1}(\lambda)$. There are two kinds of breakdown for Algorithm 8.1:

Algorithm 8.1 (non-Hermitian Lanczos algorithm)
Input: a complex matrix A, and complex vectors $\mathbf{v}, \mathbf{w}$ such that $\mathbf{w}^*\mathbf{v} \neq 0$.
Output: vectors $\mathbf{v}_0, \dots, \mathbf{v}_{n-1}$ and vectors $\mathbf{w}_0, \dots, \mathbf{w}_{n-1}$ spanning respectively $\mathcal{K}_n(A, \mathbf{v})$, $\mathcal{K}_n(A^*, \mathbf{w})$ and satisfying the biorthogonality conditions (8.2) with $\mathbf{w}_i^*\mathbf{v}_i = 1$, $i = 0, \dots, n-1$.

Initialize: $\mathbf{v}_{-1} = \mathbf{w}_{-1} = 0$, $\beta_0 = \sqrt{\mathbf{w}^*\mathbf{v}}$, $\mathbf{v}_0 = \mathbf{v}/\beta_0$, $\mathbf{w}_0 = \mathbf{w}/\bar{\beta}_0$.
For $n = 1, 2, \dots$

$\alpha_{n-1} = \mathbf{w}_{n-1}^* A \mathbf{v}_{n-1}$,
$\widehat{\mathbf{v}}_n = A\mathbf{v}_{n-1} - \alpha_{n-1}\mathbf{v}_{n-1} - \beta_{n-1}\mathbf{v}_{n-2}$,
$\widehat{\mathbf{w}}_n = A^*\mathbf{w}_{n-1} - \bar{\alpha}_{n-1}\mathbf{w}_{n-1} - \bar{\beta}_{n-1}\mathbf{w}_{n-2}$,
$\beta_n = \sqrt{\widehat{\mathbf{w}}_n^*\widehat{\mathbf{v}}_n}$,
if $\beta_n = 0$ *then stop*,
$\mathbf{v}_n = \widehat{\mathbf{v}}_n/\beta_n$,
$\mathbf{w}_n = \widehat{\mathbf{w}}_n/\bar{\beta}_n$,

end.

1. *lucky breakdown* (or *benign breakdown*), when $\widehat{\mathbf{v}}_n = 0$ or $\widehat{\mathbf{w}}_n = 0$;
2. *serious breakdown*, when $\widehat{\mathbf{v}}_n \neq \mathbf{0}$ and $\widehat{\mathbf{w}}_n \neq \mathbf{0}$, but $\widehat{\mathbf{w}}_n^*\widehat{\mathbf{v}}_n = 0$.

In the first case either $\mathcal{K}_n(A, \mathbf{v})$ is A-invariant or $\mathcal{K}_n(A^*, \mathbf{w})$ is A^*-invariant. Then the algorithm is usually stopped since an invariant subspace is often a desirable result; see, e.g., [9], [68, Section 5] and [38, Section 10.5.5]. The second case is problematic. In [83, pp. 389–391] Wilkinson showed with some examples that well conditioned matrices with well conditioned eigenvectors can produce a breakdown. Hence as Wilkinson wrote, serious breakdown "is not associated with any shortcoming in the matrix A. It can happen even when the eigenproblem of A is very well conditioned. We are forced to regard it as a specific weakness of the Lanczos method itself." The interested reader can also refer to [75], [50, p. 34], [80, Chapter IV], [69], [68, Section 7], and [43, 44, 45].

Taylor in [80] and Parlett, Taylor, and Liu in [69] first proposed the *look-ahead Lanczos algorithm*, a strategy able to deal with the breakdown problem. When $\widehat{\mathbf{w}}_n^*\widehat{\mathbf{v}}_n = 0$, the idea behind their strategy is to look for a vector $\widetilde{\mathbf{w}}_k \in \mathcal{K}_{k+1}(A^*, \mathbf{w})$, with $k > n$ big enough, so that $\widetilde{\mathbf{w}}_k^*\widehat{\mathbf{v}}_n \neq 0$ and $\widetilde{\mathbf{w}}_k^*\mathbf{v}_j = 0$ for $j = 0, \dots, n-1$. In [28] Freund, Gutknecht, and Nachtigal implemented a different look-ahead strategy considering sequences of FOPs and quasi-orthogonal polynomials. Their procedure is based on the work of Gutknecht published in [43] and later in [44]; see also the thesis [66] by Nachtigal and the description in [27] by Freund. We also refer the reader to the strategy in [9, 10] and the related work [24]. The following part will describe the basic ideas behind the look-ahead Lanczos algorithm by Freund, Gutknecht, and Nachtigal.

Consider the linear functional (8.1) and let $p_0(\lambda) \neq 0, p_1(\lambda), \dots$ be the sequence (4.2) of polynomials so that $p_{\nu(0)}(\lambda) = p_0(\lambda), p_{\nu(1)}(\lambda), \dots$ are the regular FOPs and p_n is an $(n-\nu(t))$-quasi-orthogonal polynomial for $\nu(t) < n < \nu(t+1)$, with $\nu(t+1) = \infty$ when $p_{\nu(t)}$ is the last of the regular FOPs. Moreover, consider the vectors

$$\mathbf{v}_n = p_n(A)\,\mathbf{v} \quad \text{and} \quad \mathbf{w}_n = \bar{p}_n(A^*)\,\mathbf{w}, \quad n = 0, 1, \dots,$$

and the matrices $V_n^{(t)} = [\mathbf{v}_{\nu(t)}, \dots, \mathbf{v}_{n-1}]$, $W_n^{(t)} = [\mathbf{w}_{\nu(t)}, \dots, \mathbf{w}_{n-1}]$, with $V^{(t)} = V_{\nu(t+1)}^{(t)}$ and $W^{(t)} = W_{\nu(t+1)}^{(t)}$ for simplicity of notation. Hence for $\nu(t) < n \leq \nu(t+1)$, the columns of $V_n = [V^{(0)}, \dots, V_n^{(t)}]$ and of $W_n = [W^{(0)}, \dots, W_n^{(t)}]$ are respectively basis of $\mathcal{K}_n(A, \mathbf{v})$ and $\mathcal{K}_n(A^*, \mathbf{w})$. However, instead of the biorthogonality conditions (8.2), the following block-biorthogonality conditions hold

$$W_n^{(t)*} V_\ell^{(k)} = 0, \text{ for } t \neq k, \quad \text{and} \quad W_n^{(t)*} V_n^{(t)} = \Omega_n^{(t)}, \tag{8.3}$$

with

$$\Omega_n^{(t)} = \begin{bmatrix} \mathcal{L}(p_{\nu(t)}, p_{\nu(t)}) & \dots & \mathcal{L}(p_{\nu(t)}, p_{n-1}) \\ \vdots & & \vdots \\ \mathcal{L}(p_{n-1}, p_{\nu(t)}) & \dots & \mathcal{L}(p_{n-1}, p_{n-1}) \end{bmatrix};$$

we denote $\Omega_{\nu(t+1)}^{(t)}$ by $\Omega^{(t)}$.

By Theorem 4.4 and the recurrences (4.3) if $\nu(t) < n < \nu(t+1)$, then for some complex coefficients $\mathbf{a}_n = [\alpha_{n,\nu(t)}, \dots, \alpha_{n,n-1}]$ and $\beta_n \neq 0$ the following recurrences hold

$$\beta_n \mathbf{v}_n = A\mathbf{v}_{n-1} - \sum_{j=\nu(t)}^{n-1} \alpha_{n,j} \mathbf{v}_j, \quad \text{and} \quad \bar{\beta}_n \mathbf{w}_n = A^* \mathbf{w}_{n-1} - \sum_{j=\nu(t)}^{n-1} \bar{\alpha}_{n,j} \mathbf{w}_j.$$

If $n = \nu(t+1)$ with $t \geq 0$, then for some $\beta_n \neq 0$ the recurrences (4.5) give

$$\beta_n \mathbf{v}_n = A\mathbf{v}_{n-1} - \sum_{j=\nu(t)}^{n-1} \alpha_{n,j} \mathbf{v}_j - \gamma_n \mathbf{v}_{\nu(t-1)}$$

$$\bar{\beta}_n \mathbf{w}_n = A^* \mathbf{w}_{n-1} - \sum_{j=\nu(t)}^{n-1} \bar{\alpha}_{n,j} \mathbf{w}_j - \bar{\gamma}_n \mathbf{w}_{\nu(t-1)},$$

where $\nu(-1) = -1$, $\mathbf{v}_{-1} = \mathbf{w}_{-1} = 0$, $\gamma_{\nu(1)} = 0$,

$$\gamma_n = \frac{\mathbf{w}_{\nu(t)-1}^* A\, \mathbf{v}_{n-1}}{\mathbf{w}_{\nu(t)-1}^* \mathbf{v}_{\nu(t-1)}}, \quad t \geq 1,$$

and the coefficients $\mathbf{a}_n = [\alpha_{n,\nu(t)}, \dots, \alpha_{n,n-1}]$ are given as the solution of the system

$$\Omega^{(t)}\, \mathbf{a}_n = W^{(t)*} A \mathbf{v}_{n-1};$$

see the linear system (4.6). The described recurrences can be expressed in the matrix form

$$AV_n = V_n T_n^T + \beta_{n+1} \mathbf{v}_n \mathbf{e}_n^T \quad \text{and} \quad A^* W_n = W_n T_n^* + \bar{\beta}_{n+1} \mathbf{w}_n \mathbf{e}_n^T,$$

with T_n^T the transpose of the block tridiagonal matrix T_n defined in (4.8) and T_n^* the conjugate transpose of T_n. The resulting form of the look-ahead Lanczos algorithm is given as Algorithm 8.2 and corresponds to the algorithm proposed in [28, Algorithm 3.1]; see also [27, Algorithm 5.1].

The first n iterations of Algorithm 8.2 produce the coefficients of the block tridiagonal matrix T_n. If $\nu(t) \leq n < \nu(t+1)$, then the Gauss quadrature $\mathcal{G}_{\nu(t)}$ for the linear functional (8.1) has the matrix formulation (6.6) which is determined by the matrix T_n. Hence Algorithm 8.2 produces Gauss quadratures for the linear functional (8.1). Notice that by Lemma 5.4 the matrix T_n corresponds to the Gauss quadrature $\mathcal{G}_{\nu(t)}$ for $n = \nu(t), \dots, \nu(t+1) - 1$. Nevertheless, the iterations $\nu(t)+1, \dots, \nu(t+1)$ of Algorithm 8.2 are necessary in order to get the degree of exactness of $\mathcal{G}_{\nu(t)}$. At the same time, Algorithm 8.2 also produces the triplet $(\mathbf{e}_1, T_{\nu(t)}, \mu\, m_{\nu(1)-1} \mathbf{e}_{\nu(1)})$, i.e., the minimal partial realization (7.4) (with $B = I$) of the sequence of Markov parameters defined by

$$m_j = \mathbf{w}^* A^j\, \mathbf{v}, \qquad \text{for } j = 0, 1, \dots . \tag{8.4}$$

Consider the case in which a benign breakdown does not arise and the determinants of the Hankel submatrices (2.2) composed of the moments of the linear functional (8.1) are such that

$$\Delta_{n-1} \neq 0, \quad \Delta_{n+k} = 0, \qquad \text{for } k = 0, 1, \dots, \tag{8.5}$$

known as *incurable breakdown*; see [80, page 56], [69, Section 7], [68, p. 577]. Then $p_n(\lambda)$ is the last of the regular FOPs ($n = \nu(t)$). By Theorem 5.6, the quadrature $\mathcal{G}_n$ determined by $p_n(\lambda)$ is the Gauss quadrature with maximal number of nodes (counting the multiplicities) and it is exact for every polynomial. Equivalently, let T_n be the block tridiagonal matrix obtained at the nth step of Lanczos. Theorem 6.3 gives

$$\mathcal{L}(p) = \mathbf{w}^* p(A)\, \mathbf{v} = \mathcal{G}_n(p) = \mu\, m_{\nu(1)-1}\, \mathbf{e}_1^T p(T_n)\, \mathbf{e}_{\nu(1)}, \quad p(\lambda) \in \mathcal{P}.$$

Moreover, if $f(\lambda)$ is a function so that $f(A)$ and $f(T_n)$ are well defined matrix function, then there exists a polynomial $q(\lambda)$ interpolating (in the Hermite sense) the spectra of A and T_n; see, e.g., [47, Section 1.2]. Therefore

$$\mathcal{L}(f) = \mathbf{w}^* q(A)\, \mathbf{v} = \mu\, m_{\nu(1)-1}\, \mathbf{e}_1^T q(T_n)\, \mathbf{e}_{\nu(1)} = \mathcal{G}_n(f).$$

Looking at the Lanczos algorithm as a method for getting the Gauss quadrature for a linear functional (8.1), the incurable breakdown corresponds to the solution of the problem as well as the lucky breakdown. Furthermore, the triplet $(\mathbf{e}_1, T_n, \mu\, m_{\nu(1)-1} \mathbf{e}_{\nu(1)})$ is a minimal realization of the transfer function associated with $(\mathbf{w}, A, \mathbf{v})$, i.e., it matches the Markov parameters (8.4). The previous considerations together with Theorem 7.1 give a new proof for the

Algorithm 8.2 (look-ahead Lanczos algorithm)
Input: a complex matrix A *and complex vectors* $\mathbf{v}, \mathbf{w} \neq 0$.
Output: vectors $\mathbf{v}_0, \dots, \mathbf{v}_{n-1}$ *and vectors* $\mathbf{w}_0, \dots, \mathbf{w}_{n-1}$ *spanning respectively* $\mathcal{K}_n(A, \mathbf{v})$, $\mathcal{K}_n(A^*, \mathbf{w})$ *and satisfying the block biorthogonality conditions* (8.3).

Initialize: $\nu(-1) = -1$, $\mathbf{v}_{-1} = \mathbf{w}_{-1} = 0$, $t = \nu(0) = 0$, $\eta_0 = 1$, *fix* $\beta_0 \neq 0$,
$\mathbf{v}_0 = \mathbf{v}/\beta_0$, $\mathbf{w}_0 = \mathbf{w}/\bar{\beta}_0$, $V_1^{(0)} = [\mathbf{v}_0]$, $W_1^{(0)} = [\mathbf{w}_0]$.
For $n = 1, 2, \dots,$
  $\Omega_n^{(t)} = W_n^{(t)*} V_n^{(t)}$,
  If $\Omega_n^{(t)}$ *is regular, then*
    $\gamma_n = (\mathbf{w}^*_{\nu(t)-1} A \mathbf{v}_{n-1})/\eta_t$,
    $\mathbf{a}_n = \left(\Omega_n^{(t)}\right)^{-1} W_n^{(t)} A \mathbf{v}_{n-1}$,
    $\widehat{\mathbf{v}}_n = A\mathbf{v}_{n-1} - V_n^{(t)} \mathbf{a}_n - \gamma_n \mathbf{v}_{\nu(t-1)}$,
    $\widehat{\mathbf{w}}_n = A^*\mathbf{w}_{n-1} - W_n^{(t)} \bar{\mathbf{a}}_n - \bar{\gamma}_n \mathbf{w}_{\nu(t-1)}$,
    fix $\beta_n \neq 0$,
    $\mathbf{v}_n = \widehat{\mathbf{v}}_n/\beta_n$, $\mathbf{w}_n = \widehat{\mathbf{w}}_n/\bar{\beta}_n$,
    $\eta_{t+1} = (\mathbf{w}^*_{n-1} \mathbf{v}_{\nu(t)})$,
    $t = t + 1$, $\nu(t) = n$, $V_{n+1}^{(t)} = [\mathbf{v}_n]$, $W_{n+1}^{(t)} = [\mathbf{w}_n]$,
  else if $\Omega_n^{(t)}$ *is singular, then*
    fix the vector $\mathbf{a}_n$,
    $\widehat{\mathbf{v}}_n = A\mathbf{v}_{n-1} - V_n^{(t)} \mathbf{a}_n$,
    $\widehat{\mathbf{w}}_n = A^*\mathbf{w}_{n-1} - W_n^{(t)} \bar{\mathbf{a}}_n$,
    fix $\beta_n \neq 0$,
    $\mathbf{v}_n = \widehat{\mathbf{v}}_n/\beta_n$, $\mathbf{w}_n = \widehat{\mathbf{w}}_n/\bar{\beta}_n$,
    $V_{n+1}^{(t)} = [V_n^{(t)}, \mathbf{v}_n]$, $W_{n+1}^{(t)} = [W_n^{(t)}, \mathbf{w}_n]$,
  end if
  If $\mathbf{v}_n = 0$ *or* $\mathbf{w}_n = 0$ *then*
    stop,
  end if
end for.

Mismatch Theorem based on the properties of the Gauss quadrature for linear functionals. The Mismatch Theorem was first proved in [80, Theorem 4.2] by Taylor; see also [69, p. 117], and [68, Section 7] where the theorem was connected with the minimal realization problem.

Theorem 8.3 (Mismatch Theorem) *Le* T_n *be the block tridiagonal matrix obtained at the* n*th step of Algorithm 8.2 with* A *as the input matrix and* $\mathbf{w}, \mathbf{v} \neq 0$ *as the input vectors. If the algorithm has an incurable breakdown at the* n*th*

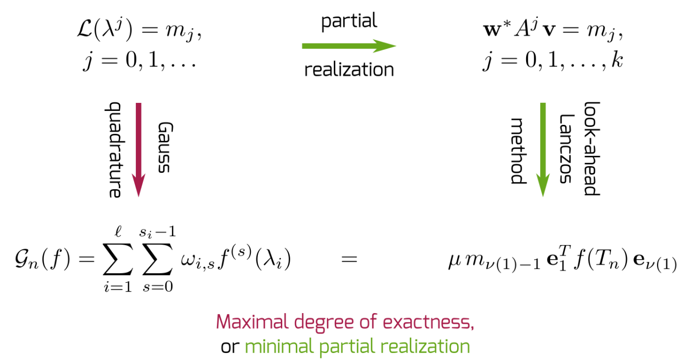


Fig. 2 Visualization of the connections between the Gauss quadrature for linear functionals, minimal partial realization, and look-ahead Lanczos algorithm.

step, i.e., the Hankel determinants corresponding to the linear functional (8.1) *satisfy* (8.5)*, then each eigenvalue of T_n (known as Ritz value) is an eigenvalue of A.*

Notice that the look-ahead Lanczos algorithm in [80] produces a block tridiagonal matrix different from the matrix T_n (4.8). However, both the matrices are minimal realization of the same sequence of numbers and therefore they are similar.

9 Conclusion

The n-node Gauss quadrature $\mathcal{G}_n$ for a linear functional $\mathcal{L}$ described in Section 5 is a straightforward extension of the quadrature introduced for real-valued linear functionals in [22, Chapter 5] to the complex case and it satisfies the following properties:

1. the Gauss quadrature $\mathcal{G}_n$ has degree of exactness *at least* $2n-1$;
2. the Gauss quadrature $\mathcal{G}_n$ exists and is unique if and only if the Hankel submatrix of moments H_{n-1} is nonsingular, i.e., $\mathit{\Delta}_{n-1} \neq 0$;
3. by Theorem 6.3 the Gauss quadrature can be written in the matrix form $\mathcal{G}_n(f) = \mu\, m_{\nu(1)-1}\, \mathbf{e}_1^T f(T_n)\, \mathbf{e}_{\nu(1)}$.

Note that such properties are weaker forms of the properties G1–G3 in Section 2.

Figure 2 summarizes the connections between the Gauss quadrature for linear functionals, minimal partial realization, and look-ahead Lanczos algorithm. On the right-hand side, the triplet $(\mathbf{w}, A, \mathbf{v})$ is a partial realization matching the first $k+1$ elements of the sequence of complex numbers $m_0, m_1, \dots$. A minimal partial realization can be obtained applying the look-ahead Lanczos algorithm to the matrix A and the vectors $\mathbf{v}, \mathbf{w}$ (this is also connected with the concept of model reduction, see, e.g., [60, Chapter 3, in particular Section 3.9]). Notice that the Lanczos algorithm applied to the partial realization (7.2) is related to the Berlekamp-Massey algorithm [5,63] (see [56], [40], and [6]). On the left-hand side, the sequence $m_0, m_1, \dots$ determines the linear functional $\mathcal{L} : \mathcal{P} \to \mathbb{C}$ by defining its moments. The functional $\mathcal{L}$ can be approximated by a Gauss quadrature. Among all the Gauss quadratures exact on $\mathcal{P}_k$, there is one with the minimal number of nodes n (counting the multiplicities). Such quadrature can be written in the matrix form

$$\mathcal{G}_n(f) = \mu\, m_{\nu(1)-1}\, \mathbf{e}_1^T f(T_n)\, \mathbf{e}_{\nu(1)},$$

i.e., it corresponds to the minimal partial realization matching $m_0, \dots, m_k$.

Sections 7 and 8 discussed the correspondence between the incurable breakdown in the look-ahead Lanczos algorithm and the minimal realization of an infinite sequence of complex numbers (and to the unique Gauss quadrature exact for every polynomial). This connection led us to a new proof for the Mismatch Theorem 8.3.

Acknowledgements We would like to thank Zdeněk Strakoš for the helpful comments and improvements suggested.

References

1. A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, vol. 6 of Advances in Design and Control, SIAM, Philadelphia, PA, 2005. With a foreword by Jan C. Willems.
2. Z. Bai, *Error analysis of the Lanczos algorithm for the nonsymmetric eigenvalue problem*, Math. Comp., 62 (1994), pp. 209–226.
3. Z. Bai, D. M. Day, and Q. Ye, *ABLE: an adaptive block Lanczos method for non-Hermitian eigenvalue problems*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 1060–1082.
4. B. Beckermann, *Complex Jacobi matrices*, J. Comput. Appl. Math., 127 (2001), pp. 17–65.
5. E. R. Berlekamp, *Algebraic coding theory*, McGraw-Hill Book Co., New York-Toronto, Ont.-London, 1968.
6. D. L. Boley, T. J. Lee, and F. T. Luk, *The Lanczos algorithm and Hankel matrix factorization*, Linear Algebra Appl., 172 (1992), pp. 109–133. Second NIU Conference on Linear Algebra, Numerical Linear Algebra and Applications (DeKalb, IL, 1991).
7. C. Brezinski, *Padé-type approximation and general orthogonal polynomials*, Internat. Ser. Numer. Math., Birkhäuser, 1980.
8. ———, *Computational aspects of linear control*, vol. 1 of Numerical Methods and Algorithms, Kluwer Acad. Publ., Dordrecht, 2002.
9. C. Brezinski, M. Redivo Zaglia, and H. Sadok, *Avoiding breakdown and near-breakdown in Lanczos type algorithms*, Numer. Algorithms, 1 (1991), pp. 261–284.
10. ———, *A breakdown-free Lanczos type algorithm for solving linear systems*, Numer. Math., 63 (1992), pp. 29–38.
11. A. Bultheel and M. Van Barel, *Linear Algebra, Rational Approximation and Orthogonal Polynomials*, vol. 6 of Stud. Comput. Math., North-Holland Publishing Co., Amsterdam, 1997.
12. P. Chebyshev, *Sur les fractions continues*, (1855). Reprinted in Oeuvres I, 11 (Chelsea, New York, 1962), pp. 203–230.

13. ———, *Le développement des fonctions à une seule variable*, (1859). Reprinted in Oeuvres I, 19 (Chelsea, New York, 1962), pp. 501–508.
14. T. S. Chihara, *On quasi-orthogonal polynomials*, Proc. Amer. Math. Soc., 8 (1957), pp. 765–767.
15. ———, *An Introduction to Orthogonal Polynomials*, vol. 13 of Mathematics and its Applications, Gordon and Breach Science Publishers, New York, 1978.
16. E. B. Christoffel, *Über die Gaußische Quadratur und eine Verallgemeinerung derselben*, J. Reine Angew. Math., 55 (1858), pp. 61–82. Reprinted in Gesammelte mathematische Abhandlungen I (B. G. Teubner, Leipzig, 1910), pp. 65–87.
17. J. Cullum, W. Kerner, and R. A. Willoughby, *A generalized nonsymmetric Lanczos procedure*, Comput. Phys. Commun., 53 (1989), pp. 19–48.
18. J. Cullum and R. A. Willoughby, *A practical procedure for computing eigenvalues of large sparse nonsymmetric matrices*, in North-Holland Mathematics Studies, vol. 127, Elsevier, 1986, pp. 193–240.
19. G. Cybenko, *An explicit formula for Lanczos polynomials*, Linear Algebra Appl., 88 (1987), pp. 99–115.
20. D. M. Day, *Semi-duality in the two-sided Lanczos algorithm*, PhD thesis, University of California, Berkeley, 1993.
21. ———, *An efficient implementation of the nonsymmetric Lanczos algorithm*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 566–589.
22. A. Draux, *Polynômes Orthogonaux Formels*, vol. 974 of Lecture Notes in Math. , Springer-Verlag, Berlin, 1983.
23. ———, *On quasi-orthogonal polynomials*, J. Approx. Theory, 62 (1990), pp. 1–14.
24. ———, *Formal orthogonal polynomials revisited. Applications*, Numer. Algorithms, 11 (1996), pp. 143–158.
25. ———, *On quasi-orthogonal polynomials of order r*, Integral Transforms Spec. Funct., 27 (2016), pp. 747–765.
26. C. Fenu, D. Martin, L. Reichel, and G. Rodriguez, *Block Gauss and Anti-Gauss Quadrature with Application to Networks*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1655–1684.
27. R. W. Freund, *The Look-Ahead Lanczos Process for Large Nonsymmetric Matrices and Related Algorithms*, in Linear Algebra for Large Scale and Real-Time Applications, M. S. Moonen, G. H. Golub, and B. L. R. de Moor, eds., vol. 232 of NATO ASI Series, Kluwer Acad. Publ., Dordrecht, 1993, pp. 137–163.
28. R. W. Freund, M. H. Gutknecht, and N. M. Nachtigal, *An implementation of the look-ahead Lanczos algorithm for non-Hermitian matrices*, SIAM J. Sci. Comput., 14 (1993), pp. 137–158.
29. R. W. Freund and M. Hochbruck, *Gauss Quadratures Associated With the Arnoldi Process and the Lanczos Algorithm*, in Linear Algebra for Large Scale and Real-Time Applications, M. S. Moonen, G. H. Golub, and B. L. R. de Moor, eds., vol. 232 of NATO ASI Series, Kluwer Acad. Publ., Dordrecht, 1993, pp. 377–380.
30. R. W. Freund and H. Zha, *A look-ahead algorithm for the solution of general Hankel systems*, Numer. Math., 64 (1993), pp. 295–321.
31. F. R. Gantmacher, *On the algebraic analysis of Krylov's method of transforming the secular equation*, Trans. Second Math. Congress, II (1934), pp. 45–48. In Russian. Title translation as in [32].
32. ———, *The Theory of Matrices. Vols. 1, 2*, Chelsea Publishing Co., New York, 1959.
33. W. Gautschi, *A survey of Gauss-Christoffel quadrature formulae*, in E. B. Christoffel (Aachen/Monschau, 1979), Birkhäuser, Basel, 1981, pp. 72–147.
34. ———, *Orthogonal Polynomials: Computation and Approximation*, Numer. Math. Sci. Comput. , Oxford University Press, New York, 2004.
35. ———, *Numerical Analysis*, Birkhäuser Boston, 2011.
36. E. Gilbert, *Controllability and observability in multivariable control systems*, SIAM J. Control, 1 (1963), pp. 128–151.
37. G. H. Golub and G. Meurant, *Matrices, Moments and Quadrature with Applications*, Princeton Ser. Appl. Math., Princeton University Press, Princeton, NJ, 2010.
38. G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins Stud. Math. Sci., Johns Hopkins University Press, Baltimore, MD, fourth ed., 2013.

39. W. B. Gragg, *Matrix interpretations and applications of the continued fraction algorithm*, Rocky Mountain J. Math., 4 (1974), pp. 213–225.
40. W. B. Gragg and A. Lindquist, *On the partial realization problem*, Linear Algebra Appl., 50 (1983), pp. 277–319.
41. A. Günnel, R. Herzog, and E. Sachs, *A note on preconditioners and scalar products in Krylov subspace methods for self-adjoint problems in Hilbert space*, Elect. Trans. Numer. Anal., 41 (2014), pp. 13–20.
42. H. Guo and R. A. Renaut, *Estimation of $u^T f(A)v$ for large-scale unsymmetric matrices*, Numer. Linear Algebra Appl., 11 (2004), pp. 75–89.
43. M. H. Gutknecht, *A completed theory of the unsymmetric Lanczos process and related algorithms. I*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 594–639.
44. ———, *A completed theory of the unsymmetric Lanczos process and related algorithms. II*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 15–58.
45. ———, *The Lanczos Process and Padé approximation*, in Proceedings of the Cornelius Lanczos International Centenary Conference (Raleigh, NC, 1993), SIAM, Philadelphia, PA, 1994, pp. 61–75.
46. M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Research Nat. Bur. Standards, 49 (1952), pp. 409–436.
47. N. J. Higham, *Functions of Matrices. Theory and Computation*, SIAM, Philadelphia, PA, 2008.
48. B. Ho and R. E. Kalman, *Effective construction of linear state-variable models from input/output functions*, at-Automatisierungstechnik, 14 (1966), pp. 545–548.
49. M. Hochbruck, *The Padé Table and its Relation to Certain Numerical Algorithms*, Habilitation thesis, Universität Tübingen, (1996).
50. A. S. Householder and F. L. Bauer, *On certain methods for expanding the characteristic polynomial*, Numer. Math., 1 (1959), pp. 29–37.
51. R. E. Kalman, *Canonical structure of linear dynamical systems*, Proc. Nat. Acad. Sci. U.S.A., 48 (1962), pp. 596–600.
52. ———, *Mathematical description linear systems*, SIAM J. Control, 1 (1963), pp. 152–192.
53. ———, *On partial realizations, transfer functions, and canonical forms*, Acta Polytech. Scand. Math. Comput. Sci. Ser., 31 (1979), pp. 9–32.
54. J. Kautský, *Matrices related to interpolatory quadratures*, Numer. Math., 36 (1980/81), pp. 309–318.
55. H. L. Krall and I. M. Sheffer, *Differential equations of infinite order for orthogonal polynomials*, Annali di Matematica Pura ed Applicata, 74 (1966), pp. 135–172.
56. S. Y. Kung, *Multivariable and multidimensional systems: analysis and design*, PhD thesis, Stanford University, Stanford, 1977.
57. K. H. Kwon and L. L. Littlejohn, *Classification of classical orthogonal polynomials*, J. Korean Math. Soc, 34 (1997), pp. 973–1008.
58. C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Research Nat. Bur. Standards, 45 (1950), pp. 255–282.
59. ———, *Solution of systems of linear equations by minimized iterations*, J. Research Nat. Bur. Standards, 49 (1952), pp. 33–53.
60. J. Liesen and Z. Strakoš, *Krylov subspace methods: principles and analysis*, Numer. Math. Sci. Comput., Oxford University Press, Oxford, 2013.
61. L. Lorentzen and H. Waadeland, *Continued Fractions with Applications*, vol. 3 of Stud. Comput. Math., North-Holland Publishing Co., Amsterdam, 1992.
62. J. Málek and Z. Strakoš, *Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs:*, SIAM Spotlights, SIAM, Philadelphia, 2015.
63. J. L. Massey, *Shift-register synthesis and* BCH *decoding*, IEEE Trans. Information Theory, IT-15 (1969), pp. 122–127.
64. G. V. Milovanović and A. S. Cvetković, *Complex Jacobi Matrices and quadrature rules*, Filomat, 17 (2003), pp. 117–134.
65. B. C. Moore, *Principal component analysis in linear systems: controllability, observability, and model reduction*, IEEE Trans. Automat. Control, 26 (1981), pp. 17–32.

66. N. M. Nachtigal, *A look-ahead variant of the Lanczos algorithm and its application to the quasi-minimal residual method for non-Hermitian linear systems*, PhD thesis, Massachusetts Institute of Technology, Cambridge, 1991.
67. C. C. Paige, I. Panayotov, and J.-P. M. Zemke, *An augmented analysis of the perturbed two-sided Lanczos tridiagonalization process*, Linear Algebra Appl., 447 (2014), pp. 119–132.
68. B. N. Parlett, *Reduction to tridiagonal form and minimal realizations*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 567–593.
69. B. N. Parlett, D. R. Taylor, and Z. A. Liu, *A look-ahead Lanczos algorithm for unsymmetric matrices*, Math. Comp., 44 (1985), pp. 105–124.
70. M. Piñar and V. Ramírez, *Matrix interpretation of formal orthogonal polynomials for nondefinite functionals*, J. Comput. Appl. Math., 18 (1987), pp. 265–277.
71. S. Pozza, M. S. Pranić, and Z. Strakoš, *Gauss quadrature for quasi-definite linear functionals*, IMA J. Numer. Anal., 37 (2017), pp. 1468–1495.
72. ———, *The Lanczos algorithm and complex Gauss quadrature*, Electron. Trans. Numer. Anal., 50 (2018), pp. 1–19.
73. S. Pozza and Z. Strakoš, *Algebraic description of the finite Stieltjes moment problem*, Linear Algebra Appl., 561 (2019), pp. 207–227.
74. M. Riesz, *Sur le problème des moments. Troisième Note*, Ark. Mat. Fys, 17 (1923), pp. 1–52.
75. H. Rutishauser, *Beiträge zur Kenntnis des Biorthogonalisierungs-Algorithmus von Lanczos*, Z. Angew. Math. Physik, 4 (1953), pp. 35–56.
76. Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, PA, second ed., 2003.
77. T. J. Stieltjes, *Recherches sur les fractions continues*, Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys., 8 (1894), pp. J. 1–122. Reprinted in Oeuvres II (P. Noordhoff, Groningen, 1918), pp. 402–566. English translation *Investigations on continued fractions* in Thomas Jan Stieltjes, Collected Papers, Vol. II (Springer-Verlag, Berlin, 1993), pp. 609–745.
78. Z. Strakoš, *Model reduction using the Vorobyev moment problem*, Numer. Algorithms, 51 (2009), pp. 363–379.
79. G. Szegö, *Orthogonal Polynomials*, vol. XXIII of Amer. Math. Soc. Colloq. Publ., American Mathematical Society, New York, 1939.
80. D. R. Taylor, *Analysis of the look ahead Lanczos algorithm*, PhD thesis, University of California, Berkeley, 1982.
81. C. H. Tong and Q. Ye, *Analysis of the finite precision bi-conjugate gradient algorithm for nonsymmetric linear systems*, Math. Comp., 69 (2000), pp. 1559–1575.
82. Y. V. Vorobyev, *Methods of Moments in Applied Mathematics*, Translated from the Russian by Bernard Seckler, Gordon and Breach Science Publishers, New York, 1965.
83. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Monographs on Numerical Analysis, The Clarendon Press Oxford University Press, New York, 1988. 1st ed. published 1963.