Arnoldi-Tikhonov Methods for Sparse Reconstruction

Silvia Gazzola†, James Nagy*, and Paolo Novati†

∗ Department of Mathematics, University of Padua, Padua, ITALY

∗ Department of Mathematics and Computer Science, Emory University, Atlanta GA, USA

gazzola@math.unipd.it, nagy@mathcs.emory.edu, novati@math.unipd.it

1. Problem Formulation

GOAL: compute an approximate solution x of

$$Ax + e = b\,,$$

where

- $A \in \mathbb{R}^{N\times N}$, ill-conditioned, N large;
- e unknown noise.

Employing Tikhonov-like regularization:

$$x \simeq x_\lambda = \arg\min_{x\in\mathbb{R}^N} \{\mathcal{J}(x) + \lambda\mathcal{R}(x)\}\,,$$

where

- $\mathcal{J}(x)$ fit-to-data term. Example:
  - $\mathcal{J}(x) = \|b - Ax\|_q^q$, $1 \le q \le 2$. In the following: $q = 2$.
- $\mathcal{R}(x)$ regularization term. Examples:
  - $\mathcal{R}(x) = \|x\|_p^p$, $1 \le p \le 2$. In particular: $p = 1$ to enforce sparsity.
  - $\mathcal{R}(x) = \|Lx\|_p^p$, $1 \le p \le 2$, $L \in \mathbb{R}^{M\times N}$ regularization matrix.
  - $\mathcal{R}(x) = \|L(x - x^*)\|_p^p$, $1 \le p \le 2$, $x^* \in \mathbb{R}^N$ initial guess.
  - $\mathcal{R}(x) = TV(x) = \|\sqrt{(D_1^h x)^2 + (D_1^v x)^2}\|_1$ (Total Variation regularization).
- $\lambda > 0$ regularization parameter.

2. Arnoldi-Tikhonov methods

GOAL: approximate $x_\lambda \simeq x_{\lambda,m} \in x^* + \mathcal{K}_m(A, r^*) = x^* + \text{span}\{r^*, Ar^*, \dots, A^{m-1}r^*\}$ (where $r^* = b - Ax^*$).
By the Arnoldi algorithm:

$$AW_m = W_{m+1}\bar{H}_m\,, \quad \text{where} \quad W_m \in \mathbb{R}^{N\times m},\ W_m^TW_m = I,\ \text{range}(W_m) = \mathcal{K}_m(A, r^*).$$

By taking $x_{\lambda,m} = x^* + W_my$ in $\min_{x\in\mathbb{R}^N}\left\{\|b - Ax\|_2^2 + \lambda\|L(x - x^*)\|_2^2\right\}$:

$$x_{\lambda,m} = x^* + W_my_{\lambda,m} \quad \text{where} \quad y_{\lambda,m} = \arg\min_{y\in\mathbb{R}^m}\left\|\begin{bmatrix}\bar{H}_m \\ \sqrt{\lambda_m}L_m\end{bmatrix}y - \begin{bmatrix}\|r^*\|_2e_1 \\ 0\end{bmatrix}\right\|_2^2.$$

- Choice of λ: if $\|e\|$ known, secant update method (i.e., discrepancy principle):

$$\lambda_m = \left|\frac{\eta\|e\| - \phi_m(0)}{\phi_m(\lambda_{m-1}) - \phi_m(0)}\right|\lambda_{m-1}\,, \quad \text{where} \quad \phi_m(\lambda) = \|b - Ax_{\lambda,m}\| = \|\|r^*\|e_1 - \bar{H}_my_{\lambda,m}\|\,.$$

- Choice of L: adaptively defined (iteratively reweighted norm approach)
  - $p = 1 : L = L_m = \text{diag}(1/\sqrt{|x_{\lambda,m-1}|})$ (componentwise).
  - $TV(x)$:

$$L = L_m = S_{m-1}D_1^{hv}\,, \quad \text{where } S_{m-1} \text{ depends on } x_{\lambda,m-1}, \quad D_1^{hv} = \begin{bmatrix}D_1\otimes I \\ I\otimes D_1\end{bmatrix} = \begin{bmatrix}D_1^h \\ D_1^v\end{bmatrix}, \ (D_1 \text{ 1st derivative})$$

References

[1] D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari. Tikhonov regularization and the L-curve for large discrete ill-posed problems. *J. Comput. Appl. Math. (2000).*

[2] S. Gazzola and J. G. Nagy. Generalized Arnoldi-Tikhonov method for sparse reconstruction. *SIAM J. Sci. Comput. (2014).*

[3] S. Gazzola and P. Novati. Automatic parameter setting for Arnoldi-Tikhonov methods. *J. Comput. Appl. Math. (2014).*

[4] Y. Saad. A flexible inner-outer preconditioned GMRES algorithm. *SIAM J. Sci. Comput. (1993).*

[5] B. Wohlberg and P. Rodriguez. An Iteratively Reweighted Norm algorithm for minimization of Total Variation functionals. *IEEE Signal Processing Letters (2007).*

[6] S. J. Wright, M. A. T. Figueiredo, and R. D. Nowak. Sparse Reconstruction by Separable Approximation. *IEEE Trans. Signal Processing (2009).*

3. The $\mathcal{R}(x) = \|x\|_1$ case

Consider the standard form transformation

$$\min_{\tilde{x}}\{\|b - \tilde{A}_m\tilde{x}\|_2^2 + \lambda\|\tilde{x} - \tilde{x}^*\|_2^2\}\,, \quad \text{where} \quad \begin{aligned}\tilde{A} &= AL^{-1},\\ \tilde{x} &= Lx,\\ \tilde{x}^* &= Lx^*.\end{aligned}$$

Since $L = L_m$ (iteration-dependent), AL^{-1} is flexibly "preconditioned".
Flexible Arnoldi algorithm:

$$AZ_m = \hat{W}_{m+1}\hat{H}_m\,, \quad \text{where} \quad Z_m = \text{span}\{L_1^{-1}\hat{w}_1, L_2^{-1}\hat{w}_2, \dots, L_m^{-1}\hat{w}_m\}\,,$$

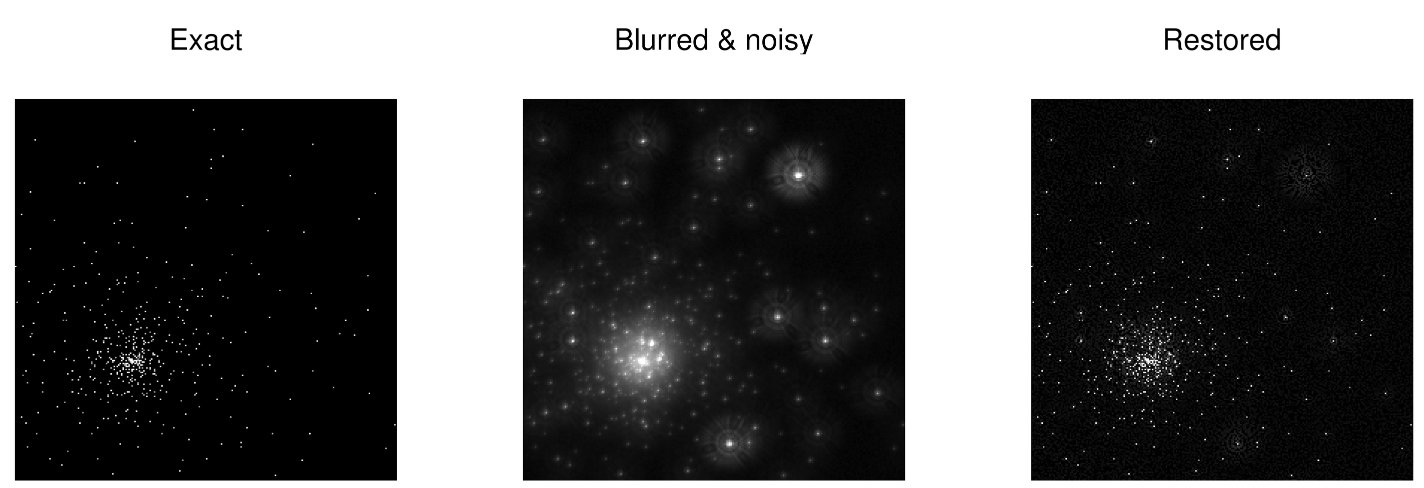
and flexible Arnoldi-Tikhonov:

$$x_{\lambda,m} = x^* + Z_my_{\lambda,m} \quad \text{where} \quad y_{\lambda,m} = \arg\min_{y\in\mathbb{R}^m}\left\|\begin{bmatrix}\hat{H}_m \\ \sqrt{\lambda}I\end{bmatrix}y - \begin{bmatrix}\|r^*\|_2e_1 \\ 0\end{bmatrix}\right\|_2^2.$$

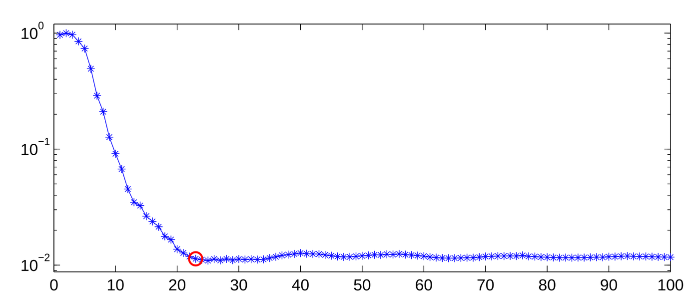
λ computed according to the secant update method.

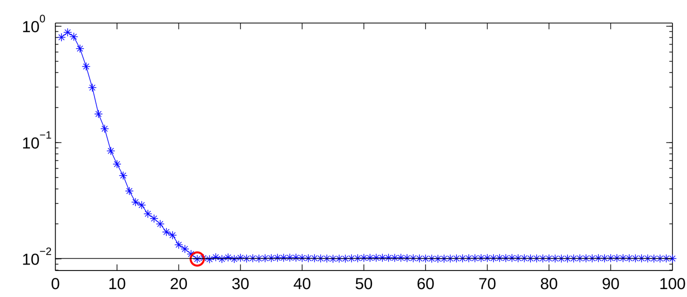
4. Numerical Experiments: $\mathcal{R}(x) = \|x\|_1$

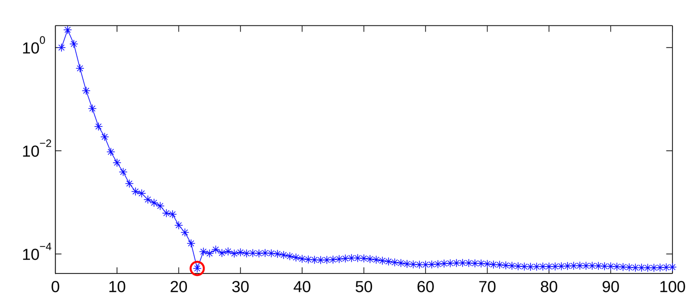
Star cluster test problem (from *Restore Tools*).



(1% Gaussian noise)

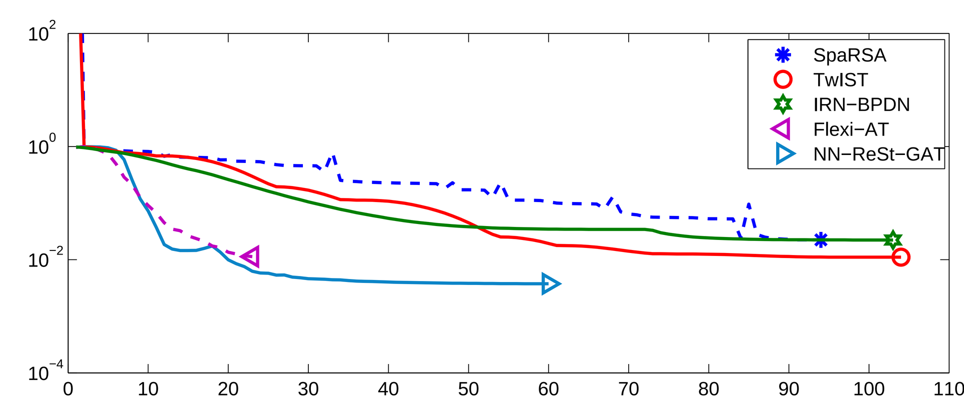






- iterations: 23;
- relative error: $1.1349\cdot 10^{-2}$;
- $\lambda = 1.1976\cdot 10^{-4}$.

Comparisons:



5. The $\mathcal{R}(x) = TV(x)$ case

Restarted Arnoldi-Tikhonov method

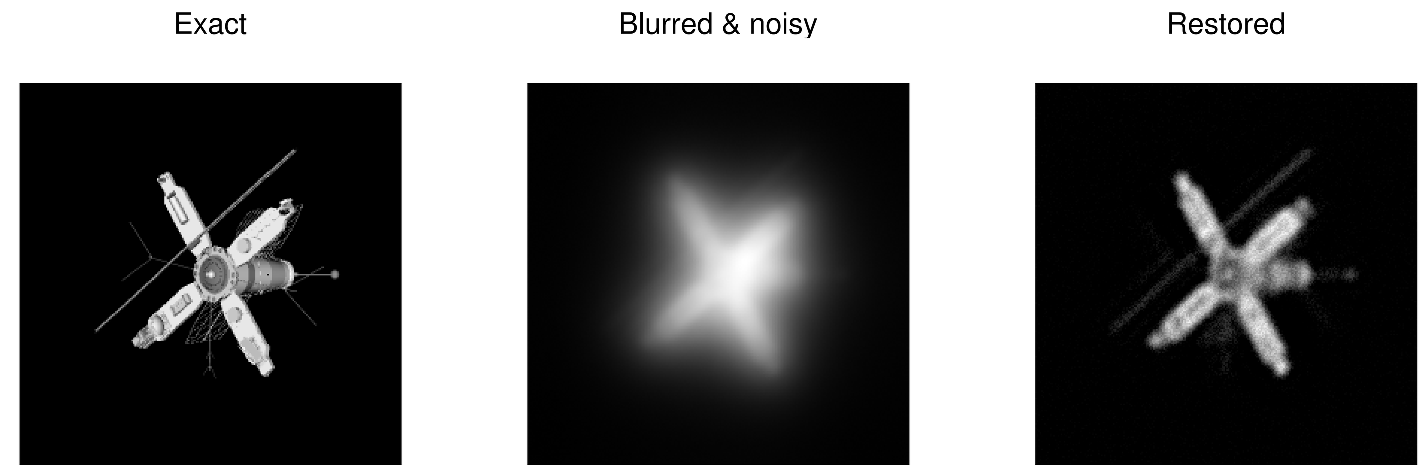
- Outer iterations: for $k = 1, \dots, m_{\text{OUT}}$
  - provide a meaningful initial guess x_k^*; consider $r_k^* = b - Ax_k^*$;
  - update the regularization matrix L_k
- Inner iterations: for $m = 1, \dots,$ until the discrepancy principle is satisfied

$$\text{compute} \quad y_{\lambda,m} = \min_{y\in\mathbb{R}^m}\left\|\begin{bmatrix}\bar{H}_m \\ \sqrt{\lambda}L_m\end{bmatrix}y - \begin{bmatrix}\|r_k^*\|_2e_1 \\ 0\end{bmatrix}\right\|_2^2$$

λ computed according to the secant update method; $x_{\lambda,m} = x_k^* + W_my_{\lambda,m}$.

6. Numerical Experiments: $\mathcal{R}(x) = TV(x)$

Satellite test problem (from *Restore Tools*).



(1% Gaussian noise)

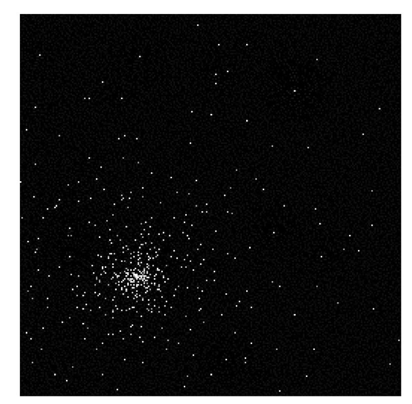
Iterations: 210 (m_{OUT}=200); relative error: $3.0092\cdot 10^{-1}$; $\lambda = 1.0966\cdot 10^{-5}$.

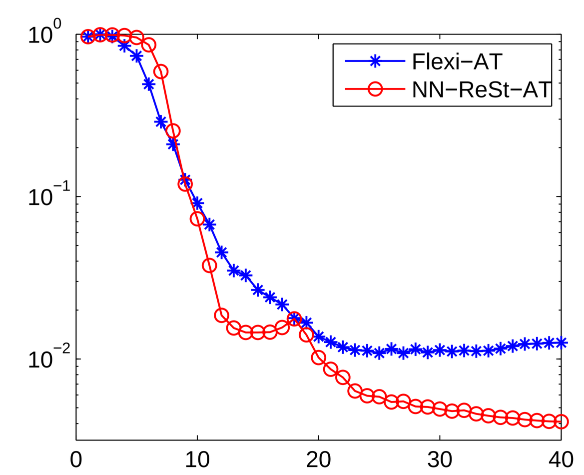
7. Enforcing nonnegativity

At each restart, projection of x_k^* into the nonnegative orthant

$$\mathbb{P} = \{x \in \mathbb{R}^N : [x]_i \ge 0 \quad \forall i = 1, \dots, N\}\,.$$

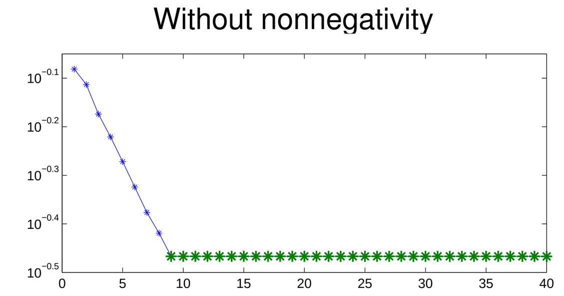
$\mathcal{R}(x) = \|x\|_1$ case

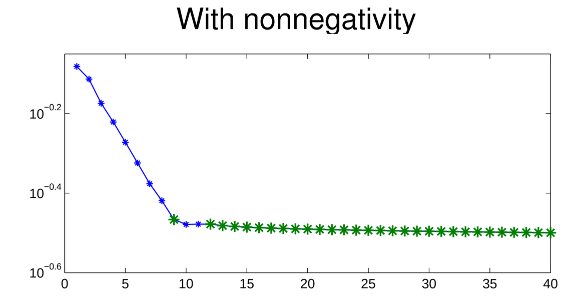




Relative errors: $1.2420\cdot 10^{-2}$ (Flexi-AT), $3.9614\cdot 10^{-3}$ (NN-ReSt-AT)

$\mathcal{R}(x) = TV(x)$ case





Relative errors: $3.4128\cdot 10^{-1}$ (without NN), $3.1658\cdot 10^{-1}$ (with NN)