



Approximation of the transfer function and of quadratic forms

Zdeněk Strakoš and Petr Tichý

Faculty of Mathematics and Physics, Charles University in Prague,
Institute of Computer Science, Czech Academy of Sciences
strakos@karlin.mff.cuni.cz, tichy@cs.cas.cz



Abstract

Model reduction in linear dynamical systems can be formulated (in a simplified form) as an approximation of the transfer function $T(\lambda) = c^*(\lambda I - A)^{-1}b$ using the reduced order matrices and vectors $A_n, I_n, c_n,$ and b_n . Approximation of the quadratic form $c^*A^{-1}b$ can seemingly be interpreted as using the previous approach with taking $\lambda = 0$. This, however, does not lead to efficient numerical algorithms for the second problem. We give a short overview of the existing approaches with emphasizing their computational efficiency and numerical stability properties.

Model reduction and transfer function

Consider a linear dynamical system

$$E \frac{dz}{dt} = Az(t) + bu(t),$$

$$y(t) = c^*z(t) + du(t),$$

where $A \in \mathbb{R}^{N \times N}, E \in \mathbb{R}^{N \times N}, b \in \mathbb{R}^N, c \in \mathbb{R}^N, d \in \mathbb{R}$ are given, $z(t) \in \mathbb{R}^N$ represents the inner variables of the system, $u(t)$ the one-dimensional input (control), and $y(t)$ the one-dimensional output. Typically, N is very large, and the basic idea of model reduction is to find an approximation by a system of the same type, but with the state-space dimension N reduced to a much smaller n . For simplicity, set $d = 0$ and $E = I$. The above description can then be replaced by

$$T(\lambda) = c^*(\lambda I - A)^{-1}b, \quad \lambda \in \mathbb{C}, \quad (1)$$

where $T(\lambda)$ is called the *transfer function*. In brief, the model reduction problem is to find $A_n, I_n, c_n,$ and b_n such that

$$T_n(\lambda) = c_n^*(\lambda I_n - A_n)^{-1}b_n, \quad \lambda \in \mathbb{C},$$

closely approximates (in an appropriate sense) $T(\lambda)$ within a given frequency range $\lambda \in \mathbb{C}_A \subset \mathbb{C}$; see [1].

A more general case

The problem of finding efficient numerical approximations to (1) arises in many applications beyond linear dynamical systems. A more general case can be written as

$$c^*F(A)b,$$

where F is a given function so that the matrix function $F(A)$ is defined. The particular case $b = c$ and $F(A) = (\lambda I - A)^{-1}$ is of a great importance in theory and practice of iterative computations, in physical chemistry and solid state physics.

Projections onto Krylov subspaces

Model reduction of linear dynamical systems based on projections onto Krylov subspaces is linked with *local approximations* of $T(\lambda)$. First consider the expansion of (1) around infinity, i.e.

$$T(\lambda) = \lambda^{-1}c^*(I - \lambda^{-1}A)^{-1}b$$

$$= \lambda^{-1}(c^*b) + \lambda^{-2}(c^*Ab) + \dots + \lambda^{-2n}(c^*A^{2n-1}b) + \dots$$

A reduced model of order n that matches the first $2n$ terms in the above expansion is known as the *minimal partial realization*.

One can also approximate (and therefore expand) $T(\lambda)$ in the neighbourhood of some finite $\lambda_0 \in \mathbb{C}$. For the simplified case $\lambda_0 = 0$ we get the expansion

$$-T(\lambda) = c^*A^{-1}(I - \lambda A^{-1})^{-1}b$$

$$= c^*A^{-1}b + \lambda(c^*A^{-2}b) + \dots + \lambda^{2n-1}(c^*A^{-2n}b) + \dots$$

(see [3, 6]). It is important to note that with the transfer function approximated in the neighbourhood of the origin, associated numerical methods that compute the model reduction are based on matching moments with the powers of A^{-1} . The computation therefore involves the *solution of the linear algebraic systems with the matrix A* . In comparison, in the model reduction based on the expansion of $T(\lambda)$ around infinity, the computation of the approximation needs only much cheaper matrix-vector multiplication with the matrix A .

Estimates in Quadratic forms

The transfer function $T(\lambda)$ gives, apart from the sign, for $\lambda = 0$ the quantity

$$c^*A^{-1}b \quad (2)$$

which in signal processing is called the *scattering amplitude*. Many applications require the approximation of this quantity. We will demonstrate that the requirements of efficiency and numerical stability of (finite precision) computations must always be taken into account. In particular, an approximation to the scalar value (2) *cannot* be efficiently computed by an application of the methods developed for approximation of the transfer function $T_n(\lambda)$. In spite of existing links, the task of computing an efficient and numerically stable approximation to (2) represents a different mathematical problem, and it requires a different approach from approximating the transfer function.

One can also suggest that $c^*A^{-1}b = c^*x$, where x solves the linear system $Ax = b$. Therefore c^*x can, in principle, be approximated by computing c^*x_n , where x_n is the n th approximation generated by a Krylov subspace method. A potential user of this approach should, however, be aware that the approximation of c^*x by the explicit numerical computation of c^*x_n can be highly inefficient even in the HPD case due to the (**hidden**) effects of rounding errors.

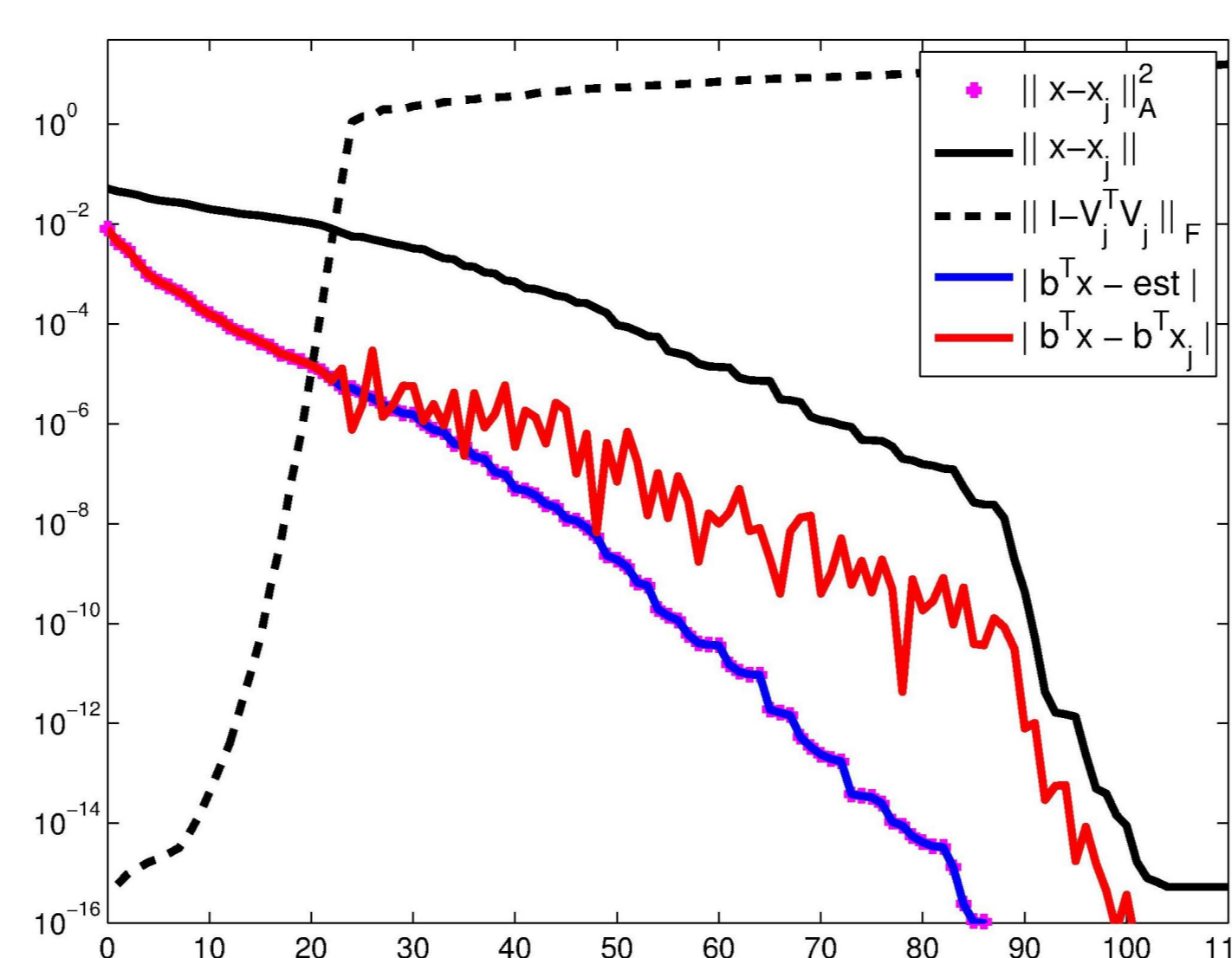
HPD case and the CG method

For A Hermitian and positive definite, and $c = b$ we have $b^*A^{-1}b = \|x\|_A^2$. In CG with $x_0 = 0$ it holds that

$$\|x\|_A^2 = b^T x_n + \|x - x_n\|_A^2,$$

$$\|x\|_A^2 = \sum_{j=0}^{n-1} \alpha_j \|r_j\|^2 + \|x - x_n\|_A^2.$$

Both formulas for approximating $\|x\|_A^2$ are mathematically equivalent, but in finite precision computations, they produce different results.

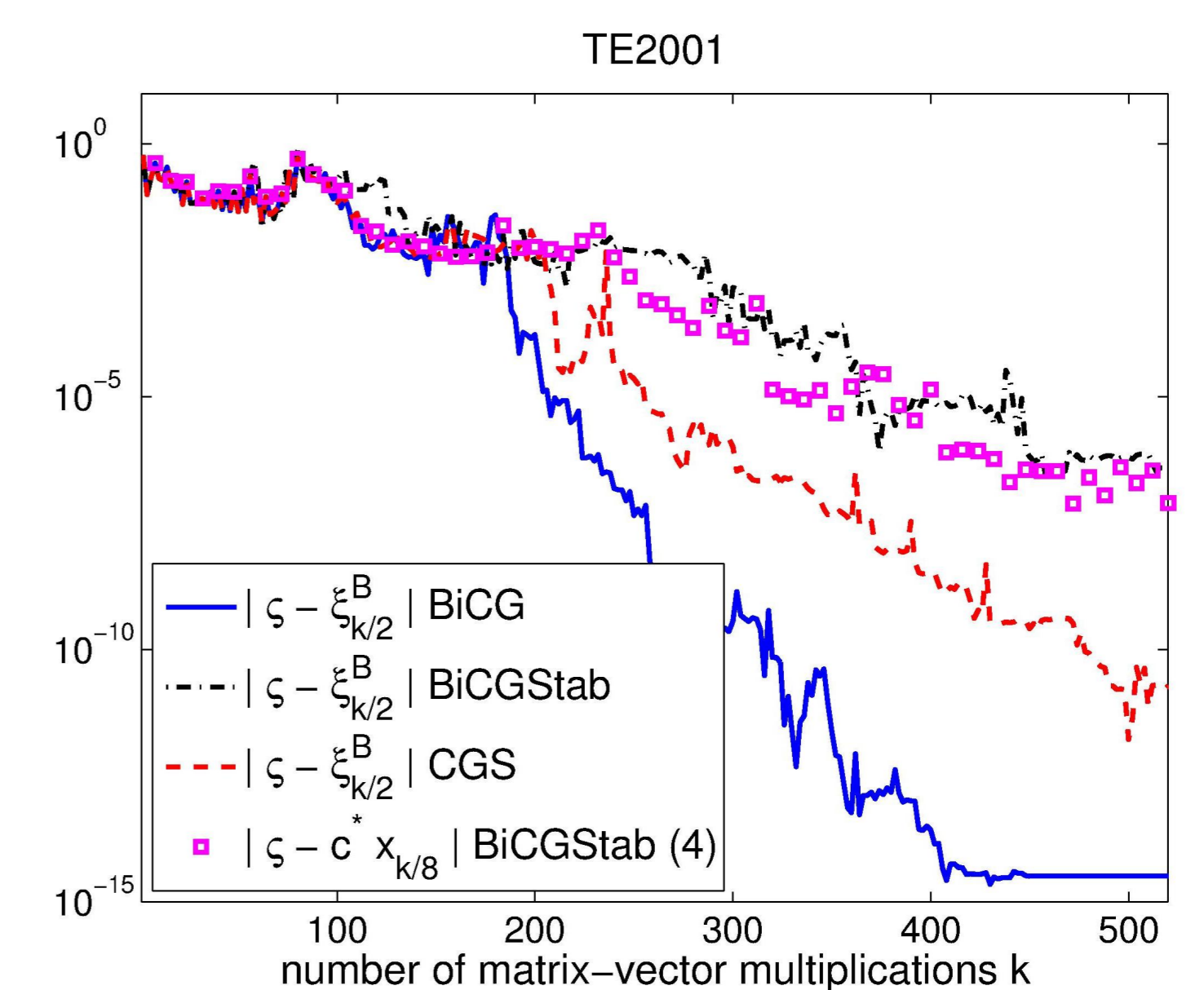


General case and comparison of methods

Similar phenomenon can be seen for non-Hermitian A and/or $b \neq c$, and the **BiCG method**. Here, the approximation is based on an analogous formula,

$$c^*A^{-1}b = \xi_n^B + s_n^*A^{-1}r_n, \quad \xi_n^B \equiv \sum_{j=0}^{n-1} \alpha_j s_j^* r_j.$$

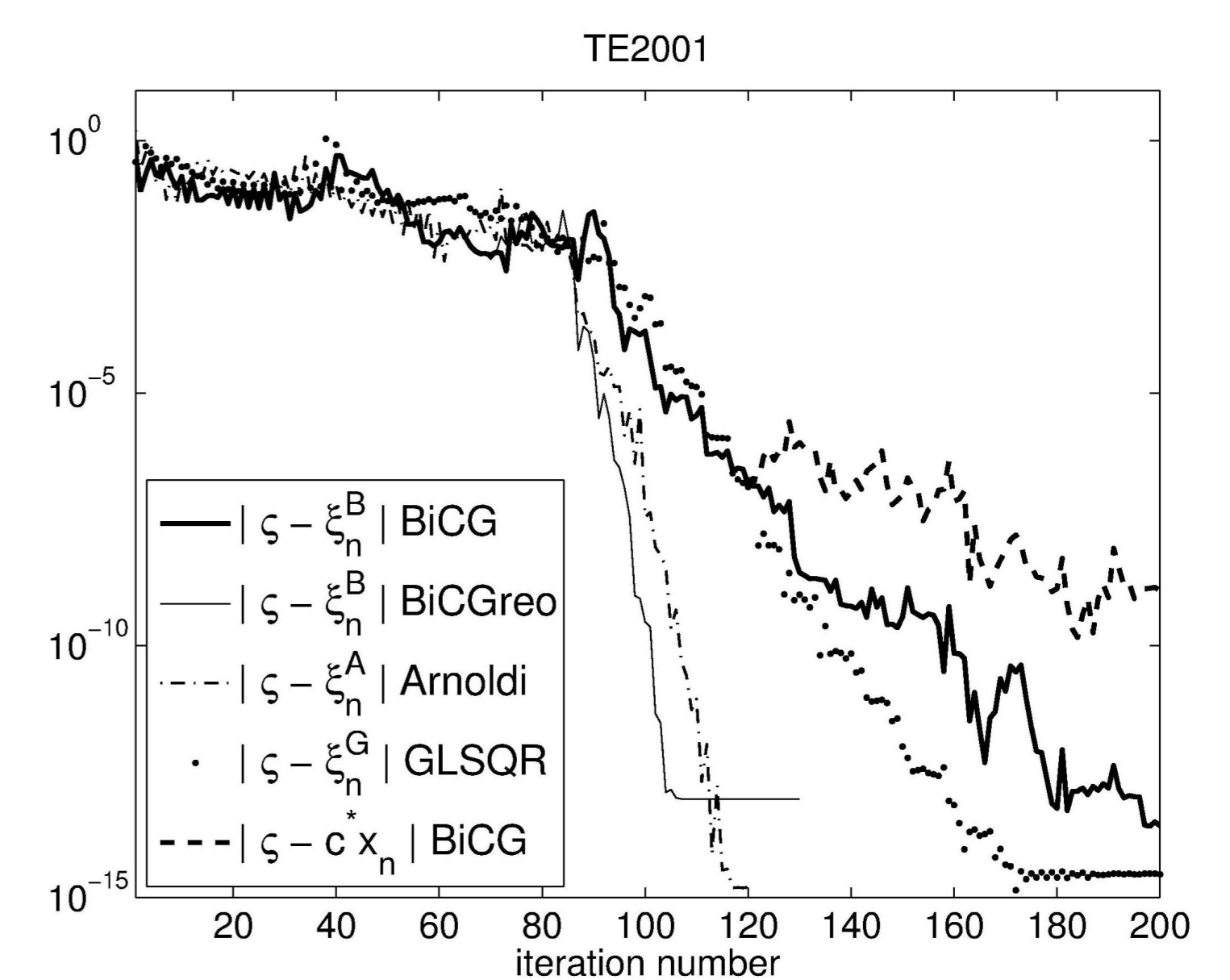
One can also use hybrid BiCG methods like CGS or BiCGStab(ℓ). While hybrid methods can be better than BiCG when approximating the solution of $Ax = b$, for approximating $c^*A^{-1}b$ BiCG is the clear winner in most cases. Indeed, in hybrid BiCG methods the BiCG coefficients are usually computed with a much lower accuracy than in BiCG.



The method with **long recurrences** like the Arnoldi algorithm can be more numerically stable but require more memory and computations, giving

$$c^*A^{-1}b \approx \xi_n^A \equiv \|b\| c^*V_n H_n^{-1} e_1,$$

where $H_n \in \mathbb{C}^{n \times n}$ and $V_n \in \mathbb{C}^{N \times n}$ are matrices from the Arnoldi algorithm.



In the experiments, we used A, b and c that arise in the problem of diffraction of light on periodic structures solved using the RCWA method; see [2].

Conclusions

There are several players in the game. Methods based on **biorthogonality** and **short recurrences** (like BiCG, BiCGStab, ...) **match at the n th step $2n$ moments** $c^*b, c^*Ab, \dots, c^*A^{2n-1}b$ (analogously to the CG case). One should, however, take into account that the biorthogonality uses an *auxiliary space* which can be very difficult to adjust to an optimal performance of the method. Methods based on **orthogonality** and **long recurrences** (like Arnoldi etc.) **match at the n th step only n moments** $c^*b, c^*Ab, \dots, c^*A^{n-1}b$, but there is no auxiliary space involved and, moreover, the methods benefit from much better stability properties. In conclusion, there is no universal winner. Efficiency of methods depend on the particular problem.

The work of the first author was supported by the ERC-CZ project LL1202 and by the GACR grant 201/09/0917. The work of the second author was supported by the GACR grant P201/13-06684 S.

References

- [1] A. C. ANTOUNAS, *Approximation of large-scale dynamical systems*, vol. 6 of Advances in Design and Control, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.
- [2] J. J. HENCH AND Z. STRAKOŠ, *The RCWA method—a case study with open questions and perspectives of algebraic computations*, Electron. Trans. Numer. Anal., 31 (2008), pp. 331–357.
- [3] R. W. FREUND, *Model reduction methods based on Krylov subspaces*, Acta Numer., 12 (2003), pp. 267–319.
- [4] J. LIESEN AND Z. STRAKOŠ, *Krylov Subspace Methods, Principles and Analysis*, Oxford University Press, ISBN 978-0-19-965541-0, 2012, 408p.
- [5] Z. STRAKOŠ AND P. TICHÝ, *On efficient numerical approximation of the bilinear form $c^*A^{-1}b$* , SIAM Journal on Scientific Computing, 33 (2011) pp. 565–587.
- [6] K. WILLCOX AND A. MEGRETSKI, *Fourier series for accurate, stable, reduced-order models in large-scale linear applications*, SIAM J. Sci. Comput., 26 (2005), pp. 944–962.