Numerische Mathematik manuscript No.
(will be inserted by the editor)

Dianne P. O'Leary · Zdeněk Strakoš · Petr Tichý

On Sensitivity of Gauss-Christoffel Quadrature

the date of receipt and acceptance should be inserted later

Abstract In numerical computations the question *how much does a function change under perturbations of its arguments* is of central importance. In this work, we investigate sensitivity of Gauss-Christoffel quadrature with respect to small perturbations of the distribution function. In numerical quadrature, a definite integral is approximated by a finite sum of functional values evaluated at given quadrature nodes and multiplied by given weights. Consider a sufficiently smooth integrated function uncorrelated with the perturbation of the distribution function. Then it seems natural that given the same number of function evaluations, the difference between the quadrature approximations is of the same order as the difference between the (original and perturbed) approximated integrals. That is perhaps one of the reasons why, to our knowledge, the sensitivity question has not been formulated and addressed in the literature, though several other sensitivity problems, motivated, in particular, by computation of the quadrature nodes and weights

The work of the first author was supported by the National Science Foundation under Grants CCR-0204084 and CCF-0514213. The work of the other two authors was supported by the Program Information Society under project 1ET400300415 and by the Institutional Research Plan AV0Z100300504.

Dianne P. O'Leary
Dept. of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742
E-mail: oleary@cs.umd.edu

Zdeněk Strakoš
Institute of Computer Science, Academy of Sciences of the Czech Republic, and Faculty of Mathematics and Physics, Charles University, Prague
E-mail: strakos@cs.cas.cz

Petr Tichý
Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, in the years 2003-2006 on leave at the Institute of Mathematics, TU Berlin, Germany
E-mail: tichy@cs.cas.cz

from moments, have been thoroughly studied by many authors. We survey existing particular results and show that *even a small perturbation of a distribution function can cause large differences in Gauss-Christoffel quadrature estimates*. We then discuss conditions under which the Gauss-Christoffel quadrature is insensitive under perturbation of the distribution function, present illustrative examples, and relate our observations to known conjectures on some sensitivity problems.

1 Introduction

The computation of orthogonal polynomials and Gauss-Christoffel quadrature draws upon several fields from classical analysis and approximation theory as well as modern numerical linear algebra. It has been intensively studied by many generations of mathematicians.

Here we consider linear functionals in the form of the Riemann-Stieltjes integral and restrict ourselves to distribution functions that are nondecreasing on a finite interval $[a, b]$ on the real line. By the k-point Gauss-Christoffel quadrature we mean the approximation of a given Riemann-Stieltjes integral

$$I_\omega(f) = \int_a^b f(x)\, d\omega(x) \tag{1}$$

by the discrete linear functional

$$I_\omega^k(f) = \sum_{j=1}^{k} \vartheta_j f(t_j)\,,$$

determined by nodes $a \le t_1 < \cdots < t_k \le b$ and positive weights $\{\vartheta_1, \ldots, \vartheta_k\}$ such that $I_\omega^k(f) = I_\omega(f)$ whenever f is a polynomial of degree at most $2k-1$ [18], [6, Section 2.7]. The recent encyclopedic book by Gautschi [23], his surveys [18,24] and the survey by Laurie [40] describe the state-of-the-art of Gauss-Christoffel quadrature computation, and can be recommended as fundamental reading for anyone interested in related problems.

In this paper we investigate sensitivity of Gauss-Christoffel quadrature with respect to small perturbations in the distribution function. Suppose we have two distribution functions $\omega(x)$ and $\tilde{\omega}(x)$ which are nondecreasing on the finite interval $[a, b]$ and close to each other. We are interested in estimating the two integrals

$$I_\omega = \int_a^b f(x)\, d\omega(x)\,, \quad I_{\tilde{\omega}} = \int_a^b f(x)\, d\tilde{\omega}(x)\,. \tag{2}$$

Although it seems natural to expect that the Gauss-Christoffel quadrature estimates of the same degree will be close when f is sufficiently smooth (and also uncorrelated with the difference between the given distribution functions), it is not clear that this is true. If we use Gauss-Christoffel quadrature to compute the estimates, then $\omega(x)$ and $\tilde{\omega}(x)$ induce different sequences of orthogonal polynomials. Therefore, the quadrature weights and nodes for the

same degree of the quadrature might be different from each other and in fact can be sensitive to small perturbations to the distribution function. Indeed, in Section 2 we present an example in which small changes in the distribution function produce large changes in the nodes, weights and quadrature approximations, even though the value of the approximated integral does not change much. This motivates our further considerations.

In Section 3, we review particular subproblems arising from different methods for computing Gauss-Christoffel quadrature formulas, with the emphasis on the sensitivity of maps from (modified) moments to the nodes and weights of the computed quadrature. For earlier results, refer to [15, p. 252 and Section 2], and for recent analysis to [1,40,23]. Despite the vast literature on related subjects, the problem of sensitivity of the Gauss-Christoffel quadrature has, to our knowledge, not been posed or examined in the literature. That problem certainly is of theoretical importance, and it is desirable to investigate its relationship with the subproblems studied in the literature. Section 4 recalls some basics about the error in Gauss-Christoffel quadrature approximations. In Section 5 we present discussion and further examples that lead to some understanding of the sensitivity of Gauss-Christoffel quadrature approximations. Section 6 gives a summary and open questions.

Our interest in this problem originated in analysis of the conjugate gradient method for solving linear systems and of the Lanczos method for solving the symmetric eigenvalue problem. The close relationship of these methods of numerical linear algebra to Gauss-Christoffel quadrature of the Riemann-Stieltjes integral has been known since their introduction; see [31, § 14-18], [59, Chapter III]. In particular, the conjugate gradient method generates a sequence of Gauss-Christoffel approximations to the piecewise constant distribution function that has jumps at the eigenvalues of the linear operator equal in magnitude to the squared components of the normalized initial residual along the corresponding eigenfunctions. Moreover, the size of the A-norm of the error at the kth step of the conjugate gradient method has a natural interpretation as the scaled remainder of the kth order Gauss-Christoffel quadrature approximation of the Riemann-Stieltjes integral; see [5] and [42, Sections 2.2 and 3.3] for a recent review of related results and bibliography. There is also an interesting relationship of the sensitivity of Gauss-Christoffel quadrature to the convergence properties of the conjugate gradient and Lanczos algorithms in finite precision arithmetic. Its detailed investigation is, however, out of the scope of this paper.

All experiments in this paper were performed using MATLAB on a computer with machine precision $\approx 10^{-16}$.

2 Motivating examples

We now present an example of a nondecreasing discontinuous distribution function $\omega(x)$ with finite points of increase, and a perturbation of this function, for which the Gauss-Christoffel quadrature estimates can be quite sensitive. We use a distribution function from [53] with the value between a and

the first point of increase zero, and points of increase $\lambda_1 < \cdots < \lambda_n$,

$$\lambda_i = \lambda_1 + \frac{i-1}{n-1}(\lambda_n - \lambda_1)\,\gamma^{n-i}, \quad i = 2, \ldots, n-1,$$

where $0 < a < \lambda_1, \lambda_n < b$ and $\gamma \in (0,1)$ is a properly chosen parameter. The sizes of the individual jumps $\delta_i,\ i = 1, \ldots, n$ are randomly generated using the MATLAB command `rand` and normalized so that

$$\int_a^b d\omega(x) = \sum_{i=1}^{n} \delta_i = 1\,.$$

We construct the related "perturbed" distribution function $\tilde{\omega}(x)$ to have two points of increase for each single point of increase of $\omega(x)$. Given a positive perturbation parameter ζ, where $\zeta \ll \lambda_1$ and $\zeta \ll \lambda_2 - \lambda_1$, we replace each point of increase λ_i of ω by two close points $\tilde{\lambda}_{2i-1} \equiv \lambda_i - \zeta$ and $\tilde{\lambda}_{2i} \equiv \lambda_i + \zeta$. We proportion the jumps $\tilde{\delta}_{2i-1}$ and $\tilde{\delta}_{2i}$ randomly (again using the MATLAB function `rand`), scaling so that $\tilde{\delta}_{2i-1} + \tilde{\delta}_{2i} = \delta_i$. For a small ζ the distribution functions ω and $\tilde{\omega}$ are close to each other.

We consider a smooth function $f(x) = x^{-1}$ and demonstrate that the difference between the Gauss-Christoffel quadrature estimates of the same degree for I_ω and $I_{\tilde{\omega}}$ can for some values of k become much larger than the difference between the integrals themselves.

In our experiment we take $\lambda_1 = 0.1$, $\lambda_n = 100$, $a = \lambda_1 - 10^{-5}$, $b = \lambda_n + 10^{-5}$, $n = 24$, $\gamma = 0.55$, and $\zeta = 10^{-8}$. The Jacobi matrices containing the recurrence coefficients of the corresponding orthogonal polynomials were computed from the spectral data using the algorithm of Gragg and Harrod implemented in the MATLAB routine `rkpw.m`; see [29,23,24].[1] The Gauss-Christoffel quadrature nodes and weights were computed as the eigenvalues and the squared first components of the corresponding normalized eigenvectors of the Jacobi matrices using the MATLAB routine `gauss.m`, see [23,24].[2]

In this first example, the Jacobi matrices could also be computed via the double-reorthogonalized Lanczos process (conjugate gradient algorithm) applied to the diagonal matrix $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with the starting vector $v_1 = [(\delta_1)^{\frac{1}{2}}, \ldots, (\delta_n)^{\frac{1}{2}}]^T$; see [54,29]. Similarly, one could use the Lanczos process (CG algorithm) on $\tilde{A} = \mathrm{diag}(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{2n})$ and $\tilde{v}_1 = [(\tilde{\delta}_1)^{\frac{1}{2}}, \ldots, (\tilde{\delta}_{2n})^{\frac{1}{2}}]^T$ to compute the perturbed nodes and weights. In this way the close relationship between the Gauss-Christoffel quadrature and the Lanczos process (conjugate gradient method) could be exploited. For small values of n the computational cost is negligible and the cost of reorthogonalization, considered in [29, p. 325], does not play a role. In the MATLAB routine `pftoqd.m` we have also implemented the algorithm by Laurie which requires no subtractions, see [38]. We emphasize that the same sensitivity phenomenon can

[1] Please note that in [23,24] the same implementation is called `lanczos.m`. Since that might cause a confusion with the implementation of the Lanczos process, we use the original name from [29, p. 328].

[2] An interested reader can find all m-files used for generating our figures, including the extended precision implementations, at `http://www.cs.cas.cz/mweb`, section "Applications".

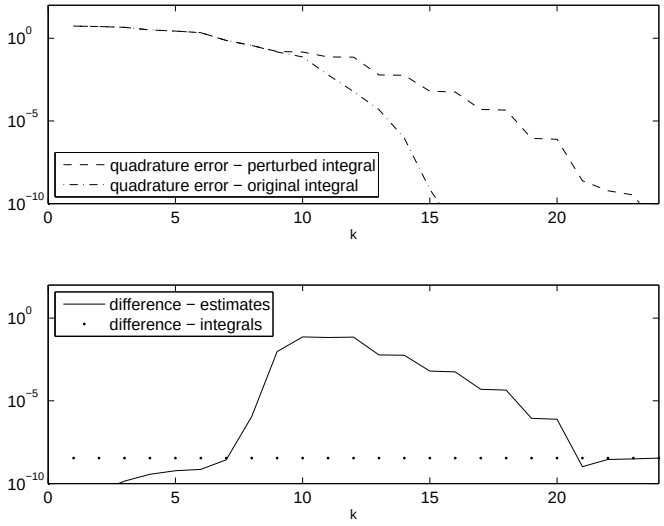

Fig. 1 Sensitivity of the Gauss-Christoffel quadrature for distribution functions with finite points of increase, $\zeta = 10^{-8}$. The top graph shows the error of the Gauss-Christoffel quadrature approximation for $f(x) = x^{-1}$ corresponding to the original stepwise distribution function ω (dash-dotted line) and to its perturbation $\tilde{\omega}$ with doubled points of increase (dashed line). The bottom graph displays the absolute value of difference in the estimates (solid line) and the difference between the approximated integrals (dots).

be observed, with differences which are here insignificant, using various computations of the recurrence coefficients from the spectral data.[3]

In the top of Fig. 1 we plot the error of the Gauss-Christoffel quadrature approximations $|E^k_\omega| \equiv |I_\omega - I^k_\omega|$ (dash-dotted line) and $|E^k_{\tilde{\omega}}| \equiv |I_{\tilde{\omega}} - I^k_{\tilde{\omega}}|$ (dashed line), and in the bottom we plot the difference between the Gauss-Christoffel approximations $|I^k_{\tilde{\omega}} - I^k_\omega|$ (solid line) and the difference between the approximated integrals $|\Delta| \equiv |I_\omega - I_{\tilde{\omega}}| \approx 3.443 \times 10^{-9}$ (dots). (Both $I_\omega \approx 5.50658692032301$ and $I_{\tilde{\omega}}$ were computed as finite sums of positive numbers to a relative accuracy close to machine precision). For $k \geq 8$ the Gauss-Christoffel approximations of the integrals $I_{\tilde{\omega}}$ and I_ω start to differ very dramatically, and the size of that difference exceeds 10^{-1} for $k = 10$. After that it is approximately equal to the error $|I_{\tilde{\omega}} - I^k_{\tilde{\omega}}|$ until that quantity drops below the size of the difference between the approximated integrals for $k = 21$.

This dramatic change in the estimates of the integral can be linked to a corresponding sensitivity in the orthogonal polynomials. Though the distribution functions ω and $\tilde{\omega}$ seem very close, the corresponding systems of orthogonal polynomials are quite different. This is illustrated in Fig. 2, which

[3] That has been confirmed independently by Dirk Laurie, who computed, with the data from the motivating example, the Jacobi matrices, nodes and weights of the quadrature to full 16 digits of accuracy using his software package [37]. Other valuable independent experiments were performed by Jarda Kautsky [33].

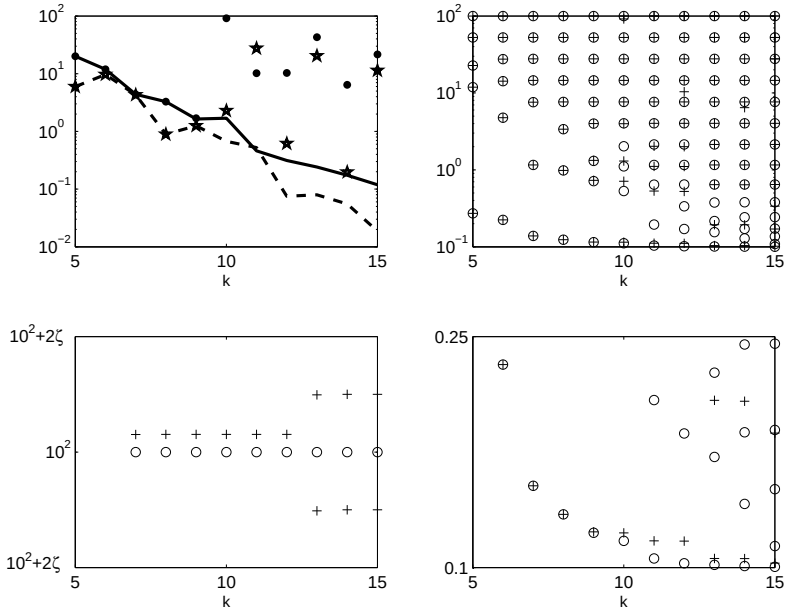

Fig. 2 Top left: Diagonal and off-diagonal entries of the Jacobi matrices corresponding to the distribution functions ω (solid line and dashed line, respectively) and $\tilde{\omega}$ (dots and stars, respectively). The other plots depict the quadrature nodes corresponding to the distribution function ω (circles) and $\tilde{\omega}$ (plusses) vs. the number of nodes k in the quadrature. Top right : all nodes. Bottom left: nodes near λ_n. Bottom right: nodes near λ_1.

shows the entries of the Jacobi matrices and the quadrature nodes (the zeros of the corresponding orthogonal polynomials) for ω, $\tilde{\omega}$ for iterations $k = 5, \ldots, 15$. In the top left part the diagonal entries of the Jacobi matrices for ω, $\tilde{\omega}$ are plotted by the solid line and by dots, respectively. Similarly, the off-diagonal elements are plotted by the dashed line and by stars, respectively. Up to $k = 7$ the computed Jacobi matrices are very close with their difference close to the square root of the machine precision. For $k = 8, 9$ the difference grows very rapidly (though in the figure the entries are still graphically indistinguishable). The corresponding entries suddenly completely depart at $k = 10$. The same is true for some quadrature nodes. Up to $k = 8$ they are graphically indistinguishable. For $k = 9$ the nodes corresponding to ω (circles) and $\tilde{\omega}$ (plusses) close to λ_1 start to visually differ, and eventually there are many fewer nodes for $\tilde{\omega}$ near λ_1 than there are for ω. The missing nodes for $\tilde{\omega}$ can be found close to λ_n, where they lie in pairs near the nodes for ω. We can see that for $\tilde{\omega}$, $\tilde{\lambda}_{2n-1}$ and $\tilde{\lambda}_{2n}$ are approximated to full accuracy starting from $k = 13$. Results are similar for different values of ζ, providing that $\zeta \ll \lambda_1$, $\zeta \ll \lambda_2 - \lambda_1$.

This first example motivates our investigation. In this paper we ask when, as illustrated in Fig. 1 and Fig. 2, Gauss-Christoffel quadrature is sensitive to small perturbations of the distribution function, and under what conditions it

is guaranteed to be insensitive. Such conditions exist, which can be verified using the following second example. Construct the perturbed distribution function $\tilde{\omega}(x)$ for $\omega(x)$ given above by placing *a single* (positive) $\tilde{\lambda}_i$ randomly in the interval of size 2ζ centered at λ_i, $i = 1, \ldots, n$, with $\tilde{\delta}_i = \delta_i$. (Here we do not specify the position of $\tilde{\lambda}_1$ and $\tilde{\lambda}_n$ relative to the centers of the intervals λ_1 and λ_n; it can be arbitrary.) Then, in contrast to the results shown in Fig. 1, the difference between the Gauss-Christoffel quadrature estimates for $f(x) = x^{-1}$ seems for all k bounded by the size of the difference between the approximated integrals $|I_\omega - I_{\tilde{\omega}}|$, independently of the choice of $0 < \zeta < 0.1$.

We will see that similar phenomena can be observed for continuous and even analytic distribution functions: Gauss-Christoffel quadrature can be highly sensitive to *some* small changes of a given distribution function, and insensitive to others. Next we describe such situations and relate them to theoretical results in the literature.

3 Literature Review

As mentioned above, although the question on sensitivity of Gauss-Christoffel quadrature has not, to our knowledge, been addressed in the literature, some related problems have been thoroughly investigated. In this section we summarize what is known about the sensitivity of generating the coefficients of the three-term recurrence satisfied by polynomials orthogonal with respect to the integral (1) and then computing the quadrature nodes and weights from the recurrence coefficients. The richness of the mathematical roots of this field is evidenced in the fact that the same problems have been described independently in many different ways and analyzed using many different techniques in literature that has little cross-reference. It would be very useful to relate in detail all of the existing results, but in this section we give just a brief overview.

3.1 Sensitivity in Computation of the Recurrence Coefficients

Analytic expressions for the recurrence coefficients are explicitly known for some classical distribution functions and the corresponding orthogonal polynomials; see, e.g., [24, p. 217], [40, p. 203], [55]. In practical applications, though, an analytic knowledge of the recurrence coefficients is exceptional, and one has to calculate them. Gautschi [24] presented four techniques. Using our terminology these are:

T1. A modified Chebyshev algorithm.
T2. Discretization of the distribution function.
T3. Computation of the recurrence coefficients for the discrete Riemann-Stieltjes integral.
T4. Computation of the recurrence coefficients for one distribution function from known coefficients for another distribution function.

The technique **T4** is not generally applicable, restricted to the case in which the original distribution function is multiplied by a rational nonnegative function [24, Section 2.5], [23, Section 2.4]. The problem of changes in orthogonal polynomials with respect to certain classes of modifications to the distribution function has been studied in many papers; see, e.g. [58,34], and [23, Section 2.4]. For a description of an old general result attributed to Markov concerning the dependence of the zeros of orthogonal polynomials on the parameter in the distribution function we refer to [57, Section 6.12, pp. 111-112]. Though such results are somewhat related to the problem of sensitivity of the Gauss-Christoffel quadrature, they are either of restricted applicability or merely qualitative. They do not lead to a general perturbation theory.

The modified Chebyshev algorithm **T1** represents an example of a more general approach based on knowledge of the recursion coefficients for some classical orthogonal polynomials determined by an auxiliary distribution function [24, Section 2.2]. Assume that the modified moments of the chosen (auxiliary) orthogonal polynomials with respect to the original distribution function can be determined *accurately*. From these moments and the *known* recurrence coefficients of the auxiliary polynomials, the modified Chebyshev algorithm determines the unknown recurrence coefficients of the desired orthogonal polynomials. The difficulties are the possibly large computational cost (not important in the context of our paper) and the possible inaccuracy in the computed results. The last difficulty has been thoroughly studied by Gautschi; see [15–17,19], and Section 2.1 of the book [23]. Subsection 2.1.3 defines the following maps:

$\mathbf{K}_k$: the map from the modified moments to the recurrence coefficients;
$\mathbf{G}_k$: the map from the modified moments to the nodes and weights of the computed quadrature;
$\mathbf{H}_k$: the map from the nodes and weights of the computed quadrature to the recurrence coefficients.

Then $\mathbf{K}_k$ can be represented as a composition of the other two maps,

$$\mathbf{K}_k = \mathbf{H}_k \circ \mathbf{G}_k \,.$$

The *condition numbers* attributed to $\mathbf{G}_k$ and $\mathbf{K}_k$ were studied in [23, Sections 2.1.4, 2.1.5 and 2.1.6, pp. 59-75]. If monomials are used as the auxiliary polynomials, the modified moments reduce to ordinary moments and the maps $\mathbf{G}_k$ and $\mathbf{K}_k$ are notoriously ill-conditioned. Even for a good choice of the auxiliary polynomials (such as the Chebyshev polynomials) and modified moments the situation is not simple. There are distribution functions for which the condition numbers are small, but there are other distribution functions for which the condition numbers grow exponentially with the number of nodes k. Moreover, the assumption that the modified moments are known accurately is difficult to satisfy; see [40, Section 3.2].

The general question of how to choose the auxiliary distribution function was analyzed by Beckermann and Bourreau in the remarkable paper [1]. They showed, among other results, that if the original and auxiliary distribution

functions have different supports, i.e. the sets of all points of their increase, [23, p. 3], then the condition number of $\mathbf{K}_k$ grows exponentially with k; see [1, Theorem 11, p. 93]. The authors further conjectured on the same page that the condition number of $\mathbf{K}_k$ is linked with the condition numbers of the matrices of modified and mixed moments.

The map $\mathbf{H}_k$ is said to be generally well-conditioned in [23, p. 59], though *numerically stable computation* of the entries of the Jacobi matrix from the quadrature nodes and weights is not easy; see [23, §3.1.1, pp. 154-155, §3.5, pp. 253-254 and Notes to §1.3, p. 50] with references to [38, Theorem on p. 168], [39], [40] and [1, Theorem 1 and Corollary 8]. (See also the last two paragraphs of this section, which explain in detail history of that problem and the fact that the algorithmic construction of Laurie in [38] also gives the perturbation result.) Further results on the condition numbers of the map $\mathbf{H}_k$ (and also of its inverse $\mathbf{H}_k^{-1}$) can be found in [1, relation (7), Section 2 and Appendix], see also [14, Section 4, pp. 190-193]. (We address the map $\mathbf{H}_k^{-1}$, in particular sensitivity of the nodes and weights and their computation from the entries of the Jacobi matrix, in Section 3.2). The approach from [14] is based on a remarkable result by Nevai on modification of the recurrence coefficients when adding a single point of increase to the given distribution function; see [45, Section 7, Lemma 15, p. 131], [14, Section 3, Lemma 1, p. 187]. For an instructive algebraic description and application of the same idea we refer to [12] and [13]. It is interesting that *essentially the same problem* of sensitivity of the entries in the Jacobi matrix to small perturbations of the nodes and weights of the corresponding distribution function (i.e. the eigenvalues and the first components of the normalized eigenvectors respectively) has recently been studied in a different way (independently of the results mentioned above) in [11]; see also [44] and the earlier paper [61]. A related more general problem of sensitivity of the Lanczos reduction has been thoroughly investigated in [47], see also [4,35].

The maps $\mathbf{K}_k, \mathbf{H}_k$ and $\mathbf{G}_k$ are interesting to study. However, as we will see in Section 5, they do not represent a relevant tool for investigation of sensitivity of Gauss-Christoffel quadrature. Their detailed discussion has been included here in order to explain the differences between the sensitivity problems studied previously and the sensitivity question posed and investigated in this paper.

The techniques **T2** and **T3** couple into one approach. The basic idea behind the discretization methods (see [24, Section 2.4], [23, Section 2.2, p. 90]) is an approximation of the given distribution function by a suitable *discrete* distribution function, computation of the recurrence coefficients for the discrete distribution function, and approximation of the desired recurrence coefficients by the computed (discrete) ones. Gautschi identified in [23, p. 90] two important issues which must be considered: the appropriate choice of discretizations and convergence of the discrete orthogonal polynomials (recurrence coefficients) to the desired ones. Both issues are tightly related. In the simple case when the original distribution function is composed of several components for which the analytic formulas for the orthogonal polynomials (Legendre, Chebyshev, ...) are known, discretization by a suitable combination of the N-point Gauss-type quadratures (Gauss-Legendre,

Gauss-Chebyshev, ...) for a sufficiently large $N \gg k$ gives the result. Assuming exact arithmetic, the first $N-1$ polynomials orthogonal with respect to the original distribution function are then also orthogonal with respect to the discretized distribution function, and the desired recurrence coefficients are determined *accurately*; see [24, p. 222]. Practical cases can be much more complicated, and finding an appropriate discretization is a rather involved procedure [23, Section 2.2.4].

There is one additional very important issue not mentioned in [23,24]. Convergence $N \to \infty$ describes the limiting case. In order to evaluate the accuracy of the methods based on discretization, one must be able to estimate the discretization error for a *finite* N. In other words, one must investigate *how fast* the discrete orthogonal polynomials converge to the desired ones, or, in a more complex way, *sensitivity* of the Gauss-Christoffel quadrature under small perturbations of the original distribution function. It seems that sensitivity is indeed a fundamental issue which cannot be omitted from consideration. If the Gauss-Christoffel quadrature is sensitive to small perturbations of the distribution function, then the computation based on the discretization may in general fail even if the discrete orthogonal polynomials, and, subsequently, the nodes and weights of the discrete quadrature, are determined accurately. A particular discretization procedure is not justified without proving that the results of the Gauss-Christoffel quadrature are insensitive with respect to the perturbation of the original distribution function represented by its discretization.

Finally, we discuss *computation* of the recurrence coefficients for the *discrete* Riemann-Stieltjes integral. This is an inverse problem: given nodes and weights of the N-point discrete Gauss-Christoffel quadrature formula, compute the entries of the corresponding Jacobi matrix.[4] In order to find the approximation to the desired k-point Gauss-Christoffel quadrature, we actually do not need the whole N by N Jacobi matrix, so we stop when we obtain its k by k left principal submatrix, $k \ll N$. In the classical language of orthogonal polynomials, the problem is solved by the discrete Stieltjes process [23, Section 2.2.3.1, p. 95]. In the language of numerical linear algebra, the Stieltjes process (implemented with modified Gram-Schmidt orthogonalization and normalization of the orthogonal polynomials) is equivalent to the Lanczos algorithm (see, e.g., [29, p. 322]), which is numerically unstable. This fact has been noted in the orthogonal polynomial literature (see, e.g., [19,20], [14, Section 2]), and reorthogonalization has been rejected as too costly [29, p. 325]. When k is small, however, the cost of reorthogonalization is negligible. Moreover, the analysis of the Lanczos algorithm behavior in finite precision arithmetic by Paige, Parlett, Scott, Simon, Greenbaum and others (reviewed, for example, in [42]) is almost unknown in the literature of orthogonal polynomial community, despite some notable work [28,3,34,2,26, 22] which emphasizes the interplay between the classical polynomial and vector algebraic formulations. The analysis can supply, at least, very convincing examples for illustrating and testing numerical instabilities.

[4] Note that the inverse problem corresponds in the literature to the map $\mathbf{H}_k$, not to $\mathbf{H}_k^{-1}$.

In order to overcome the numerical instability of the Lanczos algorithm, Gragg and Harrod suggested in their beautiful paper [29] a new algorithm based on ideas of Rutishauser. For an interesting experimental comparison, see [50, Section 2]. An alternative approach, based on the above mentioned results of Nevai [45], along with an experimental comparison, can be found in [14]. From numerical results Gragg and Harrod spotted a curious phenomenon: close nodes and weights can give two very different $k \times k$ Jacobi matrices. They concluded that the problem of reconstructing a Jacobi matrix from the weights and nodes is ill-conditioned [29, p. 330 and p. 332]. This conclusion has been examined by Laurie [38], who pointed out that the negative statement is linked to the use of the max-norm for vectors. He suggested instead measuring the perturbation of the weights in the componentwise relative sense [38, p. 179], [40, Section 6]. The main part of [38] is devoted to the constructive proof of the following statement [38, Theorem on p. 168]: given the weights and the $N-1$ positive differences between the consecutive nodes, the main diagonal entries of the corresponding Jacobi matrix (shifted by the smallest node) and the off-diagonal entries can be computed in $\frac{9}{2}N^2 + O(N)$ arithmetic operations, all of which can involve only addition, multiplication and division of positive numbers. Consequently, in finite precision arithmetic they can be computed to a *relative accuracy* no worse than $\frac{9}{2}N^2\varepsilon + O(N\varepsilon)$, where ε denotes machine precision. This result bounds also the conditioning of the problem. If the weights and the $N-1$ positive differences between the consecutive nodes are perturbed, with the size of the relative perturbations of the individual entries bounded by some small ϵ, then such perturbation can cause a relative change of the individual entries of the shifted main diagonal and of the individual off-diagonal entries of the Jacobi matrix not larger than $\frac{9}{2}N^2\epsilon + O(N\epsilon)$. The resulting algorithm combines ideas from earlier works from approximation theory, orthogonal polynomials, and numerical linear algebra.

3.2 Sensitivity and Computing of the Quadrature Nodes and Weights

Computing the quadrature nodes and weights is of great interest on its own. If the recurrence coefficients are used to construct a symmetric tridiagonal matrix with positive subdiagonals (Jacobi matrix), then, as mentioned above, the quadrature nodes are the eigenvalues and the weights are the first components of the normalized eigenvectors; see, e.g., [23, Section 3.1.1.1, pp. 152-154; Section 3.5, pp. 253-254]. In some special cases such as Gauss-Legendre quadrature, it is useful to consider also different ways of computing the quadrature nodes and weights; see [56]. It should be noted, however, that the comparison given in [56] does not refer to the recent developments in eigensolvers for Jacobi matrices recalled below. In most cases, computing the quadrature nodes and weights reduces to computing eigenvalues and the first components of eigenvectors of Jacobi matrices.

It is well known that two Jacobi matrices that are close to each other also have close eigenvalues and eigenvectors in the *absolute sense*, where the closeness is measured by the absolute values of the differences between the

corresponding individual eigenvalues and the corresponding individual eigenvectors; see, e.g., [27, Chapter 8], [34, p. 454], [11, p. 104], and [1, relation (7) and Appendix] mentioned above. For eigenvectors, the proportionality constant depends on the relative gaps between the eigenvalues of the unperturbed matrix. However, two close Jacobi matrices do not necessarily have eigenvalues that are close in a *relative* sense. A small perturbation of the entries of the Jacobi matrix can cause a large *relative* change in the eigenvalues and the eigenvector entries; see [8, pp. 71-72] and [41]. It is worth noting that Kahan has shown that small relative changes in the entries of the Cholesky factors of a *positive definite Jacobi matrix* do cause small relative changes in the eigenvalues of the Jacobi matrix [48, p. 123]; see also [7]. The thesis [8] gives also a comparison of different numerically stable algorithms for computing eigenvalues and eigenvectors of Jacobi matrices; see also the survey and comments in [40, Section 2], and the recent work [9,10,30,60]. We can conclude that the computation and perturbation theory of quadrature nodes and weights from the recurrence coefficients is well understood. The main difficulty in perturbation analysis and in computation of the Gauss-Christoffel quadrature lies in generating the recurrence coefficients.

Given this vast literature and our motivating example from the previous section, we focus our attention on the sensitivity of the quadrature formulas to changes in the distribution function.

3.3 Application to motivating examples: when larger support matters

The main difference between the first example at the beginning and the second example at the end of Section 2 consists in whether or not the number of points of increase (i.e. the 'size' of the support) is changed when ω is perturbed to form $\tilde{\omega}$. We will show that if there is no change in the number of points of increase, then a result by Laurie [38] explains the observed insensitivity of Gauss-quadrature for small enough perturbations.

Suppose we perturb the (discrete) ω of Section 2, resulting in $\tilde{\omega}$ with the *same number* of points of increase. Then, by Laurie's result, the corresponding shifted Jacobi matrices are close to each other in the componentwise *relative* sense. Using a classical perturbation result for eigensystems of symmetric matrices, the resulting Gauss-Christoffel quadrature nodes and weights for ω and $\tilde{\omega}$ must also be close to each other, with individual differences proportional to the perturbation parameter ζ. Consequently, for the (smooth and monotonic) function $f(x) = x^{-1}$ with ζ sufficiently small, the difference between the quadrature estimates must be proportional to the difference between the approximated integrals $|I_\omega - I_{\tilde{\omega}}|$.

There are two limitations of this argument. First, it does not apply to the first motivating example, since Laurie's result cannot be applied when the number of points of increase changes. Second, it does not apply to the second motivating example either, since the value of $\zeta = 10^{-8}$ was chosen too large. It does, however, provide *quantitative* sensitivity results for smaller ζ, or when ω and $\tilde{\omega}$ coincide at all points of increase, except for, say, λ_n, which

is well separated from $\lambda_1, \ldots, \lambda_{n-1}$. We next prove a result that does predict the difference in behavior of our two examples.

4 Quadrature differences in terms of approximation error

We present a slight generalization of a result found in the classic textbook of Isaacson and Keller [32] in Theorem 3 (p. 329) and in the second line of the identity (6) on p. 334.

The standard approach to Gauss quadrature of the Riemann integral and to Gauss-Christoffel quadrature of the Riemann-Stieltjes integral is based on Hermite interpolation and is attributed to Markov; see, e.g. [18, p. 82]. Here we take advantage of results based on Lagrange interpolation. This allows us to retain k free parameters in the remainder term for the kth order quadrature, which will later prove convenient in evaluation of the quadrature differences. In our exposition we follow the presentation of Gauss-Christoffel quadrature given by Lanczos in [36, Chapter VI, §10], cf. also [21, Theorem 3.2.1].

Choose k distinct points $x_1, \ldots, x_k$ inside the interval $[a, b]$, and let $q_k(x) = (x - x_1) \ldots (x - x_k)$. Then the Lagrange polynomial interpolating $f(x)$ at the points $x_1, \ldots, x_k$ can be written as

$$\mathcal{L}_k(x) = \sum_{j=1}^{k} f(x_j) \frac{q_k(x)}{q_k'(x_j)(x - x_j)},$$

and

$$f(x) = \mathcal{L}_k(x) + q_k(x) f[x_1, \ldots, x_k, x]\,,$$

where $f[x_1, \ldots, x_k, x]$ is the kth divided difference of f with respect to $x_1, \ldots, x_k, x$, see e.g. [32, Section 6.1]. We can derive a corresponding interpolatory quadrature formula

$$\int_a^b f(x)\, d\omega(x) = \sum_{j=1}^{k} \vartheta_j f(x_j) + \int_a^b q_k(x) f[x_1, \ldots, x_k, x]\, d\omega(x)\,, \qquad (3)$$

where the last term represents the error and

$$\vartheta_j = \frac{1}{q_k'(x_j)} \int_a^b \frac{q_k(x)}{(x - x_j)}\, d\omega(x)\,, \quad j = 1, \ldots, k\,. \qquad (4)$$

Up to now $x_1, \ldots, x_k$ were arbitrary distinct nodes inside $[a, b]$. The beauty of the Gauss-Christoffel quadrature is in setting the interpolatory nodes equal to the roots of the kth orthogonal polynomial corresponding to $\omega(x)$. Then we can consider k additional distinct nodes inside $[a, b]$ *which we need not even know* and show that the interpolatory quadrature on k nodes is as accurate as if $2k$ nodes had been used. This elegant consequence is summarized in the following theorem.

Theorem 1 *Consider a nondecreasing function $\omega(x)$ on a finite interval $[a,b]$. Let $p_k(x) = (x-t_1)\dots(x-t_k)$ be the kth monic orthogonal polynomial with respect to the inner product defined by the Riemann-Stieltjes integral on the interval $[a,b]$ with the distribution function $\omega(x)$. Choose k arbitrary distinct points $\mu_1,\dots,\mu_k$ in $[a,b]$. Let*

$$I_\omega = \int_a^b f(x)\,d\omega(x)\,, \tag{5}$$

where f'' is continuous on $[a,b]$, and let I_ω^k be the approximation to I_ω obtained from the k-point Gauss-Christoffel quadrature rule. Then for $m = 1,\dots,k$, the error of this approximation is given by

$$E_\omega^k(f) \equiv I_\omega - I_\omega^k \;=\; \int_a^b p_k(x) f[t_1,\dots,t_k,x]\,d\omega(x) \tag{6}$$

$$= \int_a^b p_k(x)(x-\mu_1)\dots(x-\mu_m) f[t_1,\dots,t_k,\mu_1,\dots,\mu_m,x]\,d\omega(x)\,, \tag{7}$$

where $f[t_1,\dots,t_k,\mu_1,\dots,\mu_m,x]$ is the $(k+m)$th divided difference of the function $f(x)$ with respect to the nodes $t_1,\dots,t_k,\mu_1,\dots,\mu_m,x$.

Proof Assume, for the moment, that the nodes $\mu_1,\dots,\mu_k$ are distinct from the nodes $t_1,\dots,t_k$. If we derive the quadrature rule (3)-(4) using $t_1,\dots,t_k$, then we have

$$\int_a^b f(x)d\omega(x) = \sum_{j=1}^k \vartheta_j f(t_j) + \int_a^b p_k(x) f[t_1,\dots,t_k,x]\,d\omega(x),$$

where the continuity of f' guarantees the finiteness of the divided difference as x varies. If $f(x)$ is a polynomial of degree at most $2k-1$, then $f[t_1,\dots,t_k,x]$ is a polynomial in x of degree at most $k-1$ and the rule is exact, since the orthogonality of $p_k(x)$ to all such polynomials makes the error term equal to zero. Consequently, the resulting interpolatory quadrature represents the Gauss-Christoffel quadrature. If we derive a quadrature rule using the points t_j plus the new nodes μ_i, then for $m = 1,\dots,k$,

$$\int_a^b f(x)\,d\omega(x) \;=\; \sum_{j=1}^k \hat{\vartheta} f(t_j) + \sum_{i=1}^m \hat{\xi}_i f(\mu_i)$$

$$+\; \int_a^b p_k(x)\,(x-\mu_1)\dots(x-\mu_m) f[t_1,\dots,t_k,\mu_1,\dots,\mu_m,x]\,d\omega(x)\,.$$

We observe from (4) that for $i = 1,\dots,m$ the weight $\hat{\xi}_i$ of each additional node is proportional to

$$\int_a^b p_k(x) r_i(x)\,d\omega(x) = 0\,,$$

where $r_i(x) = (x-\mu_1)\dots(x-\mu_m)/(x-\mu_i)$ is a polynomial of degree at most $k-1$, and therefore the orthogonality of $p_k(x)$ to all such polynomials results in a zero weight. Consequently, the contribution of the additional nodes $\mu_1, \dots, \mu_m$ to the integration formula vanishes, i.e.

$$\sum_{i=1}^{m} \hat{\xi}_i f(\mu_i) = 0\,.$$

It follows from uniqueness of the Gauss-Christoffel quadrature rules that $\vartheta_j = \hat{\vartheta}_j$ and the statement is proved.

If some μ_i is equal to some t_j, then replacing the Lagrange interpolant by the Hermite interpolant (cf. [43, p. 175] and [32, p. 330]), and using the continuity of f'' finishes the proof in an analogous way.

For analytic functions $f(x)$ it is possible to express the error of the Gauss-Christoffel quadrature rule without using derivatives or divided differences. Letting $p_k(x)$ be as above, the function

$$\rho_k(z) = \int_a^b \frac{p_k(x)}{z-x}\, d\omega(x)$$

is analytic in the complex plane outside the interval $[a,b]$. Suppose that $f(z)$ is analytic in a simply connected domain containing $[a,b]$ in its interior, and let Γ be a simple closed positively oriented curve in that domain encircling $[a,b]$. Then

$$E_\omega^k(f) = \frac{1}{2\pi\sqrt{-1}} \int_\Gamma K_k(z) f(z)\, dz\,, \quad K_k(z) = \frac{\rho_k(z)}{p_k(z)}\,; \tag{8}$$

see [23, Theorem 2.48], [6, p. 303, relation (4.6.18)]. This identity has been applied to estimate the error and to study its decrease with k for some particular classes of distribution functions $\omega(x)$ [18,25,23], [6, Section 4.6]. The kernel $K_k(z)$ depends through $p_k(z)$ and $\rho_k(z)$ on the given distribution function $\omega(x)$. The question of sensitivity of $E_\omega^k(f)$ with respect to perturbations of the distribution function $\omega(x)$ is thus reduced to the question of sensitivity of $K_k(z)$, where z lies on a properly chosen curve Γ in the complex plane, with respect to small perturbations of $\omega(x)$.

An application of Theorem 1 gives the following important result, an expression for the difference between the Gauss-Christoffel quadrature approximations.

Theorem 2 *Let $p_k(x) = (x-x_1)\dots(x-x_k)$ be the kth orthogonal polynomial with respect to $d\omega$ on $[a,b]$, and let $\tilde{p}_k(x) = (x-\tilde{x}_1)\dots(x-\tilde{x}_k)$ be the kth orthogonal polynomial with respect to $d\tilde{\omega}$. Denote by $\hat{p}_s(x) = (x-\xi_1)\dots(x-\xi_s)$ the least common multiple of the polynomials $p_k(x)$ and $\tilde{p}_k(x)$. If f'' is continuous on $[a,b]$, then the difference between the approximation I_ω^k to I_ω*

and the approximation $I^k_{\tilde{\omega}}$ to $I_{\tilde{\omega}}$, obtained from the k-point Gauss-Christoffel quadrature rule, is bounded as

$$|I^k_\omega - I^k_{\tilde{\omega}}| \le \left| \int_a^b \hat{p}_s(x) f[\xi_1, \ldots, \xi_s, x]\, d\omega(x) \; - \; \int_a^b \hat{p}_s(x) f[\xi_1, \ldots, \xi_s, x]\, d\tilde{\omega}(x) \right| + \left| \int_a^b f(x)\, d\omega(x) \; - \; \int_a^b f(x)\, d\tilde{\omega}(x) \right| . \tag{9}$$

Proof Consider the difference between the two Gauss quadrature approximations:

$$I^k_\omega - I^k_{\tilde{\omega}} = I_\omega - E^k_\omega - (I_{\tilde{\omega}} - E^k_{\tilde{\omega}}) = (E^k_{\tilde{\omega}} - E^k_\omega) + (I_\omega - I_{\tilde{\omega}}) . \tag{10}$$

Let the polynomials $p_k(x)$ and $\tilde{p}_k(x)$ have $k - m$ common zeros, numbered so that $x_{m+1} = \tilde{x}_{m+1}, \ldots, x_k = \tilde{x}_k$. Let $s = k + m$ and use the last equality in Theorem 1 twice. For E^k_ω set the points $t_1, \ldots, t_k$ in the theorem to be the zeros of $p_k(x)$, and set the points $\mu_1, \ldots, \mu_m$ to be the first m zeros $\tilde{x}_1, \ldots, \tilde{x}_m$ of $\tilde{p}_k(x)$. For $E^k_{\tilde{\omega}}$, set the points $t_1, \ldots, t_k$ to be the zeros of $\tilde{p}_k(x)$, and set the points $\mu_1, \ldots, \mu_m$ to be the first m zeros $x_1, \ldots, x_m$ of $p_k(x)$. The statement will immediately follow.

Note that from (10) the difference between the Gauss-Christoffel quadrature approximations is of order of the difference between the integrals (*or smaller*) if and only if the first term in the bound (9) is of order of the second term or smaller. Please note that the integrands in the first term in the bound (9) are *identical*. This simplifies the situation in comparison with a possible use of the standard quadrature error formulas known from the literature, from which it seems very difficult to get insight into the sensitivity phenomenon.

We state an analogous result for the weighted Riemann integral with nonnegative weight function that is (for simplicity) continuous on the finite interval $[a, b]$. The continuity assumption is not essential but simplifies the exposition.

Corollary 1 *Let $w(x)$ and $\tilde{w}(x)$ be nonnegative and continuous functions on the finite interval $[a, b]$; let*

$$\omega(x) = \int_a^x w(t)\, dt, \quad \tilde{\omega}(x) = \int_a^x \tilde{w}(t)\, dt, \quad x \in [a, b]$$

be the corresponding distribution functions. Then the integrals I_ω and $I_{\tilde{\omega}}$ in (2) represent the weighted Riemann integrals. Using the notation and assumptions of Theorem 2,

$$|I^k_\omega - I^k_{\tilde{\omega}}| \le \left| \int_a^b \hat{p}_s(x) f[\xi_1, \ldots, \xi_s, x](w(x) - \tilde{w}(x))\, dx \right| + \left| \int_a^b f(x)(w(x) - \tilde{w}(x))\, dx \right| . \tag{11}$$

Proof The statement follows immediately as a special case of Theorem 2.

If $f(x)$ is analytic, we can get identities which do not contain divided differences. Using the kernel expression of the error (8),

$$|I_\omega^k - I_{\tilde{\omega}}^k| \le \frac{1}{2\pi} \left| \int_\Gamma (K_k(z) - \tilde{K}_k(z)) f(z) dz \right| + |I_\omega - I_{\tilde{\omega}}|$$
$$= \frac{1}{2\pi} \left| \int_\Gamma \frac{\rho_k(z)\tilde{p}_k(z) - \tilde{\rho}_k(z)p_k(z)}{p_k(z)\tilde{p}_k(z)} f(z)\, dz \right| + |I_\omega - I_{\tilde{\omega}}| \,. \quad (12)$$

5 Discussion and numerical illustrations

The previous section presents simple bounds for the size of the difference $|I_\omega^k - I_{\tilde{\omega}}^k|$ between the results of the k-point Gauss-Christoffel quadrature which immediately follow from the identity (10) and the quadrature error formulas. As shown in Theorem 2, the crucial first term on the right hand side of (10) represents a difference between two integrals with the same integrand $\hat{p}_s(x) f[\xi_1, \ldots, \xi_s, x]$ and different distribution functions ω and $\tilde{\omega}$. We will explain why for some distribution functions ω and nearby $\tilde{\omega}$, with f sufficiently smooth and uncorrelated with the difference $\omega - \tilde{\omega}$, this term must inevitably become large, while for slightly different nearby distribution functions the term remains small. We will start with a closer look at our motivating examples from Section 2.

5.1 Discrete distribution functions: Motivating example revisited.

First, for clarity of exposition, we simplify the motivating examples from Section 2: in both examples keep the first $n-1$ points of increase of $\tilde{\omega}(x)$ equal to $\lambda_1, \ldots, \lambda_{n-1}$, with the corresponding weights $\delta_1, \ldots, \delta_{n-1}$. Thus, in both examples, $\tilde{\omega}(x)$ differs from $\omega(x)$ only near λ_n. In the first example, λ_n is replaced by *two* points of increase $\tilde{\lambda}_n = \lambda_n - \zeta$ and $\tilde{\lambda}_{n+1} = \lambda_n + \zeta$, with the (positive) weights $\tilde{\delta}_n$ respectively $\tilde{\delta}_{n+1}$, $\tilde{\delta}_n + \tilde{\delta}_{n+1} = \delta_n$. In the second example λ_n is perturbed to $\tilde{\lambda}_n = \lambda_n + \zeta$ with $\tilde{\delta}_n = \delta_n$.

For $f(x) = x^{-1}$ we get $f[\xi_1, \ldots, \xi_s, x] = (-1)^s (x\,\xi_1 \ldots \xi_s)^{-1}$, which holds, by induction, for any $s \le 2k$. Therefore the integrand in the first part of the bound (9) for the k-point quadrature simplifies to

$$g^k(x) \equiv \hat{p}_s(x) f[\xi_1, \ldots, \xi_s, x] = \frac{\hat{p}_s(x)}{x\hat{p}_s(0)} = f(x)\, \frac{\hat{p}_s(x)}{\hat{p}_s(0)} \,, \quad (13)$$

where the last term represents a polynomial having value one at zero. Using (7) we find that

$$E_{\tilde{\omega}}^k - E_\omega^k \equiv h^k$$

where

$$h^k \equiv \tilde{\delta}_n \left(g^k(\tilde{\lambda}_n) - g^k(\lambda_n)\right) + \tilde{\delta}_{n+1} \left(g^k(\tilde{\lambda}_{n+1}) - g^k(\lambda_n)\right).$$

Therefore, using (10) we have

$$I_\omega^k - I_{\tilde{\omega}}^k = h^k + \Delta\,, \quad \Delta = I_\omega - I_{\tilde{\omega}}\,.$$

In the second example, the second term in h^k is nonexistent.

In the first example, $E_{\tilde{\omega}}^k - E_\omega^k = h^k$ corresponds to the replacement of the *single* λ_n by *two* nearby points $\tilde{\lambda}_n$ and $\tilde{\lambda}_{n+1}$. With the given distribution functions ω and $\tilde{\omega}$ and for some values of $k \ll n$, the term h^k becomes much larger in magnitude than $|\Delta|$.

For small k, the Gauss-Christoffel quadrature approximation $I_{\tilde{\omega}}^k$ does not recognize $\tilde{\lambda}_n$ and $\tilde{\lambda}_{n+1}$ as two distinct points, and h^k is small. For larger k, λ_n becomes closely approximated by the largest node from the Gauss-Christoffel quadrature approximation I_ω^k of I_ω, and $g^k(\lambda_n)$ becomes very small. At the same time, $\tilde{\lambda}_n$ and $\tilde{\lambda}_{n+1}$ are approximated by a *single* quadrature node from $I_{\tilde{\omega}}^k$, placed in between them. Then $g^k(x)$ has in between $\tilde{\lambda}_n$ and $\tilde{\lambda}_{n+1}$ *two* roots, with one of them very close to λ_n. As k grows, this will soon become not enough to keep h^k small, since $g^k(\tilde{\lambda}_n)$ and $g^k(\tilde{\lambda}_{n+1})$ will grow in magnitude and are of the same sign, while $g^k(\lambda_n)$ is small due to the closeness of the quadrature node for I_ω^k to λ_n. Consequently, the differences $g^k(\tilde{\lambda}_n) - g^k(\lambda_n)$ and $g^k(\tilde{\lambda}_{n+1}) - g^k(\lambda_n)$ will also grow in magnitude and are of the same sign. Inevitably, for some value of k, $I_{\tilde{\omega}}^k$ has to place a *second* node, so that both $\tilde{\lambda}_n$ and $\tilde{\lambda}_{n+1}$ are sufficiently closely approximated and the size of the term h^k is kept under control. For that k, then, compared to the quadrature formula for I_ω^k, the quadrature formula for $I_{\tilde{\omega}}^k$ has one fewer node in some other part of the interval of integration. Therefore $|I_{\tilde{\omega}}^k - I_\omega^k|$ will suddenly become large. The missing node appears in the $(k+1)$st step of the Gauss-Christoffel quadrature approximation. Therefore, from then on, although $|I_{\tilde{\omega}}^k - I_\omega^k|$ may not be small, the difference shifted by one step, i.e., $|I_{\tilde{\omega}}^{k+1} - I_\omega^k|$, is small.

The situation is illustrated in Figure 3. In the top part the quadrature errors E_ω^k and $E_{\tilde{\omega}}^k$ are plotted by the solid and dashed line, respectively. They cannot be visually distinguished until $k = 9$. Starting from $k = 11$, the convergence of $I_{\tilde{\omega}}^k$ is *delayed* by one step in comparison to I_ω^k. Entries of the corresponding Jacobi matrices behave in an interesting way, which is illustrated by plotting the off-diagonal entries in the bottom part of the figure, with the solid line corresponding to I_ω^k and the dashed line to $I_{\tilde{\omega}}^k$. Until $k = 9$ the lines coincide. For $k = 10, 11$ the corresponding entries separate, and, starting from $k = 12$, the dashed line is just delayed (shifted to the right) by one step.

The dash-dotted lines in both parts of Figure 3 correspond to an additional example where λ_n is replaced by four close points $\tilde{\lambda}_n = \lambda_n - \zeta$, $\tilde{\lambda}_{n+1} = \lambda_n - \zeta/3$, $\tilde{\lambda}_{n+2} = \lambda_n + \zeta/3$, $\tilde{\lambda}_{n+3} = \lambda_n + \zeta$, while the new points share the original weight δ_n. The situation is fully analogous. Starting from $k = 14$ and $k = 18$, the convergence of $I_{\tilde{\omega}}^k$ is delayed by two and three steps, respectively.

The behavior of $g^k(x)$ is illustrated in Figure 4. For clarity we plot $\mathrm{sign}(g^k(x)) \log_{10}(1 + |g^k(x)|)$. The left part plots the behavior in the whole interval of integration for $k = 10$. The right part plots the behavior near λ_n

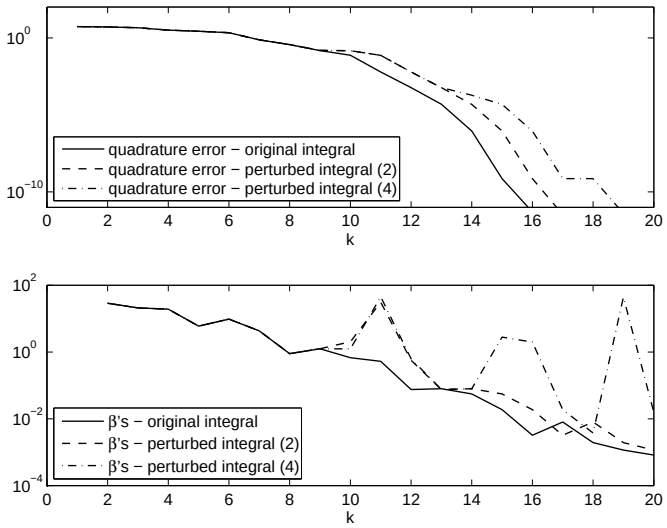

Fig. 3 Sensitivity of the Gauss-Christoffel quadrature for distribution functions with finite points of increase which differ only near λ_n, $\zeta = 10^{-8}$. The top graph shows the error of the Gauss-Christoffel quadrature approximation for $f(x) = x^{-1}$ corresponding to the original stepwise distribution function ω (solid line), to its perturbation $\tilde{\omega}$ with two points of increase near λ_n (dashed line), and to its perturbation $\tilde{\omega}$ with four points of increase near λ_n (dash-dotted line). The bottom graph displays the off-diagonal entries of the corresponding Jacobi matrices.

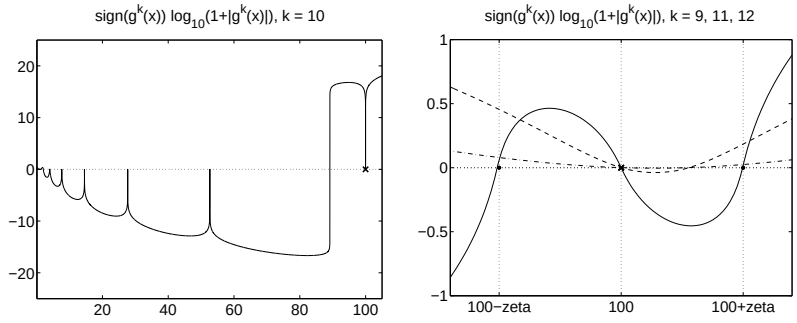

Fig. 4 The behavior of $g^k(x)$, see (13) (for better graphical view we plot $\mathrm{sign}(g^k(x)) \log_{10}(1+|g^k(x)|)$). The left part shows the behavior in the whole interval of integration for $k = 10$. The points of increase $\tilde{\lambda}_n$ and $\tilde{\lambda}_{n+1}$ are approximated by a single node of $I^k_{\tilde{\omega}}$ between them (the node close to 90 is still far away). The right part displays the behavior near λ_n for $k = 9$ (dash-dotted line), $k = 11$ (dashed line) and $k = 12$ (solid line). For $k = 12$ both nodes $\tilde{\lambda}_n$ and $\tilde{\lambda}_{n+1}$ are very closely approximated by the nodes of $I^k_{\tilde{\omega}}$.

for $k = 9, 11, 12$. For $k = 9$ the line is close to the horizontal axis. For $k = 11$ we can observe the increasing gradient of $g^k(x)$ ($\zeta = 10^{-8}$), and for $k = 12$ both $\tilde{\lambda}_n$ and $\tilde{\lambda}_{n+1}$ are closely approximated by quadrature nodes of $I^k_{\tilde{\omega}}$.

This phenomenon is closely related to the fact that the presence of close eigenvalues affects the rate of convergence of the conjugate gradient method; see the beautiful explanation given by van der Sluis and van der Vorst [51, 52]. Similarly, it is closely related to the convergence of the Rayleigh quotient in the power method and to the so-called 'misconvergence phenomenon' in the Lanczos method see [49,46]. In exact arithmetic in the presence of very close eigenvalues, a Ritz value in the Lanczos and the CG method initially converges to the cluster as fast as if the cluster were replaced by a single eigenvalue with the combined weight. Within a few further steps it converges very fast to one of the eigenvalues, with another Ritz value converging simultaneously to approximate the rest of the cluster. In the presence of more than two eigenvalues in a cluster, the story repeats until all eigenvalues in a cluster are approximated by individual Ritz values.

Now we consider the second modified example, where λ_n is perturbed to $\tilde{\lambda}_n = \lambda_n + \zeta$, $\tilde{\delta}_n = \delta_n$. Then $I^k_{\tilde{\omega}}$ converges to $I_{\tilde{\omega}}$ with the same speed as I^k_{ω} to I_{ω}, there is no delay, and the fact that $|E^k_{\omega} - E^k_{\tilde{\omega}}|$ is small can be *proved* using the result by Laurie [38]; see Section 3.3.

In the original motivating examples from Section 2 the situation is quite analogous, with the effects described on the simplified examples now taking place (for different values of k) near $\lambda_n, \lambda_{n-1}, \ldots$. A steep increase of $|I^k_{\omega} - I^k_{\tilde{\omega}}|$ significantly above $|\Delta|$ is well pronounced in the presence of well-separated rightmost points $\lambda_n, \lambda_{n-1}, \ldots$, because they are fast approximated to high accuracy by the quadrature nodes. The phenomenon is almost independent of the position of the eigenvalues within the individual clusters (here $0 < \zeta < 0.1$ in order to ensure $\tilde{\lambda}_1 > 0$); see a similar statement in [51, Section 6.7, point (d), p. 559]. When λ_n is well-separated, the phenomenon *must* take place even for very small ζ.

The sensitivity of the Gauss-Christoffel quadrature is a consequence of the fact that $\tilde{\omega}$ has *more points of increase* (here two) close to the single points of increase of ω. The Gauss-Christoffel quadrature is sensitive because the number of points $\{\tilde{\lambda}_1, \ldots, \tilde{\lambda}_m\}$ in the support of $\tilde{\omega}$ is *larger* than the number in the support $\{\lambda_1, \ldots, \lambda_n\}$ of ω. More precisely, $\tilde{\omega}$ has more points of increase in the area where the gradient of $g^k(x)$ becomes very large as k increases. The second example, with the same number of points of increase, shows that moving each point of increase slightly does not cause sensitivity if the number of points is kept the same.

5.2 A continuous analog of the motivating example.

Consider the analytic function $\Phi(x;\sigma) \equiv \sum_{i=1}^{n} \delta_i \varphi(x;\sigma,\lambda_i)$, where $0 < a < \lambda_1 < \cdots < \lambda_n < b$, $\delta_1, \ldots, \delta_n$ are as above and

$$\varphi(x;\sigma,t) \equiv \left[1 + e^{-\frac{x-t}{\sigma}}\right]^{-1} \tag{14}$$

is the strictly increasing sigmoid function with values between 0 and 1. Define the distribution function

$$\Omega(x;\sigma) \equiv c_0\,\Phi(x;\sigma)\,, \quad \int_a^b d\Omega(x;\sigma) = 1\,, \tag{15}$$

where c_0 is the normalization constant. Clearly, $\Omega(x;\sigma)$ approximates the step function from the motivating example:

$$\lim_{\sigma\to 0} \Omega(x;\sigma) = \omega(x)\,,$$

for all $a \le x \le b$ except for $x = \lambda_i, i = 1,\ldots,n$, and the value of the parameter σ determines how closely $\Omega(x;\sigma)$ approximates $\omega(x)$.

In order make our computations accurate, we use the following linearization of $\Omega(x;\sigma)$. Divide the interval $[t-10\sigma, t+10\sigma]$ into $2m$ equal subintervals, with $m = 50$. Define $\hat{\varphi}(x;\sigma,t)$ to be the piecewise linear continuous function that interpolates $\Omega(x:\sigma)$ at the endpoints of the subintervals and is constant on $(-\infty, t-10\sigma]$ and $[t+10\sigma,\infty)$. Then, using $\hat{\Phi}(x;\sigma) \equiv \sum_{i=1}^n \delta_i \hat{\varphi}(x;\sigma,\lambda_i)\,,$ we obtain a linearized distribution function

$$\hat{\Omega}(x;\sigma) \equiv c_1\,\hat{\Phi}(x;\sigma)\,, \quad \int_a^b d\hat{\Omega}(x;\sigma) = 1, \tag{16}$$

with c_1 the normalization constant.

The Riemann-Stieltjes integral $I_{\hat{\Omega}}(x^{-1}) = \int_a^b x^{-1} d\hat{\Omega}(x;\sigma)$ can be computed analytically. The recurrence coefficients of the orthogonal polynomials were computed by the double-reorthogonalized Lanczos process, with the corresponding integrals computed numerically. Using the partitioning described above and the fact that $\hat{\Omega}(x;\sigma)$ is linear on each subinterval, we conveniently use on each subinterval the Gauss-Legendre quadrature of sufficient order, implemented in MATLAB by Laurie in the file `r_jacobi.m`; see [24, Section 2.1]. For determining the quadrature nodes and weights we then use the standard approach implemented in the file `gauss.m` by Gautschi [23, pp. 153-154], [24, Section 2.4]. We use $\sigma = 10^{-8}$ and $\sigma = 10^{-6}$, $a = \lambda_1 - 10^{-5} = 10^{-1} - 10^{-5}$ and $b = \lambda_n + 10^{-5} = 100 + 10^{-5}$. Results for the original distribution function $\hat{\Omega}(x;10^{-8})$ and its *perturbation* $\hat{\Omega}(x;10^{-6})$, analogous to Figures 1 and 2, are presented in Figures 5 and 6. We can observe the same phenomena as in the motivating example, and the explanation is analogous. Since now the distribution functions are continuous, *many* quadrature nodes are eventually placed close to the rightmost $\lambda_n, \lambda_{n-1}, \ldots$ for both $\sigma = 10^{-8}$ and $\sigma = 10^{-6}$.

We emphasize that the observed Gauss-Christoffel quadrature sensitivity is a consequence of the fact that the support of $\hat{\Omega}(x;10^{-6})$, which is the union of intervals of length 2×10^{-5} around the points λ_i, is *larger* than the corresponding support of $\hat{\Omega}(x;10^{-8})$. If the supports were different (with the difference of a similar scale as before) but of the *same size*, no sensitivity would occur. Indeed, computation confirms that if $\tilde{\Omega}(x;10^{-8})$ is a *perturbation* of the original distribution function $\hat{\Omega}(x;10^{-8})$ obtained by shifting

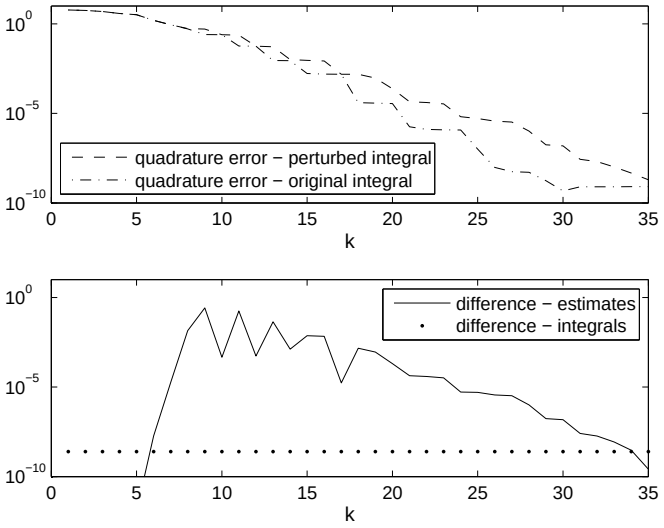

Fig. 5 Sensitivity of the Gauss-Christoffel quadrature for the continuous distribution function $\hat{\Omega}(x;\sigma)$. The top graph shows the error of the Gauss-Christoffel quadrature approximation for $f(x) = x^{-1}$ corresponding to the original distribution function with $\sigma = 10^{-8}$ (dash-dotted line) and to its perturbation with $\sigma = 10^{-6}$ (dashed line). The bottom graph displays the absolute value of difference in the estimates (solid line) and the difference between the approximated integrals (dots).

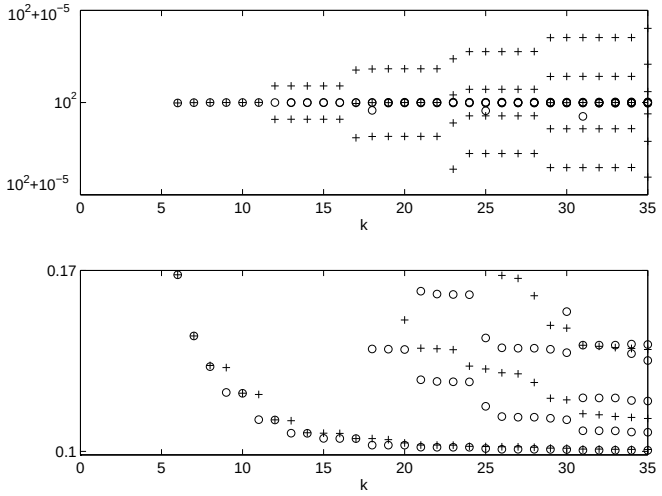

Fig. 6 Quadrature nodes corresponding to the distribution function $\hat{\Omega}(x;\sigma)$ with $\sigma = 10^{-8}$ (circles) and $\sigma = 10^{-6}$ (plusses) in two subintervals close to λ_n (top) and λ_1 (bottom). The horizontal axis is the number of nodes k in the quadrature.

the individual sigmoids randomly 10^{-6} to the left or right, with subsequent normalization, then the quadrature nodes and weights change proportionally to the shifts of the individual sigmoids. The size of the difference between the Gauss-Christoffel quadrature estimates for $f(x) = x^{-1}$ and $\hat{\Omega}(x; 10^{-8})$ and $\tilde{\Omega}(x; 10^{-8})$ remains below or close to the size of the difference between the estimated integrals. In short, in agreement with our discussion above, no sensitivity of the Gauss-Christoffel quadrature appears.

5.3 Discussion: Relationship to modified moments

We will explain that the sensitivity of the Gauss-Christoffel quadrature described above cannot be analyzed by investigation of modified moments. Our point is that the Gauss-Christoffel quadrature can be highly sensitive to some small changes of the original distribution function but insensitive to others, and this principal difference cannot be captured by the conditioning of the map $\mathbf{K}_k$ from the modified moments to the recurrence coefficients studied by Gautschi [23] and Beckermann and Bourreau [1]; see Section 3.1. In order to justify our claim, we will use the example with continuous distribution functions given above.

Using the previous notation, consider the original distribution function $\Omega_0(x) \equiv \hat{\Omega}(x; 10^{-8})$ and two *perturbations* $\Omega_1(x) \equiv \hat{\Omega}(x; 10^{-6})$, $\Omega_2(x) \equiv \tilde{\Omega}(x; 10^{-8})$. We will now consider $\Omega_1(x)$ and $\Omega_2(x)$ two different *auxiliary distribution functions* in the sense of the modified Chebyshev algorithm; see Section 3.1. We know that the Gauss-Christoffel quadrature is sensitive to change from Ω_0 to Ω_1, and insensitive to change from Ω_0 to Ω_2. We might intuitively expect that the sensitivity in the first case would be reflected by the ill-conditioning of the map $\mathbf{K}_k^{(1)}$, which corresponds to the original distribution function Ω_0 and the auxiliary distribution function Ω_1, and that the insensitivity is in the second case would perhaps be accompanied by well-conditioning of the map $\mathbf{K}_k^{(2)}$, which corresponds to the original distribution function Ω_0 and the auxiliary distribution function Ω_2. But this is not true. The support of Ω_0 is different from the supports of Ω_1 and Ω_2, and using [1, Theorem 11, p. 93], we find that *both* maps $\mathbf{K}_k^{(1)}$ and $\mathbf{K}_k^{(2)}$ are notoriously ill-conditioned.

In order to illustrate this numerically, we consider the conjecture [1, p. 93] that there is a link between the condition number of $\mathbf{K}_k$ and that of the matrix of mixed moments of the polynomials orthogonal with respect to the original and auxiliary distribution functions (where the mixed moments are computed using the original distribution function; see [1, the matrix of transmission coefficients $T_n(\sigma, s)$ on p. 93]. The mixed moments appear in the modified Chebyshev algorithm as intermediate quantities; see [23, p. 76, relation (2.1.101)]). With a reference to the habilitation thesis of Beckermann, it is argued that the condition number GM_k of the matrix of modified moments and the condition number MM_k of the matrix of mixed moments grow exponentially if the supports of the original and the auxiliary distribution functions do not coincide. This is illustrated in Figure 7. Here GM_k is plotted by the solid line, MM_k by the dashed line. The left part corresponds

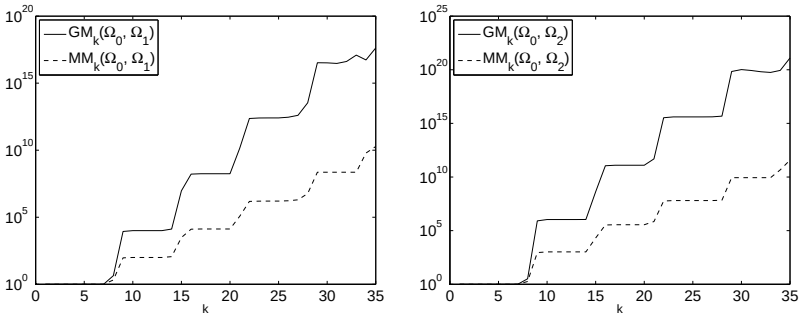

Fig. 7 Condition numbers of the matrix of the modified moments (GM_k, solid line) and of the matrix of mixed moments (MM_k, dashed line). The left graph corresponds to the distribution functions Ω_0 and Ω_1, and the right graph to the distribution functions Ω_0 and Ω_2.

to the distribution functions Ω_0 and Ω_1, the right part to Ω_0 and Ω_2. We can see the condition numbers GM_k and MM_k are growing essentially exponentially with k in *both* cases. (The staircase character of the plots is yet to be analyzed.) Using the conjecture in [1, p. 93], the fast growth of MM_k can be linked with the ill-conditioning of the maps $\mathbf{K}_k^{(1)}$ and $\mathbf{K}_k^{(2)}$.

In conclusion, the Gauss-Christoffel quadrature for a given distribution function can be *insensitive* to some perturbations despite the corresponding large MM_k and the corresponding ill-conditioning of the map $\mathbf{K}_k$.

5.4 Analytic distribution functions with different support.

The phenomena described above can also be observed with analytic distribution functions. We present experiments with distribution function $\Omega(x;\sigma)$; see (15). The recurrence coefficients of the corresponding orthogonal polynomials are again computed by the double reorthogonalized Lanczos process, where for the numerical computation of the required integrals we use the MATLAB adaptive Lobatto quadrature `quadl`. The quadrature nodes and weights are then determined as above using the code `gauss.m`. We set $a = 0.1$ and $b = 100.2$. In order to reduce numerical errors below a noticeable level we take $\lambda_1 = 0.3, \lambda_n = 100, n = 4, \gamma = 0.55$, and consider the original distribution function $\Omega(x;0.04)$ and its *perturbation* $\Omega(x;0.08)$. Figure 8 shows results of the Gauss-Christoffel quadrature estimates for $k = 1,\ldots,10$, $f(x) = x^{-1}$ (top) and $f(x) = 1+\sin(x)$ (bottom). The sensitivity of the Gauss-Christoffel quadrature is here less pronounced than before. Still it is observable.

5.5 Analytic distribution functions with the same support.

For slightly perturbed analytic functions with the same support the difference between the Gauss quadrature approximations $|I_{\tilde{\omega}}^k - I_\omega^k|$ is typically of

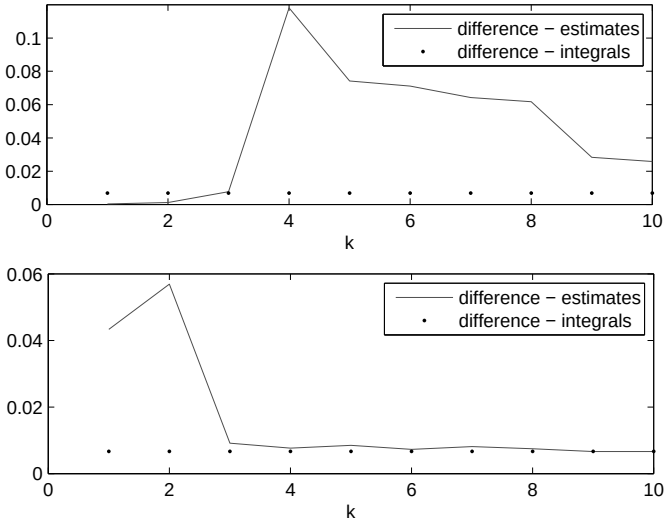

Fig. 8 Sensitivity of the Gauss-Christoffel quadrature for the analytic distribution function $\Omega(x;\sigma)$. The figure shows the absolute value of the difference in the quadrature estimates (solid line) and the difference between the approximated integrals (dots) for $f(x) = x^{-1}$ (top) and $f(x) = 1 + \sin(x)$ (bottom), corresponding to the original distribution function with $\sigma = 0.04$ and to its perturbation with $\sigma = 0.08$.

the order $|I_{\tilde{\omega}} - I_{\omega}|$. For small values of k, the errors of the corresponding estimates are much larger than the difference between them. Eventually the estimates must separate because they aim at approximating different integrals. The value k for which the two estimates separate is essentially determined by the difference between the approximated integrals. The roots of the corresponding orthogonal polynomials are typically very stable. We performed experiments, e.g., for the weight function $w(x) = \sqrt{x - x^2}$ for the shifted Chebyshev polynomials of the second kind, for the highly oscillatory weight function $w(x) = 1 + \cos(10\pi x)$, and for the Jacobi weight functions $w(x) = (1 - x)^{\alpha}(1 + x)^{\beta}$ with various values of the exponents and various perturbations.

We observed two characteristics in these experiments. First, the rate of decrease of the quadrature error was exponential, which can be explained using the Cauchy integrating kernels; see (8). Second, when perturbation of the distribution function preserves its support (here the whole interval), the quadrature is not sensitive. For an interesting example where the preservation of the support is linked with the analysis of the conditioning of the map $\mathbf{K}_k$ we refer to [1, Example 15, p. 96].

6 Conclusions

Literature about Gauss-Christoffel quadrature and about its computational aspects is extensive. This paper raises the following points which seem, however, new:

1. Gauss-Christoffel quadrature for a small number of quadrature nodes can be highly sensitive to small changes in the distribution function. In particular, the difference between the corresponding quadrature approximations (using the same number of quadrature nodes) can be many orders of magnitude larger than the difference between the integrals being approximated.
2. This sensitivity in Gauss-Christoffel quadrature can be observed for discontinuous, continuous, and even analytic distribution functions, and for analytic integrands uncorrelated with changes in the distribution functions and with no singularity close to the interval of integration.
3. The sensitivity of the Gauss-Christoffel quadrature illustrated in this paper is related to the difference in the *size* of the support of the original and of the perturbed distribution functions. For a discrete distribution function, the size is the number of points of increase, and for a continuous distribution function it is the length (measure) of the union of intervals containing points at which the distribution function increases. In general, different supports of the *same size* do not exhibit sensitivity in quadrature results.
4. The sensitivity of Gauss-Christoffel quadrature cannot be explained using existing analysis based on modified moments. In our examples, if the support of the original distribution function differs in size from the support of the auxiliary (perturbed) distribution function, then the matrices of both modified and mixed moments become highly ill-conditioned. The same is true if the supports are different but of the same size. But only in the case of different size of the supports are the recurrence coefficients (i.e., the entries of the Jacobi matrix) and the Gauss-Christoffel quadrature estimates highly sensitive to the perturbation.

Many open questions remain. We give several *examples* of sensitivity of the Gauss-Christoffel quadrature. It would certainly be of great interest to describe the *classes of problems* for which the Gauss-Christoffel quadrature is sensitive to small perturbations of the distribution function, and determine which of them are of practical importance. Application of these results to theory of the conjugate gradient and Lanczos methods in finite precision arithmetic will be considered in our future work. Another highly relevant question is how to measure differences between distribution functions.

Acknowledgements Numerous comments and insightful suggestions of the anonymous referees were very helpful in revising the original version of the manuscript. The authors are indebted to Dirk Laurie and Jarda Kautsky for their interest in this work, for their questions, thoughts and for performing independent numerical experiments.

References

1. Beckermann, B., Bourreau, E.: How to choose modified moments? J. Comput. Appl. Math. **98**(1), 81–98 (1998)
2. Boley, D., Golub, G.H.: A survey of matrix inverse eigenvalue problems. Inverse Problems **3**(4), 595–622 (1987)
3. de Boor, C., Golub, G.H.: The numerically stable reconstruction of a Jacobi matrix from spectral data. Linear Algebra Appl. **21**(3), 245–260 (1978)
4. Carpraux, J.F., Godunov, S.K., Kuznetsov, S.V.: Condition number of the Krylov bases and subspaces. Linear Algebra Appl. **248**, 137–160 (1996)
5. Dahlquist, G., Golub, G.H., Nash, S.G.: Bounds for the error in linear systems. In: Semi-infinite programming (Proc. Workshop, Bad Honnef, 1978), *Lecture Notes in Control and Information Sci.*, vol. 15, pp. 154–172. Springer, Berlin (1979)
6. Davis, P.J., Rabinowitz, P.: Methods of numerical integration, second edn. Computer Science and Applied Mathematics. Academic Press Inc., Orlando, FL (1984)
7. Demmel, J., Kahan, W.: Accurate singular values of bidiagonal matrices. SIAM J. Sci. Statist. Comput. **11**(5), 873–912 (1990)
8. Dhillon, I.S.: A new $O(n^2)$ algorithm for the symmetric tridiagonal eigenvalue/eigenvector problem. Ph.D. thesis, University of California, Berkeley, USA (1997)
9. Dhillon, I.S., Parlett, B.N.: Orthogonal eigenvectors and relative gaps. SIAM J. Matrix Anal. Appl. **25**(3), 858–899 (electronic) (2003)
10. Dhillon, I.S., Parlett, B.N.: Multiple representations to compute orthogonal eigenvectors of symmetric tridiagonal matrices. Linear Algebra Appl. **387**, 1–28 (2004)
11. Druskin, V., Borcea, L., Knizhnerman, L.: On the sensitivity of Lanczos recursions to the spectrum. Linear Algebra Appl. **396**, 103–125 (2005)
12. Elhay, S., Golub, G.H., Kautsky, J.: Updating and downdating of orthogonal polynomials with data fitting applications. SIAM J. Matrix Anal. Appl. **12**(2), 327–353 (1991)
13. Elhay, S., Golub, G.H., Kautsky, J.: Jacobi matrices for sums of weight functions. BIT **32**(1), 143–166 (1992)
14. Fischer, H.J.: On generating orthogonal polynomials for discrete measures. Z. Anal. Anwendungen **17**(1), 183–205 (1998)
15. Gautschi, W.: Construction of Gauss-Christoffel quadrature formulas. Math. Comp. **22**, 251–270 (1968)
16. Gautschi, W.: On the construction of Gaussian quadrature rules from modified moments. Math. Comp. **24**, 245–260 (1970)
17. Gautschi, W.: Questions on numerical condition related to polynomials. In: Recent Advances in Numerical Analysis (Madison, 1978), pp. 45–72. Academic Press, New York (1978)
18. Gautschi, W.: A survey of Gauss-Christoffel quadrature formulae. In: P.L. Butzer, F. Fehér (eds.) E. B. Christoffel: The Influence of His Work on Mathematics and Physics, pp. 72–147. Birkhäuser, Basel (1981)
19. Gautschi, W.: On generating orthogonal polynomials. SIAM J. Sci. Statist. Comput. **3**(3), 289–317 (1982)
20. Gautschi, W.: Is the recurrence relation for orthogonal polynomials always stable? BIT **33**(2), 277–284 (1993)
21. Gautschi, W.: Numerical analysis. Birkhäuser Boston Inc., Boston, MA (1997). An introduction
22. Gautschi, W.: The interplay between classical analysis and (numerical) linear algebra—a tribute to Gene H. Golub. Electron. Trans. Numer. Anal. **13**, 119–147 (electronic) (2002)
23. Gautschi, W.: Orthogonal polynomials: computation and approximation. Numerical Mathematics and Scientific Computation. Oxford University Press, New York (2004). Oxford Science Publications
24. Gautschi, W.: Orthogonal polynomials (in Matlab). J. Comput. Appl. Math. **178**(1-2), 215–234 (2005)

25. Gautschi, W., Varga, R.S.: Error bounds for Gaussian quadrature of analytic functions. SIAM J. Numer. Anal. **20**(6), 1170–1186 (1983)
26. Golub, G.H.: Matrix computation and the theory of moments. In: Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), pp. 1440–1448. Birkhäuser, Basel (1995)
27. Golub, G.H., Van Loan, C.F.: Matrix computations, *Johns Hopkins Series in the Mathematical Sciences*, vol. 3, second edn. Johns Hopkins University Press, Baltimore, MD (1989)
28. Golub, G.H., Welsch, J.H.: Calculation of Gauss quadrature rules. Math. Comp. 23 (1969), 221-230; addendum, ibid. **23**(106, loose microfiche suppl), A1–A10 (1969)
29. Gragg, W.B., Harrod, W.J.: The numerically stable reconstruction of Jacobi matrices from spectral data. Numer. Math. **44**(3), 317–335 (1984)
30. Grosser, B., Lang, B.: On symmetric eigenproblems induced by the bidiagonal SVD. SIAM J. Matrix Anal. Appl. **26**(3), 599–620 (electronic) (2005)
31. Hestenes, M.R., Stiefel, E.: Methods of conjugate gradients for solving linear systems. J. Research Nat. Bur. Standards **49**, 409–436 (1953) (1952)
32. Isaacson, E., Keller, H.B.: Analysis of numerical methods. John Wiley & Sons Inc., New York (1966)
33. Kautský, J.: Personal communication (October 2006)
34. Kautský, J., Golub, G.H.: On the calculation of Jacobi matrices. Linear Algebra Appl. **52/53**, 439–455 (1983)
35. Kuznetsov, S.V.: Perturbation bounds of the Krylov bases and associated Hessenberg forms. Linear Algebra Appl. **265**, 1–28 (1997)
36. Lanczos, C.: Applied Analysis. Prentice Hall, Inc., Inc., Englewood Cliffs, New Jersey (1956)
37. Laurie, D.: Personal communication (October 2006)
38. Laurie, D.P.: Accurate recovery of recursion coefficients from Gaussian quadrature formulas. J. Comput. Appl. Math. **112**(1-2), 165–180 (1999)
39. Laurie, D.P.: Questions related to Gaussian quadrature formulas and two-term recursions. In: Applications and computation of orthogonal polynomials (Oberwolfach, 1998), *Internat. Ser. Numer. Math.*, vol. 131, pp. 133–144. Birkhäuser, Basel (1999)
40. Laurie, D.P.: Computation of Gauss-type quadrature formulas. J. Comput. Appl. Math. **127**(1-2), 201–217 (2001). Numerical analysis 2000, Vol. V, Quadrature and orthogonal polynomials
41. Li, R.C.: Relative perturbation theory. III. More bounds on eigenvalue variation. Linear Algebra Appl. **266**, 337–345 (1997)
42. Meurant, G., Strakoš, Z.: The Lanczos and conjugate gradient algorithms in finite precision arithmetic. Acta Numerica **15**, 471–542 (2006)
43. Milne-Thomson, L.M.: The Calculus of Finite Differences. MacMillan & Co., New York (1960)
44. Natterer, F.: Discrete Gel'fand-Levitan theory, www page of the author (1999)
45. Nevai, P.G.: Orthogonal polynomials. Mem. Amer. Math. Soc. **18**(213), v+185 (1979)
46. O'Leary, D.P., Stewart, G.W., Vandergraft, J.S.: Estimating the largest eigenvalue of a positive definite matrix. Math. Comp. **33**(148), 1289–1292 (1979)
47. Paige, C.C., Van Dooren, P.: Sensitivity analysis of the Lanczos reduction. Numer. Linear Algebra Appl. **6**(1), 29–50 (1999). Czech-US Workshop in Iterative Methods and Parallel Computing, Part I (Milovy, 1997)
48. Parlett, B.N., Dhillon, I.S.: Relatively robust representations of symmetric tridiagonals. In: Proceedings of the International Workshop on Accurate Solution of Eigenvalue Problems (University Park, PA, 1998), vol. 309, pp. 121–151 (2000)
49. Parlett, B.N., Simon, H., Stringer, G.: On estimating the largest eigenvalue with the Lanczos algorithm. Maths. of Comp. **38**, 153–165 (1982)
50. Reichel, L.: Fast QR decomposition of Vandermonde-like matrices and polynomial least squares approximation. SIAM J. Matrix Anal. Appl. **12**(3), 552–564 (1991)

51. van der Sluis, A., van der Vorst, H.: The rate of convergence of conjugate gradients. Numer. Math. **48**, 543–560 (1986)
52. van der Sluis, A., van der Vorst, H.: The convergence behavior of Ritz values in the presence of close eigenvalues. Linear Algebra Appl. **88**, 651–694 (1987)
53. Strakoš, Z.: On the real convergence rate of the conjugate gradient method. Linear Algebra Appl. **154–156**, 535–549 (1991)
54. Strakoš, Z., Tichý, P.: On error estimation in the conjugate gradient method and why it works in finite precision computations. Electron. Trans. Numer. Anal. **13**, 56–80 (electronic) (2002)
55. Stroud, A.H., Secrest, D.: Gaussian quadrature formulas. Prentice-Hall Inc., Englewood Cliffs, N.J. (1966)
56. Swarztrauber, P.N.: On computing the points and weights for Gauss-Legendre quadrature. SIAM J. Sci. Comput. **24**(3), 945–954 (electronic) (2002)
57. Szegö, G.: Orthogonal Polynomials. Colloquium Publications, vol. XXIII. American Mathematical Society, New York (1939)
58. Uvarov, V.B.: The connection between systems of polynomials that are orthogonal with respect to different distribution functions. Ž. Vyčisl. Mat. i Mat. Fiz. **9**, 1253–1262 (1969)
59. Vorobyev, Y.V.: Methods of moments in applied mathematics. Translated from the Russian by Bernard Seckler. Gordon and Breach Science Publishers, New York (1965)
60. Willems, P., Lang, B., Vömel, C.: Computing the bidiagonal SVD using multiple relatively robust representations. SIAM J. Matrix Anal. Appl. **28**(4), 907–926 (2006)
61. Xu, S.F.: A stability analysis of the Jacobi matrix inverse eigenvalue problem. BIT **33**(4), 695–702 (1993)