Electronic Transactions on Numerical Analysis.
Volume 47, pp. 179–196, 2017.

ISSN 1068–9613.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

VECTOR ESTIMATES FOR $f(A)$b VIA EXTRAPOLATION*

MARILENA MITROULI† AND PARASKEVI ROUPA†

Abstract. Let $A \in \mathbb{R}^{p\times p}$ be a diagonalizable matrix and f a smooth function. We are interested in the problem of approximating the action of $f(A)$ on a vector $\mathbf{b} \in \mathbb{R}^p$, i.e., $f(A)\mathbf{b}$, without explicitly computing the matrix $f(A)$. In the present work, we derive families of one-term, two-term, and three-term inexpensive approximations to the quantity $f(A)\mathbf{b}$ via an extrapolation procedure. For a given diagonalizable matrix A, the proposed families of vector estimates allow us to approximate the form $W^T f(A)U$, for any matrices $W, U \in \mathbb{R}^{p\times m}$, $1 \leq m \ll p$, not necessarily biorthogonal. We present several numerical examples to illustrate the effectiveness of our method for several functions f for both the quantity $f(A)\mathbf{b}$ and the form $W^T f(A)U$.

Key words. f(A)**b**, vector estimates, vector moments, extrapolation, diagonalizable matrices

AMS subject classifications. 65F15, 65F30, 65F60, 65B05, 15A18

1. Introduction. Let $A \in \mathbb{R}^{p\times p}$ be a diagonalizable matrix, $\mathbf{b}$ a real vector of length p, and f a smooth function defined on the spectrum of the matrix A. The aim of this work is the approximation of the matrix vector product $f(A)\mathbf{b}$ without explicitly computing the matrix $f(A)$ by an extrapolation procedure. Furthermore, we apply estimates of $f(A)\mathbf{b}$ for approximating expressions of the form $W^T f(A)U$ for any thin matrices $W, U \in \mathbb{R}^{p\times m}$, $1 \leq m \ll p$. The motivation for calculating the quantity $f(A)\mathbf{b}$ arises from applications in which it is neither necessary to approximate the whole matrix $f(A)$ nor feasible to compute $f(A)$ explicitly, especially when A is a large sparse matrix. Particularly, in applications, even when A is sparse, the matrix $f(A)$ is dense as long as A does not have a special structure (e.g., diagonal, triangular, block diagonal/triangular, etc.). Therefore, it is not feasible to deal with the whole matrix $f(A)$ when A is of large dimension since this is too expensive.

There are specific situations in which the action of $f(A)$ on a vector $\mathbf{b}$ is desired. In particular, $f(A)\mathbf{b}$ often appears in applications which originate from partial differential equations [17]. It also arises in lattice quantum chromodynamics computations in chemistry and physics; see [10] and references therein. In these applications, the sign function, sign(A), is used, and the given matrix A is very large, sparse, and complex Hermitian. Furthermore, it is often useful to compute $f(A)\mathbf{b}$ with $f(A) = A^{1/2}$ and A being a symmetric positive definite matrix. In problems arising in population dynamics and in neutron transport, the numerical solution of stochastic differential equations which contain the quantity $A^{1/2}\mathbf{b}$ is needed [1]. In case of the given matrix A being also sparse, $A^{1/2}\mathbf{b}$ appears in sampling from a Gaussian process distribution [9]. In network analysis, the quantity $\exp(A)\mathbf{b}$ determines the total communicability, which serves as a global measure of how well the nodes in a graph can exchange information. The total communicability of a node i can be defined as the ith entry of the vector $\exp(A)\mathbf{1}$, where $\mathbf{1}$ is a vector with all elements equal to one, i.e., $TC(i) := (\exp(A)\mathbf{1})_i$ [2, 3]. Also, expressions of the form $\exp(-\tau A)\mathbf{b}$, where A is a nonnegative definite matrix, appear in predicting the time evolution of electrical circuits and in computing the transient solution of Markov chains [19].

Among the developed methods for the $f(A)\mathbf{b}$-problem are those based on Krylov subspaces (see [17, 16] and references therein); others employ rational approximations [10] and polynomial approximations [9, 19]. The aim of this work is to introduce an alternative approach that utilizes extrapolation for estimating the quantity $f(A)\mathbf{b}$. Specifically, we produce

*Received February 15, 2017. Accepted October 2, 2017. Published online on November 15, 2017. Recommended by G. Meurant.
†University of Athens, Department of Mathematics, Panepistimiopolis, 15784 Athens, Greece
({mmitroul, parask_roupa}@math.uoa.gr).

families of approximations of $f(A)\mathbf{b}$ for a diagonalizable matrix $A \in \mathbb{R}^{p\times p}$ and a vector $\mathbf{b} \in \mathbb{R}^p$ via an extrapolation procedure. The derived family of one-term vector estimates φ_z depends on a parameter $z \in \mathbb{C}$. Optimal values $(z_{opt})_i$ for each entry of φ_z and bounds for these values are obtained. Moreover, families of two-term $\{\hat{\varphi}_{n,k}, n, k \in \mathbb{Z}\}$ and three-term $\{\tilde{\varphi}_{n,k,\ell}, n, k, \ell \in \mathbb{Z}\}$ vector estimates are generated. The behaviour of these families is tested by several numerical examples, and very satisfactory relative errors are reported.

The paper is organized as follows. In Section 2, families of one-term, two-term, and three-term vector estimates for the quantity $f(A)\mathbf{b}$ are derived. In Section 3, approximations of expressions of the form $W^T f(A)U$ are found by using the previously derived estimates for $f(A)\mathbf{b}$. Numerical examples are reported in Section 4, and Section 5 contains concluding remarks.

Throughout the paper, $(\cdot,\cdot)$ is the Euclidean inner product, for a vector $\mathbf{x} \in \mathbb{R}^p$ we denote by $\|\mathbf{x}\|$ the Euclidean vector norm, and x_i denotes the ith entry of the vector $\mathbf{x}$. For a matrix $A \in \mathbb{R}^{p\times p}$, $\|A\|_2$ is the spectral norm, and $\kappa_2(A)$ is the spectral condition number of the matrix A. I_k denotes the identity matrix of order k, e_i is the ith column of the identity matrix of suitable size, and $\mathbf{1}$ is a vector with all elements equal to one. The superscript ^{T} denotes the transpose, the symbol $\simeq$ means "approximately equal to", and the symbol $\ll$ means "much smaller than".

2. Vector estimates for $f(A)\mathbf{b}$.

In this section, we outline how the quantity $f(A)\mathbf{b}$ can be approximated by an extrapolation procedure. This kind of extrapolation procedure was introduced by Brezinski in [4] for approximating the norm of the error for the solution of linear systems, and it was extended in [7] and [8]. In [5, 6] an extrapolation procedure was developed for an approximation of the trace of powers of positive self-adjoint linear operators and an approximation of the inverse of a linear operator on a Hilbert space, respectively. Families of estimates for the bilinear form $\mathbf{x}^T A^{-1}\mathbf{y}$ for any nonsingular matrix were derived in [14], and in [13] families of estimates for the bilinear form $\mathbf{y}^* f(A)\mathbf{x}$ for a Hermitian matrix were given.

In the present work, we derive families of one-term, two-term, and three-term vector estimates for $f(A)\mathbf{b}$, where $A \in \mathbb{R}^{p\times p}$ is a diagonalizable matrix, $\mathbf{b}$ is a real vector of length p, and f is a smooth function defined on the spectrum of A.

We assume that the matrix A has the factorization

$$A = Q\Lambda Q^{-1},$$

where

$$Q = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_p \end{bmatrix} \in \mathbb{C}^{p\times p} \text{ is nonsingular}, \qquad \Lambda = \mathrm{diag}[\lambda_1, \dots, \lambda_p] \in \mathbb{C}^{p\times p},$$

$$Q^{-1} = \begin{bmatrix} \hat{\mathbf{q}}_1^T \\ \hat{\mathbf{q}}_2^T \\ \dots \\ \hat{\mathbf{q}}_p^T \end{bmatrix} \in \mathbb{C}^{p\times p}, \qquad \mathbf{q}_i, \hat{\mathbf{q}}_i \in \mathbb{C}^{p\times 1}, \quad i = 1, 2, \dots, p.$$

The diagonal elements λ_i of the matrix Λ are the eigenvalues of A, which are real or appear in complex conjugate pairs. The column vectors of Q are the right eigenvectors of A, and the row vectors of Q^{-1} are the left eigenvectors of the matrix A. Since $Q^{-1}Q = I_p$, it holds that $\hat{\mathbf{q}}_i^T \mathbf{q}_j = \delta_{ij}$ for $i, j = 1, 2, \dots, p$, where δ_{ij} is the Kronecker delta. For a function f defined on the spectrum of the matrix A, the matrix $f(A) \in \mathbb{C}^{p\times p}$ can be expressed as

$$f(A) = Qf(\Lambda)Q^{-1} = \sum_{j=1}^{p} f(\lambda_j)\mathbf{q}_j\hat{\mathbf{q}}_j^T.$$

For any integer $r \in \mathbb{Z}$ and a vector $\mathbf{b} \in \mathbb{R}^p$, we define the vector moments $\mathbf{v_r} \in \mathbb{R}^p$ of the matrix A as follows:

$$\mathbf{v_r} = A^r \mathbf{b}.$$

Each entry $v_{r,i}$ of the vector moment $\mathbf{v_r}$ can be written as the following sum

$$v_{r,i} = e_i^T \mathbf{v_r} = e_i^T A^r \mathbf{b} = \sum_{j=1}^{p} \lambda_j^r (e_i, \mathbf{q}_j)(\hat{\mathbf{q}}_j, \mathbf{b}) = \sum_{j=1}^{p} \lambda_j^r \alpha_{j,i}\beta_j, \tag{2.1}$$

where $\alpha_{j,i} = (e_i, \mathbf{q}_j)$, $\beta_j = (\hat{\mathbf{q}}_j, \mathbf{b})$, $i, j = 1, 2, \ldots, p$. The function vector moment $\mathbf{v_f} \in \mathbb{C}^p$ of the matrix A is defined as

$$\mathbf{v_f} = f(A)\mathbf{b}.$$

For the entries $v_{f,i}$, $i = 1, 2, \ldots, p$, of the function vector moment $\mathbf{v_f}$, we have

$$v_{f,i} = e_i^T f(A)\mathbf{b} = \sum_{j=1}^{p} f(\lambda_j)(e_i, \mathbf{q}_j)(\hat{\mathbf{q}}_j, \mathbf{b}) = \sum_{j=1}^{p} f(\lambda_j)\alpha_{j,i}\beta_j. \tag{2.2}$$

An approximation of $\mathbf{v_f}$ can be obtained without computing the factorization of A. Let us keep k terms in the summation (2.2), that is,

$$\mathbf{v_f} = \begin{bmatrix} v_{f,1} \\ v_{f,2} \\ \vdots \\ v_{f,p} \end{bmatrix} \simeq \begin{bmatrix} \sum_{j=1}^k f(\tilde{\lambda}_{j,1})\tilde{\alpha}_{j,1}\tilde{\beta}_{j,1} \\ \sum_{j=1}^k f(\tilde{\lambda}_{j,2})\tilde{\alpha}_{j,2}\tilde{\beta}_{j,2} \\ \vdots \\ \sum_{j=1}^k f(\tilde{\lambda}_{j,p})\tilde{\alpha}_{j,p}\tilde{\beta}_{j,p} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^k f(\tilde{\lambda}_{j,1})\tilde{m}_{j,1} \\ \sum_{j=1}^k f(\tilde{\lambda}_{j,2})\tilde{m}_{j,2} \\ \vdots \\ \sum_{j=1}^k f(\tilde{\lambda}_{j,p})\tilde{m}_{j,p} \end{bmatrix}. \tag{2.3}$$

The unknowns $\tilde{\lambda}_{j,i}$ and $\tilde{m}_{j,i} = \tilde{\alpha}_{j,i}\tilde{\beta}_{j,i}$ can be determined by imposing as interpolation conditions the relation (2.1) for various nonnegative values of r, i.e.,

$$\mathbf{v_r} = \begin{bmatrix} v_{r,1} \\ v_{r,2} \\ \vdots \\ v_{r,p} \end{bmatrix} \simeq \begin{bmatrix} \sum_{j=1}^k \tilde{\lambda}_{j,1}^r \tilde{\alpha}_{j,1}\tilde{\beta}_{j,1} \\ \sum_{j=1}^k \tilde{\lambda}_{j,2}^r \tilde{\alpha}_{j,2}\tilde{\beta}_{j,2} \\ \vdots \\ \sum_{j=1}^k \tilde{\lambda}_{j,p}^r \tilde{\alpha}_{j,p}\tilde{\beta}_{j,p} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^k \tilde{\lambda}_{j,1}^r \tilde{m}_{j,1} \\ \sum_{j=1}^k \tilde{\lambda}_{j,2}^r \tilde{m}_{j,2} \\ \vdots \\ \sum_{j=1}^k \tilde{\lambda}_{j,p}^r \tilde{m}_{j,p} \end{bmatrix}. \tag{2.4}$$

2.1. One-term vector estimates. By applying the summations of (2.3) for $k = 1$, each entry of the function vector moment $\mathbf{v_f}$ can be approximated by

$$v_{f,i} \simeq f(\tilde{\lambda}_{1,i})\tilde{m}_{1,i}, \qquad i = 1, 2, \ldots, p,$$

where the unknowns $\tilde{\lambda}_{1,i}$, $\tilde{m}_{1,i}$ are determined from the interpolation conditions (2.4) for $r = 0, 1, 2$, i.e.,

$$\begin{aligned} v_{0,i} &\simeq \tilde{m}_{1,i}, \\ v_{1,i} &\simeq \tilde{\lambda}_{1,i}\tilde{m}_{1,i} \simeq \tilde{\lambda}_{1,i} v_{0,i}, \\ v_{2,i} &\simeq \tilde{\lambda}_{1,i}^2 \tilde{m}_{1,i} \simeq \tilde{\lambda}_{1,i}^2 v_{0,i}. \end{aligned}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

By solving the above system of equations, we arrive at the following compact formula for $\tilde{\lambda}_{1,i}$:

$$\tilde{\lambda}_{1,i} = v_{0,i}^{j-1} v_{1,i}^{1-2j} v_{2,i}^{j}, \qquad j = 0, \tfrac{1}{2}, 1, \; i = 1, 2, \ldots, p. \tag{2.5}$$

We notice that in formula (2.5), if j takes any complex number, i.e., $j = \gamma + i\delta$, then this results again in the same value $\tilde{\lambda}_{1,i}$. More specifically, for $j = \gamma + i\delta$, it holds that

$$v_{0,i}^{\gamma+i\delta-1} v_{1,i}^{1-2(\gamma+i\delta)} v_{2,i}^{\gamma+i\delta} = (\tilde{m}_{1,i})^{\gamma+i\delta-1} (\tilde{\lambda}_{1,i}\tilde{m}_{1,i})^{1-2\gamma-2i\delta} (\tilde{\lambda}_{1,i}^2 \tilde{m}_{1,i})^{\gamma+i\delta} = \tilde{\lambda}_{1,i}.$$

Therefore, we obtain a family of one-term vector estimates φ_z for $\mathbf{v_f}$.

PROPOSITION 2.1.

1. *A family of one-term vector estimates $\{\varphi_z, \; z \in \mathbb{C}\}$ for the function vector moment $\mathbf{v_f}$ is given by*

$$\varphi_{z,i} = f\left(v_{0,i}^{z-1} v_{1,i}^{1-2z} v_{2,i}^{z}\right) v_{0,i}, \qquad z \in \mathbb{C}, \; i = 1, 2, \ldots, p. \tag{2.6}$$

2. *The family of one-term vector estimates (2.6) satisfies the relation*

$$\varphi_{z,i} = f\left(\rho_i^z \frac{v_{1,i}}{v_{0,i}}\right) v_{0,i}, \qquad z \in \mathbb{C}, \tag{2.7}$$

where $\rho_i = v_{0,i}v_{2,i}/v_{1,i}^2$ and $v_{0,i}v_{1,i} \neq 0$, for $i = 1, 2, \ldots, p$. If $v_{0,i} = 0$ or $v_{1,i} = 0$, then we can apply the formula (2.6) for $z = 1$ or $z = \frac{1}{2}$, respectively.

Proof. 1. By replacing the values of $\tilde{\lambda}_{1,i}$ in the expression $v_{f,i} \simeq f(\tilde{\lambda}_{1,i})v_{0,i}$, the formula $\varphi_{z,i} = f\left(v_{0,i}^{z-1} v_{1,i}^{1-2z} v_{2,i}^{z}\right) v_{0,i}$ is obtained for all $i = 1, 2, \ldots, p$.

2. It holds that

$$\begin{aligned}
\varphi_{z,i} &= f\left(v_{0,i}^{z-1} v_{1,i}^{1-2z} v_{2,i}^{z}\right) v_{0,i} = f\left(\frac{v_{0,i}^z v_{2,i}^z}{v_{1,i}^{2z}} \cdot \frac{v_{1,i}}{v_{0,i}}\right) v_{0,i} \\
&= f\left(\left(\frac{v_{0,i}v_{2,i}}{v_{1,i}^2}\right)^z \cdot \frac{v_{1,i}}{v_{0,i}}\right) v_{0,i} = f\left(\rho_i^z \frac{v_{1,i}}{v_{0,i}}\right) v_{0,i},
\end{aligned}$$

where $\rho_i = v_{0,i}v_{2,i}/v_{1,i}^2$. Therefore, $\varphi_{z,i} = f\left(\rho_i^z \frac{v_{1,i}}{v_{0,i}}\right) v_{0,i}$. ∎

The family of one-term vector estimates φ_z depends on a parameter $z \in \mathbb{C}$. The following lemma specifies the existence of optimal values $(z_{opt})_i$, $i = 1, 2, \ldots, p$, which lead to an exact approximation of each entry of the vector moment $\mathbf{v_f}$.

LEMMA 2.2. *Let $A \in \mathbb{R}^{p \times p}$ be a diagonalizable matrix, f an invertible function, and $\rho_i = v_{0,i}v_{2,i}/v_{1,i}^2$.*

1. *If $\rho_i \neq 1, \; i = 1, 2, \ldots, p$, then there exists an optimal value $(z_{opt})_i$ given by*

$$(z_{opt})_i = \frac{\log\left(f^{-1}\left(\frac{v_{f,i}}{v_{0,i}}\right)\frac{v_{0,i}}{v_{1,i}}\right)}{\log(\rho_i)}$$

such that $\varphi_{(z_{opt})_i,i}$ gives the exact value of $v_{f,i}$.

2. *If $\rho_i = 1$, then the optimal value $(z_{opt})_i$ can be any complex number, i.e., $\varphi_{z,i} = v_{f,i}$, for all $z \in \mathbb{C}$, $i = 1, 2, \ldots, p$.*

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Proof. 1. It holds that, for all $i = 1, 2, \dots, p$,

$$\begin{aligned}
\varphi_{(z_{opt})_i,i} = v_{f,i} &\overset{(2.7)}{\Rightarrow} f\left(\rho_i^{(z_{opt})_i}\frac{v_{1,i}}{v_{0,i}}\right)v_{0,i} = v_{f,i} \Rightarrow f\left(\rho_i^{(z_{opt})_i}\frac{v_{1,i}}{v_{0,i}}\right) = \frac{v_{f,i}}{v_{0,i}}\\
&\Rightarrow \rho_i^{(z_{opt})_i}\frac{v_{1,i}}{v_{0,i}} = f^{-1}\left(\frac{v_{f,i}}{v_{0,i}}\right) \Rightarrow \rho_i^{(z_{opt})_i} = f^{-1}\left(\frac{v_{f,i}}{v_{0,i}}\right)\frac{v_{0,i}}{v_{1,i}}\\
&\Rightarrow (z_{opt})_i = \frac{\log\left(f^{-1}\left(\frac{v_{f,i}}{v_{0,i}}\right)\frac{v_{0,i}}{v_{1,i}}\right)}{\log(\rho_i)},
\end{aligned}$$

where $\rho_i = v_{0,i}v_{2,i}/v_{1,i}^2 \neq 1$.

2. If $\rho_i = 1$, then the relation (2.7) can be written as $\varphi_{z,i} = f\left(\frac{v_{1,i}}{v_{0,i}}\right)v_{0,i}$, which is independent of the parameter $z \in \mathbb{C}$. ∎

By the Cauchy-Schwarz inequality, we obtain a bound for each optimal value $(z_{opt})_i$, which is given in the following lemma.

LEMMA 2.3. *Let $A \in \mathbb{R}^{p\times p}$ be a diagonalizable matrix and f an increasing real-valued function. If $v_{0,i} > 0$, $v_{1,i} > 0$, and $\rho_i > 1$, then an upper bound for the optimal value $(z_{opt})_i$ is given by*

$$(z_{opt})_i \leq \frac{\log\left(f^{-1}\left(\kappa_2(Q)f(\rho(A))\frac{\|\mathbf{b}\|}{v_{0,i}}\right)\frac{v_{0,i}}{v_{1,i}}\right)}{\log(\rho_i)},$$

where $\kappa_2(Q)$ is the spectral condition number of Q, the matrix of eigenvectors of A, and $\rho(A)$ is the spectral radius of A.

Proof. It holds that $\|f(A)\|_2 \leq \kappa_2(Q)\cdot f(\rho(A))$ [16, p. 102]. By the Cauchy-Schwarz inequality, we write

$$\begin{aligned}
v_{f,i} = (e_i, f(A)\mathbf{b}) \leq |(e_i, f(A)\mathbf{b})| &\leq \|e_i\|\cdot\|f(A)\mathbf{b}\| \leq \|f(A)\|_2\cdot\|\mathbf{b}\|\\
&\leq \kappa_2(Q)\cdot f(\rho(A))\cdot\|\mathbf{b}\|.
\end{aligned}$$

Since $v_{0,i} > 0$ and f is an increasing function, we have

$$f^{-1}\left(\frac{v_{f,i}}{v_{0,i}}\right) \leq f^{-1}\left(\kappa_2(Q)\cdot f(\rho(A))\cdot\frac{\|\mathbf{b}\|}{v_{0,i}}\right).$$

Considering that $v_{1,i} > 0$, we obtain

$$\begin{aligned}
&f^{-1}\left(\frac{v_{f,i}}{v_{0,i}}\right)\frac{v_{0,i}}{v_{1,i}} \leq f^{-1}\left(\kappa_2(Q)\cdot f(\rho(A))\cdot\frac{\|\mathbf{b}\|}{v_{0,i}}\right)\frac{v_{0,i}}{v_{1,i}}\\
&\quad\Rightarrow \log\left(f^{-1}\left(\frac{v_{f,i}}{v_{0,i}}\right)\frac{v_{0,i}}{v_{1,i}}\right) \leq \log\left(f^{-1}\left(\kappa_2(Q)\cdot f(\rho(A))\cdot\frac{\|\mathbf{b}\|}{v_{0,i}}\right)\frac{v_{0,i}}{v_{1,i}}\right)\\
&\quad\Rightarrow (z_{opt})_i \leq \frac{\log\left(f^{-1}\left(\kappa_2(Q)\cdot f(\rho(A))\cdot\frac{\|\mathbf{b}\|}{v_{0,i}}\right)\frac{v_{0,i}}{v_{1,i}}\right)}{\log(\rho_i)}
\end{aligned}$$

since $\rho_i > 1$. ∎

REMARK 2.4. If the function f is decreasing, $v_{0,i}$, $v_{1,i}$ are not both positive, and ρ_i is either greater or less than 1, then a similar result as that of Lemma 2.3 can be obtained.

REMARK 2.5. If the matrix A is normal, a sharper bound for the optimal value $(z_{opt})_i$ is obtained. In particular, it holds that $\kappa_2(Q) = 1$, hence, it is not necessary to compute the matrix of eigenvectors Q.

Next, we provide a connection between the one-term vector estimate (2.6) for $z = 0$ with an estimate which can be obtained by the nonsymmetric Lanczos algorithm and the nonsymmetric Gauss quadrature rules with one iteration [15]. In particular, the following proposition holds for any matrix.

PROPOSITION 2.6. *Let $A \in \mathbb{R}^{p\times p}$ be a nonsymmetric matrix. Each element of the vector estimate $\varphi_{\mathbf{0}}$ is equal to the estimate which is obtained by the nonsymmetric Lanczos algorithm and the nonsymmetric Gauss quadrature with one iteration for the bilinear form $e_i^T f(A)\mathbf{b}$ considering that $(e_i, \mathbf{b}) \neq 0$.*

Proof. We apply the nonsymmetric Lanczos algorithm with initial vectors $\mathbf{u} = \frac{\mathbf{b}}{(e_i,\mathbf{b})}$ and $\tilde{\mathbf{u}} = e_i$ [15, p. 43], which satisfy the relation $(\tilde{\mathbf{u}}, \mathbf{u}) = 1$. By keeping one iteration of this algorithm, the Jacobi matrix J_1 has only one element, i.e., $J_1 = \left[\, \tilde{\mathbf{u}}^T A\mathbf{u} \,\right]$. It holds that

$$\tilde{\mathbf{u}}^T A\mathbf{u} = e_i^T A\mathbf{b}\frac{1}{(e_i,\mathbf{b})} = \frac{v_{1,i}}{v_{0,i}}.$$

Therefore,

$$e_i^T f(A)\mathbf{b} = (e_i,\mathbf{b}) f(\tilde{\mathbf{u}}^T A\mathbf{u}) = v_{0,i} f(\frac{v_{1,i}}{v_{0,i}}) = \varphi_{0,i},$$

for all $i = 1, 2, \ldots, p$. ∎

The family of one-term vector estimates φ_z provides an exact approximation in case that the optimal values $(z_{opt})_i$, $i = 1, 2, \ldots, p$, can be specified a priori. In order to avoid this difficulty, we keep more terms in the summations (2.4) and derive estimates with more than one term.

2.2. Two-term vector estimates. By applying the summation in (2.3) for $k = 2$, each entry of the function vector moment $\mathbf{v_f}$ can be approximated by

$$v_{f,i} \simeq f(\tilde{\lambda}_{1,i})\tilde{m}_{1,i} + f(\tilde{\lambda}_{2,i})\tilde{m}_{2,i}, \qquad i = 1, 2, \ldots, p,$$

where the unknowns $\tilde{\lambda}_{j,i}$, $\tilde{m}_{j,i}$, $j = 1, 2$, are determined by the interpolation conditions (2.4) for $r = 0, 1, 2$.

PROPOSITION 2.7. *The two-term estimates for the vector moments $\mathbf{v_n}$ satisfy the following difference equation of order two,*

$$v_{n+1,i} - r_i v_{n,i} + q_i v_{n-1,i} = 0, \qquad n \in \mathbb{Z},$$

where $r_i = \tilde{\lambda}_{1,i} + \tilde{\lambda}_{2,i} \in \mathbb{C}$, $q_i = \tilde{\lambda}_{1,i}\tilde{\lambda}_{2,i} \in \mathbb{C}$, for all $i = 1, 2, \ldots, p$.

Proof. It holds that $v_{n+1,i} \simeq \tilde{\lambda}_{1,i}^{n+1}\tilde{m}_{1,i} + \tilde{\lambda}_{2,i}^{n+1}\tilde{m}_{2,i}$.

$$\begin{aligned}
r_i v_{n,i} &= \left(\tilde{\lambda}_{1,i} + \tilde{\lambda}_{2,i}\right)\left(\tilde{\lambda}_{1,i}^{n}\tilde{m}_{1,i} + \tilde{\lambda}_{2,i}^{n}\tilde{m}_{2,i}\right)\\
&= \tilde{\lambda}_{1,i}^{n+1}\tilde{m}_{1,i} + \tilde{\lambda}_{1,i}\tilde{\lambda}_{2,i}^{n}\tilde{m}_{2,i} + \tilde{\lambda}_{2,i}\tilde{\lambda}_{1,i}^{n}\tilde{m}_{1,i} + \tilde{\lambda}_{2,i}^{n+1}\tilde{m}_{2,i},\\
q_i v_{n-1,i} &= \left(\tilde{\lambda}_{1,i}\tilde{\lambda}_{2,i}\right)\left(\tilde{\lambda}_{1,i}^{n-1}\tilde{m}_{1,i} + \tilde{\lambda}_{2,i}^{n-1}\tilde{m}_{2,i}\right)\\
&= \tilde{\lambda}_{2,i}\tilde{\lambda}_{2,i}^{n}\tilde{m}_{1,i} + \tilde{\lambda}_{1,i}\tilde{\lambda}_{2,i}^{n}\tilde{m}_{2,i}.
\end{aligned}$$

From these relations, we observe that the difference equation $v_{n+1,i} - r_i v_{n,i} + q_i v_{n-1,i} = 0$ is satisfied. ∎

ETNA
Kent State University and
Johann Radon Institute (RICAM)

We consider the system of equations

$$v_{n+1,i} - r_i v_{n,i} + q_i v_{n-1,i} = 0, \qquad v_{n+2+k,i} - r_i v_{n+1+k,i} + q_i v_{n+k,i} = 0, \qquad n,k \in \mathbb{Z}.$$

The solution of this system is given by

$$r_i = \frac{v_{n-1,i}v_{n+2+k,i} - v_{n+1,i}v_{n+k,i}}{v_{n-1,i}v_{n+1+k,i} - v_{n,i}v_{n+k,i}}, \qquad q_i = \frac{v_{n,i}v_{n+2+k,i} - v_{n+1,i}v_{n+1+k,i}}{v_{n-1,i}v_{n+1+k,i} - v_{n,i}v_{n+k,i}}.$$

Also, for all $i = 1, 2, \dots, p$, we have

$$\tilde{\lambda}_{1,i} = \frac{r_i + \sqrt{r_i^2 - 4q_i}}{2} \quad \text{and} \quad \tilde{\lambda}_{2,i} = \frac{r_i - \sqrt{r_i^2 - 4q_i}}{2}. \tag{2.8}$$

The solution of the system of the corresponding interpolation conditions is given by the following formulae, for all $i = 1, 2, \dots, p$,

$$\tilde{m}_{1,i} = \frac{1}{\tilde{\lambda}_{2,i} - \tilde{\lambda}_{1,i}}(\tilde{\lambda}_{2,i}v_{0,i} - v_{1,i}), \qquad \tilde{\lambda}_{1,i} \neq \tilde{\lambda}_{2,i}, \tag{2.9}$$

$$\tilde{m}_{2,i} = \frac{1}{\tilde{\lambda}_{2,i} - \tilde{\lambda}_{1,i}}(v_{1,i} - \tilde{\lambda}_{1,i}v_{0,i}), \qquad \tilde{\lambda}_{1,i} \neq \tilde{\lambda}_{2,i}. \tag{2.10}$$

As a result, the following proposition provides a family of two-term vector estimates $\hat{\varphi}_{n,k}$ for the function vector moment $\mathbf{v_f}$.

PROPOSITION 2.8. *A family of two-term vector estimates $\{\hat{\varphi}_{n,k},\ n,k \in \mathbb{Z}\}$ for the function vector moment $\mathbf{v_f}$ is given by*

$$(\hat{\varphi}_{n,k})_i = f(\tilde{\lambda}_{1,i})\tilde{m}_{1,i} + f(\tilde{\lambda}_{2,i})\tilde{m}_{2,i}, \qquad n,k \in \mathbb{Z},\ i = 1, 2, \dots, p, \tag{2.11}$$

where $\tilde{\lambda}_{1,i}$, $\tilde{\lambda}_{2,i}$, $\tilde{m}_{1,i}$, and $\tilde{m}_{2,i}$ are defined by the formulae (2.8), (2.9), and (2.10), respectively.

REMARK 2.9. Let $(\hat{\rho}_{n,k})_i = (v_{n-1,i}v_{k+1,i})/(v_{n,i}v_{k,i})$, for $i = 1, 2, \dots, p$. If for some i it holds that $v_{n-1,i} = v_{n+k,i} = 0$ or $(\hat{\rho}_{n,n+k})_i = 1$ or $r_i^2 = 4q_i$, then the formula (2.11) does not give an estimate for that specific selection of the pair (n,k) of parameters. In case that $n = k = 1$, we find $(\hat{\rho}_{1,1})_i = (v_{0,i}v_{2,i})/v_{1,i}^2$ which is equal to ρ_i used in the one-term vector estimates.

2.3. Three-term vector estimates. By keeping three terms in the summations in (2.3), each entry of the vector moment $\mathbf{v_f}$ can be approximated by

$$v_{f,i} \simeq f(\tilde{\lambda}_{1,i})\tilde{m}_{1,i} + f(\tilde{\lambda}_{2,i})\tilde{m}_{2,i} + f(\tilde{\lambda}_{3,i})\tilde{m}_{3,i}, \qquad i = 1, 2, \dots, p,$$

where the unknowns $\tilde{\lambda}_{j,i}$, $\tilde{m}_{j,i}$, $j = 1, 2, 3$, are determined by the interpolation conditions (2.4) for $r = 0, 1, 2$.

PROPOSITION 2.10. *The three-term estimates for the vector moments $\mathbf{v_n}$ satisfy the following difference equation of order three, i.e.,*

$$v_{n+2,i} - s_i v_{n+1,i} + t_i v_{n,i} - g_i v_{n-1,i} = 0, \qquad n \in \mathbb{Z},$$

where

$$\begin{aligned} s_i &= \tilde{\lambda}_{1,i} + \tilde{\lambda}_{2,i} + \tilde{\lambda}_{3,i}, \\ t_i &= \tilde{\lambda}_{1,i}\tilde{\lambda}_{2,i} + \tilde{\lambda}_{1,i}\tilde{\lambda}_{3,i} + \tilde{\lambda}_{2,i}\tilde{\lambda}_{3,i}, \qquad \textit{and} \\ g_i &= \tilde{\lambda}_{1,i}\tilde{\lambda}_{2,i}\tilde{\lambda}_{3,i}. \end{aligned} \tag{2.12}$$

Proof. It holds that $v_{n+2,i} \simeq \tilde{\lambda}_{1,i}^{n+2}\tilde{m}_{1,i} + \tilde{\lambda}_{2,i}^{n+2}\tilde{m}_{2,i} + \tilde{\lambda}_{3,i}^{n+2}\tilde{m}_{3,i}$,

$$\begin{aligned}
s_i v_{n+1,i} &= \left(\tilde{\lambda}_{1,i} + \tilde{\lambda}_{2,i} + \tilde{\lambda}_{3,i}\right)\left(\tilde{\lambda}_{1,i}^{n+1}\tilde{m}_{1,i} + \tilde{\lambda}_{2,i}^{n+1}\tilde{m}_{2,i} + \tilde{\lambda}_{3,i}^{n+1}\tilde{m}_{3,i}\right) \\
&= \tilde{\lambda}_{1,i}^{n+2}\tilde{m}_{1,i} + \tilde{\lambda}_{1,i}\tilde{\lambda}_{2,i}^{n+1}\tilde{m}_{2,i} + \tilde{\lambda}_{1,i}\tilde{\lambda}_{3,i}^{n+1}\tilde{m}_{3,i} + \tilde{\lambda}_{2,i}\tilde{\lambda}_{1,i}^{n+1}\tilde{m}_{1,i} \\
&\quad + \tilde{\lambda}_{2,i}^{n+2}\tilde{m}_{2,i} + \tilde{\lambda}_{2,i}\tilde{\lambda}_{3,i}^{n+1}\tilde{m}_{3,i} + \tilde{\lambda}_{3,i}\tilde{\lambda}_{1,i}^{n+1}\tilde{m}_{1,i} + \tilde{\lambda}_{3,i}\tilde{\lambda}_{2,i}^{n+1}\tilde{m}_{2,i} + \tilde{\lambda}_{3,i}^{n+2}\tilde{m}_{3,i}, \\
t_i v_{n,i} &= \left(\tilde{\lambda}_{1,i}\tilde{\lambda}_{2,i} + \tilde{\lambda}_{1,i}\tilde{\lambda}_{3,i} + \tilde{\lambda}_{2,i}\tilde{\lambda}_{3,i}\right)\left(\tilde{\lambda}_{1,i}^{n}\tilde{m}_{1,i} + \tilde{\lambda}_{2,i}^{n}\tilde{m}_{2,i} + \tilde{\lambda}_{3,i}^{n}\tilde{m}_{3,i}\right) \\
&= \tilde{\lambda}_{2,i}\tilde{\lambda}_{1,i}^{n+1}\tilde{m}_{1,i} + \tilde{\lambda}_{1,i}\tilde{\lambda}_{2,i}^{n+1}\tilde{m}_{2,i} + \tilde{\lambda}_{1,i}\tilde{\lambda}_{2,i}\tilde{\lambda}_{3,i}^{n}\tilde{m}_{3,i} + \tilde{\lambda}_{3,i}\tilde{\lambda}_{1,i}^{n+1}\tilde{m}_{1,i} \\
&\quad + \tilde{\lambda}_{1,i}\tilde{\lambda}_{3,i}\tilde{\lambda}_{2,i}^{n}\tilde{m}_{2,i} + \tilde{\lambda}_{1,i}\tilde{\lambda}_{3,i}^{n+1}\tilde{m}_{3,i} + \tilde{\lambda}_{2,i}\tilde{\lambda}_{3,i}\tilde{\lambda}_{1,i}^{n}\tilde{m}_{1,i} \\
&\quad + \tilde{\lambda}_{3,i}\tilde{\lambda}_{2,i}^{n+1}\tilde{m}_{2,i} + \tilde{\lambda}_{2,i}\tilde{\lambda}_{3,i}^{n+1}\tilde{m}_{3,i}, \\
g_i v_{n-1,i} &= \left(\tilde{\lambda}_{1,i}\tilde{\lambda}_{2,i}\tilde{\lambda}_{3,i}\right)\left(\tilde{\lambda}_{1,i}^{n-1}\tilde{m}_{1,i} + \tilde{\lambda}_{2,i}^{n-1}\tilde{m}_{2,i} + \tilde{\lambda}_{3,i}^{n-1}\tilde{m}_{3,i}\right) \\
&= \tilde{\lambda}_{2,i}\tilde{\lambda}_{3,i}\tilde{\lambda}_{1,i}^{n}\tilde{m}_{1,i} + \tilde{\lambda}_{1,i}\tilde{\lambda}_{3,i}\tilde{\lambda}_{2,i}^{n}\tilde{m}_{2,i} + \tilde{\lambda}_{1,i}\tilde{\lambda}_{2,i}\tilde{\lambda}_{3,i}^{n}\tilde{m}_{3,i}.
\end{aligned}$$

These relations imply that the difference equation $v_{n+2,i} - s_i v_{n+1,i} + t_i v_{n,i} - g_i v_{n-1,i} = 0$ is satisfied. □

In order to find the parameters s_i, t_i, g_i, we consider the system of equations, for $n, k, \ell \in \mathbb{Z}$,

$$\begin{aligned}
v_{n+2,i} - s_i v_{n+1,i} + t_i v_{n,i} - g_i v_{n-1,i} &= 0, \\
v_{n+k+3,i} - s_i v_{n+k+2,i} + t_i v_{n+k+1,i} - g_i v_{n+k,i} &= 0, \\
v_{n+\ell+4,i} - s_i v_{n+\ell+3,i} + t_i v_{n+\ell+2,i} - g_i v_{n+\ell+1,i} &= 0.
\end{aligned}$$

The solution to these equations is found by applying the Symbolic Math Toolbox of MATLAB. In the symbolically formulae we have made appropriate simplifications and finally obtain the following expressions:

$$\begin{aligned}
s_i &= \frac{v_{n,i}v_{n+k,i}v_{n+\ell+4,i} - v_{n,i}v_{n+\ell+1,i}v_{n+k+3,i}}{w_i} \\
&\quad + \frac{-v_{n+k,i}v_{n+\ell+2,i}v_{n+2,i} + v_{n+k+1,i}v_{n+\ell+1,i}v_{n+2,i}}{w_i} \\
&\quad + \frac{-v_{n+k+1,i}v_{n-1,i}v_{n+\ell+4,i} + v_{n+\ell+2,i}v_{n-1,i}v_{n+k+3,i}}{w_i} \\
t_i &= \frac{v_{n+1,i}v_{n+k,i}v_{n+\ell+4,i} - v_{n+1,i}v_{n+\ell+1,i}v_{n+k+3,i}}{w_i} \\
&\quad + \frac{-v_{n+k,i}v_{n+\ell+3,i}v_{n+2,i} + v_{n+k+2,i}v_{n+\ell+1,i}v_{n+2,i}}{w_i} \\
&\quad + \frac{-v_{n+k+2,i}v_{n-1,i}v_{n+\ell+4,i} + v_{n+\ell+3,i}v_{n-1,i}v_{n+k+3,i}}{w_i}, \\
g_i &= \frac{v_{n,i}v_{n+k+2,i}v_{n+\ell+4,i} - v_{n,i}v_{n+\ell+3,i}v_{n+k+3,i}}{w_i} \\
&\quad + \frac{-v_{n+1,i}v_{n+k+1,i}v_{n+\ell+4,i} + v_{n+1,i}v_{n+\ell+2,i}v_{n+k+3,i}}{w_i} \\
&\quad + \frac{v_{n+k+1,i}v_{n+\ell+3,i}v_{n+2,i} - v_{n+k+2,i}v_{n+\ell+2,i}v_{n+2,i}}{w_i},
\end{aligned}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

where

$$\begin{aligned} w_i = v_{n,i}v_{n+k,i}v_{n+\ell+3,i} - v_{n,i}v_{n+k+2,i}v_{n+\ell+1,i} \\ - v_{n+1,i}v_{n+k,i}v_{n+\ell+2,i} + v_{n+1,i}v_{n+k+1,i}v_{n+\ell+1,i} \\ - v_{n+k+1,i}v_{n+\ell+3,i}v_{n-1,i} + v_{n+k+2,i}v_{n+\ell+2,i}v_{n-1,i}. \end{aligned}$$

The formulae for the parameters $\tilde\lambda_{1,i}$, $\tilde\lambda_{2,i}$, and $\tilde\lambda_{3,i}$, which are obtained by symbolically solving the system of equations (2.12), are given by

$$\tilde\lambda_{1,i} = \frac{s_i}{3} - \frac{(A_p)_i}{3\cdot 2^{\frac{4}{3}}} + \frac{3t_i - s_i^2}{3\cdot 2^{\frac{2}{3}}(A_p)_i} - 2^{-\frac{2}{3}}3^{-1/2}\sqrt{-\left(\frac{3t_i - s_i^2 + 2^{-\frac{2}{3}}(A_p)_i^2}{(A_p)_i}\right)^2}, \tag{2.13}$$

$$\tilde\lambda_{2,i} = \frac{s_i}{3} - \frac{(A_p)_i}{3\cdot 2^{\frac{4}{3}}} + \frac{3t_i - s_i^2}{3\cdot 2^{\frac{2}{3}}(A_p)_i} + 2^{-\frac{2}{3}}3^{-1/2}\sqrt{-\left(\frac{3t_i - s_i^2 + 2^{-\frac{2}{3}}(A_p)_i^2}{(A_p)_i}\right)^2}, \tag{2.14}$$

$$\tilde\lambda_{3,i} = (B_p)_i + \frac{s_i}{3} + \frac{s_i^2 - 3t_i}{9(B_p)_i}, \tag{2.15}$$

where

$$(A_p)_i = \left(27g_i + 3\sqrt{3}\sqrt{4t_i^3 - t_i^2s_i^2 - 18t_ig_is_i + 27g_i^2 + 4g_is_i^3} - 9t_is_i + 2s_i^3\right)^{\frac{1}{3}} \quad \text{and}$$

$$(B_p)_i = \left(\frac{g_i}{2} - \frac{t_is_i}{6} + \sqrt{\left(\frac{s_i^3}{27} - \frac{t_is_i}{6} + \frac{g_i}{2}\right)^2 + \left(\frac{t_i}{3} - \frac{s_i^2}{9}\right)^3} + \frac{s_i^3}{27}\right)^{\frac{1}{3}}.$$

The solution of the system of the corresponding interpolation conditions is given by

$$\tilde m_{1,i} = \frac{v_{2,i} - \tilde\lambda_{2,i}v_{1,i} - \tilde\lambda_{3,i}v_{1,i} + \tilde\lambda_{2,i}\tilde\lambda_{3,i}v_{0,i}}{\left(\tilde\lambda_{1,i} - \tilde\lambda_{2,i}\right)\left(\tilde\lambda_{1,i} - \tilde\lambda_{3,i}\right)}, \tag{2.16}$$

$$\tilde m_{2,i} = -\frac{v_{2,i} - \tilde\lambda_{1,i}v_{1,i} - \tilde\lambda_{3,i}v_{1,i} + \tilde\lambda_{1,i}\tilde\lambda_{3,i}v_{0,i}}{\left(\tilde\lambda_{1,i} - \tilde\lambda_{2,i}\right)\left(\tilde\lambda_{2,i} - \tilde\lambda_{3,i}\right)}, \tag{2.17}$$

$$\tilde m_{3,i} = \frac{v_{2,i} - \tilde\lambda_{1,i}v_{1,i} - \tilde\lambda_{2,i}v_{1,i} + \tilde\lambda_{1,i}\tilde\lambda_{2,i}v_{0,i}}{\left(\tilde\lambda_{1,i} - \tilde\lambda_{3,i}\right)\left(\tilde\lambda_{2,i} - \tilde\lambda_{3,i}\right)}. \tag{2.18}$$

Therefore, we obtain a family of three-term vector estimates $\tilde\varphi_{n,k,\ell}$ for the function vector moment $\mathbf{v_f}$.

PROPOSITION 2.11. *A family of three-term vector estimates $\{\tilde\varphi_{n,k,\ell},\ n,k,\ell\in\mathbb{Z}\}$ for the function vector moment $\mathbf{v_f}$ is given by*

$$(\tilde\varphi_{n,k,\ell})_i = f(\tilde\lambda_{1,i})\tilde m_{1,i} + f(\tilde\lambda_{2,i})\tilde m_{2,i} + f(\tilde\lambda_{3,i})\tilde m_{3,i}, \qquad n,k,\ell\in\mathbb{Z}, \tag{2.19}$$

where $\tilde\lambda_{1,i}$, $\tilde\lambda_{2,i}$, $\tilde\lambda_{3,i}$, $\tilde m_{1,i}$, $\tilde m_{2,i}$, $\tilde m_{3,i}$ are defined by (2.13), (2.14), (2.15), (2.16), (2.17), (2.18)*, respectively, for $i = 1,2,\dots,p$.*

REMARK 2.12. If for some i, it holds that $v_{n+1,i} = v_{n+k+2,i} = v_{n+\ell+3,i} = 0$ or $(\hat\rho_{n,n+k})_i = (\hat\rho_{n+\ell+2,n-1})_i = (\hat\rho_{n+k+1,n+\ell+1})_i = 1$, then the formula (2.19) does not provide estimates for the specific selection of the triplet (n,k,ℓ) of parameters.

3. Estimates for $W^Tf(A)U$. In this section, we are interested in approximately evaluating expressions of the form

$$W^Tf(A)U,$$

where $A \in \mathbb{R}^{p\times p}$ is a diagonalizable matrix, f is a function defined on the spectrum of the matrix A, and

$$\begin{aligned} W &= \begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \dots & \mathbf{w}_m \end{bmatrix}, \qquad \mathbf{w}_i \in \mathbb{R}^p, \quad i=1,2,\dots,m, \\ U &= \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_m \end{bmatrix}, \qquad \mathbf{u}_i \in \mathbb{R}^p, \quad i=1,2,\dots,m, \quad \text{with } 1\le m \ll p. \end{aligned}$$

There are also approaches for approximating the form $W^Tf(A)U$, where the block W^TU is of full rank. In [12, 18] the form $W^Tf(A)U$ is estimated based on block Gauss and block anti-Gauss quadrature rules. The initial matrices are assumed to be biorthogonal, i.e., $W^TU = I_m$, or if not, then the matrices W, U can be biorthogonalized by the SVD. The present extrapolation-based method works for arbitrary matrices U, W, and biorthogonality is not required. We write the general form as

$$\begin{aligned} W^Tf(A)U &= \begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \dots & \mathbf{w}_m \end{bmatrix}^T f(A) \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_m \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{w}_1^T \\ \mathbf{w}_2^T \\ \vdots \\ \mathbf{w}_m^T \end{bmatrix} \begin{bmatrix} f(A)\mathbf{u}_1 & f(A)\mathbf{u}_2 & \dots & f(A)\mathbf{u}_m \end{bmatrix}. \end{aligned}$$

We notice that each matrix vector product $f(A)\mathbf{u}_i$, $i = 1, 2, \dots, m$, can be approximately found by the one-term, two-term, and three-term vector estimates φ_z in (2.6), $\hat{\varphi}_{n,k}$ in (2.11), and $\tilde{\varphi}_{n,k,\ell}$ in (2.19), respectively. Specifically, we can formulate the following proposition:

PROPOSITION 3.1. *Families of one-term, two-term, and three-term vector estimates for the quantity $W^Tf(A)U$ can be obtained by applying to each entry $W^Tf(A)\mathbf{u}_i$ the one-term, two-term, and three-term vector estimates $W^T\varphi_z$, $W^T\hat{\varphi}_{n,k}$, $W^T\tilde{\varphi}_{n,k,\ell}$, respectively.*

REMARK 3.2. In the same way, we can evaluate matrix functionals of the form $\mathbf{x}^Tf(A)\mathbf{y}$, where $\mathbf{x}, \mathbf{y}$ are real vectors of length p. This approach approximates matrix functionals of this form without using the polarization identity, which was employed in other works such as in [13, 14, 15].

4. Numerical implementation. In this section, we discuss the implementation of our estimates by applying them to certain products $f(A)\mathbf{b}$ and matrix forms $W^Tf(A)U$. First, we analyze the complexity of the obtained formulae.

4.1. Computation complexity of the estimates. The computational complexity of the vector estimates for the quantity $f(A)\mathbf{b}$ and the form $W^Tf(A)U$ is presented in Table 4.1. In particular, the family of one-term vector estimates φ_z for $f(A)\mathbf{b}$ requires the calculation of only two matrix-vector products (mvps). Moreover, the complexity of the families of two-term $\hat{\varphi}_{n,k}$ and three-term $\tilde{\varphi}_{n,k,\ell}$ vector estimates depends on the values of the parameters. The computational complexity of the family $\tilde{\varphi}_{n,k,\ell}$ is of order $\mathcal{O}\left(\mu p^2\right)$, where $\mu = \max\{n+k+3, n+\ell+4\}$, for a dense matrix A of order p. In case that the given matrix is banded with bandwidth s, the complexity is of order $\mathcal{O}(sp)$. The computational complexity of the vector estimates for the form $W^Tf(A)U$ for thin matrices $W, U \in \mathbb{R}^{p\times m}$, $1 \le m \ll p$, depends on m, and the order is $\mathcal{O}\left(mp^2\right)$. When the given matrix A is banded with bandwidth s, the complexity is of order $\mathcal{O}(msp)$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 4.1
Computational complexity of the vector estimates for $f(A)\mathbf{b}$ and $W^T f(A)U$.

Matrix A	$f(A)\mathbf{b}$		
	φ_z	$\hat{\varphi}_{n,k}$	$\tilde{\varphi}_{n,k,\ell}$
dense	$\mathcal{O}\left(2p^2\right)$	$\mathcal{O}\left((n+k+2)p^2\right)$	$\mathcal{O}\left(\mu p^2\right)$
banded	$\mathcal{O}\left(2sp\right)$	$\mathcal{O}\left((n+k+2)sp\right)$	$\mathcal{O}\left(\mu sp\right)$
	$W^T f(A)U$		
	$W^T\varphi_z$	$W^T\hat{\varphi}_{n,k}$	$W^T\tilde{\varphi}_{n,k,\ell}$
dense	$\mathcal{O}\left(2mp^2\right)$	$\mathcal{O}\left(m(n+k+2)p^2\right)$	$\mathcal{O}\left(m\mu p^2\right)$
banded	$\mathcal{O}\left(2msp\right)$	$\mathcal{O}\left(m(n+k+2)sp\right)$	$\mathcal{O}\left(m\mu sp\right)$

TABLE 4.2
Relative errors for approximating $f(A)\mathbf{b}$ using the one-term vector estimate φ_0.

function $f(A)$	A^{-1}	$\exp(A)$	$\sqrt{A}$
Rel. Error	8.6856e-17	1.3646e-15	5.4705e-15

For the families of two-term and three-term vector estimates $\{\hat{\varphi}_{n,k}, n, k \in \mathbb{Z}\}$ and $\{\tilde{\varphi}_{n,k,\ell}, n, k, \ell \in \mathbb{Z}\}$, a good choice for the pair or triplet of parameters are integers of small values resulting in an inexpensive computation as seen in Table 4.1. By setting the first parameter, for example, equal to 1, we can vary the values of the other three parameters, usually from zero to 2 or 3 and compute a range of values for the estimates. By taking the mean value of these approximations, we obtain a better estimation of the desired quantity.

4.2. Numerical examples. Subsequently, we present several numerical examples that illustrate the performance of the derived families of vector estimates. All computations were performed using MATLAB (R2015a) 64-bit on an Intel Core i7 computer with 16 Gb DDR4 RAM. In all the examples, we present tables with the relative errors for the approximation by the families of vector estimates. The exact values are determined by evaluating the matrix $f(A)$ using MATLAB matrix functions such as expm, sqrtm, logm, etc. The sign function of A is computed by the formula $\text{sign}(A) = A(A^2)^{-1/2}$. Throughout the examples, $rand(p, 1)$ denotes a random vector of length p with positive elements which is drawn from the uniform distribution, and $randn(p, 1)$ denotes a random vector of length p which is drawn from the normal distribution.

Example 1: One-term vector estimates for $\boldsymbol{f(A)}\mathbf{b}$. We test a symmetric, positive definite, and orthogonal matrix $A = B^T B$ of order $p = 700$, where B is the nearly orthogonal symmetric eigenvector matrix of the second-order difference matrix, which is generated by the MATLAB gallery function $B = gallery('orthog', 700)$*. The elements of B are given by $B_{ij} = \sqrt{\frac{2}{p+1}}\sin(\frac{ij\pi}{p+1})$. Let $\mathbf{b}$ be a vector of order $p = 700$ which is drawn from the normal distribution. We estimate $f(A)\mathbf{b}$ for three different functions f by using the one-term vector estimate φ_0. Since the matrix A is orthogonal, the optimal value for the family parameter is $z = 0$. This is an extension of the result in [14, Remark 4]. The results are reported in Table 4.2.

Example 2: Two-term vector estimates for $\boldsymbol{A^{1/2}}\mathbf{b}$. We consider the nonsymmetric matrix $A = dw256B$ of order $p = 512$, which arises in electromagnetic problems and is selected from the SuiteSparse Matrix Collection [11]. This matrix is diagonalizable with

*The MATLAB gallery: http://www.mathworks.com/help/matlab/ref/gallery.html

TABLE 4.3
Relative errors for approximating $A^{1/2}\mathbf{b}$ for the matrix $A = dw256B$ of order $p = 512$ by the family of two-term vector estimates for various values of the parameters n and k.

(n,k)	Rel. Error [$\mathbf{b} = rand(512{,}1)$]	Rel. Error [$\mathbf{b} = randn(512{,}1)$]
(1,0)	9.9679e-4	7.7755e-4
(1,1)	1.9075e-3	1.1618e-3
(1,3)	3.6696e-3	1.3892e-3
(1,-2)	3.6671e-4	8.3626e-4
(0,4)	3.5097e-3	4.4683e-3
MRE	2.0901e-3	1.7266e-3

TABLE 4.4
Relative errors for approximating $\exp(A)\mathbf{b}$ for the Poisson matrix of order $p = 1600$ by the family of three-term vector estimates for various values of the parameters n, k and ℓ.

(n,k,ℓ)	Rel. Error [$b_i = tan(i)$]	Rel. Error [$\mathbf{b} = rand(1600, 1)$]
(1,0,0)	1.6454e-6	1.6563e-10
(1,2,0)	2.8671e-4	5.8612e-5
(1,0,3)	4.0753e-8	6.7314e-6
(1,2,5)	1.9537e-6	4.5347e-6
(1,2,3)	9.2279e-2	1.5779e-5

positive eigenvalues and well conditioned ($\kappa_2(A) = 3.7328$). We estimate the quantity $A^{1/2}\mathbf{b}$ for two different vectors $\mathbf{b}$. We have chosen the vector $\mathbf{b}$ to be drawn from the uniform distribution (second column of Table 4.3) and also from the normal distribution (third column of Table 4.3). The approximation of these quantities is done by the family of two-term vector estimates $\hat{\varphi}_{n,k}$ (2.11) for several choices of the parameters n and k. The results are presented in Table 4.3.

As we can notice in Table 4.3, by using the two-term vector estimates $\hat{\varphi}_{n,k}$ for different n, k, we achieve relative errors of order $\mathcal{O}(10^{-3})$ or $\mathcal{O}(10^{-4})$. In the last row of this table, we record the mean relative error (MRE) of the derived estimates for different n, k.

Example 3: Three-term vector estimates for $\exp(A)\mathbf{b}$. We consider the Poisson matrix of order $p = 1600$ [15]. This matrix is symmetric, positive definite, block tridiagonal (sparse) with condition number $\kappa_2(A) = 6.8062e2$ and arises from the five-point finite difference approximation of the Poisson equation on a unit square with an $m \times m$ mesh, $m = \sqrt{p} = 40$. The Poisson matrix is of the form $A = tridiag(-I_m, T_m, -I_m)$, where each block $T_m = tridiag(-1, 4, -1)$ has dimension $m = 40$. The matrix A in this test example is the Poisson matrix multiplied by a factor of 0.02, which is obtained by the MATLAB gallery function $A = 0.02 * gallery('poisson', 40)$. Also, we choose two different vectors $\mathbf{b}$, namely, we consider the vector $\mathbf{b}$ such that the ith entry of $\mathbf{b}$ is equal to the tangent of the corresponding index i, i.e., $b_i = tan(i)$, $i = 1, 2, \ldots, p$ (second column of Table 4.4). In the second case, the vector $\mathbf{b}$ is drawn from the uniform distribution (third column of Table 4.4).

We approximate the quantity $\exp(A)\mathbf{b}$ for these two vectors $\mathbf{b}$. The estimation of these quantities is done by the family of three-term vector estimates $\tilde{\varphi}_{n,k,\ell}$ (2.19) for several choices of the parameters n, k, and ℓ. The results are presented in Table 4.4.

As we can notice in Table 4.4, by using the three-term vector estimates $\tilde{\varphi}_{n,k,\ell}$ for different values of the parameters n, k, ℓ, we can achieve satisfactory relative errors. In particular, the

TABLE 4.5
Relative errors for approximating $f(A)\mathbf{b}$ by the families of two-term and three-term vector estimates.

matrix A	vector $\mathbf{b}$	function $f(A)$	$\hat{\varphi}_{1,0}$	$\tilde{\varphi}_{1,0,0}$
kms(800,0.2)	$rand$(800,1)	log(A)	1.8984e-2	8.2597e-4
kms(800,0.2)	$randn$(800,1)	sqrt(A)	1.3775e-3	4.8433e-5
$rand$(300)/100	$rand$(300,1)	exp(A)	3.0348e-2	5.8355e-8
$rand$(300)/100	**1**	sqrt(A)	1.3289e-2	1.3277e-2
ex1	**1**	exp(A)	2.4133e-8	8.3008e-12
ex1	$randn$(216,1)	exp(A)	5.3671e-7	2.8318e-11
Chem97ZtZ	**1**	sqrt(A)	7.2613e-3	6.4037e-3
Trefethen_500	$randn$(500,1)	log(A)	6.8993e-3	6.0093e-3
Trefethen_500	$rand$(500,1)	sign(A)	4.1969e-4	6.4524e-4

resulting relative errors are of order $\mathcal{O}(10^{-8})$ and $\mathcal{O}(10^{-10})$ if the initial vector $\mathbf{b}$ is a random vector or $b_i = tan(i)$, $i = 1, 2, \ldots, p$.

Example 4: Comparison of two-term and three-term vector estimates. In this example, we compare the difference in performance between the two-term and the three-term vector estimates. In particular, we approximate the quantity $f(A)\mathbf{b}$ for several matrices A, vectors $\mathbf{b}$, and different functions f. We evaluate $f(A)\mathbf{b}$ by the two-term vector estimate $\hat{\varphi}_{1,0}$ and the three-term vector estimate $\tilde{\varphi}_{1,0,0}$. We test with the kms matrix of order $p = 800$. This matrix is symmetric, positive definite, and Toeplitz with elements $A_{ij} = 0.2^{|i-j|}$, and it can be found in the MATLAB gallery. Also, we test with the diagonalizable matrix $A \in \mathbb{R}^{300\times 300}$ whose elements are uniformly distributed [18]. This matrix is nonsymmetric and indefinite with complex eigenvalues. It is generated by the MATLAB command $A = rand(300)/100$. The other test matrices used are selected from the SuiteSparse Matrix Collection [11]. The matrices ex1 (nonsymmetric) and Chem97ZtZ (symmetric) are diagonalizable with dimension $p = 216$ and $p = 2541$, respectively. The last test matrix is Trefethen_500, which is a symmetric, positive definite matrix of order $p = 500$. The results are presented in Table 4.5. In particular, in the last two columns of Table 4.5, we report the relative errors that are obtained with the two-term vector estimate $\hat{\varphi}_{1,0}$ and the three-term vector estimate $\tilde{\varphi}_{1,0,0}$, respectively.

In Table 4.5, we observe that the three-term vector estimates $\tilde{\varphi}_{1,0,0}$ yield better relative errors than the two-term version $\hat{\varphi}_{1,0}$.

Example 5: An application of $A^{1/2}\mathbf{b}$. We consider a covariance matrix A of order $p = 2000$ whose entries are $A_{ii} = 1 + i^{\alpha}$ and $A_{ij} = \frac{1}{|i-j|^{\beta}}$, where $\alpha, \beta \in \mathbb{R}$. We denote this matrix as $A = covariance(p, \alpha, \beta)$. It is symmetric, positive definite of the form $A = XX^T$, where X is the data matrix. In statistics, one of the most common problems concerns sampling from a multivariate Gaussian distribution with a positive definite covariance matrix A [9]. In this kind of problems, the product $A^{1/2}\mathbf{b}$ appears, where $\mathbf{b} = randn(2000, 1)$. We test the behaviour of the vector estimates $\hat{\varphi}_{1,0}$ and $\tilde{\varphi}_{1,0,0}$ for different covariance matrices by varying the values of the parameters α, β. The results are presented in Table 4.6. As we can notice in this table, the behaviour of the estimates is almost the same for the tested covariance matrices, and the corresponding relative errors of these estimates are satisfactory.

Example 6: Estimating the form $W^T \exp(A)U$, $W^TU = I_m$. We consider the nonsymmetric, diagonalizable matrix $A = rand(300)/100$ as described in Example 4. We approximate the form

$$W^T \exp(A)U,$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 4.6

Relative errors for approximating $A^{1/2}\mathbf{b}$ for covariance matrices by the families of two- and three-term vector estimates.

(α,β)	$\hat{\varphi}_{1,0}$	$\tilde{\varphi}_{1,0,0}$
$(1,1)$	9.8098e-4	2.3706e-4
$(1,3/4)$	2.5919e-4	2.3214e-4
$(1/2,4)$	1.2601e-4	1.0777e-5

TABLE 4.7

Relative errors for approximating the form $W^T\exp(A)U$ with A nonsymmetric and $W^TU=I_2$.

$\hat{\varphi}_{n,k}$		$\tilde{\varphi}_{n,k,\ell}$	
(n,k)	Rel. Error	(n,k,ℓ)	Rel. Error
$(1,0)$	4.6408e-5	$(1,0,0)$	9.0910e-9
$(1,1)$	9.0938e-5	$(1,-4,0)$	4.7983e-6
$(1,2)$	8.3419e-5	$(1,0,1)$	2.5599e-8
$(1,3)$	8.3175e-5	$(1,0,2)$	2.5774e-8
$(1,7)$	8.3169e-5	$(1,7,3)$	1.6071e-6

where [18]

$$W=\begin{bmatrix} e_1, & 2e_1+3e_2 \end{bmatrix}\in\mathbb{R}^{300\times 2},\quad U=\begin{bmatrix} e_1-\frac{2}{3}e_2, & \frac{1}{3}e_2 \end{bmatrix}\in\mathbb{R}^{300\times 2},\quad W^TU=I_2.$$

In this example, we use the family of two-term vector estimates $W^T\hat{\varphi}_{n,k}$ (2.11) and the family of three-term vector estimates $W^T\tilde{\varphi}_{n,k,\ell}$ (2.19) for several parameter values. In Table 4.7, the corresponding relative errors are reported. As we can notice in Table 4.7, a fair accuracy can be achieved by using either the two-term or the three-term vector estimates. Nevertheless, the family of three-term vector estimates gives better results. Specifically, the order of the relative errors for $W^T\hat{\varphi}_{n,k}$ is $\mathcal{O}(10^{-5})$, and the order of the relative errors for $W^T\tilde{\varphi}_{n,k,\ell}$ varies from $\mathcal{O}(10^{-6})$ to $\mathcal{O}(10^{-9})$.

Example 7: Estimating the form $\boldsymbol{W^Tf(A)U}$, $\boldsymbol{W^TU\neq I_m}$. This example is similar to one in [18]. Let $A\in\mathbb{R}^{100\times 100}$ be a symmetric, positive definite matrix with equidistantly-spaced eigenvalues in the interval $[\frac{5}{2},4]$. We generate this matrix by the relation $A=Q^TDQ$, where D is the diagonal matrix which contains the given equidistant eigenvalues of A and Q is an orthogonal matrix obtained by the MATLAB command $Q=orth(randn(p,p))$. We are interested in approximating the form $W^Tf(A)U$, where

$$W=\begin{bmatrix} 3e_1, & \frac{5}{2}e_2+e_3 \end{bmatrix}\in\mathbb{R}^{100\times 2},\ U=\begin{bmatrix} e_1+\frac{1}{2}e_3, & e_2+e_3 \end{bmatrix}\in\mathbb{R}^{100\times 2},\ W^TU\neq I_2.$$

We test five different functions f. We approximate this form by the two-term vector estimate $W^T\hat{\varphi}_{1,0}$ and the three-term vector estimate $W^T\tilde{\varphi}_{1,0,0}$. We report the corresponding relative errors in Table 4.8. As can be noticed from this table, we can achieve satisfactory relative errors by both the two-term and three-term estimates.

4.3. Comparison with other methods. In this section we compare the behaviour of the derived families of vector estimates for the matrix vector product $f(A)\mathbf{b}$ and the form $W^Tf(A)U$. In particular, we compare the proposed approximations of the product $f(A)\mathbf{b}$ with the estimates obtained by the Arnoldi method [16] or a polynomial approximation which is described in [9]. Also, we evaluate the form $W^Tf(A)U$ approximately by using the block

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 4.8
Relative errors for approximating the form $W^T f(A)U$ with A symmetric positive definite and $W^T U \neq I_2$.

$f(A)$	$\hat{\varphi}_{1,0}$	$\tilde{\varphi}_{1,0,0}$
A^{-1}	2.6614e-4	2.7289e-5
$\exp(A)$	1.1694e-3	9.8659e-6
$\log(A)$	5.8018e-5	2.7359e-6
$\sqrt{A}$	1.0499e-5	3.3539e-7
sign(A)	4.3491e-16	1.4225e-12

TABLE 4.9
Relative errors for approximating $f(A)\mathbf{b}$ by the family of three-term vector estimates and the derived estimates via the Arnoldi algorithm.

matrix A	$\mathbf{b}$	$f(A)$	Arnoldi approx.	$\tilde{\varphi}_{1,0,0}$	speedup
ex1	$rand$	$\exp(A)$	1.8002e-12 ($k_a = 6$)	1.0731e-11	1.2
ex1	$randn$	$\sqrt(A)$	3.0297e-2 ($k_a = 10$)	2.9920e-2	2
dw256B	$randn$	$\log(A)$	2.1910e-2 ($k_a = 3$)	6.0449e-2	0.6
covariance(100,1,1)	$randn$	$\sqrt{A}$	2.1202e-4 ($k_a = 14$)	2.3245e-4	2.8
(kms,100,0.2)	$randn$	$\exp(A)$	3.6915e-6 ($k_a = 5$)	1.4288e-6	1

Gauss and anti-Gauss algorithm which is described in [12]. The Matlab implementation of this method can be found in `http://bugs.unica.it/~gppe/soft/#blgaussexp`.

Example 8: $f(A)\mathbf{b}$: Arnoldi method vs. extrapolation. In this example, we compare the vector estimates for $f(A)\mathbf{b}$ with the approximations by the Arnoldi algorithm. The results are reported in Table 4.9, and the matrices in this table are described in the above examples. Specifically, we record the number k_a of required iterations of the Arnoldi algorithm to achieve the same order of relative errors as those of the three-term vector estimates $\tilde{\varphi}_{1,0,0}$. As we can see in this table, the number of Arnoldi iterations depends on the matrix, for example, the covariance matrix with parameters $\alpha = \beta = 1$ requires 14 iterations to achieve almost the same accuracy as the extrapolation estimate. Nevertheless, it is worth mentioning that the Arnoldi method can achieve better approximations as the number of the iterations increases, but then the complexity increases as well. In the last column of Table 4.9 we record the speedup which is defined as the ratio of the number of required mvps of the Arnoldi approximation, i.e., k_a mvps, divided by the number of mvps for $\tilde{\varphi}_{1,0,0}$, i.e., 5 mvps. The recorded value expresses the speedup of the extrapolation method.

Example 9: $f(A)\mathbf{b}$: Polynomial approximation vs. extrapolation. In this example we compare the behaviour of the vector estimates with a polynomial approximation of $f(A)\mathbf{b}$ for symmetric positive definite matrices that was introduced by Chen et al. in [9]. Specifically, we estimate the product $f(A)\mathbf{b}$ for two different functions f with $\mathbf{b} = randn(p, 1)$ using the two-term $\hat{\varphi}_{1,0}$ and the three-term $\tilde{\varphi}_{1,0,0}$ vector estimates. Since $\mathbf{b}$ is a random vector, we run the algorithm for computing $\hat{\varphi}_{1,0}$, $\tilde{\varphi}_{1,0,0}$ ten times, and we calculate the mean relative error. We compare the results of these vector estimates with the corresponding relative errors reported in Tables 6.1 and 6.4 in [9] for some symmetric, positive definite matrices from the SuiteSparse Matrix Collection [11].

As we can see in Table 4.10, the proposed extrapolation procedure and the polynomial approximation achieve almost the same order of accuracy, but the polynomial approximation requires $k_p = 200$ iterations for these matrices. It is worth mentioning that this polynomial approximation yields very satisfactory estimates for the product $\exp(A)\mathbf{b}$, [9].

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 4.10
Relative errors for approximating $f(A)\mathbf{b}$ by the families of two- and three-term vector estimates and a polynomial approximation.

matrix A	$f(A)$	polyn. approx.	$\hat{\varphi}_{1,0}$	$\tilde{\varphi}_{1,0,0}$
Trefethen_2000	$\log(A)$	1.8060e-4	3.3217e-3	3.1668e-3
plbuckle	$\log(A)$	1.0433e-2	6.9828e-2	5.4545e-2
plbuckle	$\sqrt{A}$	2.87e-4	6.8554e-2	3.1275e-2
nasa1824	$\sqrt{A}$	1.26e-3	1.2821e-1	9.5264e-2
nasa1824	$\log(A)$	2.6332e-2	1.4220e-1	1.3022e-1

TABLE 4.11
Relative errors for approximating $W^T f(A)U$ by the families of two- and three-term vector estimates and block Gauss/anti-Gauss quadrature.

matrix A	block (anti) Gauss	$W^T\hat{\varphi}_{1,0}$	$W^T\tilde{\varphi}_{1,0,0}$
$rand$(300)/100	3.1969e-6 [8.8105e-3 sec]	3.0993e-4 [1.1209e-2 sec]	5.0143e-9 [9.5020e-3 sec]
$A = QDQ^T$	1.8063e-6 [9.2280e-3 sec]	1.1683e-3 [4.3060e-3 sec]	5.9752e-6 [9.4030e-3 sec]

Example 10: $W^T\exp(A)U$: Gauss/anti-Gauss quadrature vs. extrapolation. We approximate the form $W^T\exp(A)U$, where

$$W = \begin{bmatrix} e_1, & 2e_1 + 3e_2 \end{bmatrix} \in \mathbb{R}^{p\times 2} \quad \text{and} \quad U = \begin{bmatrix} e_1 - \frac{2}{3}e_2, & \frac{1}{3}e_2 \end{bmatrix} \in \mathbb{R}^{p\times 2},$$

by using block Gauss and anti-Gauss quadrature and the proposed extrapolation procedure. We test with two matrices in order to compare the behaviour of the estimates. In particular, we use the nonsymmetric, diagonalizable matrix $rand(300)/100$ from Example 4 and the symmetric matrix $A = Q^T DQ$ with fixed eigenvalues as described in Example 7. The results are presented in Table 4.11. The execution time in seconds is given in brackets in this table. As we can see from the results, we achieve a fair accuracy with both the methods but the accuracy depends on the matrix, too. The approach based on block Gauss and anti-Gauss quadrature generally provides accurate values in short time. The proposed extrapolation procedure attains fair estimates in comparable execution time as well. We notice that in some cases the three-term vector estimates yield a better accuracy in shorter time.

5. Concluding remarks. In the present paper, we developed families of one-term φ_z, two-term $\hat{\varphi}_{n,k}$, and three-term vector estimates $\tilde{\varphi}_{n,k,\ell}$ for the action of the matrix $f(A)$ on a given vector $\mathbf{b} \in \mathbb{R}^p$, i.e., $f(A)\mathbf{b}$, for any diagonalizable matrix $A \in \mathbb{R}^{p\times p}$. The family of one-term vector estimates φ_z requires only two mvps, but the computation depends on the knowledge of an appropriate value $z \in \mathbb{C}$, which is difficult to determine. On the other hand, the families of two-term $\{\hat{\varphi}_{n,k}, n, k \in \mathbb{Z}\}$ and three-term $\{\tilde{\varphi}_{n,k,\ell}, n, k, \ell \in \mathbb{Z}\}$ vector estimates require more mvps with a complexity of quadratic order, but they allow for a broader selection range of parameters values and attain very good relative errors. As it was shown in the examples, good choices for the pair of parameters (n, k) in the two-term family $\hat{\varphi}_{n,k}$ and for the triplet of parameters (n, k, ℓ) in the three-term family $\tilde{\varphi}_{n,k,\ell}$, are small integers, which also provide low complexity.

Moreover, the generated families of vector estimates for $f(A)\mathbf{b}$ were applied for approximations of $W^T f(A)U$ for any matrices $U, W \in \mathbb{R}^{p\times m}$, $1 \le m \ll p$, not necessarily biorthogonal. Furthermore, the described approach approximates matrix functionals of the form $\mathbf{x}^T f(A)\mathbf{y}, \mathbf{x}, \mathbf{y} \in \mathbb{R}^p$ without requiring the polarization identity with half the complexity of comparable methods. In most other approaches [13, 14, 15], the polarization identity is always employed. The computed examples illustrate the effectiveness of the produced vector estimates for $f(A)\mathbf{b}$ and the form $W^T f(A)U$.

In conclusion, the advantage of the extrapolation vector estimates, when compared with other methods, is that they are cheaper to evaluate and they can be easily implemented in vectorized form. Their use is attractive when a direct computation is expensive and for problems where high accuracy is not required.

The application of the proposed families of estimates in a parallel environment, which could improve the speed of the computations, is under consideration.

Acknowledgments. The authors would like to thank professor Michele Benzi for useful advice and comments. The authors are grateful to the reviewers of this paper whose valuable remarks improved significantly this work.

The work of the second author (P. R.) was implemented with a scholarship from State Scholarships Foundation (IKY) which was funded by the Act ”Scholarship Program for postgraduate studies of a second cycle of studies“ from resources of the OP ”Human Resources Development, Education and Lifelong Learning“, 2014–2020 with the cofinancing of the European Social Fund (ESF) and the Greek State.

REFERENCES

[1] E. J. ALLEN, J. BAGLAMA, AND S. K. BOYD, *Numerical approximation of the product of the square root of a matrix with a vector*, Linear Algebra Appl., 310 (2000), pp. 167–181.
[2] F. ARRIGO AND M. BENZI, *Updating and downdating techniques for optimizing network communicability*, SIAM J. Sci. Comput., 38 (2016), pp. B25–B49.
[3] M. BENZI AND C. KLYMKO, *Total communicability as a centrality measure*, J. Complex Netw., 1 (2013), pp. 124–149.
[4] C. BREZINSKI, *Error estimates for the solution of linear systems*, SIAM J. Sci. Comput., 21 (1999), pp. 764–781.
[5] C. BREZINSKI, P. FIKA, AND M. MITROULI, *Moments of a linear operator on a Hilbert space, with applications to the trace of the inverse of matrices and the solution of equations*, Numer. Linear Algebra Appl., 19 (2012), pp. 937–953.
[6] ———, *Estimations of the trace of power of positive self-adjoint operators by extrapolation of the moments*, Electron. Trans. Numer. Anal., 39 (2012), pp. 144–155. http://etna.ricam.oeaw.ac.at/vol.39.2012/pp144-155.dir/pp144-155.pdf
[7] C. BREZINSKI, G. RODRIGUEZ, AND S. SEATZU, *Error estimates for linear systems with applications to regularization*, Numer. Algorithms, 49 (2008), pp. 85–104.
[8] ———, *Error estimates for the regularization of least squares problems*, Numer. Algorithms, 51 (2009), pp. 61–76.
[9] J. CHEN, M. ANITESCU, AND Y. SAAD, *Computing $f(A)b$ via least squares polynomial approximations*, SIAM J. Sci. Comput., 33 (2011), pp. 195–222.
[10] P. I. DAVIES AND N. J. HIGHAM, *Computing $f(A)b$ for matrix functions*, in QCD and Numerical Analysis III, A. Boriçi, A. Frommer, B. Joó, A. Kennedy, and B. Pendleton, eds., Lecture Notes in Computational Science and Engineering, 47, Springer, Berlin, 2005, pp. 15–24,.
[11] T. DAVIS AND Y. HU, *The SuiteSparse Matrix Collection*, https://sparse.tamu.edu/.
[12] 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.
[13] P. FIKA AND M. MITROULI, *Estimation of the bilinear form $y^* f(A)x$ for Hermitian matrices*, Linear Algebra Appl., 502 (2016), pp. 140–158.
[14] P. FIKA, M. MITROULI, AND P. ROUPA, *Estimates for the bilinear form $x^T A^{-1} y$ with applications to linear algebra problems*, Electron. Trans. Numer. Anal., 43 (2014/15), pp. 70–89. http://etna.ricam.oeaw.ac.at/vol.43.2014-2015/pp70-89.dir/pp70-89.pdf
[15] G. H. GOLUB AND G. MEURANT, *Matrices, Moments and Quadrature with Applications*, Princeton University Press, Princeton, 2010.
[16] N. J. HIGHAM, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, 2008.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[17] N. J. HIGHAM AND A. H. AL-MOHY, *Computing matrix functions*, Acta Numer., 19 (2010), pp. 159–208.
[18] L. REICHEL, G. RODRIGUEZ, AND T. TANG, *New block quadrature rules for the approximation of matrix functions*, Linear Algebra Appl., 502 (2016), pp. 299–326.
[19] B. N. SHEEHAN, Y. SAAD, AND R. B. SIDJE, *Computing* $\exp(-\tau A)b$ *with Laguerre polynomials*, Electron. Trans. Numer. Anal., 37 (2010), pp. 147–165.
http://etna.ricam.oeaw.ac.at/vol.37.2010/pp147-165.dir/pp147-165.pdf