Sparse reconstruction by means of the standard Tikhonov regularization

Shuai Lu and Sergei V. Pereverzev
Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Science, Altenbergerstrasse 69, A-4040 Linz, Austria (shuai.lu@oeaw.ac.at, sergei.pereverzyev@oeaw.ac.at)

Abstract. It is a common belief that Tikhonov scheme with \( \| \cdot \|_{L^2} \)-penalty fails in sparse reconstruction. We are going to show, however, that this standard regularization can help if the stability measured in \( L^1 \)-norm will be properly taken into account in the choice of the regularization parameter.

The crucial point is that now a stability bound may depend on the bases with respect to which the solution of the problem is assumed to be sparse.

We discuss how such a stability can be estimated numerically and present the results of computational experiments giving the evidence of the reliability of our approach.

1. Introduction

In this paper, we will discuss a practically important problem on the recovery of an element of interest which has a sparse expansion with respect to a preassigned basis. Such a problem often arise in scientific context, ranging from image reconstruction and restoration to wavelet denosing [1], to inverse bifurcation analysis [3] etc.

In a rather general form the problems can be represented as an operator equation

\[
Ax = y
\]

with a linear operator \( A \in \mathcal{L}(X,Y) \) acting between Hilbert spaces \( X \) and \( Y \) and having a non-closed range \( R(A) \). This non-closedness is reflected in the discontinuity of the inverse operator \( A^{-1} \), if it exists. In general, the best-approximate solution \( A^\dagger y \), where \( A^\dagger \) is the Moore-Penrose inverse of \( A \), does not depend continuously on the right-hand side \( y \). In real application, usually only noisy data \( y^\delta \) are available such that

\[
\|y - y^\delta\|_Y \leq \delta.
\]

We consider the problem of a sparsity-promoting regularization as it is presented in the paper [1] by Daubechies, Defrise and De Mol. The focus in this problem is to recover \( x^\dagger = A^\dagger y \) from (1), (2) under the assumption that it has a sparse expansion

\[
x^\dagger = \sum_i \hat{x}_i \phi_i, \quad \#\{\hat{x}_i \neq 0\} \text{ sufficient small},
\]
on some given system \{\phi_i\} of linear independent elements \phi_i \in X. We define sparsity of \(x^\dagger\) by the presence of a small number of large coefficients \(\hat{x}_i\) in (3) and zeroes elsewhere, although a \textit{priori} we do not know either the number of non-zero coefficients, or their indices.

Generally speaking, compared with some classical ill-posed problems (c.f. [2]), in sparse reconstruction we now need to recover the exact solution \(x^\dagger\) as an element of some space \(Z_\rho\) promoting sparsity and equipped with a distance \(\rho = \rho(u_1, u_2), u_1, u_2 \in Z_\rho\). The ordinary Tikhonov method, for example, can be employed in such a recovery under the conditions that it produces the approximations \(x_\alpha^\delta = \arg \min \{\|Ax - y^\delta\|^2_Y + \alpha\|x\|^2_Z\}\) which also belong to \(Z_\rho\), and for noise free data these approximations \(x_\alpha\) converge to \(x^\dagger\) in \(Z_\rho\) as \(\alpha\) tends to zero, i.e. for \(\alpha > 0, x_\alpha \in Z_\rho\), and

\[
\lim_{\alpha \to 0} \rho(x^\dagger, x_\alpha) = 0. \tag{4}
\]

Note that in applications we have in mind such conditions are not restrictive at all.

Keeping in mind that a non-Hilbert space \(Z_\rho\) is not built into the construction of Tikhonov method, the performance of regularized approximations \(x_\alpha^\delta\) in \(Z_\rho\) can only be controlled through the choice of the regularization parameter \(\alpha\). At the same time, the existing parameter choice strategies does not take into account our wish to regularize a problem in a non-Hilbert space \(Z_\rho\).

Indeed, the idea of the discrepancy principle, for example, is to choose the regularization parameter \(\alpha\) via \(\alpha_{Disc} := \sup\{\alpha > 0 : \|Ax_\alpha^\delta - y^\delta\|_Y \leq c\delta\}\), where \(c\) is some fixed constant. Clearly, the space \(Z_\rho\) has no impact on such a choice of the regularization parameter. Therefore, one should not expect that a regularized solution \(x_\alpha^\delta\) with \(\alpha = \alpha_{Disc}\) will perform well in the space \(Z_\rho\).

At this point we would like to have a retrospect on the recently introduced criteria in inverse problems theory named balancing principle. In its general form given in [6] the balancing principle is based on an error decomposition in the space \(Z_\rho\) due to the triangle inequality

\[
\rho(x^\dagger, x_\alpha^\delta) \leq \rho(x^\dagger, x_\alpha) + \rho(x_\alpha, x_\alpha^\delta),
\]

and on some reliable estimation for the stability term \(\rho(x_\alpha, x_\alpha^\delta)\) of the form

\[
\rho(x_\alpha, x_\alpha^\delta) \leq \psi(\delta, \alpha), \tag{5}
\]

where \(\psi(\delta, \alpha)\) is assumed to be a decreasing function of \(\alpha\). Then in accordance with the balancing principle a regularization parameter \(\alpha = \alpha_+\) is selected from a sequence \(\Sigma_N = \{\alpha_j\}_{j=1}^N\) as follows

\[
\alpha_+ = \max\{\alpha_i \in \Sigma_N : \forall j = 1, 2, \cdots, i-1, \rho(x_\alpha^\delta, x_\alpha^\delta_j) \leq 4\psi(\delta, \alpha_j)\}. \tag{6}
\]

To draw a conclusion from this parameter choice we note that in view of (4) one can always find a non-decreasing function \(\varphi : [0, \alpha_N] \to [0, \infty)\) such that \(\varphi(0) = 0\) and for any \(\alpha \in [0, \alpha_N]\)

\[
\rho(x^\dagger, x_\alpha) \leq \varphi(\alpha). \tag{7}
\]

As in [6], any of such functions \(\varphi\) will be called \textit{admissible} for \(x^\dagger\) and can be used as a measure for the rate of convergence in (4). Then under a rather general assumption concerning the function \(\psi\) in (5) the conclusion from the parameter choice (6) can be drawn in the form of the following error bound

\[
\rho(x^\dagger, x_\alpha^\delta) \leq c \min\{\varphi(\alpha) + \psi(\delta, \alpha), \alpha \in \Sigma_N, \varphi \text{ is admissible}\}, \tag{8}
\]

where \(c\) is a constant depending only on \(\psi\) and \(\Sigma_N\).
Inequality (8) has a clear interpretation. Using (5), (7) the error of an approximate solution \( x_\alpha^\delta \) can be estimated as

\[
\rho(x^\dagger, x_\alpha^\delta) \leq \varphi(\alpha) + \psi(\delta, \alpha). \tag{9}
\]

Then the parameter choice \( \alpha = \alpha_+ \) allows us to reach (up to a constant factor \( c \)) the best error bound of the form (9) that in principle can be obtained for \( \alpha \in \Sigma_N \).

We would like to stress that the parameter choice strategy (6) is based on the function \( \psi \) alone. One does not need to know an admissible function corresponding to the best convergence rate in (4), while the information about the stability, as given by \( \psi \) in (5), is extremely important. The balancing principle (6) can be implemented in any metric space and for any regularization method provided such an information is available.

If we are going to recover a solution directly in the Hilbert space \( X \), where \( \rho(u_1, u_2) = \| u_1 - u_2 \|_X \), then for standard regularization methods the function \( \psi(\delta, \alpha) \) in (5) can be taken as

\[
\psi(\delta, \alpha) = k\delta\alpha^{-1/2}, \tag{10}
\]

where the constant \( k \) depends only on a method in hand. For Tikhonov regularization one can take \( k = 1/2 \). We refer to Chapter 4 of the book [2] by Engl, Hanke and Neubauer for further details.

Thus, in the Hilbert space setting the stability estimation (5) is problem independent. We will show in Section 3 and 4 that it is not the case for a regularization implementation in a non-Hilbert space \( Z_\rho \).

2. Discretized Tikhonov regularization for sparse reconstruction

It is worth to notice that in practice we are able to handle only a finite section of an expansion (3). Therefore, in reality one tries to recover a sparse structure of a projection \( P_M x^\dagger = \sum_{i=1}^{M} \hat{x}_i \phi_i \). Here \( P_M \) is the orthogonal projector from \( X \) onto \( \text{span}\{\phi_i\}_{i=1}^{M} \). Note that \( P_M x^\dagger \) solves the equation \( Ax = Ax^\dagger - A(I - P_M)x^\dagger \). Moreover, a system \( \{\phi_i\} \) has usually a reasonable approximation property such that \( \|(I - P_M)A^*\|_{Y \to X} \to 0 \) as \( M \to \infty \). Then for sufficiently large \( M \) one has \( \|(I - P_M)A^*\|_{Y \to X} \leq \delta/\|x^\dagger\|_X \), and \( \|y^\delta - (Ax^\dagger - A(I - P_M)x^\dagger)\|_Y \leq \|Ax^\dagger - y^\delta\|_Y + \|x^\dagger\|_X \|(I - P_M)A^*\|_{Y \to X} \leq 2\delta \).

It means that a level of a noise in the right-hand side of the equation

\[
AP_M x = y^\delta \tag{11}
\]

is of the same order of magnitude as in \( y^\delta \), and it can be used for a recovery of the sparse structure in \( P_M x^\dagger \) from (11).

If \( \{\phi_i\} \) is an orthonormal system then the sparsity of \( P_M x^\dagger \) can be recovered from (11) by minimizing the functional

\[
D_\alpha(A, y^\delta, \{\phi_i\}, x) = \|AP_M x - y^\delta\|_Y^2 + \alpha \sum_{i=1}^{M} |\langle x, \phi_i \rangle_X|, \tag{12}
\]

as it has been suggested in the paper [1] by Daubechies, Defrise and De Mol. The minimization of (12) is not so easy. In [1] the functional (12) has been replaced by a sequence of surrogate functionals which are easier to minimize. At the same time, the quality of the recovery via minimizer of (12) depends on the choice of \( \alpha \). In [1] it has been suggested to choose \( \alpha = \alpha(\delta) \) in such a way that \( \alpha(\delta) \to 0 \) and \( \delta^2/\alpha(\delta) \to 0 \) as \( \delta \to 0 \). In contrast to (6) it is not clear how \( \alpha \) should be chosen for a fixed noise level \( \delta \).
Now we are going to discuss an indirect implementation of a sparsity-promoting regularization within the framework of ordinary Tikhonov method. Applying it to (11) we obtain a regularized approximation \( x_{\alpha,M}^\delta \) that can be written as

\[
x_{\alpha,M}^\delta = (\alpha I + P_M A^* A P_M)^{-1} P_M A^* y^\delta = \sum_{i=1}^M \hat{e}_{i,M} \phi_i,
\]

where the vector \( \hat{e}_M = (\hat{e}_{1,M}, \hat{e}_{2,M}, \ldots, \hat{e}_{M,M}) \) of coefficients solves a system of linear algebraic equations

\[
\alpha \hat{e}_M + B \hat{e}_M = b_\delta
\]

with a matrix \( B = \{\langle A \phi_i, A \phi_j \rangle_Y \}_{i,j=1}^M \) and a vector \( b_\delta = \{ \langle A \phi_1, y^\delta \rangle_Y, \langle A \phi_2, y^\delta \rangle_Y, \ldots, \langle A \phi_M, y^\delta \rangle_Y \} \), \((\cdot, \cdot)_Y \) is the inner product in a Hilbert space \( Y \). It remains to choose \( \alpha \). The reader is encouraged to consult [2] for more detailed information on Tikhonov regularization and discretized Tikhonov regularization.

3. A sample-based simulation algorithm

To make the further discussion more concrete we consider an example, where \( A \) is the linear integral operator \( Ax(t) = \int_0^1 a(t,s) x(s) ds, \quad t \in [0,1] \), with the Green’s function

\[
a(t,s) = \begin{cases} 
  t(1-s), & s \ge t, \\
  s(1-t), & s \le t,
\end{cases}
\]

as a kernel. In inverse problems community this operator is frequently used as a prototype example (see e.g. a recent paper [8] by Neubauer, or the paper [5] by Mair and Ruymgaart). Moreover, among orthonormal systems \( \{\phi_i\} \) discussed in the paper [1] by Daubechies, Defrise and De Mol, we choose the simplest one, where \( \phi_i = \phi_i^M(t) \) are \( L_2 \)-orthonormalized characteristic functions of the intervals \( [\frac{i-1}{M}, \frac{i}{M}], i = 1, 2, \ldots, M \). For such a system the second term in (12) is just equal to \( \alpha \| P_M x \|_{L_1} \), that means we consider the minimizer of (12) as an element of the space \( L_1 = L_1(0,1) \).

It hints at the use of \( L_1 \)-distance \( \rho(x_{\alpha,M}^\delta, x_{\alpha,M}^{\delta'}) = \| x_{\alpha,M}^\delta - x_{\alpha,M}^{\delta'} \|_{L_1} \) in the balancing principle (6) to promote a sparsity of Tikhonov regularized approximation indirectly. The corresponding norm is thus utilized in the balancing principle as the indicator of solution properites we want to construct.

To implement the parameter choice rule (6) for such a distance \( \rho \) one needs to know the function \( \psi(\delta, \alpha) \) in the stability estimation (5). In view of obvious inequalities

\[
\rho(x_{\alpha,M}^\delta, x_{\alpha,M}^{\delta'}) = \| x_{\alpha,M}^\delta - x_{\alpha,M}^{\delta'} \|_{L_1} \le \| x_{\alpha,M}^\delta - x_{\alpha,M}^{\delta'} \|_{L_2} \le \delta \| (\alpha I + P_M A^* A P_M)^{-1} P_M A^\dagger \| \le \frac{\delta}{2 \sqrt{\alpha}}
\]

one is tempted to use the rule (6) with the function \( \psi(\delta, \alpha) = \frac{1}{2} \delta \alpha^{-1/2} \) as in (10).

Unfortunately the stability estimation (15) is too rough to allow the reconstruction of a sparse structure. It can be seen from Figure 1 (left) displaying the graph (dashed line) of the exact solution \( x^\dagger = P_M x^\dagger = 6(\phi_3^M + \phi_4^M + \phi_5^M + \phi_6^M) \), \( M = 100 \), together with the graph (dotted line) of \( x_{\alpha,M}^\delta \) given by (13), where perturbed data \( y^\delta \) corresponds to \( \delta = 10^{-4} \) (details of a noise simulation are presented below), and \( \alpha = 8.1860 \times 10^{-5} \) is chosen from

\[
\Sigma_{100} = \{ \alpha_i = \alpha_1 q^{i-1}, i = 1, 2, \ldots, 100, \quad \alpha_1 = 10^{-14}, \quad q > 1 \}
\]
in accordance with (6) for $\psi(\delta, \alpha) = \frac{1}{2}\delta \alpha^{-1/2}$ as in (15). It is clear that no sparse structure can be reconstructed from such $x_{\alpha}^\delta$. By the way, a similar situation appears in the case of $\alpha = \alpha, \delta_{\text{is}}$ chosen in accordance with the discrepancy principle mentioned above.

At the same time, in Figure 1 (right) one can see also the graph (solid line) of $x_{\alpha,M}^\delta$ with $\alpha = 2.9463 \times 10^{-9}$ chosen from (16) in accordance with (6) where

$$\psi(\delta, \alpha) = 10 \delta \alpha^{-0.3}. \tag{17}$$

This time the sparse structure of $x_{\alpha,M}^\delta$ is similar to $x^\dagger$ and it can be clearly seen.

We briefly explain the reason to take $\psi(\delta, \alpha)$ in the form of (17). The expression (17) was obtained by a heuristic sample-based method. We took the exact solution $x^\dagger = 8\phi_4^{50} + 10\phi_4^{50}$ and calculate $y_0 = Ax^\dagger$ and the vector

$$b_0 = (\langle A\phi_1^{50}, y_0 \rangle_{L_2}, \langle A\phi_2^{50}, y_0 \rangle_{L_2}, \ldots, \langle A\phi_{50}^{50}, y_0 \rangle_{L_2}).$$

Then Tikhonov regularized approximations $x_{\alpha,50}$ were calculated in accordance with (13), (14) for all $\alpha$ from (16) using the vector $b_0$ as noise free data. Then a normalized random vector $\xi \in \mathbb{R}^{50}$ with uniformly distributed random components was given by a random number generator, and noisy vector $b_\delta = b_0 + \delta \xi$ was simulated. This vector was used in (14) to construct $x_{\alpha,50}^\delta$ for all $\alpha$ from (16). As a result, one was able to calculate the norms $d(\alpha_i) = \|x_{\alpha_i,50} - x_{\alpha_i,50}^\delta\|_{L_1}$ for $\alpha_i \in \Sigma_{100}$. Then in the expression $\psi(\delta, \alpha) = c\delta \alpha^{-v}$ the parameters $c$ and $v$ were estimated by the standard nonlinear fitting to approximate the data $d(\alpha_i)$ by the values of $\psi(\delta, \alpha_i)$, $\alpha_i \in \Sigma_{100}$, $q = 1.3$. This procedure was repeated for $\delta = 10^{-4}, 10^{-5}, 10^{-6}$ and for the exact solution $x^\dagger = 8\phi_4^{50} + 10\phi_4^{50}$. Corresponding values of $c$ and $v$ are presented in the Table 1. If function $\psi(\delta, \alpha) = c\delta \alpha^{-v}$ is used in the stability estimation of the form (5) then the parameters $c$ and $v$ play different roles. The parameter $v$ can be seen as a degree of instability: the large $v$ the more a noisy approximation deviates from its noise free counterpart.

The role of $c$ is not of the same importance: it basically scales a noise level. Keeping in mind such interpretation we take the approximated mean value of $v$ from the Table 1, which is about $v = 0.3$, and $c = 10$. It gives us the function (17).

Although such a stability estimation has been obtained for the system $\{\phi_i^{50}\}_{i=1}^{150}$, it allows a reconstruction of the sparse structure with respect to other systems of piece-wise constant functions such as $\{\phi_i^{100}\}_{i=1}^{100}$ or $\{\phi_i^{25}\}_{i=1}^{25}$. It can be seen from Figure 2 (left) and Figure 2 (right), where the exact solutions $x^\dagger = \phi_3^{100} + 8\phi_{58}^{100}$ and $x^\dagger = 6\phi_{10}^{25} + 7\phi_{12}^{25} + 8\phi_{14}^{25}$ (dashed lines) are displayed together with their approximations $x_{\alpha,100}^\delta$, $\alpha = 4.9793 \times 10^{-9}$, and $x_{\alpha,25}^\delta$,
Table 1. Sample-based simulation results for different noise level $\delta$ as introduced in Section 3.

| $\delta$  | $c$   | $v$    |
|-----------|-------|--------|
| 0.0001    | 8.0611| 0.32031|
| 0.00001   | 17.365| 0.28558|
| 0.000001  | 10.119| 0.30889|

For $x^1 = 8 \times \phi_{10}^{50} + 10 \times \phi_{6}^{50}$. $\delta = 0.0001$ $c = 8.0611$ $v = 0.32031$

$\delta = 0.00001$ $c = 17.365$ $v = 0.28558$

$\delta = 0.000001$ $c = 10.119$ $v = 0.30889$

Figure 2. Orthonormal basis: reconstruction with the stability estimation (17) used in the standard Tikhonov regularization. The exact solutions are $x^1 = \phi_3^{100} + 8\phi_{98}^{100}$ (left) and $x^1 = 6\phi_{10}^{25} + 7\phi_{12}^{25} + 8\phi_{13}^{25}$ (right). In both figures, the dashed line is the exact solution and the solid line is the reconstruction.

$\alpha = 3.2118 \times 10^{-10}$ (solid lines). Here $\delta = 10^{-4}$ (5 – 10% of the relative error), and the regularization parameters have been chosen in accordance with the balancing principle (6) corresponding to $L_1$-distance and $\psi$ as in (17). In both cases a sparse structure can be easily recognized.

4. Numerical tests for a non-orthonormal basis

Let us consider now the Abel integral operator

$$Ax(t) = \int_0^t \frac{x(s)}{\sqrt{t-s}} \, ds, \quad t \in [0, 1], \quad (18)$$

which is also used in the inverse problems theory as a prototype example (see, e.g., [2], p.9). Moreover, we also change a system $\{\phi_i\}$ and consider the recovery of a sparse structure with respect to the system of piece-wise linear B-splines

$$\phi_i(t) = \phi_i^M(t) = \begin{cases} 
M(t - \frac{i-1}{M}), & t \in \left[\frac{i-1}{M}, \frac{i}{M}\right], \\
M\left(\frac{i+1}{M} - t\right), & t \in \left[\frac{i}{M}, \frac{i+1}{M}\right], \\
0, & t \notin \left[\frac{i-1}{M}, \frac{i+1}{M}\right],
\end{cases} \quad i = 1, 2, \ldots, M - 1.$$  

This system is also discussed in the context of a sparsity recovery (see, e.g., the dissertation [4] by Malioutov). It is not an orthogonal system, but the version (13), (14) of the ordinary Tikhonov regularization can be also used in the considered case without changes. We just need
a stability estimation to implement the balancing principle (6) with $L_1$-distance for $x_\delta^\theta = x^\delta_{\alpha_1, M}$, $\alpha_i \in \Sigma_{100}$.

Using the heuristic sample-based method for $x^\dagger = 3\phi_{12}^{100}$, as it has been described in Section 3, we come up with $L_1$-stability estimation of the form (5), where

$$\psi(\delta, \alpha) = 0.1\delta\alpha^{-0.15}. \quad (19)$$

Comparing it with (17) one can observe that this time a degree of instability $v$ is two times smaller than $v = 0.3$ obtained for the integral operator with Green’s kernel. It looks reasonable, since the inversion of such operator is known to be more ill-posed than the Abel inversion.

To test a reliability of $L_1$-stability estimation given via the function (19) we incorporate this function into the balancing principle (6) and use it for recovering a sparse structure with respect to another system of B-spline $\{\phi_{100}^i\}$.

Typical results are presented in Figure 3 (left) and Figure 3 (middle), where the graphs (dashed lines) of the exact solutions $x^\dagger = 3\phi_{10}^{100} + 6\phi_{42}^{100}$ and $x^\dagger = 3\phi_{10}^{100} + 6\phi_{12}^{100}$ are displayed together with their Tikhonov approximations $x^\delta_{\alpha_1, 100}$, $\alpha = 1.6929 \times 10^{-7}$, and $x^\delta_{\alpha, 100}$, $\alpha = 2.7264 \times 10^{-7}$. Here $\delta = 10^{-4}$ and the regularization parameters have been chosen from $\Sigma_{100}$ ($\alpha_1 = 10^{-10}, q = 1.1$) in accordance with the balancing principle based on $L_1$-distance and $\psi$ as in (19).

Note that a sparsity reconstruction presented in Figure 3 (left) and Figure 3 (middle) is of the same quality as in the tests by Malioutov (see Fig. 4.1-4.3 there), where a regularization via minimization of a Tikhonov type functional with $l_1$-penalty $\Sigma|x^\theta|_1$ has been used, which is similar to (12).

It is also interesting to note that the test presented in Figure 3 (middle) is rather hard, since the modes $\phi_{100}^{10}$ and $\phi_{12}^{10}$ are very close to each other (narrow band problem). The quality of the reconstruction given by the standard scheme (13), (14) is acceptable, but it can be essentially improved by a simple self-regularization, where no additional regularization is needed.

It is based on the observation that in all our tests the approximation $x^\dagger_{\alpha, M}$ with properly chosen regularization parameters had only a few relatively large coefficients $x^\dagger_{i, M}$ in the representation (3), and they can be easily distinguished.

Let $\theta$ be a subset of $\{1, 2, \ldots, M\}$ containing the indices of these large coefficients, and $P_\theta$ be the orthoprojector onto $\text{span}\{\phi_{10}^M\}_{i \in \theta}$. Let $x^\theta = \sum_{i \in \theta} x_i \phi_{10}^M$ be the solution of the equation $P_\theta A^* P_\theta x = P_\theta A^* y^\theta$. The vector $\hat{x}^\theta = \{\hat{x}_i\}_{i \in \theta}$ of its coefficients can be found from a linear algebraic system $B_\theta \hat{x}^\theta = b^\theta$, where the matrix $B_\theta$ and the vector $b^\theta$ contain the components
of $B$ and $b_\delta$ from (14) with indices $(i,j)$ and $i$ such that $i,j \in \theta$ (no additional calculations are needed).

The approximation $x_\theta$ is based on the previous regularized approximation $x_{\alpha,M}^\delta$, but usually it allows much better resolution. Figure 3 (right) presents an example of a perfect sparsity reconstruction given by $x_\theta$ which is based on the approximation shown in Figure 3 (middle) suggesting $\theta = \{8, 9, \ldots, 14\}$.

5. Conclusion

It is well-known that the Tikhonov scheme was designed for a regularization in Hilbert spaces such as $L_2$. In the seminal paper [1] discussing the problem of the reconstruction of a sparse structure the authors suggest the use of a Tikhonov scheme, where the original $L_2$-penalty term is changed for a penalty in $L_p$-norm, $1 \leq p < 2$. They explain it by the fact that in $L_p$-norm, $1 \leq p < 2$, a sparse structure can be seen more clearly than in $L_2$-space. At the same time, it should be remind that the standard Tikhonov scheme is used not only in Hilbert spaces. For example, in the classical book [9] this scheme is used in $C$-space for reconstructing continuous function from its noisy Fourier coefficients. Following the logic of [1] one would employ for such a reconstruction the Tikhonov scheme with $C$-penalty. But in the papers [10], [7] it has been shown that the use of the standard Tikhonov scheme with a Hilbert space penalty allows order-optimal reconstruction in $C$-norm. Thus, the range of applications of the standard Tikhonov method is much wider than a Hilbert space regularization. The present paper can be seen as the first attempt to understand the role of this classical scheme in the reconstruction of the sparse structure. We show that controlling the performance of the standard Tikhonov scheme in $L_1$-space by means of the balancing principle based on $L_1$-stability bounds, one can obtain rather accurate reconstruction of the sparse structure. By the way, in our second example the reconstruction is of the same quality as in [4], where a Tikhonov scheme with a non-Hilbert space penalty has been used. It turns out that in contrast to the Hilbert space setting, $L_1$-stability bounds used in our examples are problem dependent. We are going to clarify this dependence in our future studies.

Acknowledgement

This research is supported by the Austrian Fonds Zur Förderung der Wissenschaftlichen Forschung (FWF), Grant P20235-N18.

References

[1] Daubechies I, Defrise M, De Mol C 2004 An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, Comm. Pure Appl. Math. 57, 1413-1457.

[2] Engl H W, Hanke M, Neubauer A 1996 Regularization of Inverse Problems, Kluwer Academic Publishers.

[3] Lu J, Engl H W, Machne R, Schuster P 2007 Inverse bifurcation analysis of a model for mammalian G1/S regulatory module, Lecture Notes in Bioinformatics, 4414, 168-184.

[4] Malioutov D M 2003 A sparse signal reconstruction perspective for source localization with sensor arrays, Massachusetts Institute of Technology, Dissertation.

[5] Mair B A, Ruymgaart F H 1996 Statistical inverse estimation in Hilbert scales, SIAM J. App. Math., 56, 1424-1444.

[6] Mathé P, Pereverzev S V 2006 Regularization of some linear ill-posed problems with discretized random noisy data, Math. Comp. 75, 1913–1929.

[7] Mathé P and Pereverzev S V: Stable summation of orthogonal series with noisy coefficients, J. Approxim. Theory, 118, 66-80 (2002).

[8] Neubauer A 2006 Estimation of discontinuous solutions of ill-posed problems via adaptive grid regularization, J. Inv. Ill-posed problems, 14, 705-716.

[9] Tikhonov A N and Arsenin V Y 1977 Solution of ill-posed problems, John Wiley & Sons, New York.

[10] Tsybakov A B 1998 Pointwise and sup-norm sharp adaptive estimation of functions on the Sobolev classes, Ann. Statist., 26, 2420-2469.