Level choice in truncated total least squares

Diana M. Sima, Sabine Van Huffel

Electrical Engineering Department, Katholieke Universiteit Leuven, Belgium

Abstract

In the context of ill-conditioned linear systems, regularization techniques where a hyperparameter is present are often employed. In particular, a truncation level must be chosen in the case of truncation-based methods, since these methods aim at cutting off the contribution of noise or rounding errors in an expansion such as the singular value decomposition. The method of truncated total least squares is an alternative to the truncated singular value decomposition method, but these two methods often yield very similar solutions. We focus on an advantage of the truncated total least squares formulation for selecting the truncation level. It is based on the observation that the truncated total least squares solution becomes more dominated by noise or errors when the truncation level is overestimated. Model selection methods that are often applied in the context of truncated singular value decomposition are modified in this paper for the context of truncated total least squares.

Key words: truncated singular value decomposition, truncated total least squares, filter factors, effective number of parameters, model selection, generalized cross validation

1 Introduction

In this paper, we study truncation methods for the regularization of linear discrete ill-posed problems. Ill-posed problems are problems where the solution does not depend continuously on the input data, where arbitrarily small perturbations in the input data produce arbitrarily large changes in the solution. Discrete ill-posed problems might arise from the discretization of continuous problems, such as integral or differential models. Linear ill-posed problems are linear systems $Ax \approx b$, where the coefficient matrix A of dimension $m \times n$

Email addresses: diana.sima@esat.kuleuven.be (Diana M. Sima), sabine.vanhuffel@esat.kuleuven.be (Sabine Van Huffel).

($m > n$) and the right-hand side vector b of length m are the given data, while the vector x is the solution to be computed.

Extensive characterizations of discrete ill-posed problems of the type $Ax \approx b$ exist in the literature (see the book [10] and the references therein). It is known, for instance, that the properties of such a system can be understood from the singular value decomposition [7] of A. In an ill-posed system, there is no clear gap in the decay of the singular values of A. This makes it difficult to decide the best value that can be used as numerical rank. Moreover, it usually happens that the singular vectors corresponding to smaller and smaller singular values have increasing complexity (meaning that they contain more and more sign changes, oscillations). If the least squares or total least squares solution of $Ax \approx b$ is computed, then these oscillations are emphasized in the solution by the noise in the data or by the effects of working in finite precision arithmetic.

Among the numerous methods of regularization that can be used for approaching discrete ill-posed problems [10], we revisit in this paper one of the simplest: truncation. The aim of this study is the analysis of methods for choosing the truncation level, by adapting several classical methods for model selection. The specific conditions under which the methods here will be most favorable are the errors-in-variables assumptions that both the matrix A and the right-hand side vector b are contaminated by errors. Typically, the methods for choosing a regularization parameter or truncation level use the assumption that b is the sum of a b^{exact} and a b^{noise}, where b^{exact} is in such a way that $Ax = b^{\text{exact}}$ is compatible (and has as solution the true solution x^{exact}), and b^{noise} is a vector with zero mean i.i.d. random elements. This scenario is the basis for many well-known methods for choosing hyperparameters for the regularization of discrete ill-posed problems, such as the discrepancy principle, cross validation and generalized cross validation, information criteria, the L-curve (see [10] for an overview) and also for newer analyses such as the residual exploitation technique in [11].

In the case when the coefficient matrix A itself is the sum of an exact part and a noise part, we can in principle identify the level of the noise in A from its singular value decomposition; then we can truncate the subspace corresponding to the singular values embedded in the noise. This can be achieved with a simple discrepancy principle method (which requires knowledge of the noise magnitude), but also with methods that estimate the noise level directly from data.

But in ill-posed problems, we are not interested in identifying a noise level and separating the "exact" from the "noisy" data in $[A\ b]$. Instead, we are interested in a good reconstruction of the vector x, a solution that is not contaminated by the noise or the round-off errors. To this end, we propose a modification of the generalized cross validation criterion and of Akaike's infor-

mation criterion, which will use the *filter factors* that pertain to the truncated total least squares solution. In this way, we aim at complementing the discussion in [4] – the paper that firstly proposed and analyzed the truncated total least squares as a regularization method; in [4], the methods for choosing the truncation level are either the L-curve, or the minimization of $\|Ax - b\|$. The advantage of our modified generalized cross validation and generalized information criterion is that they work even if the errors-in-variables assumption is not satisfied. This comes from a computational aspect: when there is noise in the data, we make it "work" in our advantage by monitoring at which truncation level it starts to take over the computed solution. When there is no added noise in the data, we make use of the *rounding errors* that normally spoil the accuracy of numerical computations, but will have a similar helpful role, when we solve discrete ill-posed problems with truncation methods.

2 Truncation methods for linear ill-posed problems

2.1 Truncation methods for linear estimation

Truncation methods are an effective way of eliminating unwanted subspaces (such as the noise subspace) from given data sets, using the singular value decomposition (SVD) [7] or other factorizations of the sort. For instance, if for a certain application the noise level is *a priori* known, then it is reasonable to eliminate all information that is below the noise level. It may happen that useful information from the problem is also embedded in that noise; this is an intuitive explanation why *truncation methods* might be too drastic, and thus they are not always the most appropriate methods for all problem settings.

Algorithms 1 and 2 outline the procedures of truncation in the least squares (LS) and, respectively, total least squares (TLS [6, 15]) settings.

Algorithm 1 Truncated SVD (TSVD).

1: Compute the SVD of $A = U'\Sigma'{V'}^\top$, where U' is $m \times n$, V' is $n \times n$ and ${U'}^\top U' = {V'}^\top V' = V'{V'}^\top = I_n$, and Σ' is an $n \times n$ diagonal matrix with the singular values $\sigma'_1 \geq \sigma'_2 \geq \ldots \geq \sigma'_n \geq 0$ on the diagonal;
2: Choose an appropriate truncation level $k \leq n$;
3: Compute the truncated SVD solution

$$x_{\mathrm{TSVD},k} = V'{\Sigma'_k}^\dagger {U'}^\top b, \qquad \text{with } \Sigma'_k = \mathrm{diag}\{\sigma'_1, \ldots, \sigma'_k, \underbrace{0, \ldots, 0}_{n-k}\}. \tag{1}$$

For truncating in the least squares context, the truncated SVD (TSVD) algorithm cuts off the contribution of the $n-k$ singular triplets of A corresponding

Algorithm 2 Truncated TLS (TTLS).

1: Compute the SVD of $[A\ b] = U\Sigma V^\top$, where U is $m \times (n+1)$, V is $(n+1)\times(n+1)$ and $U^\top U = V^\top V = VV^\top = I_{n+1}$, and Σ is an $(n+1)\times(n+1)$ diagonal matrix with the singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq \sigma_{n+1} \geq 0$ on the diagonal;
2: Choose an appropriate truncation level $k \leq n$.
3: Let $[A_k\ b_k] := U\Sigma_k V^\top$, where Σ_k is Σ with the smallest $n-k$ singular values set to zero, $\Sigma_k = \mathrm{diag}\{\sigma_1, \ldots, \sigma_k, \sigma_{k+1}, 0, \ldots, 0\}$;
4: Solve in the TLS sense the truncated problem $A_k x \approx b_k$ and obtain the truncated TLS solution

$$x_{\mathrm{TTLS},k} = -V_{12}^k (V_{22}^k)^\dagger = -V_{12}^k (V_{22}^k)^T / \|V_{22}^k\|^2, \tag{2}$$

where we partition V as (with $l = n-k+1$):

$$V = \begin{array}{c} \overset{k}{\longleftrightarrow}\ \overset{l}{\longleftrightarrow} \\ \begin{bmatrix} V_{11}^k & V_{12}^k \\ V_{21}^k & V_{22}^k \end{bmatrix} \begin{array}{l} \updownarrow n \\ \updownarrow 1 \end{array} \end{array} \tag{3}$$

to the smallest $n-k$ singular values of A. In the total least squares context, the truncated total least squares (TTLS) algorithm solves for the minimum norm solution x of a rank deficient TLS problem $A_k x \approx b_k$, where $[A_k\ b_k]$ is the best rank $k+1$ approximation of $[A\ b]$.

In Algorithm 2, there is a necessary condition that the vector V_{22}^k is not all zero. V_{22}^k is zero in *nongeneric* cases (see [15]), but it can be nearly zero in close-to-nongeneric cases, which can happen in ill-posed problems. As noted in [4], we can always decrease k enough such that (the longer) V_{22}^k has large enough norm. On the other hand, if V_{22}^k is nearly zero (for some k), this will be an indication that we are using a too big truncation level k.

For later use, we remind the interlacing property for singular values [7], which garantees the following ordering of the singular values of A and $[A\ b]$:

$$\sigma_1 \geq \sigma_1' \geq \sigma_2 \geq \sigma_2' \geq \cdots \geq \sigma_n \geq \sigma_n' \geq \sigma_{n+1}. \tag{4}$$

2.2 Linear ill-posed problems

Before going further, we illustrate with an example of an ill-posed problem the fact that it is not advisable to use the classical estimation methods (least squares, total least squares) directly.

A matrix A_0 and a vector b_0 with real valued elements and dimensions $m = 200$, $n = 30$, are simulated using the example `ilaplace` from the Regularization Tools [9]. This example that originally comes from [16] constructs A_0 as a discretization of the inverse Laplace transform, using Gauss-Laguerre quadrature for the discretization of the inverse Laplace integral operator. In [16] and in Hansen's toolbox [9], the sampling points where the integral equation is evaluated were as many as the quadrature points, thus yielding a square discretized system. We slightly modify this setting in order to construct rectangular systems, by allowing more sampling points than quadrature points.

In this simulation, $A_0 x \approx b_0$ has a certain exact solution vector x_0, satisfying the equation $A_0 x_0 = b_0$. We construct an incompatible noisy example $Ax \approx b$, by adding Gaussian white noise with a standard deviation of 1e-5 to all elements of A_0 and b_0. Figure 1 shows the singular values of A_0 as they decay without noticeable gap towards machine precision. For comparison, Figure 1

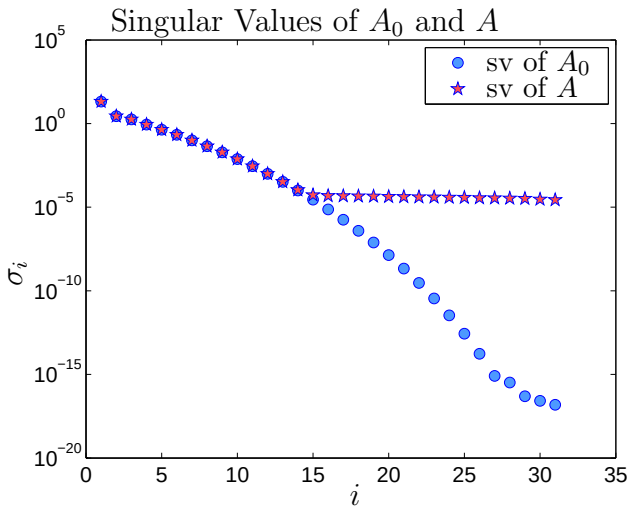


Fig. 1. Singular values of A_0 decaying without gap towards zero, where the data A_0 comes from a discrete ill-posed problem. Singular values of the noisy data A level off at the noise level.

also shows the decaying singular values of the noisy A and it is noticeable that the smallest singular values stagnate at a certain level, determined by the standard deviation of the added noise.

The conditioning of the matrix A (about 2e+5) is much better than that of A_0 (about 3e+17), but this is a misleading improvement. In fact, the added noise has a disastrous effect on the problem of estimating x using, *e.g.*, the LS or TLS method.

Figure 2 shows the true solution x_0 together with the LS and TLS solutions computed from the data $[A\ b]$. Note that both estimated solutions are very noisy and are not reconstructing the true solution well. The TLS solution is

noticeably noisier than the LS solution.

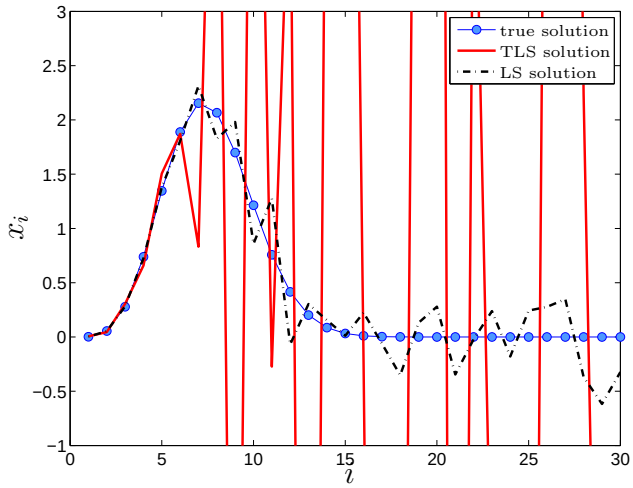


Fig. 2. The true solution x_0 is chosen to be the discretization of a smooth function. However, the LS and TLS solutions x computed from the noisy data $[A\ b]$ are influenced by the noise and are far from being reasonable approximations of the underlying true solution x_0.

Next, we aim to illustrate that truncation is a good regularization method for the type of problem described before. We used the SVD of the noisy A and, respectively, $[A\ b]$ and we computed truncated SVD and truncated TLS solutions ($x_{\mathrm{TSVD},k}$ and $x_{\mathrm{TTLS},k}$) for all possible values of k, from 1 to $n = 30$. We plot in Figure 5 several reconstructed solutions, together with the exact x_0. Some facts are to be noted:

- the "non-truncated" solutions $x_{\mathrm{TSVD},n}$ and $x_{\mathrm{TTLS},n}$ are identical to the LS, respectively, TLS solutions, and the truncated solutions with large k indices are also very much influenced by the noise and resemble the LS/TLS solutions;
- the truncated solutions with very small k level are too simplistic to capture the characteristics of the true solution x_0;
- there are several values of k, such as $k = 10$ in Figure 5, for which the truncated SVD and truncated TLS solutions are very good reconstructions of the true solution x_0.

The noiseless problem $A_0x = b_0$ is also inaccurately solved numerically, unless we use truncation. The range of good truncation levels is wider than in the noisy case, but the best reconstruction of x_0 is obtained for $k = 20$, for which singular values below 1e-10 are discarded from the SVD of A or, respectively, $[A\ b]$.

The aim of regularization by truncation is thus to appropriately identify a good truncation level, and to construct a truncated solution that can capture

the essential features of the unknown true solution, without explicit knowledge about the true solution, and even without *a priori* knowledge about the magnitude of the noise in the data.

3 Choice of the truncation level

3.1 Classical criteria for TSVD

For the truncated SVD formulation, and assuming that only b is noisy, many classical model selection methods exist. Some basic ideas are common to all methods; for instance, the goodness of fit is traded-off with respect to the parsimony of the model. We shortly enumerate several well-known criteria.

The discrepancy principle

This method requires knowledge about the statistical properties of the noise that perturbs the data [13]. Specifically, let's assume that the noise is i.i.d. with zero mean and variance σ^2. The norm of the residual error is computed for the candidate truncation level k and the following criterion is monitored:

$$\min k \text{ such that } \quad \|Ax_{\mathrm{TSVD},k} - b\|^2 \leq m\sigma^2, \tag{5}$$

thus, a truncation level is sought that provides a residual error of the same magnitude as the noise in the data.

The L-curve

The norm of the truncated solution $\|x_{\mathrm{TSVD},k}\|$ is plotted against the residual error norm $\|Ax_{\mathrm{TSVD},k} - b\|$ in log-log scale for various k's, and the *corner* in log-log scale is chosen [8]. A variant is the *rotated L-curve* [14], where the product $\|Ax_{\mathrm{TSVD},k} - b\|\|x_{\mathrm{TSVD},k}\|$ is minimized for k. In log-log scale, the plot of this function looks like a rotated L-curve, such that the L-curve's corner becomes a minimum.

Generalized cross validation

Generalized cross validation (GCV) [3, 5, 17] is a method derived from the leave-one-out cross validation criterion. It can be written as the following

minimization problem:

$$\min_k \frac{\|Ax_{\mathrm{TSVD},k} - b\|^2}{(m-k)^2}. \tag{6}$$

Akaike's information criterion

For TSVD, the classical Akaike's information criterion [1] becomes simply a logarithmic version of the GCV:

$$\min_k \ \log\left(\|Ax_{\mathrm{TSVD},k} - b\|\right) + \frac{k}{m}. \tag{7}$$

3.2 The TTLS case

In the TTLS case, the consideration that A can be noisy implies several changes in the elements required by the model selection criteria.

The residual error norm in this case is given by

$$\|r_k^{\mathrm{TLS}}\|_F^2 := \left\|[A\ b] - [\widehat{A}_k\ \widehat{b}_k]\right\|_F^2 \tag{8}$$

where $[\widehat{A}_k\ \widehat{b}_k]$ denotes the best rank-k approximation of $[A\ b]$. Thus, in the **L-curve** criterion, we plot $\|r_k^{\mathrm{TLS}}\|_F$ (instead of $\|Ax_{\mathrm{TTLS},k} - b\|$) against $\|x_{\mathrm{TTLS},k}\|$ in log-log scale.

The **discrepancy principle** could also be generalized such that it uses this residual error measure. If the noise in A and b is zero mean i.i.d. with variance σ^2, then the estimate of a 'denoised' rank-k approximation can be computed with the criterion:

$$\min k \text{ such that } \|r_k^{\mathrm{TLS}}\|_F^2 \leq m(n+1)\sigma^2. \tag{9}$$

The factor $m(n+1)$ takes care of the fact that there are $m(n+1)$ noisy elements in A and b. This simple generalization of the classical discrepancy principle does not share the statistical motivation of its counterpart in the LS case. In the LS sense, the LS residual preserves the (Gaussian) distribution of the assumed noise added to b^{exact}. In the TLS sense, the low rank residual matrix $[A\ b] - [\widehat{A}_k\ \widehat{b}_k]$ does not preserve the statistical properties of the added noise on the true A^{exact} and b^{exact} data that generates A and b.

In signal processing, there are many situations when the dimension of a *signal subspace* of a noisy data matrix is sought. (In this context, we think of $[A\ b]$ as a single data matrix C consisting of a true mean matrix C^{true} plus a noisy error matrix C^{noise}.) A method of rank determination that aims at separating

the signal subspace and the *noise subspace* is based on signal-to-noise ratio considerations. The signal-to-noise ratio of the best rank k approximation of $[A\ b]$ has the following definition [15, p. 89]:

$$\mathrm{SNR}_j := \frac{\sigma_j([A_0\ b_0])^2}{m\sigma^2}, \tag{10}$$

where $\sigma_j([A_0\ b_0])$ is the jth singular value of the exact but unobservable $m \times (n+1)$ matrix $[A_0\ b_0]$, and σ^2 is the variance of the i.i.d. zero mean additive noise in $[A\ b]$. Then, the optimal choice for the truncation level k^{opt}, such that the SVD components with signal-to-noise ratio smaller than 1 are discarded is [15, p. 89]: choose k^{opt} as the number of singular values σ_j that satisfy $\sigma_j > \sqrt{2m}\sigma$.

When the true noise variance is unknown, the squared smallest singular value, divided by the row dimension m, is sometimes used as an estimate. Thus we can often successfully apply the criterion: choose k as the number of singular values σ_j that satisfy $\sigma_j > \sqrt{2}\sigma_{\min}([A\ b])$. (For the optimality of this choice see also [2].) The disentanglement between the signal subspace and the noise subspace will be accurate when there is a clear distinction between signal and noise (which translates into a clear gap in the singular values of the matrix). The method can still be applied when the singular values decay without clear gap.

Other choices of the signal subspace dimension used in practice are versions of Akaike's information criterion and of the minimum description length principle [18]. We do not go into the details of these methods, since the interest of all these methods is a good reconstruction of the signal subspace (equivalent to a lower rank reduction), but our interest in the context of the discrete ill-posed $Ax \approx b$ problem is primarily the estimation of a good x solution, without enforcing strong statistical assumptions and without developing an asymptotic method.

3.3 Filter factors

The basis of our proposed truncation level choice comes from the concept of *filter factors* of a regularized model.

The TSVD solution has a simple interpretation in terms of filter factors. If we set

$$f'_1 = f'_2 = \ldots = f'_k = 1, \quad \text{and} \quad f'_{k+1} = \ldots = f'_n = 0, \tag{11}$$

then the TSVD solution with truncation level k is simply:

$$x_{\mathrm{TSVD},k} = \sum_{i=1}^{n} f_i' \frac{{u_i'}^\top b}{\sigma_i'} v_i', \tag{12}$$

where (we remind the notation from Algorithm 1), (u_i', σ_i', v_i') for $i = 1, \ldots, n$ are the singular triplets of the matrix A. Thus, using only filter factors equal to one and zero, the TSVD solution is a filtered version of the ordinary least squares solution.

In general, a regularized solution to $Ax \approx b$ can be written as

$$x^{\mathrm{reg}} = \sum_{i=1}^{n} f_i \frac{{u_i'}^\top b}{\sigma_i'} v_i', \tag{13}$$

with $f_i \in [0, 1]$. The filter factor formulation relates a regularized solution to the least squares solution of $Ax \approx b$. Moreover, the predicted right hand side that corresponds to x^{reg} is:

$$b^{\mathrm{reg}} := Ax^{\mathrm{reg}} = U'\Sigma'{V'}^\top x^{\mathrm{reg}} = U'\mathrm{diag}\{f_1, \ldots, f_n\}{U'}^\top b. \tag{14}$$

In [4], it is proved that the TTLS solution $x_{\mathrm{TTLS},k}$ can be expressed in the filter factor formulation, using the filter factor values:

$$f_i = \sum_{\substack{j=k+1 \\ \sigma_j \neq \sigma_i'}}^{n+1} \frac{v_{n+1,j}^2}{\|V_{22}^k\|^2} \left(\frac{{\sigma_i'}^2}{{\sigma_i'}^2 - {\sigma_j}^2} \right), \qquad i = 1, \ldots, n. \tag{15}$$

Here, the symbols without the prime sign correspond to the SVD of $[A\, b]$ given in Algorithm 2. Note that using these filter factors we are able to express the TTLS solution with truncation level k in terms of the SVD of the matrix A.

A theorem in [4] provides the following bounds on the TTLS filter factors:
Lemma 1 (Th. 4.2, [4])**.** *Assume $\sigma_k' \neq \sigma_{k+1}$. The first k filter factors f_i form a monotonically increasing sequence and satisfy*

$$1 \leq f_i \leq 1 + \frac{\sigma_{k+1}^2}{{\sigma_i'}^2 - \sigma_{k+1}^2}, \qquad i = 1, \ldots, k \tag{16}$$

while the last $n - k$ filter factors satisfy

$$0 \leq f_i \leq \|V_{22}\|^{-2} \frac{{\sigma_i'}^2}{\sigma_k^2 - {\sigma_i'}^2}, \qquad i = k+1, \ldots, n. \tag{17}$$

These results show that $f_1, \ldots, f_k$ are larger but close to 1, while $f_{k+1}, \ldots, f_n$ are small, provided that $\|V_{22}^k\|$ is not close to zero and that there are big

enough gaps in the ordering $\sigma'_i > \sigma_{k+1}$ (for $i = 1, \ldots, k$) and $\sigma_k > \sigma'_i$ (for $i = k+1, \ldots, n$). In particular, f_k is close enough to one and f_{k+1} drops to a small value if there is a large gap in the middle of the following interlacing relationship:

$$\sigma_k \geq \sigma'_k \gg \sigma_{k+1} \geq \sigma'_{k+1}. \tag{18}$$

In this situation, the TTLS solution resembles the TSVD solution with the same truncation level k.

However, whenever some of the filter factors for TTLS are greater than one, the TTLS solution is partly a "de-regularized" solution. We provide the following lemma containing opposite inequality relations, which show how large the filter factors may get if the above mentioned conditions on the magnitude of $\|V_{22}^k\|$ and the separation of certain singular values are not met.

Lemma 2. *Assume $\sigma'_k \neq \sigma_{k+1}$. Then the first k filter factors f_i satisfy*

$$f_i \geq 1 + \frac{\sigma_{n+1}^2}{\sigma_i'^2 - \sigma_{n+1}^2}, \qquad i = 1, \ldots, k, \tag{19}$$

and the last $n-k$ filter factors satisfy

$$f_i \geq \frac{\|V_{21}^k\|^2}{\|V_{22}^k\|^2} \left(\frac{\sigma_i'^2}{\sigma_1^2 - \sigma_i'^2} \right), \qquad i = k+1, \ldots, n. \tag{20}$$

Proof. In relation (15), we use the ordering of the singular values $\sigma_j \geq \sigma_{n+1}$, for $j \leq n$, together with the fact that $\sum_{k+1}^{n+1} v_{n+1,j}^2 = \|V_{22}^k\|^2$, and the knowledge that $v_{n+1,j} = 0$ whenever there is an i such that $\sigma_j = \sigma'_i$ [15, Th. 3.14]. Then

$$\begin{aligned} f_i &= \sum_{j=k+1}^{n+1} \frac{v_{n+1,j}^2}{\|V_{22}^k\|^2} \left(\frac{\sigma_i'^2}{\sigma_i'^2 - \sigma_j^2} \right) \geq \sum_{j=k+1}^{n+1} \frac{v_{n+1,j}^2}{\|V_{22}^k\|^2} \left(\frac{\sigma_i'^2}{\sigma_i'^2 - \sigma_{n+1}^2} \right) = \frac{\sigma_i'^2}{\sigma_i'^2 - \sigma_{n+1}^2} \\ &= 1 + \frac{\sigma_{n+1}^2}{\sigma_i'^2 - \sigma_{n+1}^2}, \end{aligned}$$

proving (19). For the second inequality we use another characterization of the filter factors, given also in [4], namely,

$$f_i = \|V_{22}^k\|^{-2} \sum_{j=1}^{k} v_{n+1,j}^2 \left(\frac{\sigma_i'^2}{\sigma_j^2 - \sigma_i'^2} \right).$$

As opposed to (15), this formulation makes it clear even for $i > k$ that, due to interlacing, all terms are positive in the sum, since the indices j and i are smaller, respectively, larger than k. Then it is easy to majorate σ_j by σ_1 in all denominators in the sum and obtain (20). ∎

As a consequence of (19), we see that if the difference between $\sigma_i'^2$ and σ_{n+1}^2 is very small compared to the value of σ_{n+1}^2, then f_i can become much larger than 1. This situation can happen in ill-posed problems if we let the truncation level k be too large. In summary, the scenario when $f_i \gg 1$ (for an $i \leq k$) involves:

- $Ax \approx b$ is close-to-nongeneric, with the smallest singular value of A being almost equal to the smallest singular value of $[A\ b]$;
- the smallest singular value of A (and that of $[A\ b]$) is close to being multiple, with multiplicity r larger than $n - k$;
- i (with $i \leq k$) is also large enough such that it falls into the set of indices corresponding to the pseudo-multiple singular values, *i.e.*, $n - r \leq i \leq k$.

Note that when there is additive i.i.d. random noise with variance σ^2 present in A and b, the smallest singular value of A (and $[A\ b]$) becomes pseudo-multiple, at a value determined by the variance of the noise. (See Figure 1.) Thus the items above will hold true.

The implications of the second lower bound, (20), cover the case when we may run into problems even if there are no pseudo-multiple smallest singular values. If k is chosen such that V_{22}^k is almost zero, the f_{k+1} filter factor can become significantly larger than zero, depending on the relative difference in magnitude orders of the singular value ratio involved, and the last row of V involved. Although we cannot strictly conclude that, *e.g.*, f_{k+1} will be large, we can conjecture that for a too big k value in a close-to-nongeneric problem, the behavior of the filter factors $f_{k+1}, f_{k+2}, \ldots$ is *unpredictable*.

Example

Continuing with the noisy ill-posed example of §2.2, Figure 6 plots the filter factors computed for a range of truncation levels k. Notice that for small truncation levels k ($k \leq 12$) the filter factors follow very well the characterization given by [4] (*i.e.*, the filter factors with indices $i = 1, \ldots, k$ stay close to 1 and the filter factors with indices $i = k+1, \ldots, n$ stay close to zero). However, when we gradually increase the truncation level k, there is a sudden point where this characterization no longer holds. From then on (*e.g.*, $k \geq 16$), the filter factors with indices i near the truncation level k increase dramatically, spoiling their regularizing property and illustrating that the ill-posedness comes into play.

3.4 Effective number of parameters, modified generalized cross validation and generalized information criteria

Although there might be other ways to make use of the properties of the TTLS filter factors, we choose to incorporate the information that they provide into

well-known methods. The idea is to identify the value of k where the filter factors change their nice behavior into the erratic growth behavior, since in this way we are likely to find a good truncation level for TTLS (and TSVD as well).

The sum of the filter factors in the regularized solution expression (13) defines the so-called *effective number of parameters* in the model; this concept typically appears in model selection methods such as the generalized cross validation [3] and the generalized Akaike information criterion [12].

There is a simple derivation of the definition of the effective number of parameters as the sum of the filter factors; it involves yet another concept – the *hat matrix*, which is a transformation matrix that relates the given right-hand side b to the corrected right-hand side of the regularized model, b^{reg} in (14). In regularization methods where the reconstructed b^{reg} can be expressed as $b^{\text{reg}} = H^{\text{reg}}b$, for some square matrix H^{reg}, there is a natural way to use this relation in the selection of regularization hyperparameters. On one hand, the fitting error between b^{reg} and b is written in closed form as $\|(I_m - H^{\text{reg}})b\|$, and on the other hand the trace (or sometimes the determinant) of the coefficient matrix $I_m - H^{\text{reg}}$ is a measure of model simplicity, which can be traded-off with respect to the goodness of fit. The number $\text{Tr}(I_m - H^{\text{reg}})$ plays the role of the *number of degrees of freedom* of a regularized model, and $\text{Tr}(H^{\text{reg}})$ is in this context the effective number of parameters.

In general regularization models, the effective number of parameters might differ from the true numbers of variables involved in the model. In a linear regression model, however, the number of effective parameters is exactly equal to the number of *free* model parameters; in TSVD, this number is k, the truncation level. Indeed, the TSVD solution $x_{\text{TSVD},k}$, though of length n, corresponds to a corrected right-hand side $\widehat{b}_k := Ax_{\text{TSVD},k}$, which lies in a k-dimensional linear subspace, and it is an orthogonal projection of b: $\widehat{b}_k = U'_k {U'_k}^\top b$. (Here, U'_k denotes the matrix of the first k left singular vectors of A.) The hat matrix in this case is simply $H_{\text{TSVD},k} := U'_k {U'_k}^\top$, whose trace is equal to k. Thus, the number of effective parameters that describe the TSVD model is k.

In the more general filter factor model described by (13), the hat matrix has also a very simple expression:

$$H^{\text{reg}} = U' \text{diag}\{f_1, \ldots, f_n\} {U'}^\top, \tag{21}$$

yielding the effective number of parameters to be equal to

$$\begin{aligned} \text{Tr}(H^{\text{reg}}) &= \text{Tr}(U' \text{diag}\{f_1, \ldots, f_n\} {U'}^\top) = \text{Tr}(\text{diag}\{f_1, \ldots, f_n\} {U'}^\top U') \\ &= \text{Tr}(\text{diag}\{f_1, \ldots, f_n\} I_n) = \sum_{i=1}^{n} f_i. \end{aligned} \tag{22}$$

We denote the effective number of parameters for the TTLS case, computed as the sum of the TTLS filter factors 15, by p_k^{eff}; this value can be plugged-in into the following (equivalent to each other!) modified criteria:

- generalized cross validation:

$$\min_k \frac{\|Ax_{\text{TTLS},k} - b\|^2}{\left(m - p_k^{\text{eff}}\right)^2}, \tag{23}$$

- generalized information criterion (see also [12])

$$\min_k \ \log(\|Ax_{\text{TTLS},k} - b\|) + \frac{p_k^{\text{eff}}}{m}. \tag{24}$$

Note that the parameter p_k^{eff} can also be computed by using an alternative formula for the filter factors. When an $x_{\text{TTLS},k}$ is already computed, its corresponding filter factors are just scaled linear coefficients in the linear expansion of $x_{\text{TTLS},k}$ in the basis given by the right singular vectors of A. It is easy to see that from

$$x_{\text{TTLS},k} = \sum_{i=1}^{n} f_i \frac{u_i'^{\top} b}{\sigma_i'} v_i' \tag{25}$$

we get

$$f_i = \sigma_i' \frac{v_i'^{\top} x_{\text{TTLS},k}}{u_i'^{\top} b}. \tag{26}$$

4 Examples

We use simulated ill-posed problems such as the problem described in §2.2, from the Regularization Tools [9]. However, not all ill-posed problems provided in the Regularization Tools can be solved satisfactorily using truncation methods. We explain this by the fact that not all types of ill-posedness are removable by simple truncation, and more complicated regularization methods should be used in those cases. The studied truncation methods favor solutions of small norms, which might not always be the most appropriate property.

The results in this section test the performance of one of the truncation level selection methods, namely the generalized cross validation criterion proposed for TTLS in §3.4, (23). For the ill-posed example that we used from the beginning (see §2.2), the proposed GCV function (23) for choosing the truncation level is compared with the classical GCV criterion (6) that applies to the TSVD method. In the left plot of Figure 3, we notice that the modified GCV has a more clear minimum around the $k = 12$ value, which is actually the intuitively optimal value for this noisy problem – the k value that minimizes

$\|x_{\mathrm{TTLS},k} - x_0\|$. On the other hand, the classical GCV has a flat minimum; this would imply that choosing a larger k (*e.g.*, $k = 20$) is almost as good. But by checking Figure 5, it is obvious that $k = 20$ is a very bad choice.

The left plot of Figure 3 illustrates the comparison (in log scale) between the effective numbers of parameters that are computed in the two scenarios. For the classical GCV (6), the effective number of parameters for truncation level k is equal to k itself.

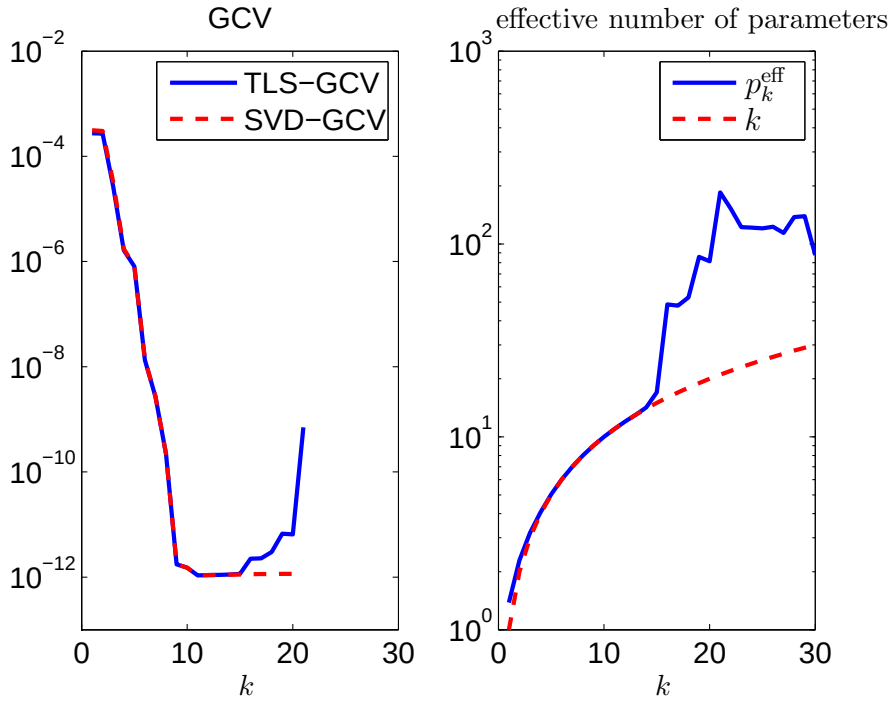


Fig. 3. The left plot shows the GCV function in the TTLS and the classical TSVD cases. The right plot compares the effective number of parameters in the TTLS case and in the TSVD case. Note that p_k^{eff} is approximately equal to k for a certain range of k values.

We have performed an extensive study to compare the GCV criterion proposed for TTLS and the classical GCV applied to the TSVD method. We used 8 problems from the Regularization Tools, namely:

1. `ilaplace(n,1)` 2. `ilaplace(n,3)` 3. `baart` 4. `shaw`
5. `phillips` 6. `foxgood` 7. `deriv2` 8. `wing`

We considered the case when both A and b contain additive noise or when only b contains additive noise; the relative noise level was 0.001. For both situations, 100 noisy realizations were considered for each of the 8 example problems. Then, each problem was solved using either TTLS or TSVD, and the truncation level was chosen with the corresponding GCV criterion. Moreover, we define for each simulation example the *best* truncation level as being the

truncation level that provides the smallest error between the truncated and the true solution. Thus, for TSVD we obtain this best truncation level as the k that minimizes $\|x_{\mathrm{TSVD},k} - x_0\|$, while for TTLS we choose it by minimizing for k the value of $\|x_{\mathrm{TTLS},k} - x_0\|$. In principle, the two "best" truncation levels in the TSVD and TTLS case can be different, but the simulations proved that they coincided in more than 90% of the simulations. The truncation levels chosen with the GCV criteria (in the TSVD and TTLS cases) are compared with the best truncation levels in Figure 4. The GCV-chosen truncation level coincides with the best truncation level, or it is at a distance less than 3 to it, in the following percentages:

	TTLS	TSVD	TTLS	TSVD
noisy A and b	36%	26%	79%	65%
noise only on b	37%	32%	75%	70%
	coinciding		≤ 2 away	

Figure 4 illustrates this information by plotting the ratios between the truncation levels computed by GCV and the best truncation levels. The case with most outliers is when there is noise on A and b and we use truncated SVD.

5 Conclusion

We addressed approximate linear modeling problems in the case when the data is coming from ill-posed problems, and we used truncation methods as a form of regularization.

We discussed model selection techniques for choosing truncation levels, and we emphasized that by considering the truncated total least squares problem we can exploit the following observation: when the truncation level is overestimated, the noise or the rounding errors take over the computed solution in a much higher degree than for the classical truncated SVD solution. The filter factors associated with the truncated total least squares solution are not necessarily equal to one and zero (as in the truncated SVD case). For this reason, they provide more information that can be used in criteria for choosing the truncation level; we illustrated the merits of a modified generalized cross validation method.

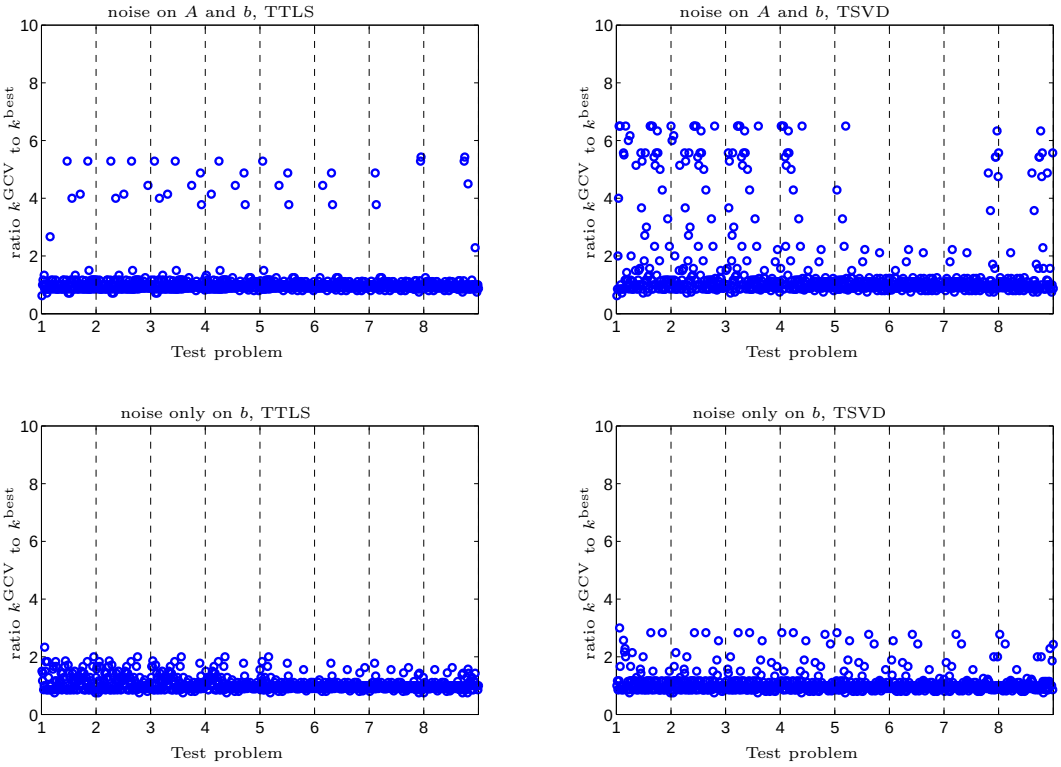


Fig. 4. Comparison between the two generalized cross validation (GCV) criteria, corresponding to the TTLS and the TSVD methods. For each test problem (from 1 to 8) we use 100 random noise realizations. Each plot shows the ratio between the value of the truncation level k^{GCV} chosen by GCV and the best k^{best} value minimizing the distance of the truncated solution to the true solution. Plotted values equal to 1 are thus optimal. The four plots correspond to: (top) noisy A and b, (bottom) noiseless A, noisy b; (left) $Ax \approx b$ solved with truncated total least squares; (right)$Ax \approx b$ solved with truncated singular value decomposition.

Acknowledgments

Dr. Diana M. Sima is a research assistant and Dr. Sabine Van Huffel is a full professor at the Katholieke Universiteit Leuven, Belgium.
This work was supported by: **Research Council KUL:** GOA-AMBioRICS, CoE EF/05/006 Optimization in Engineering, IDO 05/010 EEG-fMRI, several PhD/postdoc & fellow grants; **Flemish Government: FWO:** PhD/postdoc grants, projects, G.0407.02 (support vector machines), G.0360.05 (EEG, Epileptic), G.0519.06 (Non-invasive brain oxygenation), FWO-G.0321.06 (Tensors/Spectral Analysis), G.0341.07 (Data fusion), research communities (ICCoS, ANMMM); **IWT:** PhD Grants; **Belgian Federal Science Policy Office:** IUAP P5/22 ('Dynamical Systems and Control: Computation, Identification and Modelling'); **EU:** BIOPATTERN (FP6-2002-IST 508803), ETUMOUR (FP6-2002-LIFESCIHEALTH 503094), Healthagents (IST-

2004-27214), FAST (FP6-MC-RTN-035801); **ESA:** Cardiovascular Control (Prodex-8 C90242)

References

[1] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

[2] R. Beran. Adaptive estimation of a mean matrix: Total least squares versus total shrinkage. Available from: http://www.stat.ucdavis.edu/~beran/tlsfin.pdf, 2006.

[3] P. Craven and G. Wahba. Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, 31:377–403, 1979.

[4] R. D. Fierro, G. H. Golub, P. C. Hansen, and D. P. O'Leary. Regularization by truncated total least squares. *SIAM Journal on Scientific Computing*, 18(1):1223–1241, 1997.

[5] G. H. Golub, M. Heath, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21:215–223, 1979.

[6] G. H. Golub and C. F. Van Loan. An analysis of the total least squares problem. *SIAM Journal on Numerical Analysis*, 17:883–893, 1980.

[7] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, Maryland, third edition, 1996.

[8] P. C. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 32:561–580, 1992.

[9] P. C. Hansen. Regularization Tools, a Matlab package for analysis of discrete regularization problems. *Numerical Algorithms*, 6:1–35, 1994.

[10] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems. Numerical Aspects of Linear Inversion*. SIAM, Philadelphia, 1998.

[11] P. C. Hansen, M. Kilmer, and R. H. Kjeldsen. Exploiting residual information in the parameter choice for discrete ill-posed problems. *BIT*, 46:41–59, 2006.

[12] S. Konishi and G. Kitagawa. Generalised information criteria in model selection. *Biometrika*, 83(4):875–890, 1996.

[13] V. A. Morozov. On the solution of functional equations by the method of regularization. *Soviet. Math. Dokl.*, 7:414–417, 1966.

[14] T. Reginska. A regularization parameter in discrete ill-posed problems. *SIAM J. Sci. Comput.*, 17(3):740–749, 1996.

[15] S. Van Huffel and J. Vandewalle. *The Total Least Squares Problem: Computational Aspects and Analysis*, volume 9 of *Frontiers in Applied Mathematics*. SIAM, Philadelphia, 1991.

[16] J. M. Varah. Pitfalls in the numerical solution of linear ill-posed problems. *SIAM Journal on Scientific and Statistical Computing*, 4(2):164–

176, 1983.

[17] G. Wahba. *Spline Models for Observational Data*. CBMS-NSF Regional Conference Series in Applied Mathematics 59. SIAM, 1990.

[18] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Trans. Accoust. Speech Signal Processing*, ASSP-33(2):387–392, 1985.

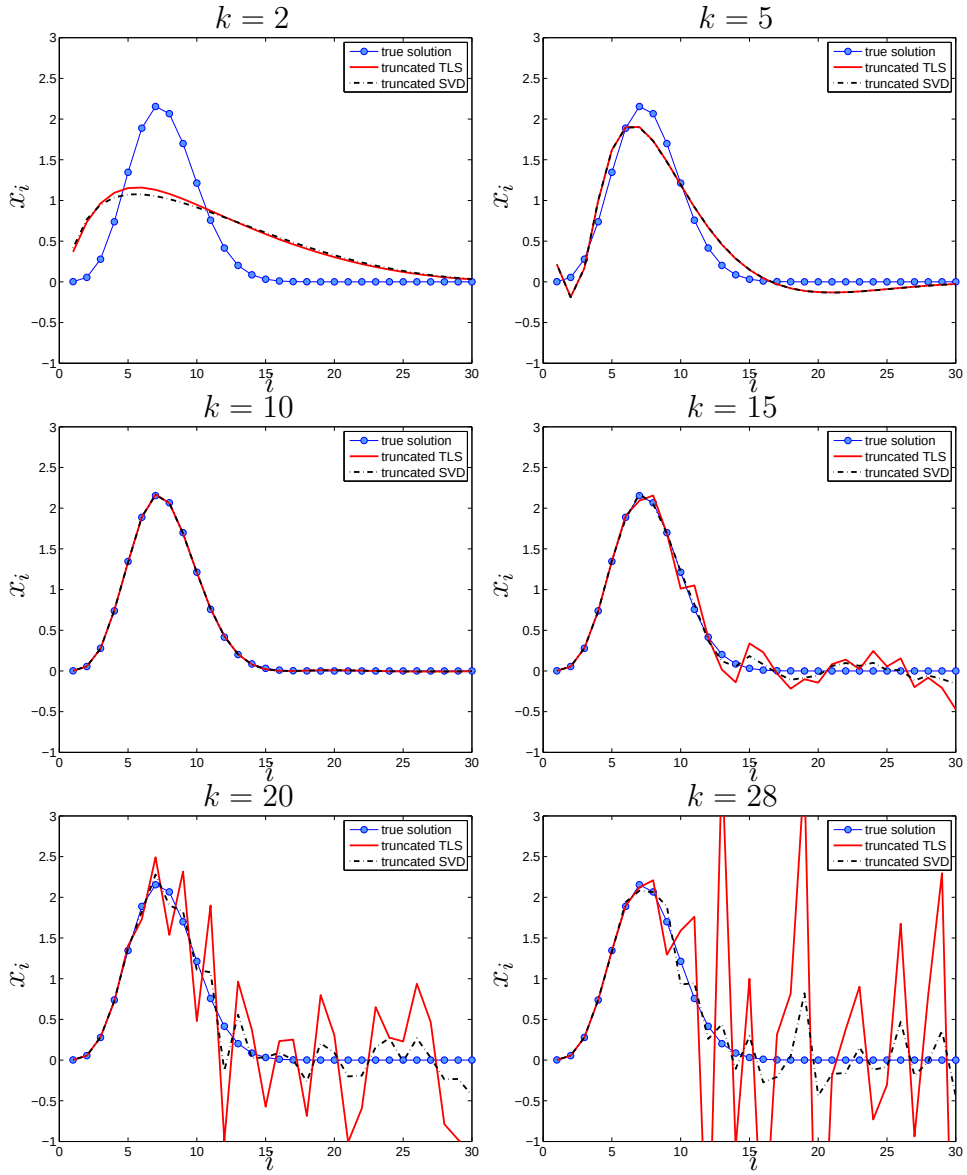


Fig. 5. Several truncated TLS and SVD solutions x_k, together with the true solution x_0, for the noisy simulation data described in §2.2. For both TTLS and TSVD solutions, the best truncation level is $k = 12$, minimizing $\|x_k - x_0\|/\|x_0\|$ to the value of 1e-3.

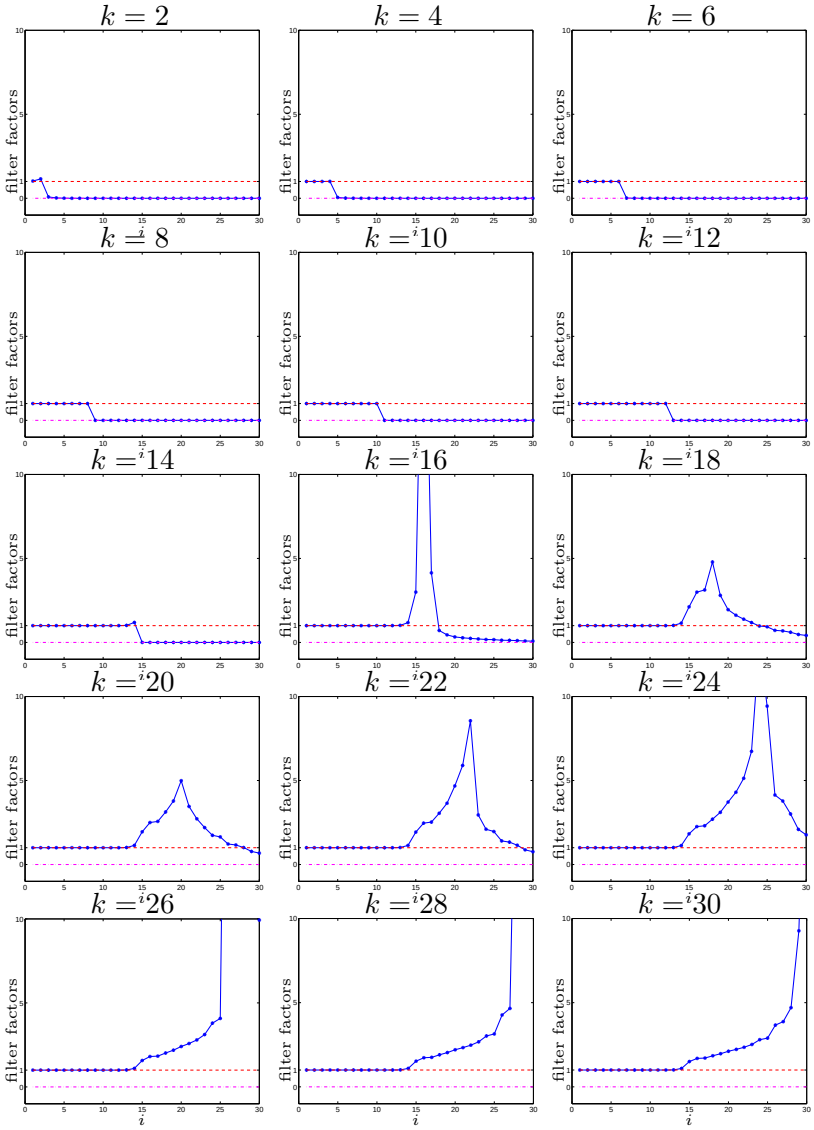


Fig. 6. The filter factors associated with the TTLS solutions for a representative range of k values between 1 and $n = 30$. The interrupted horizontal lines mark the zero and one levels. The numerical example is the same noisy example as in §2.2.