Regularization by the noise revealing Golub-Kahan iterative bidiagonalization

Iveta Hnětynková, Martin Plešinger, Zdeněk Strakoš

Charles University in Prague

Academy of Sciences of the Czech republic, Prague

EUCCO 2013, Chemnitz UT, July 2013

Outline

1. **Problem formulation**
2. Regularization by the Golub-Kahan iterative bidiagonalization
3. Noise approximation
4. Regularization and denoising
5. Summary and future work

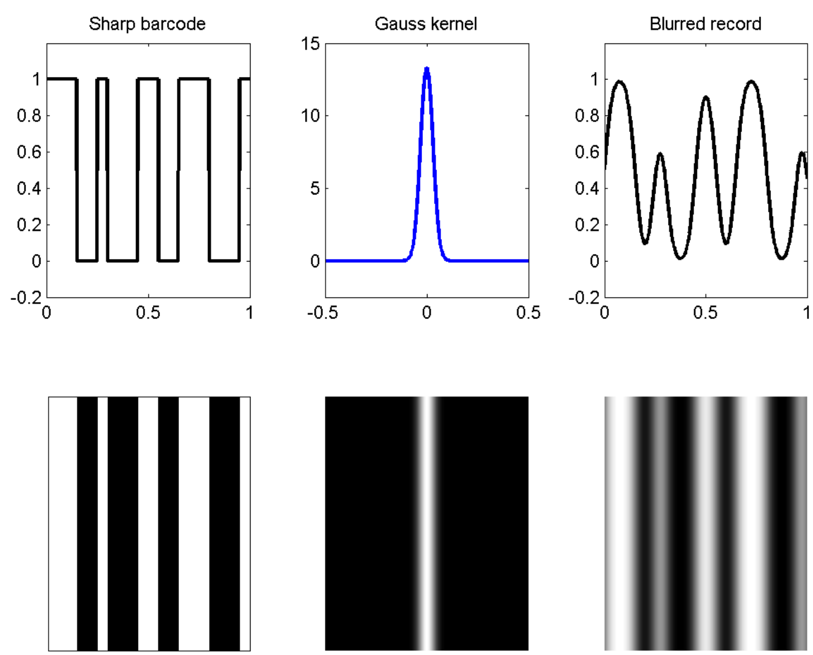
Example - Fredholm integral of the first kind:

Given the continuous smooth kernel $K(s,t)$ and the (mesuared) data $g(s)$, the aim is to find the (source) function $f(t)$ such that

$$g(s) = \int_{I_t} K(s,t) f(t) \mathrm{d}t.$$

Fredholm integral has smoothing property, i.e. high frequency components in g are dampened compared to f.

Barcode reading:



Consider a discretized inverse problem

$$Ax \approx b, \quad b = b^{\text{exact}} + b^{\text{noise}}, \quad A \in \mathbb{R}^{N \times M}, \quad x \in \mathbb{R}^{M}, \quad b \in \mathbb{R}^{N}$$

polluted by noise (measurement, discretization, rounding errors, ...) with unknown noise level

$$\delta^{\text{noise}} \equiv \| b^{\text{noise}} \| / \| b^{\text{exact}} \| \, .$$

Usual properties:

- the problem is ill-posed,
- A is a discretization of a smoothing operator,
- singular values σ_j of A decay gradually,
- singular vectors u_j, v_j of A represent increasing frequencies,
- b^{exact} is smooth and satisfies the discrete Picard condition (DPC),
- $\|b^{\text{exact}}\| \gg \|b^{\text{noise}}\| \, .$

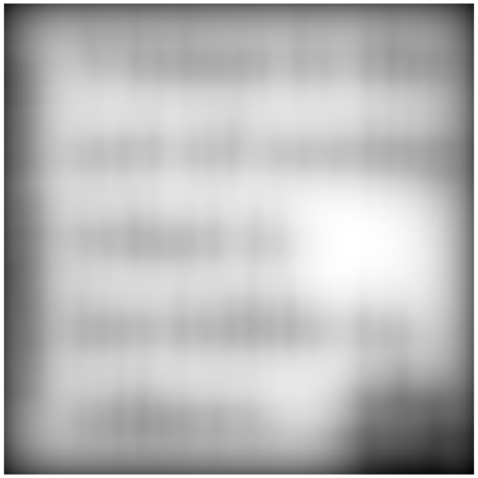
We want to approximate

$$x^{\text{exact}} = A^{\dagger} b^{\text{exact}} \, .$$

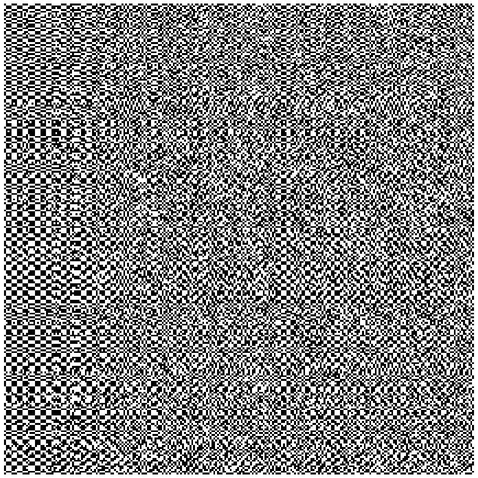
Image deblurring problem: Original image, white noise contaminated image and the "naive" solution $x^{\text{naive}} \equiv A^{\dagger}b$:

x = true image

b = blurred, noisy image



x = inverse solution



The SVD components of the naive solution

$$x^{\text{naive}} \equiv A^\dagger b = \underbrace{\sum_{j=1}^{l} \frac{u_j^T b^{\text{exact}}}{\sigma_j} v_j}_{x^{\text{exact}}} + \underbrace{\sum_{j=1}^{l} \frac{u_j^T b^{\text{noise}}}{\sigma_j} v_j}_{\text{amplified noise}}$$

$$+ \underbrace{\sum_{j=l+1}^{M} \frac{u_j^T b^{\text{exact}}}{\sigma_j} v_j}_{x^{\text{exact}}} + \underbrace{\sum_{j=l+1}^{M} \frac{u_j^T b^{\text{noise}}}{\sigma_j} v_j}_{\text{amplified noise}}$$

corresponding to small σ_j's are dominated by amplified (white) noise.

Exact data satisfy DPC: On average, $|u_j^T b^{\text{exact}}|$ decay faster than the singular values σ_j of A, $j = 1, \ldots, N$.

White noise components $|(b^{\text{noise}}, u_j)|$ do not exhibit any trend.

Outline

1. Problem formulation
2. **Regularization by the Golub-Kahan iterative bidiagonalization**
3. Noise approximation
4. Regularization and denoising
5. Summary and future work

Golub-Kahan iterative bidiagonalization (GK) of A:

Given $w_0 = 0\,,\ s_1 = b\,/\,\beta_1\,,$ where $\beta_1 = \|b\|\,,$ for $j = 1, 2, \ldots$

$$\begin{aligned} \alpha_j\, w_j &= A^T s_j - \beta_j\, w_{j-1}\,, & \|w_j\| = 1\,, \\ \beta_{j+1}\, s_{j+1} &= A\, w_j - \alpha_j\, s_j\,, & \|s_{j+1}\| = 1\,. \end{aligned}$$

Let $S_k = [s_1, \ldots, s_k]$, $W_k = [w_1, \ldots, w_k]$ be the associated matrices with orthonormal columns. Denote

$$L_k = \begin{bmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & \beta_k & \alpha_k \end{bmatrix}, \qquad L_{k+} = \begin{bmatrix} L_k \\ e_k^T \beta_{k+1} \end{bmatrix}$$

the bidiagonal matrices containing the normalization coefficients.

Regularization based on GK:

First the problem is projected onto a lower dimensional Krylov subspace

$$\mathcal{K}_k(A^TA, A^Tb) = Span\{A^Tb,\, (A^TA)A^Tb, \ldots,\, (A^TA)^{k-1}A^Tb\}$$

(regularization by projection with k representing the regularization parameter), giving

$$L_{k+}\, y \approx \beta_1\, e_1.$$

The projected problem is solved either directly (e.g. in LSQR, CGLS) or some inner regularization is applied (in hybrid methods); see e.g. [Hansen – 11, 98], [Kilmer, Hansen, Español – 06], [Kilmer, O'Leary – 01], [Fiero, Golub, Hansen, O'Leary – 97], [O'Leary, Simmons – 81].

Outline

1. Problem formulation
2. Regularization by the Golub-Kahan iterative bidiagonalization
3. **Noise approximation**
4. Regularization and denoising
5. Summary and future work

In [H., Plešinger, Strakoš: '09] noise propagation in GK has been analyzed; see also theses [Vasilík: '11], [Michenková: '13].

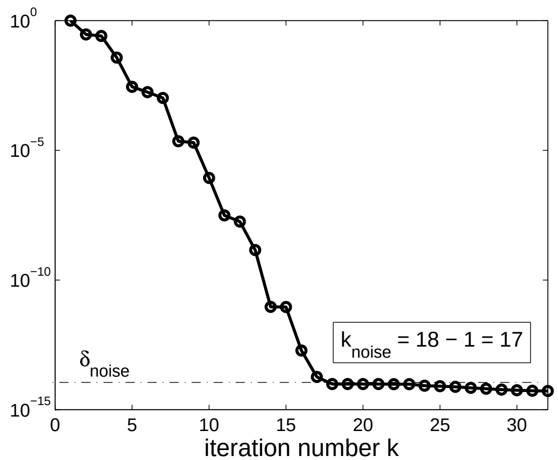
Result: The size of the first component of the left singular vector $p_1^{(k)}$ corresponding to the smallest singular value of L_k decreases. At some iteration k_{noise} (the noise revealing iteration) it sharply starts to (almost) stagnate close to the noise level, i.e.

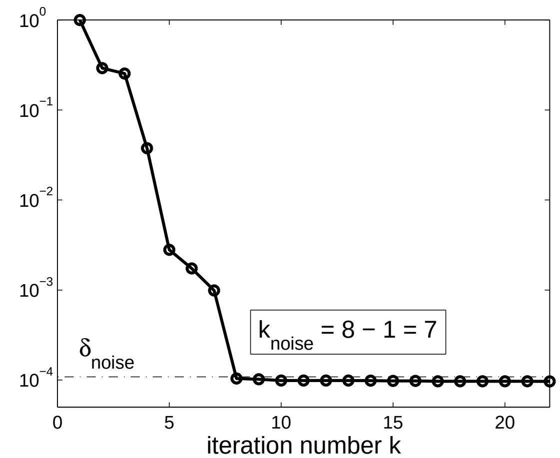
$$\delta_{\text{noise}} \approx |(p_1^{(k_{\text{noise}})}, e_1)| \,.$$

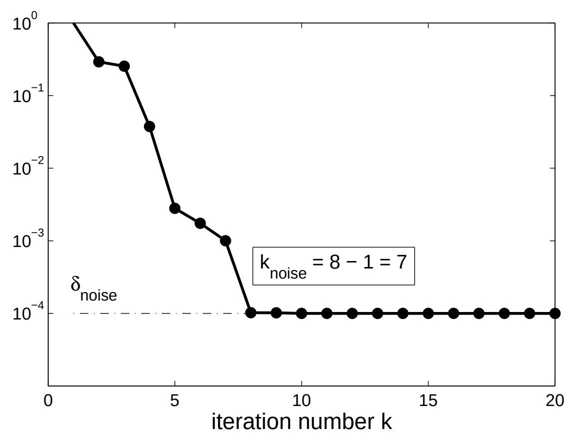
Moreover, the bidiagonalization vector $s_{k_{\text{noise}}}$ is fully dominated by (the high frequency part of) noise. Thus

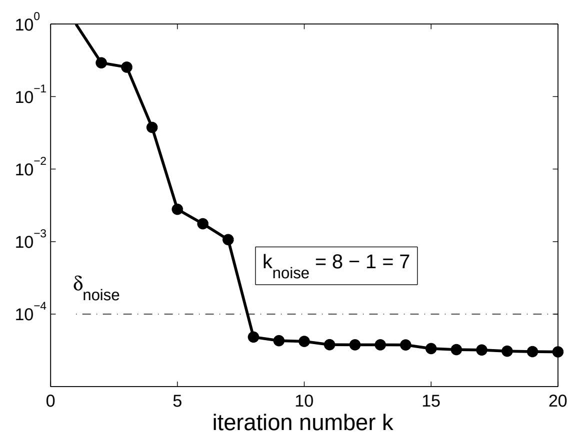
$$b^{\text{noise}} \approx \|b^{\text{noise}}\| \;\; s_{k_{\text{noise}}} \approx \|b\| \, |(p_1^{(k_{\text{noise}})}, e_1)| \, s_{k_{\text{noise}}} \,.$$

Components $|(p_1^{(k)}, e_1)|$, $k = 1, 2, \ldots$ for Shaw from [RegToolbox], with white noise $\delta_{\text{noise}} = 10^{-14}, 10^{-4}$ (top); high frequency violet, and low frequency Brown noise, $\delta_{\text{noise}} = 10^{-4}$ (bottom):









Noise level δ_{noise} in the data, and the estimated noise level (average values computed using 1000 randomly chosen white noise vectors b^{noise}):

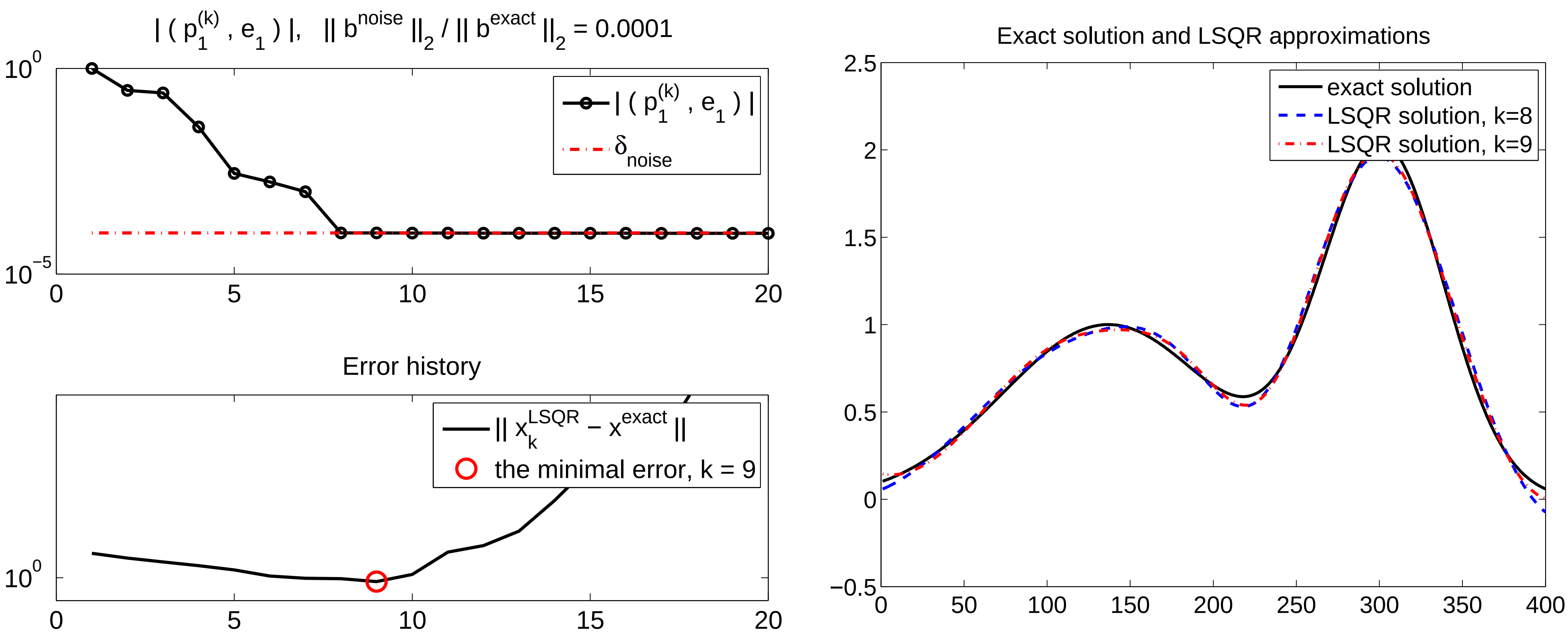
SHAW(400)				
δ_{noise}	1×10^{-14}	1×10^{-6}	1×10^{-4}	1×10^{-2}
estimate	1.80×10^{-14}	1.31×10^{-6}	1.01×10^{-4}	1.03×10^{-2}
k_{noise}	16	9	7	4
ILAPLACE(100,1)				
δ_{noise}	1×10^{-13}	1×10^{-7}	1×10^{-2}	1×10^{-1}
estimate	9.12×10^{-14}	1.34×10^{-7}	1.02×10^{-2}	1.11×10^{-1}
k_{noise}	22	15.30	6.02	2

Remark: k_{noise} can be detected automatically, see [H., Plešinger, Strakoš: '09], [Vasilík: '11 - thesis].

Outline

1. Problem formulation
2. Regularization by the Golub-Kahan iterative bidiagonalization
3. Noise approximation
4. **Regularization and denoising**
5. Summary and future work

Components $|(p_1^{(k)}, e_1)|$, $k = 1, 2, \ldots$ (top), error history of LSQR solutions (bottom), and the best LSQR reconstructions, Shaw with $\delta_{\text{noise}} = 10^{-4}$:

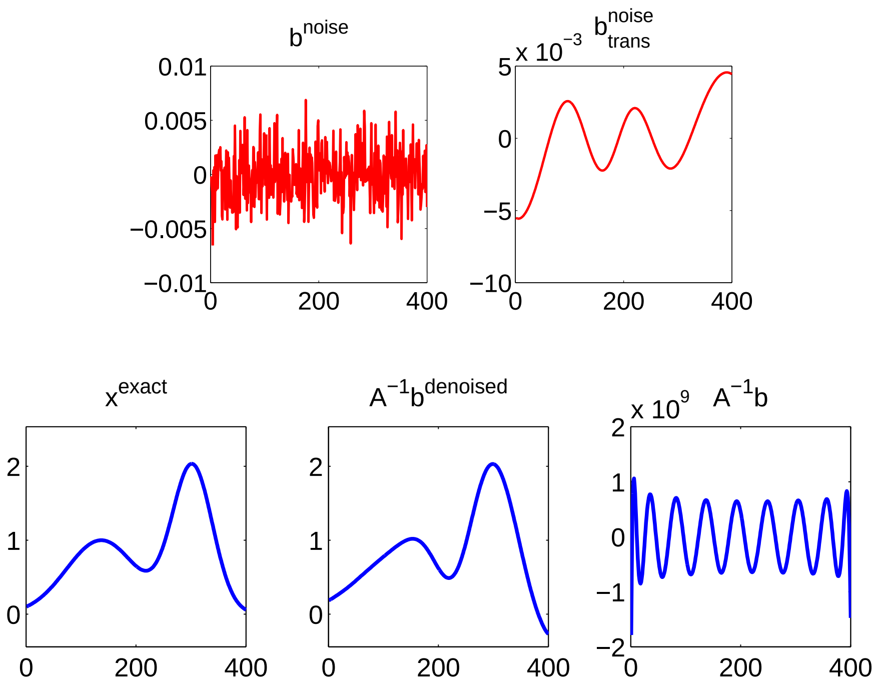


Denoising: In the iteration k_{noise}, we have the approximation

$$b^{\text{noise}} \approx \|b\| \, |(p_1^{(k_{\text{noise}})}, e_1)| \, s_{k_{\text{noise}}} .$$

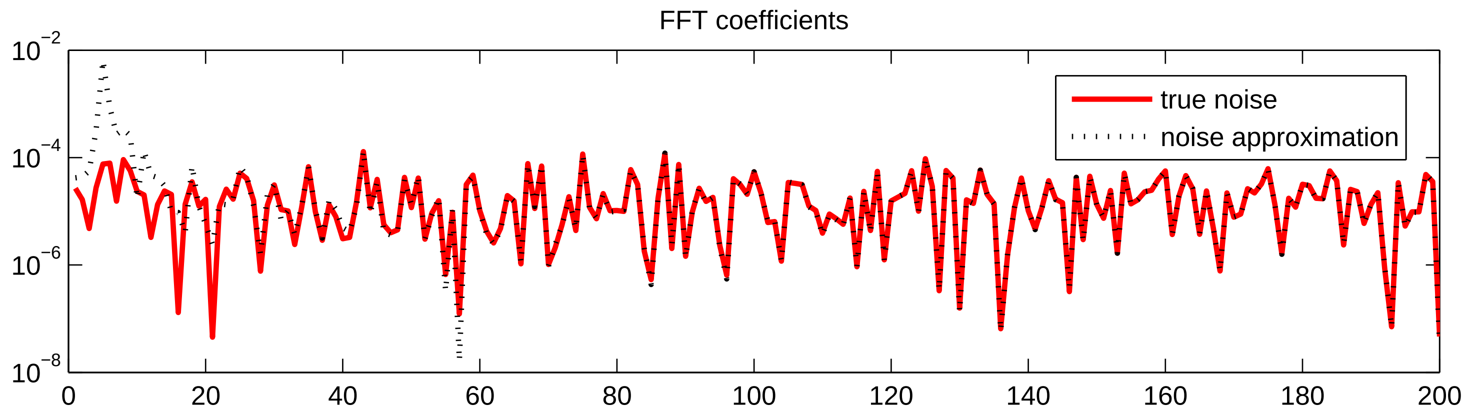
We can subtract it from b. [Michenková: '13 - thesis].

Original noise, and noise remaining after denoising, Shaw:



Repeating the process gives a better approximation of original noise.

Fourier coeffs. of the original and approximated noise vector after 5 repeats, Shaw with $\delta_{\text{noise}} = 10^{-4}$:



Singular values of A, and spectral coeffs. of the original and denoised observation vectors, Shaw with $\delta_{\text{noise}} = 10^{-4}$:

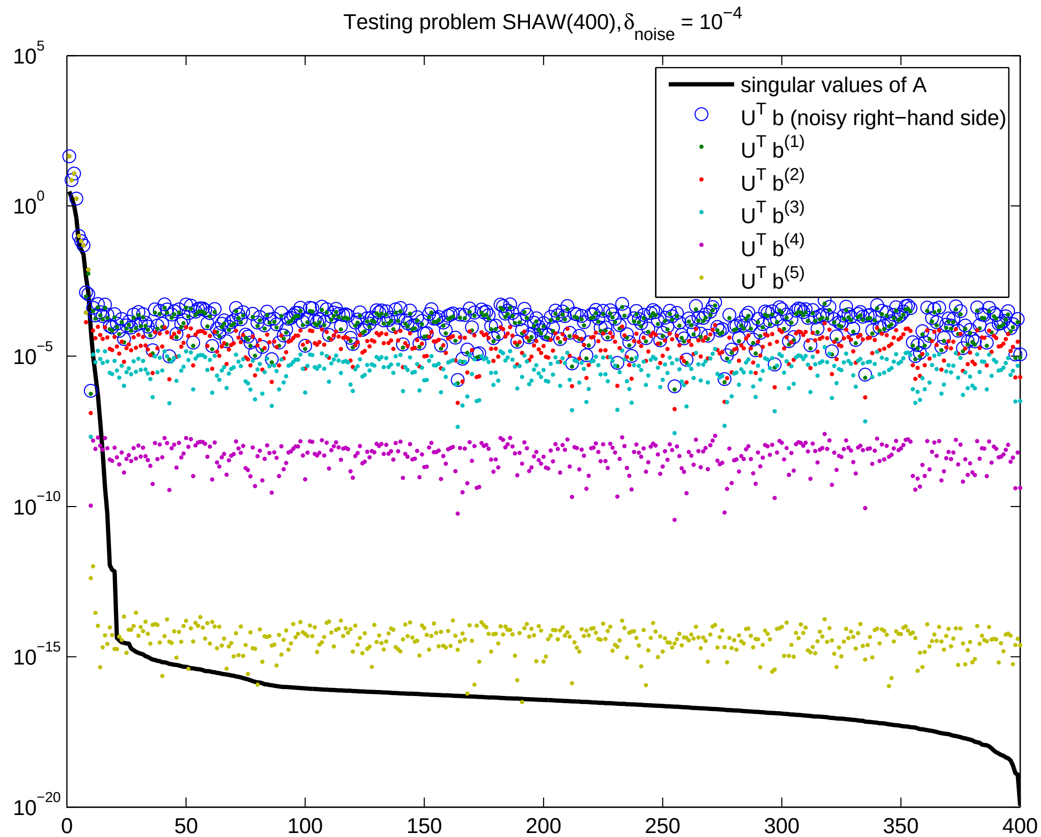
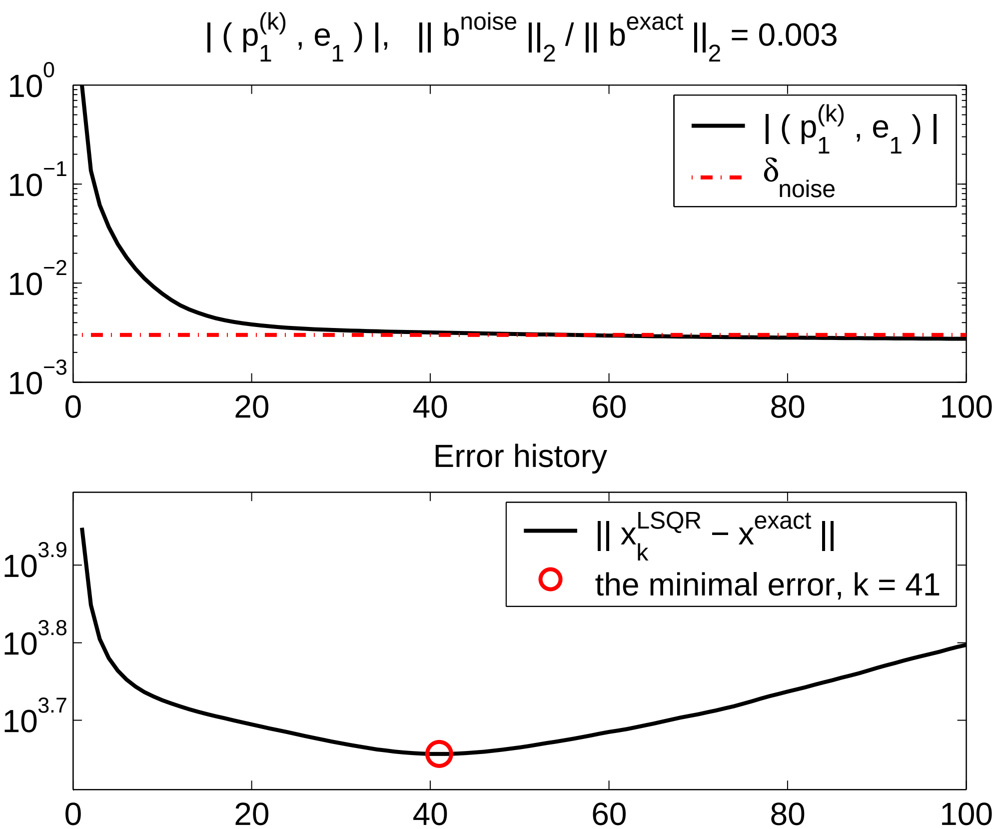


Image deblurring problem: image size 324×470 **pixels, problem dimension** $N = 152280$**, the exact solution (left) and the noisy right-hand side (right),** $\delta_{\text{noise}} = 3 \times 10^{-3}$**:**

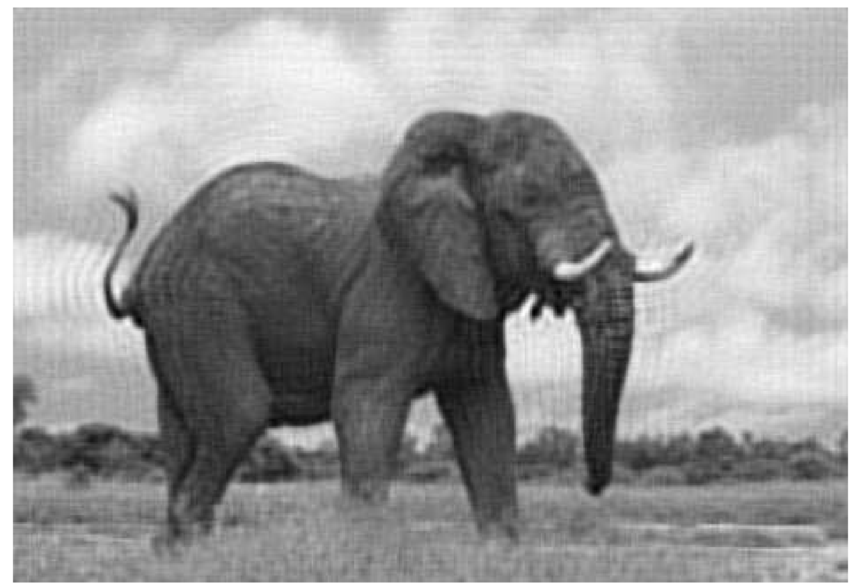
x^{exact}

$b^{\text{exact}} + b^{\text{noise}}$

Components $|(p_1^{(k)}, e_1)|$, $k = 1, 2, \ldots$ (top), error history of LSQR solutions (bottom), and the best LSQR reconstruction, GK without reorthogonalization:



LSQR reconstruction with minimal error, x_{41}^{LSQR}



Hybrid LSQR:

Stopping GK for $k \leq k_{\text{noise}}$, some information is not absorbed in the problem yet and the result is sometimes unsatisfactory.

Stopping GK for $k > k_{\text{noise}}$, the bidiagonal problem inherits a part of the ill-posedness of the original problem, inner regularization must be applied. The discrepancy principle can be used to stop GK.

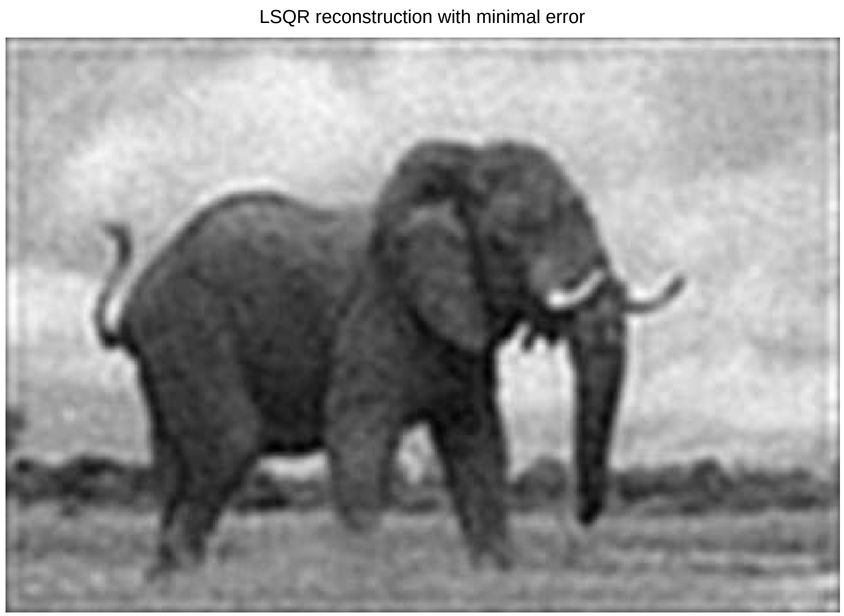
Discrepancy principle:

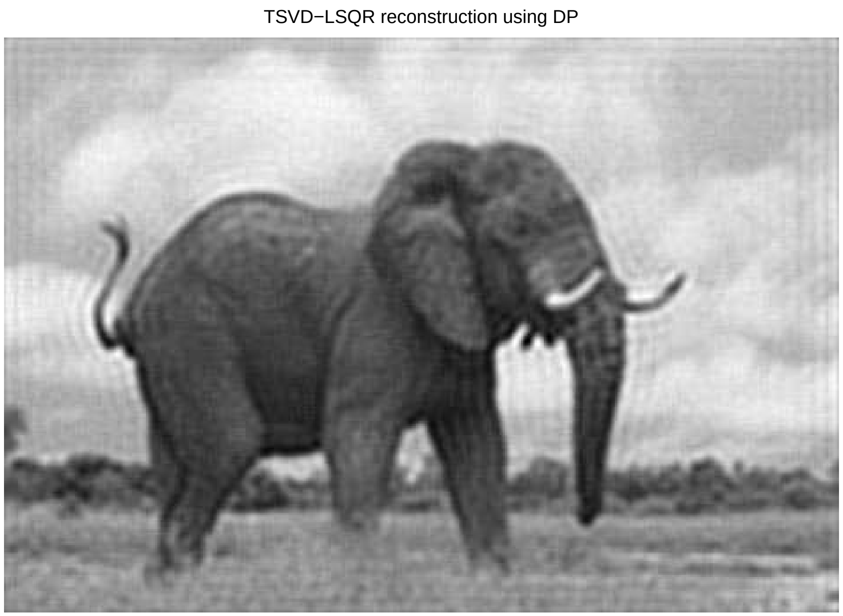
Bidiagonalization is stopped for the smallest k where

$$\|b - Ax_k^{\text{TSVD},r}\| \;=\; \alpha \; \|b^{\text{noise}}\| \;\approx\; \alpha \; \delta \; \|b\|$$

for some truncation parameter $1 \leq r \leq k$, where δ is the estimate of the noise level and α is a given real parameter.

Here $\delta_{\text{noise}} = 1 \times 10^{-1}$: the best LSQR solution (left) and the hybrid TSVD-LSQR solution stopped by discrepancy principle based on the noise level estimate (right):





Outline

1. Problem formulation
2. Regularization by the Golub-Kahan iterative bidiagonalization
3. Noise approximation
4. Regularization and denoising
5. **Summary and future work**

Summary: Using GK, cheap and accurate estimate of the noise level and noise vector can be obtained. Then one can, e.g.

- stop LSQR at the iteration k_{noise};
- stop a hybrid method based on the discrepancy principle;
- try to denoise the right-hand side b and solve the problem again;
- ...

Future work:

- large scale problems (determining k_{noise});
- rigorous justification of the finite precision behavior;
- applications in regularization and denoising.