Two Topics from Theory of Linear Approximation Problems

Doktorandský den ÚI AV ČR 2006 – Monínec

Martin Plešinger & Zdeněk Strakoš

Outline

1. Core problem theory
2. Stopping criteria in ill-posed problems
3. Multiple right-hand side
4. Summary and future work

1. Core problem theory

For a given orthogonally invariant linear approximation problem

$$Ax \approx b, \qquad A \in \mathbb{R}^{n \times m}, \quad b \in \mathbb{R}^{n},$$

there exist orthogonal matrices P, Q such that

$$P^T \left[\, b \,\middle|\, A \,\right] \left[\begin{array}{c|c} 1 & \\ \hline & Q \end{array}\right] = \left[\begin{array}{c||c|c} b_1 & A_{11} & \\ \hline & & A_{22} \end{array}\right].$$

The original problem is split into two independent subproblems.

The first subproblem $A_{11}\,x_1 \approx b_1$ has **minimal dimensions** and contains **all necessary and sufficient information** for solving the original problem. This subproblem is called **core problem** within the original approximation problem.

[Pl., St., DD2005], [Paige, St., 2006].

Many different transformations give the minimally dimensioned subproblem,
e. g., the **upper bidiagonalization of the extended matrix** $[\,b\,|\,A\,]$.

[Paige, St., 2006], different proof in [Hnětynková, St., 2006], [Hnětynková, Pl., St., 2006].

The core problem theory:

It completes the total least squares theory initiated by the paper *[Golub, Van Loan, 1980]*.
It explains and clarifies the nongeneric solution concept (in single right-hand side case) established in the book *[Van Huffel, Vandewalle, 1991]*.

Application of the core problem theory:

Application in close-to-nongeneric problems, ill-posed and rank deficient problems.

Generalization of the core problem theory:

Extension of the core problem theory on the multiple right-hand side problems – instead of the n vector b, the n by d matrix B is considered.

2. Stopping criteria in ill-posed problems

Let b represent exact data polluted by an **unknown noise**,

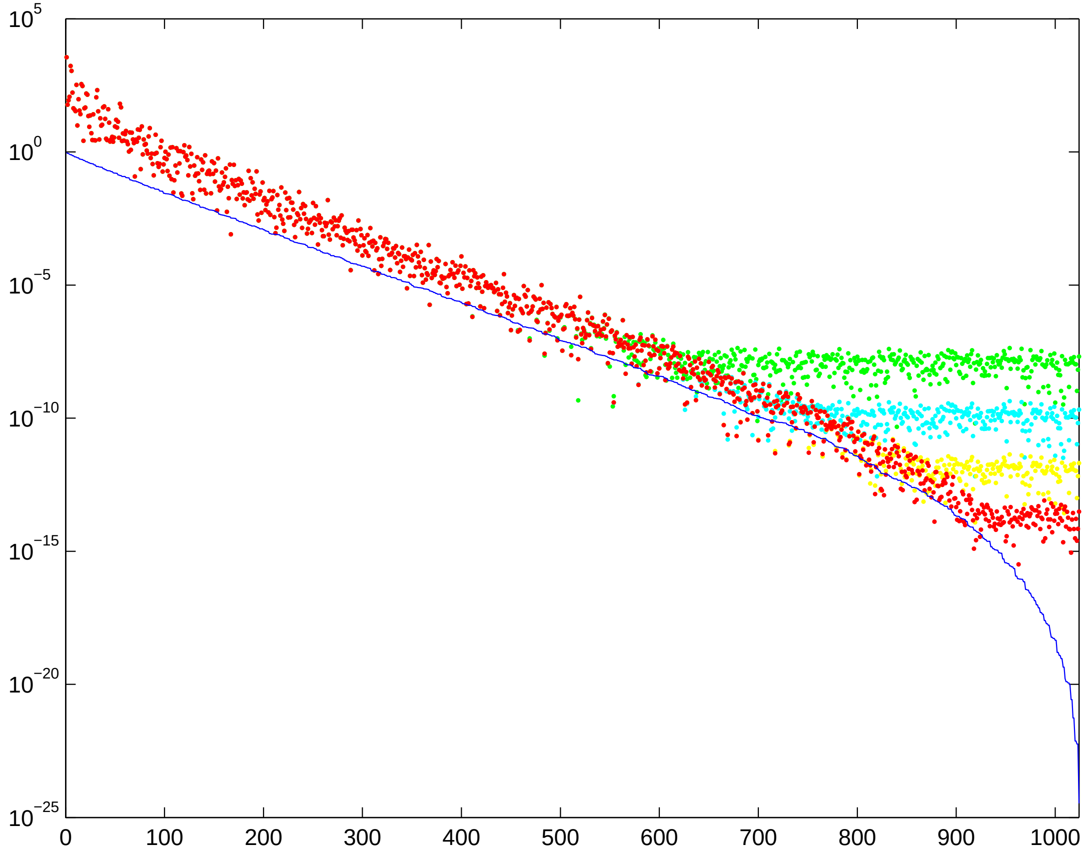
$$b = b_{\text{exact}} + b_{\text{noise}}, \qquad \|b_{\text{exact}}\|_2 \gg \|b_{\text{noise}}\|_2 .$$

In ill-posed problems singular values of A gradually decay to zero without noticeable gap.

Using, e. g., upper bidiagonalization of $[\, b \,|\, A \,]$ and some **stopping criteria** only an approximation $\tilde{A}_{11}\, \tilde{x}_1 \approx \tilde{b}_1$ of the exact core problem may be obtained. If the approximation contains noise, then this noise magnified by small singular values of $\tilde{A}_{11}$ during the (pseudo)-inversion of $\tilde{A}_{11}$ may cover all meaningful information in the solution,

$$\|\tilde{x}_{1,\text{exact}}\|_2 \leq \|\tilde{A}_{11}^{\dagger}\, \tilde{b}_{1,\text{noise}}\|_2 .$$

Singular values of A gradually decay to zero.



The core problem computation is usually based on the **Golub-Kahan bidiagonalization algorithm** (**GK**) applied on A and started from $s_1 = b/\|b\|_2$, or equivalently, by the upper bidiagonalization of $[\, b \,|\, A \,]$.

GK produces two sets of vectors $\{ s_j \}$ and $\{ w_j \}$ and the bidiagonal matrix L_j of growing dimensions. The core problem is obtained by the first zero bidiagonal element in L_j (in exact arithmetic).

The question is: **How to define the stopping criteria for GK if the ill-posed problem with noisy right-hand side is considered?**

(*Rounding errors may have similar effect as a noise contamination, when the finite precision arithmetic is used.*)

GK starts with the normalized noisy right-hand side $s_1 = b / \|b\|_2$, thus vectors s_j has to contain some information about the noise.

Our idea is: **An information about the noise level can be revealed from the vectors s_j generated by GK.**

Technique: **The noise level may be found using the Fourier analysis of the vectors s_j generated by GK.**

We choose two different Fourier basis. The first is the **basis of the left singular vectors $\hat{u}_j$ of A**. It is the most natural basis useful for the theoretical analysis but is not applicable in practical computations.

The second is the **trigonometrical basis**. It is well applicable in practical computations – the fast Fourier transform algorithm (**FFT**).

An example:

Consider the problem SHAW(400) from *[Hansen, RTools]* with a noisy right-hand side (the noise was artificially added)

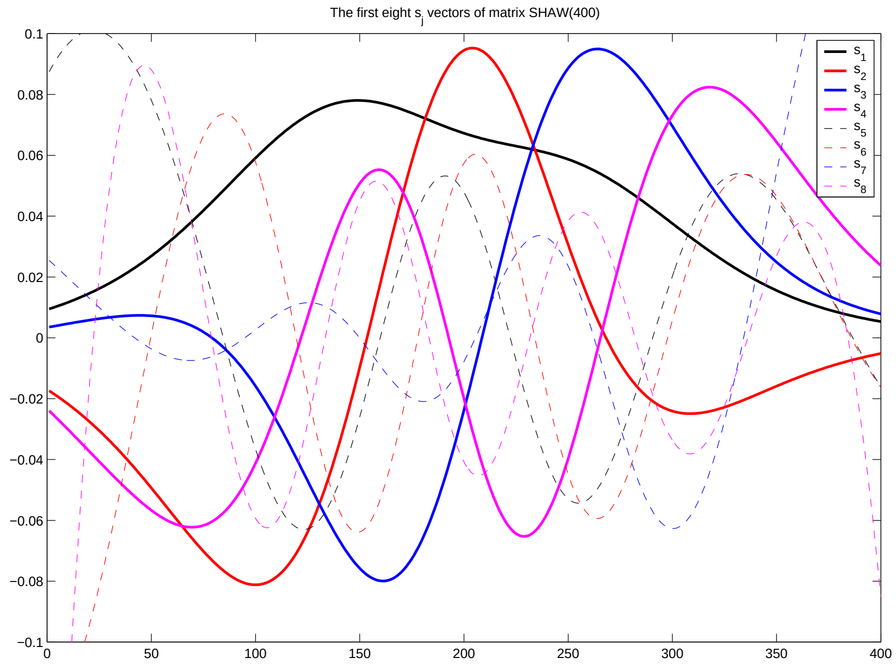
$$46.6225 = \| b_{\text{exact}} \|_2 \gg \| b_{\text{noise}} \|_2 = 10^{-12} .$$

We study the noise-contaminated vectors s_j in the noise-free basis of the left singular vectors $\widehat{U} = [\hat{u}_1, \ldots, \hat{u}_n]$, and in the frequency domain,

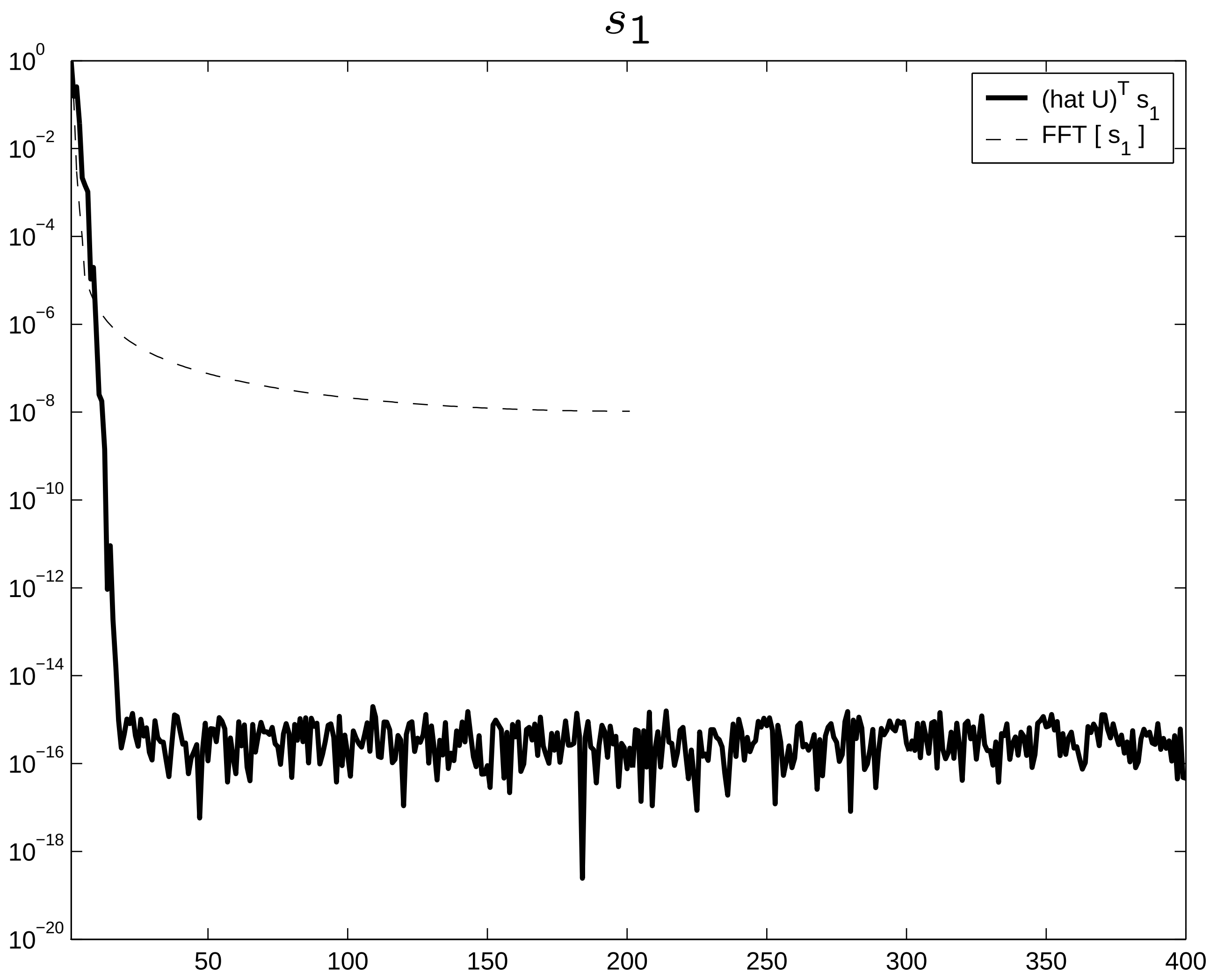
$$(\widehat{U}^T s_j), \quad \text{and} \quad \mathscr{F}[s_j], \qquad j = 1, 2, \ldots,$$

where $\mathscr{F}$ denotes the FFT operator.

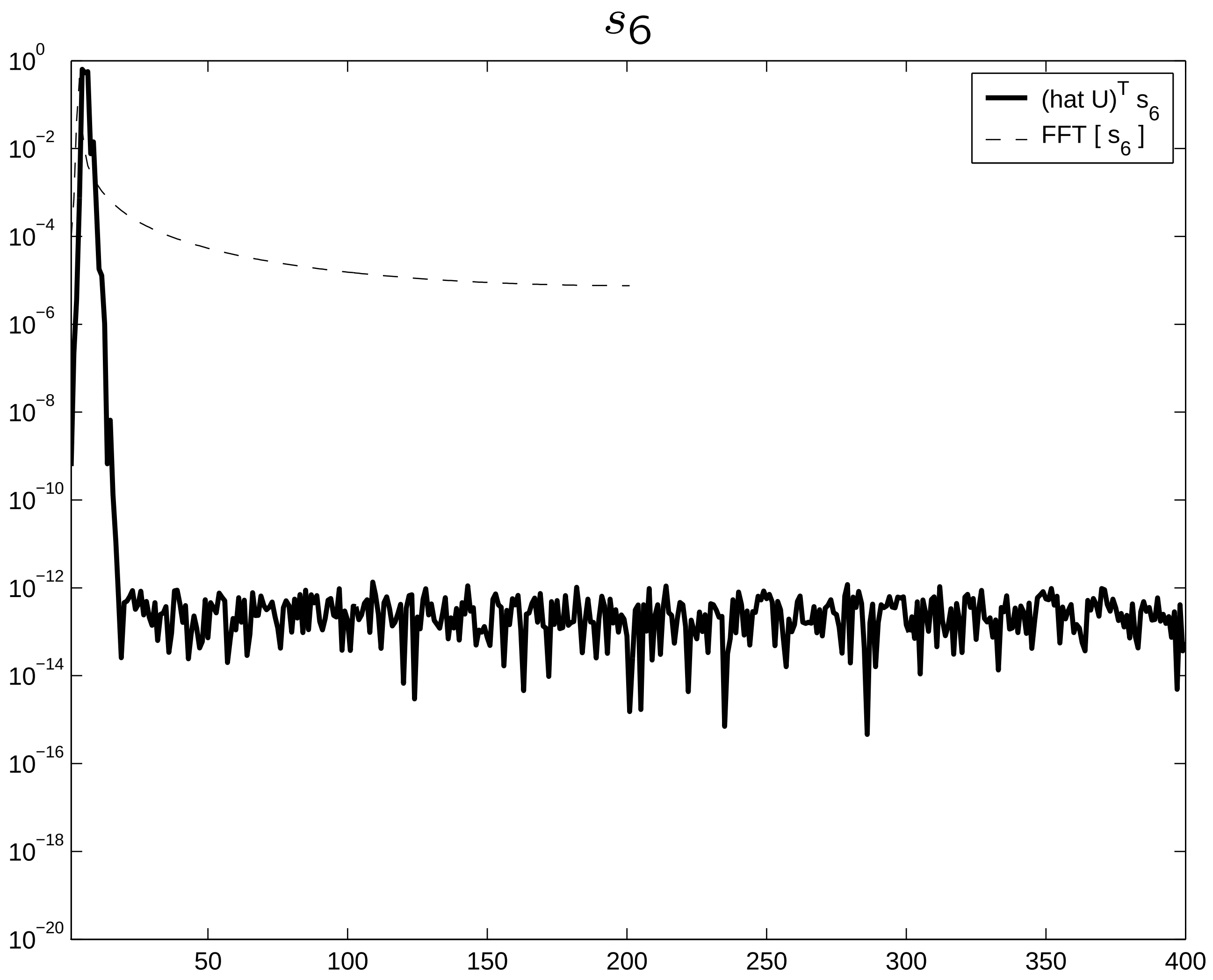
The vector s_1 is **dominated by low frequencies**, thus it has dominant projection in the direction of the left singular vector $\widehat{u}_1$ and possibly several next vectors. Analogously s_2, s_3, ...



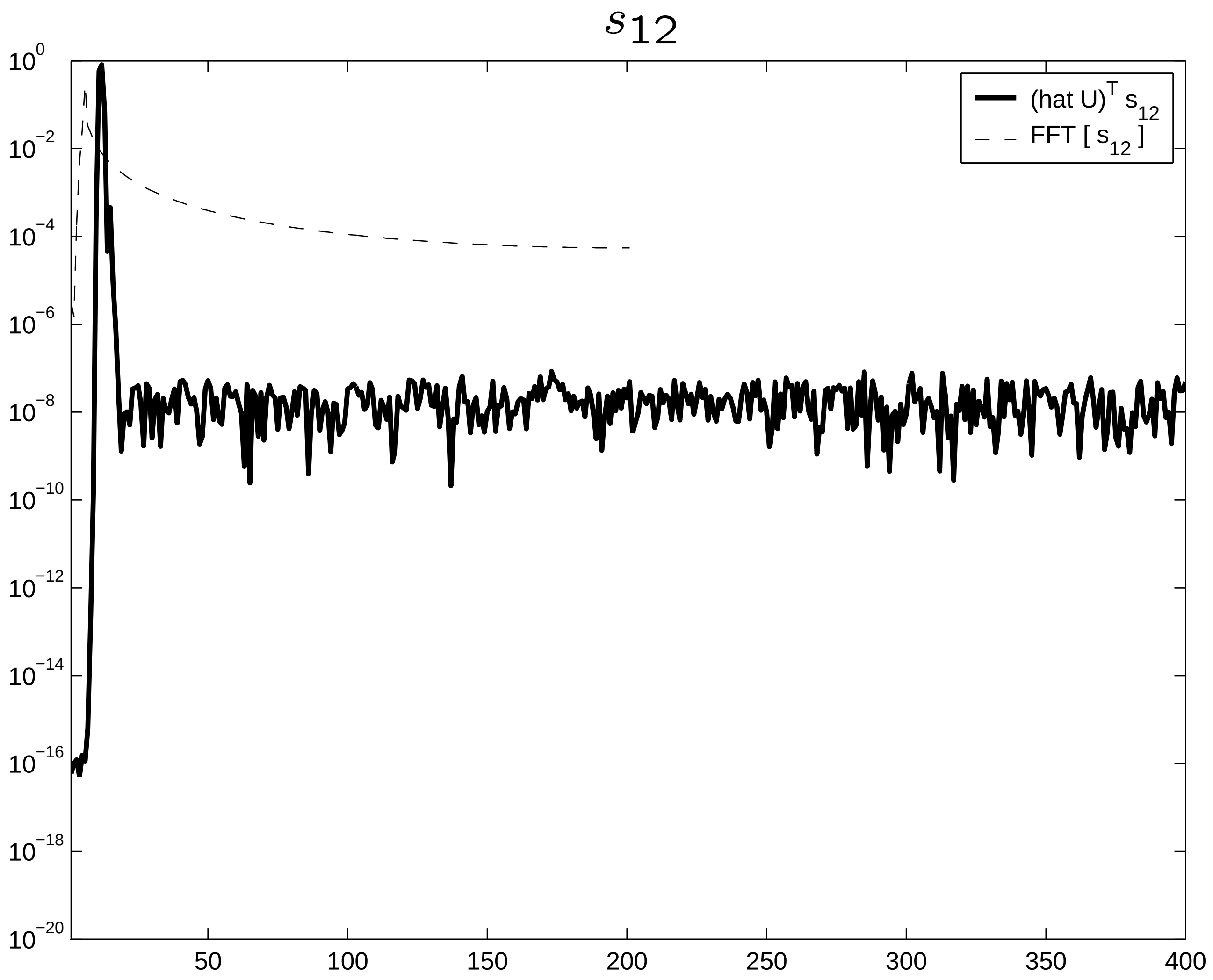
For some index $j = k$ the low frequencies information is projected out from s_k by orthogonalization against the previous vectors s_j, $j = 1, 2, 3, \ldots, k-1$, and the noise is revealed.



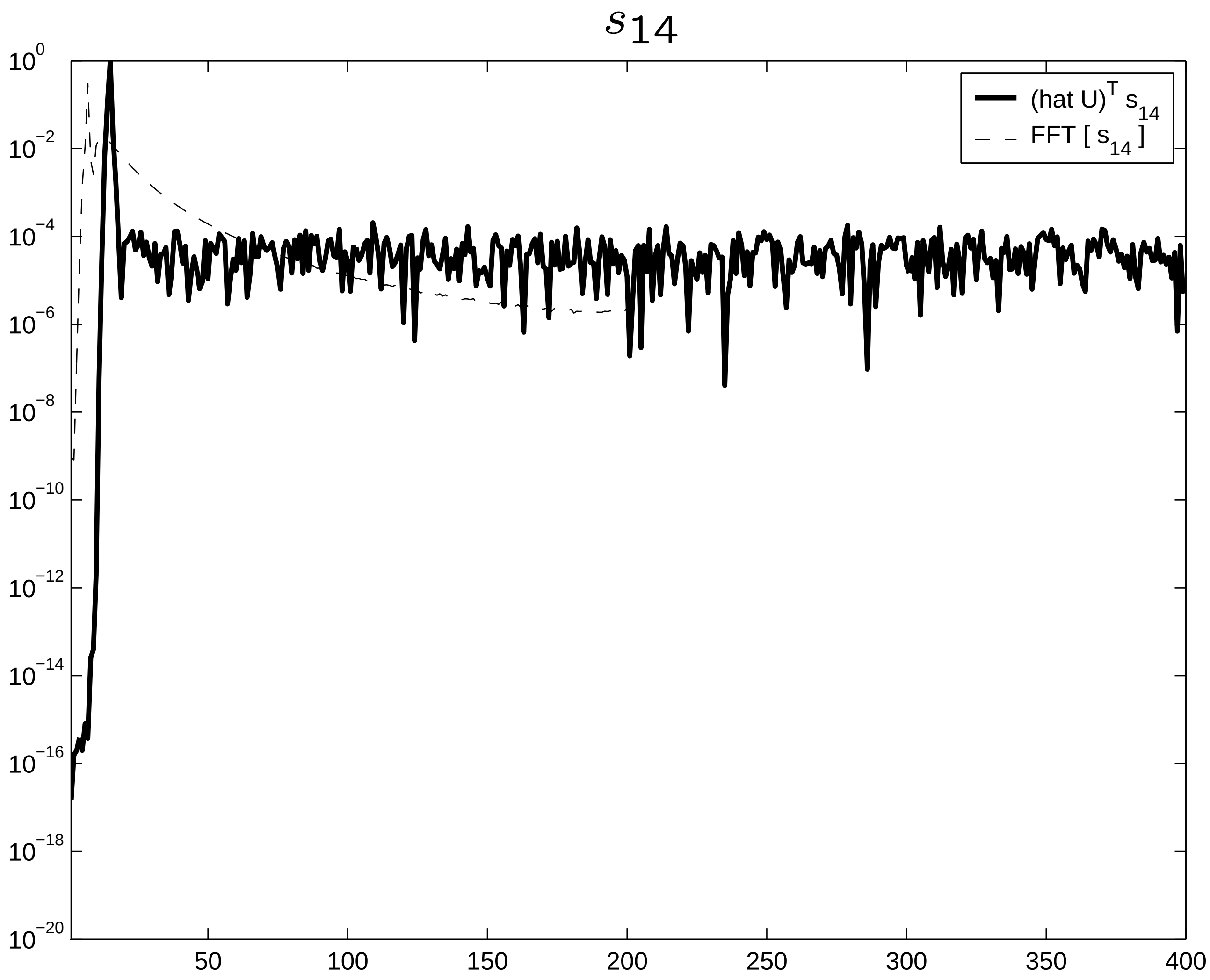
s_1
(hat U)^T s_1
FFT [s_1]
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
10^{-14}
10^{-16}
10^{-18}
10^{-20}
50
100
150
200
250
300
350
400



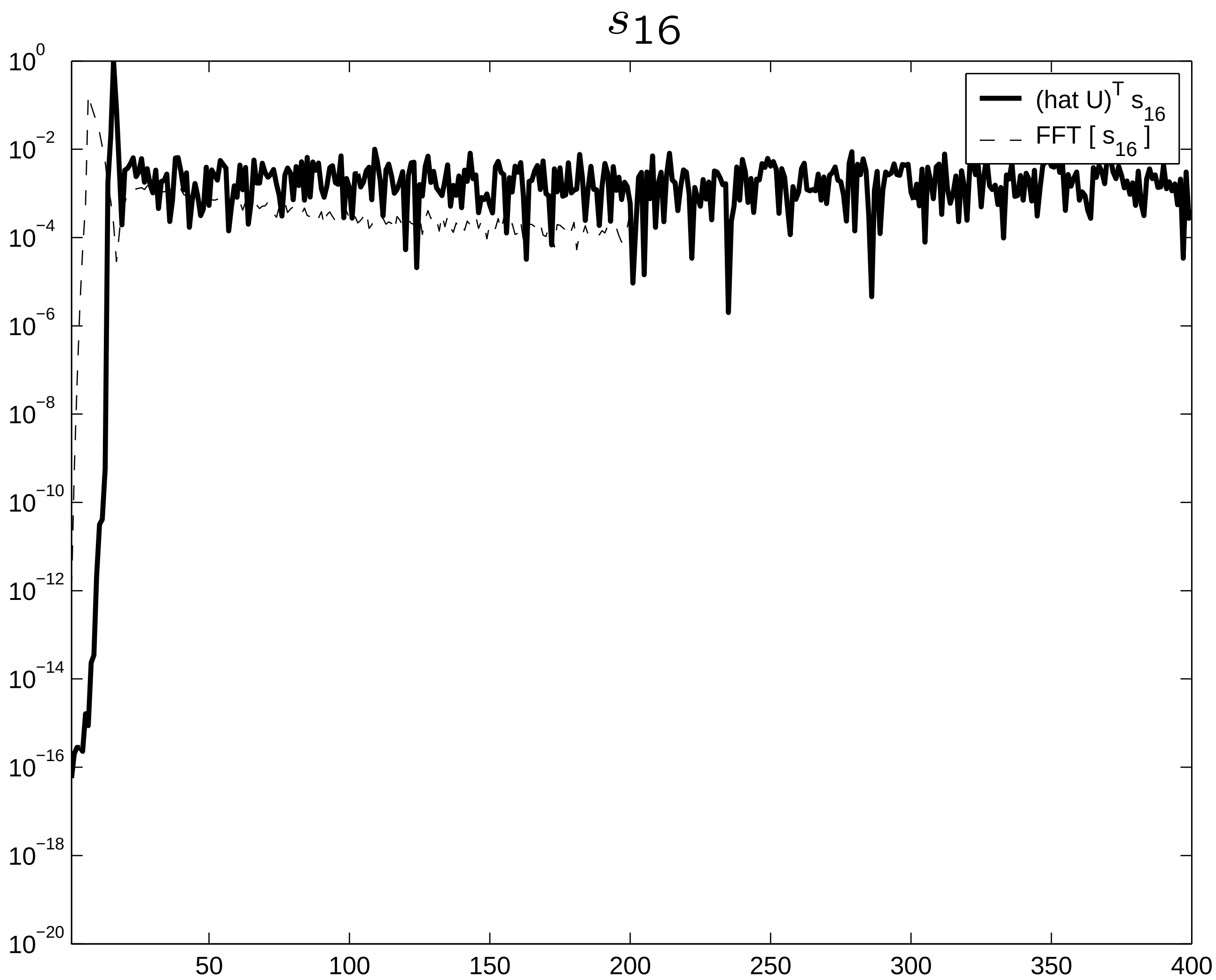
s_6
(hat U)^T s_6
FFT [s_6]
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
10^{-14}
10^{-16}
10^{-18}
10^{-20}
50
100
150
200
250
300
350
400

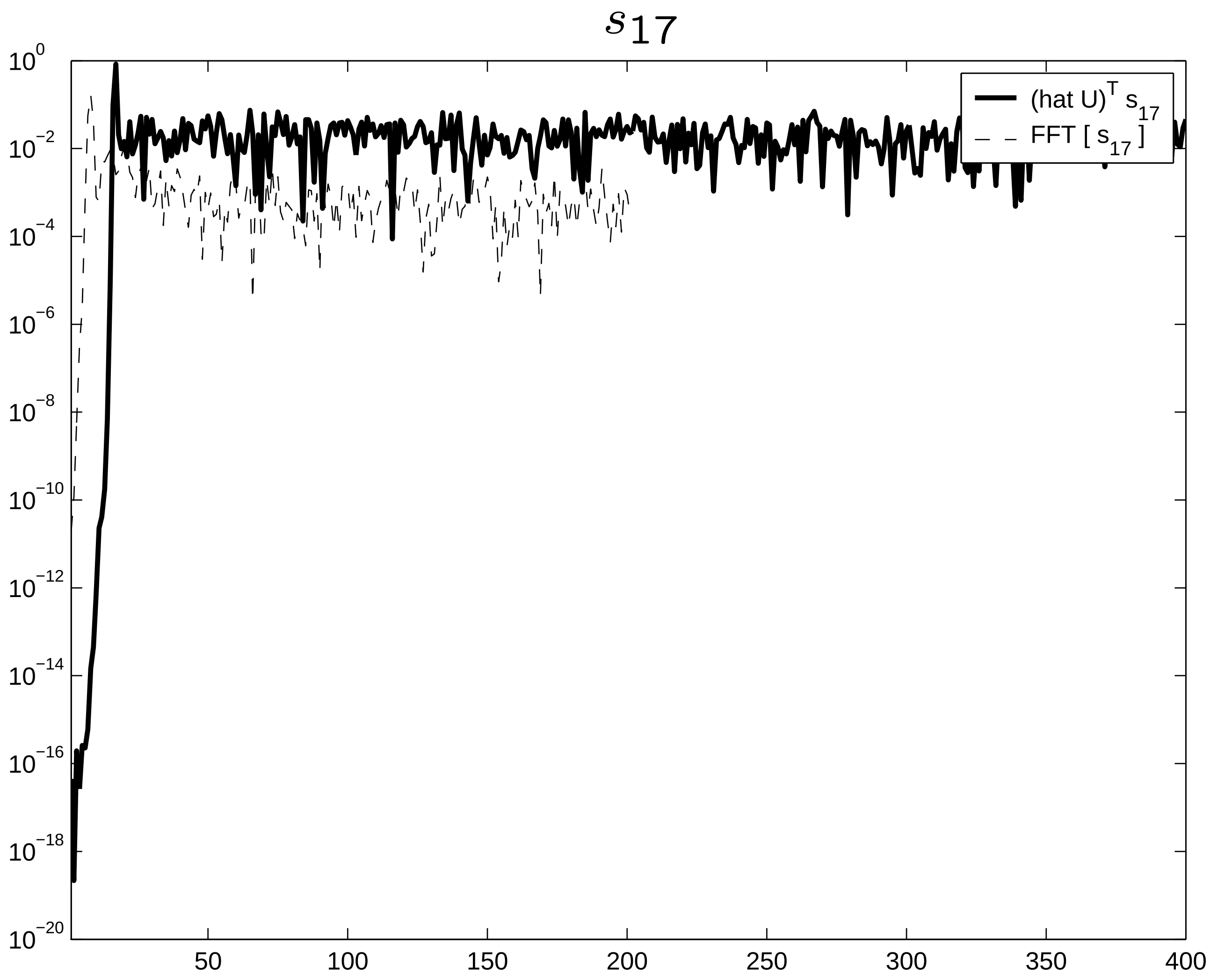


s_{12}
(hat U)^T s_12
FFT [s_12]
10^0
10^−2
10^−4
10^−6
10^−8
10^−10
10^−12
10^−14
10^−16
10^−18
10^−20
50
100
150
200
250
300
350
400



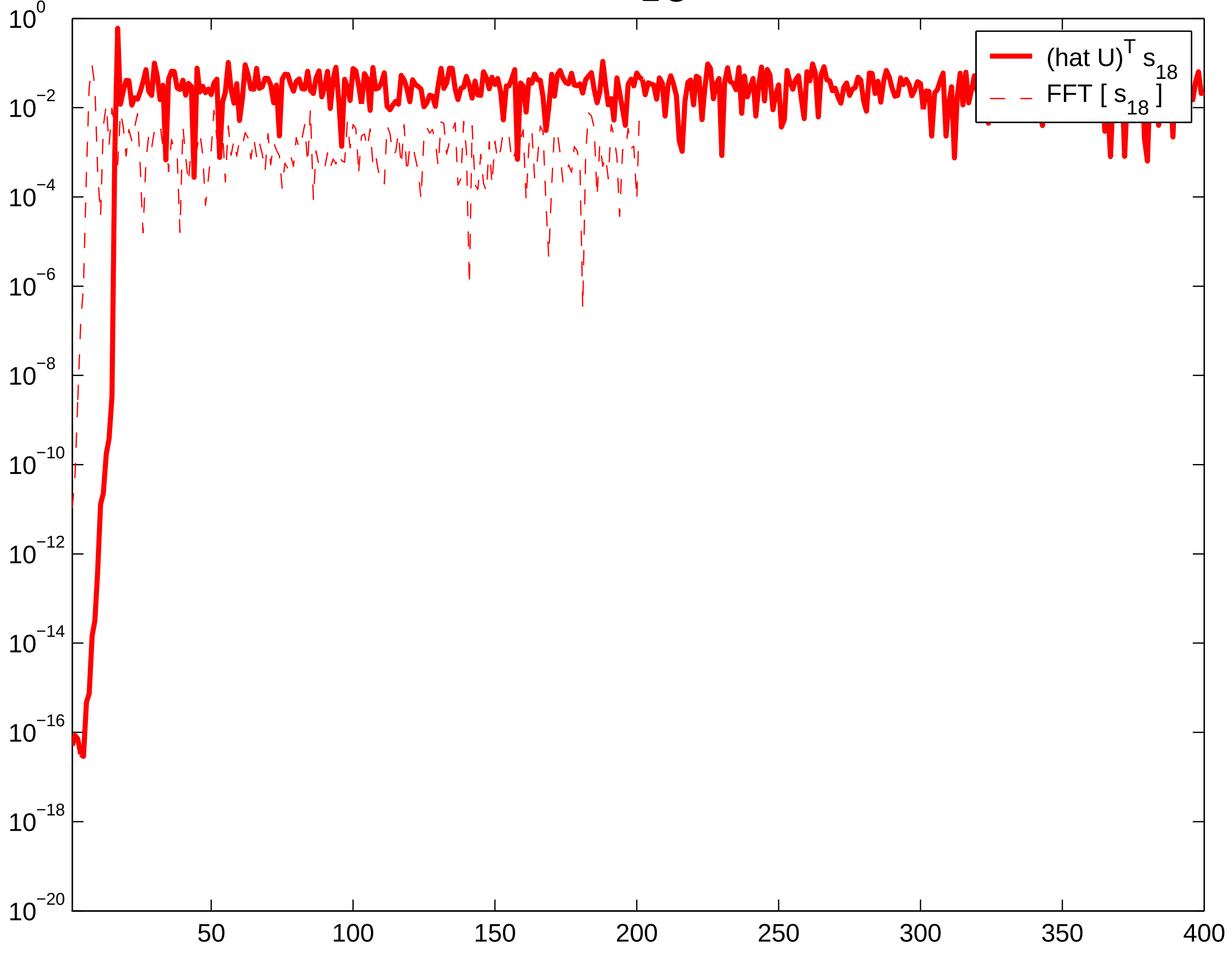
s_{14}
(hat U)^T s_14
FFT [s_14]
10^0
10^−2
10^−4
10^−6
10^−8
10^−10
10^−12
10^−14
10^−16
10^−18
10^−20
50
100
150
200
250
300
350
400


s_{16}
(hat U)^T s_{16}
FFT [s_{16}]
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
10^{-14}
10^{-16}
10^{-18}
10^{-20}
50
100
150
200
250
300
350
400

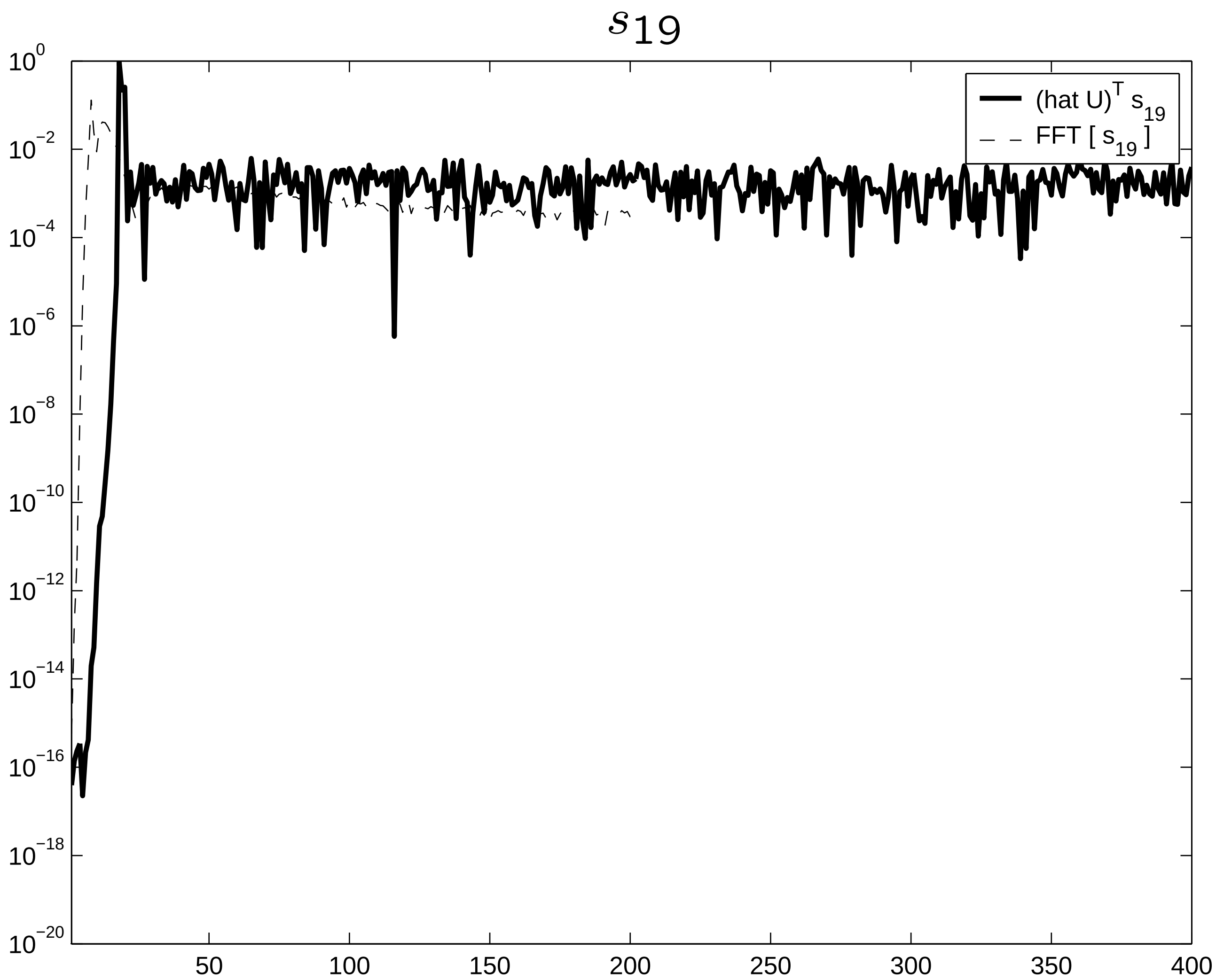


s_{17}
(hat U)^T s$_{17}$
FFT [s$_{17}$]
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
10^{-14}
10^{-16}
10^{-18}
10^{-20}
50
100
150
200
250
300
350
400

s_{18}



(hat U)^T s_18
FFT [s_18]
10^0
10^-2
10^-4
10^-6
10^-8
10^-10
10^-12
10^-14
10^-16
10^-18
10^-20
50
100
150
200
250
300
350
400


s_{19}
(hat U)^T s_{19}
FFT [s_{19}]
10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
10^{-14}
10^{-16}
10^{-18}
10^{-20}
50
100
150
200
250
300
350
400

Vector s_{18} is fully dominated by noise –

the **noise level is revealed**.

Now we get explicit information when the noise begins to cover useful information in the data. The solution of the original problem $A\,x = b$ computed through the core problem approximation with

$$\tilde{A}_{11} = L_j\,,$$

for $j > k = 18$ can be significantly polluted by the noise.

(*In the* 19*th step, the noise is partially projected out because vectors* s_j *has to be mutually orthonormal.*)

This work was presented on GAMM-SIAM Conference, July 2006, Düsseldorf. Similar idea is used in *[Hansen, Kilmer, Kjeldsen, 2006]*.

3. Multiple right-hand side

Consider orthogonally invariant linear approximation problem

$$A\,X \approx B\,, \qquad A \in \mathbb{R}^{n\times m}\,, \quad B \in \mathbb{R}^{n\times d}\,.$$

We are looking for orthogonal matrices P, Q, R such that

$$P^T \left[\, B \mid A \,\right] \left[\begin{array}{c|c} R & \\ \hline & Q \end{array}\right] = \left[\begin{array}{c||c|c} B_1 & A_{11} & \\ \hline & & A_{22} \end{array}\right]\,,$$

and $A_{11}\,X_1 \approx B_1$ has **minimal dimensions** and contains **all necessary and sufficient information** for solving the original problem.

The original problem is split into two independent subproblems and the first subproblem

The situation is much more complicated than in the single right-hand side case, e. g., **the TLS solution for general data** $[\,B\,|\,A\,]$ **is still not well defined**, see *[Van Huffel, Vandewalle, 1991]*.

The intuitive approach is presented in *[Björck, 2005]*, *[Björck, 2006]*.

First fragments of analysis of this problem are in
[Sima, Van Huffel, 2006], *[Sima, PhD2006]*.

Our analysis based on SVDs of B and A was presented on 4th International Workshop on TLS and EIV Modeling, August 2006, Leuven.

4. Summary and future work

Summary:

- Stopping criteria in solving ill-posed problems (*presented on international conference, work still in progress*).
- Generalization of core theory in the multiple right-hand side case (*presented on int. workshop, work still in progress*).
- Stable and fast implementation of two bidiagonalization algorithms (*presented in proceedings of DD ÚI AV ČR 2006*).

Future work:

- Application of our ideas in solving ill-posed problems arising from image deblurring.
- Definition of the core problem in the multiple right-hand side case (eventually definition of solution for general data $[\,B\,|\,A\,]$).
- Implementation of stable solver for ill-posed problems based on the core problem theory and our stopping criterion.

THANK YOU

FOR YOUR ATTENTION

References

- Björck – **Bidiagonal Decomposition and Least Squares**, Presentation, Canberra, 2005.
- Björck – **A Band-Lanczos Generalization of Bidiagonal Decomposition**, Presentation, Conference in Honor of G. Dahlquist, Stockholm, 2006.
- Golub, Van Loan – **An analysis of the total least squares problem**, Numer. Anal., 1980.
- Hansen – **Rank-Deficient and Discrete Ill-Posed Problems**, SIAM Monographs Math. Modeling Comp., 1998.
- Hansen – **Matlab package: REGULARIZATION TOOLS 3.2**, http://www2.imm.dtu.dk/˜pch/Regutools/index.html.
- Hansen, Kilmer, Kjeldsen – **Exploiting residual information in the parameter choice for discrete ill-posed problems**, BIT, 2006.
- Hnětynková, Plešinger, Strakoš – **Lanczos tridiagonalization, Golub-Kahan bidiagonalization and core problem**, to appear in PAMM.

- Hnětynková, Strakoš – **Lanczos tridiagonalization and core problems**, to appear in Linear Alg. Appl.
- Paige, Strakoš – **Core problem in linear algebraic systems**, SIAM J. Matrix Anal. Appl., 2006
- Plešinger, Strakoš – **Core problem v úlohách nejmenšich čtverců**, in Proceedings of DD ÚI AV ČR, 2005
- Sima, Van Huffel – **Core problems in $AX \approx B$**, Technical Report, Faculty of Engineering, Katholieke Universiteit, Leuven, 2006.
- Sima – **Regularization techniques in model fitting and parameter estimation**, PhD. thesis, Faculty of Engineering, Katholieke Universiteit, Leuven, 2006.
- Van Huffel, Vandewalle – **The total least squares problem: Computational aspects and analysis**, SIAM Publications, 1991.