

Research Article

Use of Tikhonov Regularization to Improve the Accuracy of Position Estimates in Inertial Navigation

Tuukka Nieminen, Jari Kangas, and Lauri Kettunen

Institute of Electromagnetics, Tampere University of Technology, P.O. Box 692, 33101 Tampere, Finland

Correspondence should be addressed to Tuukka Nieminen, tuukka.nieminen@tut.fi

Received 26 November 2010; Accepted 12 February 2011

Academic Editor: Abbas Mohammed

Copyright © 2011 Tuukka Nieminen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Inertial navigation problems are often understood as *initial value problems*. However, there are many applications where *boundary value problems* naturally arise. In these situations, it has been shown that the *finite element method* can be efficiently used to compute accurate position and velocity estimates. We will propose that finite element method complemented with *Tikhonov regularization*—a basic tool for inverse problems—is a powerful combination for further accuracy improvements. The proposed method provides a straightforward way to exploit prior information of various types and is subject to rigorous optimality results. Use and accuracy of the proposed method are demonstrated with examples.

1. Introduction

The term *inertial navigation* is often associated with the term *initial value problem* (IVP). This kind of an association is justified in many cases, but not always: there are situations, where an inertial navigation problem may just as naturally be posed as a *boundary value problem* (BVP) [1]. In case of a BVP, all measurements within the interesting time interval are exploited, whereas an IVP only exploits measurements up to a certain time within the time interval.

We will discuss BVPs concerning the position of the object, given the accelerations in the navigation frame. Instead of the “full equations”, where rotation of Earth is taken into account, we consider “simplified equations”, where this is neglected [1]. In this case, to obtain the position (\mathbf{p}) of the object, we need to solve the following problem: given $\mathbf{a}(t) : \mathbb{R} \rightarrow \mathbb{R}^3$, find $\mathbf{p}(t) : \mathbb{R} \rightarrow \mathbb{R}^3$ such that \mathbf{r} satisfies

$$\frac{d^2 \mathbf{p}(t)}{dt^2} = \mathbf{a}(t). \quad (1)$$

The problem of resolving accelerations $\mathbf{a}(t)$ given the specific force and angular rate measurements is a different matter, not considered herein. What makes (1) a BVP is that the necessary constraints are not given at same time [2]. More specifically, we will consider two-point BVPs where the

necessary constraints are given at the beginning and at the end of the interesting time interval [3].

In [1], a custom-made FEM model is derived to treat two-point BVPs of the form (1) with various choices of boundary conditions. The resulting system of linear equations is shown in [1] to be of the form

$$\begin{bmatrix} \mathbf{A}_x & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_y & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_z \end{bmatrix} \begin{bmatrix} \mathbf{r}_x \\ \mathbf{r}_y \\ \mathbf{r}_z \end{bmatrix} = \begin{bmatrix} \mathbf{b}_x \\ \mathbf{b}_y \\ \mathbf{b}_z \end{bmatrix}. \quad (2)$$

In (2), $\mathbf{r}_x, \mathbf{r}_y, \mathbf{r}_z \in \mathbb{R}^N$, and \mathbf{r}_x , for example, is a vector containing the x -coordinate of the object at N different time instances. Vectors $\mathbf{b}_x, \mathbf{b}_y$, and \mathbf{b}_z are determined by $\mathbf{a}(t)$. Submatrices $\mathbf{A}_x, \mathbf{A}_y$, and \mathbf{A}_z are symmetric, positive definite and with the basis functions considered in [1], tridiagonal $N \times N$ matrices. In the following, a shorthand notation for (2) is used. It is written in the form

$$\mathbf{A}\mathbf{r} = \mathbf{b} + \boldsymbol{\epsilon}, \quad (3)$$

where vector $\boldsymbol{\epsilon} \in \mathbb{R}^{3N}$ is added to emphasize the presence of *measurement errors*, generally of unknown properties. Based on the above discussion, the matrix $\mathbf{A} \in \mathbb{R}^{3N \times 3N}$ is known to be nonsingular. For a 2D problem, the dimensions of

system (3) are $2N$ and, respectively, for 1D, N . Unless otherwise noted, the following discussion assumes dimensions to be $3N$.

The goal of this paper is to estimate \mathbf{r} as accurately as possible, given \mathbf{A} , \mathbf{b} , and *additional information* of some form. Sometimes, also some properties of $\boldsymbol{\epsilon}$ can be known but in general, we do not make such an assumption. We will be using the term “additional information” throughout this paper, so let us now clarify this concept: “additional information” refers to all available information that can be represented in the linear form

$$\mathbf{D}\mathbf{r} = \mathbf{e} + \boldsymbol{\theta}, \quad (4)$$

where $\mathbf{D} \in \mathbb{R}^{M \times 3N}$ and $\mathbf{e} \in \mathbb{R}^M$. In (4), the term $\boldsymbol{\theta}$ denotes uncertainty, just like $\boldsymbol{\epsilon}$ in (3). The integer M is independent of N and indicates how many individual equations the additional information contains. Notice that properties of \mathbf{D} are generally unknown.

In [1], problems of the form (3) are discussed. There, it is assumed that $\boldsymbol{\epsilon}$ contains mainly deterministic error components, caused by deficient sensors. Then, one can exploit provided additional information of type (4) by modeling errors of these deficient sensors, which results in smaller position and velocity errors via reduction of $\|\boldsymbol{\epsilon}\|_2$ (where index 2 refers to the 2-norm). While this approach was shown to work well in many cases, it is adequate only for situations where the types of the most significant sensor errors are known. In practice, the approach proposed in [1] also limits the attainable accuracy in situations where M is significantly larger than the number of parameters in the chosen error model. The motivation of this paper is to present a new technique to exploit additional information, without the limitations of the technique proposed in [1].

The technique proposed herein is based on *Tikhonov regularization* [4–6], independently developed also by Phillips [7, 8]. As the name suggests, the basic idea of the method is to *regularize* the solution \mathbf{r} of (3). This is required in order to reduce the effects of the error term $\boldsymbol{\epsilon}$. Tikhonov regularization has been studied extensively during the last few decades, mainly in the field of *inverse problems* [8]. Thus, the properties of the method are well known. In the statistical literature, Tikhonov regularization is known as *ridge regression* [8–10].

We apply a tool of inverse problem theory to inertial navigation, because *inertial navigation problems are, in fact, inverse problems*. That is, we are basically resolving the time-parametrized trajectory of an object based on erroneous observations of specific force and angular rate. This reasoning is supported also by the fact that inertial navigation problems are particularly sensitive to measurement errors, which is the hallmark of inverse problems.

This paper is organized as follows: first, we will discuss the background of this study in Section 1.1 and make some remarks about this study in Section 1.2. In Section 2, Tikhonov regularization is presented. Some practical aspects of the proposed method are discussed in Section 3. Sections 4 and 5 are devoted to numerical examples and the conclusions are presented in Section 6.

1.1. Background. Problems resembling the ones considered here have previously been resolved using the means of *fixed interval smoothing* [11–14]. The most frequently used methods of solving these problems are *two-filter smoother* [15] and *Rauch-Tung-Striebel smoother* [16]. These are exploited, for example, in [14, 17–19]. In both cases, differential equation (1) is realized as an IVP. The basic idea of these methods is to run the filter in the forward direction as a “predictor” in phase one and then to run the filter in the backward direction as a “corrector” in phase two. In this approach, each additional constraint is assumed to be associated to a single time instance.

The proposed BVP formulation does not make a distinction between forward and backward directions. Instead of advancing one step at a time, the BVP formulation considers the whole time period, and the solution is obtained in one phase. In addition to the different realization of (1), the key difference is the form of (4), which allows additional information to involve an arbitrary number of time instances. This makes it possible and straightforward to exploit wide-ranging types of additional information. Another particularly useful feature of the proposed method is its capability of finding a reasonable balance between (3) and (4) without knowledge of $\|\boldsymbol{\epsilon}\|_2$ and $\|\boldsymbol{\theta}\|_2$. This property is discussed in Section 2.3.

The similarity between fixed interval smoothing and the BVP approach is that they can exploit information with varying reliability and have similar optimality results. Some optimality results of the proposed method are presented in Section 2.2. Although there are situations, where both methods could be exploited, we will not compare these methods quantitatively.

1.2. Basic Assumptions and Notation. We assume that some additional information is always available, that is, $M > 0$. If this is not the case, the presented method will reduce back to the one considered already in [1]. We will insist on additional information given in the form (4). A more general setting can be found, for example, in [20].

From time to time, we will need to print the entries of matrices and vectors. They are denoted $A_{i,j}$ (for matrix \mathbf{A}) and b_j (for vector \mathbf{b}), where i and j are positive integers within the allowed range. Symbols i and j are reserved only for this use. Similarly, notation t_i refers to the i th time instance, where the time starts at t_1 and ends at t_N . Step size is defined as $h_i = t_{i+1} - t_i$, where $1 \leq i \leq N - 1$. Without losing generality, for simplicity we treat the time step as if it was a constant. Finally, a normally distributed random variable x with expectation (or mean) value $E(x)$ and variance σ^2 is denoted as $x \sim \mathcal{N}(E(x), \sigma^2)$.

2. An Introduction to Tikhonov Regularization

Tikhonov regularized solution of (3) and (4) is the solution of

$$\mathbf{r}_\lambda = \underset{\mathbf{r}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{r} - \mathbf{b}\|_2^2 + \lambda \|\mathbf{D}\mathbf{r} - \mathbf{e}\|_2^2, \quad (5)$$

where $\lambda > 0$ is called the *regularization parameter*. It is used to weight (3) with respect to (4). Value $\lambda = 1$, for example,

indicates that both equations are weighted equally. With $\lambda \rightarrow 0$, $\|\mathbf{A}\mathbf{r}_\lambda - \mathbf{b}\| \rightarrow 0$ and the regularized solution \mathbf{r}_λ approaches the solution of (3) (with $\boldsymbol{\epsilon} = \mathbf{0}$). Respectively, with $\lambda \rightarrow \infty$, the regularized solution \mathbf{r}_λ approaches a solution satisfying (4).

Let us point out that in the literature, Tikhonov regularization is seldom presented in the form (5). Most often, it is assumed that the dimensions of \mathbf{A} and \mathbf{D} are equal. Moreover, in many formulations, \mathbf{D} is set to be a difference operator or the identity matrix (\mathbf{I}) and/or $\mathbf{e} = \mathbf{0}$ [8, 21, 22].

2.1. Existence and Uniqueness of the Solution. The questions regarding problem (5) are

- (1) is there a solution to (5)?
- (2) is the solution unique?
- (3) how can one find the solution?

Answer to question one follows easily from the fact that the cost function of (5) is nonnegative for every $\lambda > 0$, indicating that the minimization problem has at least one solution. To answer questions two and three, let us denote the value of the cost function by

$$Q_\lambda(\mathbf{r}) = \|\mathbf{A}\mathbf{r} - \mathbf{b}\|_2^2 + \lambda \|\mathbf{D}\mathbf{r} - \mathbf{e}\|_2^2. \quad (6)$$

As known, function $Q_\lambda(\mathbf{r})$ has a local extremum in a critical point of (6), in other words, at a point where the Gâteaux derivative

$$\left. \frac{d}{du} Q_\lambda(\mathbf{r} + u\mathbf{w}) \right|_{u=0} = 0 \quad \forall \mathbf{w} \neq \mathbf{0}, \quad (7)$$

where $u \in \mathbb{R}$ and $\mathbf{w} \in \mathbb{R}^{3N}$. To find critical point(s) of (6), let us substitute (6) into (7). For the first term of (6), we get

$$\begin{aligned} & \left. \frac{d}{du} \|\mathbf{A}(\mathbf{r} + u\mathbf{w}) - \mathbf{b}\|_2^2 \right|_{u=0} \\ &= \left. \frac{d}{du} \langle \mathbf{A}\mathbf{r} + u\mathbf{A}\mathbf{w} - \mathbf{b}, \mathbf{A}\mathbf{r} + u\mathbf{A}\mathbf{w} - \mathbf{b} \rangle \right|_{u=0} \\ &= 2\langle \mathbf{A}\mathbf{r} - \mathbf{b}, \mathbf{A}\mathbf{w} \rangle, \end{aligned} \quad (8)$$

where $\langle \cdot, \cdot \rangle$ denotes an inner product. Similarly for the latter term of (6),

$$\left. \frac{d}{du} \|\mathbf{D}(\mathbf{r} + u\mathbf{w}) - \mathbf{e}\|_2^2 \right|_{u=0} = 2\langle \mathbf{D}\mathbf{r} - \mathbf{e}, \mathbf{D}\mathbf{w} \rangle. \quad (9)$$

By substituting these into (7) and dividing by 2, we have

$$\langle \mathbf{A}\mathbf{r} - \mathbf{b}, \mathbf{A}\mathbf{w} \rangle + \lambda \langle \mathbf{D}\mathbf{r} - \mathbf{e}, \mathbf{D}\mathbf{w} \rangle = 0. \quad (10)$$

By exploiting the properties of the inner product, it follows that

$$\langle (\mathbf{A}^T\mathbf{A} + \lambda\mathbf{D}^T\mathbf{D})\mathbf{r} - (\mathbf{A}^T\mathbf{b} + \lambda\mathbf{D}^T\mathbf{e}), \mathbf{w} \rangle = 0. \quad (11)$$

Since this has to hold for all $\mathbf{w} \neq \mathbf{0}$, we finally get

$$\mathbf{r}_\lambda = (\mathbf{A}^T\mathbf{A} + \lambda\mathbf{D}^T\mathbf{D})^{-1} (\mathbf{A}^T\mathbf{b} + \lambda\mathbf{D}^T\mathbf{e}). \quad (12)$$

Critical point of (6) is unique as long as the matrix $(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{D}^T\mathbf{D})$ is nonsingular. This is true if and only if the null spaces (i.e., kernels) of matrices \mathbf{A} and \mathbf{D} intersect trivially: $\ker(\mathbf{A}) \cap \ker(\mathbf{D}) = \{\mathbf{0}\}$ [8]. Given that \mathbf{A} is nonsingular, this is the case independently of \mathbf{D} . Thus, the critical point of (6) is always unique. Because the minimization problem (5) is known to have a solution, the found critical point must be located at the global minimum of the cost function $Q_\lambda(\mathbf{r})$.

Alternatively, the solution of (5) can be obtained by finding the least squares solution of

$$\begin{bmatrix} \mathbf{A} \\ \lambda^{1/2}\mathbf{D} \end{bmatrix} \mathbf{r}_\lambda = \begin{bmatrix} \mathbf{b} \\ \lambda^{1/2}\mathbf{e} \end{bmatrix}, \quad (13)$$

which is equivalent to (12). For later reference, system (12) is called *normal equations* and (13) *augmented system*.

2.2. Optimality Results of r_λ . Let us now examine the proposed method from the Bayesian point of view. For this, consider the special case of (5), where $M = 3N$ and $\mathbf{D} = \mathbf{I}$:

$$\mathbf{r}_\lambda = \underset{\mathbf{r}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{r} - \mathbf{b}\|_2^2 + \lambda \|\mathbf{r} - \mathbf{e}\|_2^2. \quad (14)$$

Furthermore, consider that $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_\epsilon^2\mathbf{I})$ and $\mathbf{r} \sim \mathcal{N}(\mathbf{e}, \sigma_\theta^2\mathbf{I})$, where σ_ϵ^2 and σ_θ^2 represent variances of $\boldsymbol{\epsilon}$ and θ , respectively. Here, \mathbf{r} is considered to be a random variable. Then, according to [23], the maximum a posteriori estimator of \mathbf{r} is the solution of

$$\underset{\mathbf{r}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{r} - \mathbf{b}\|_2^2 + \frac{\sigma_\epsilon^2}{\sigma_\theta^2} \|\mathbf{r} - \mathbf{e}\|_2^2. \quad (15)$$

That is, the choice $\lambda = \sigma_\epsilon^2/\sigma_\theta^2$ yields the best estimate of \mathbf{r} in the sense that the likelihood of \mathbf{r} is maximized.

Similar optimality results are known also for more general situations, where for example, $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_\epsilon^2\mathbf{C}\mathbf{C}^T)$. For an extensive discussion of these situations, see [8, 24–26]. To the best of our knowledge, it is an open question whether similar results apply in case of the general form of (5).

2.3. The Choice of λ . In this section, we will provide an introduction to the most relevant systematical methods of finding λ .

According to [8], parameter-choice methods can be divided into two classes depending on their assumptions about the error norm $\|\boldsymbol{\epsilon}\|_2$:

- (1) methods based on knowledge of $\|\boldsymbol{\epsilon}\|_2$,
- (2) methods that do not require $\|\boldsymbol{\epsilon}\|_2$.

Out of these two classes of methods, our main interest is in class two. The reason for this is that knowledge of $\|\boldsymbol{\epsilon}\|_2$ is, in general, a too strict requirement to make. For reference, the most widespread parameter-choice method of class one is the *discrepancy principle* [8].

In class two, there are three popular and widely used methods. These are the *quasioptimality criterion* [8, 27], *generalized cross-validation* [8, 28], and *L-curve criterion* [8, 29, 30]. Let us next take a closer look at the L-curve method, which will be used here to choose the value of λ .

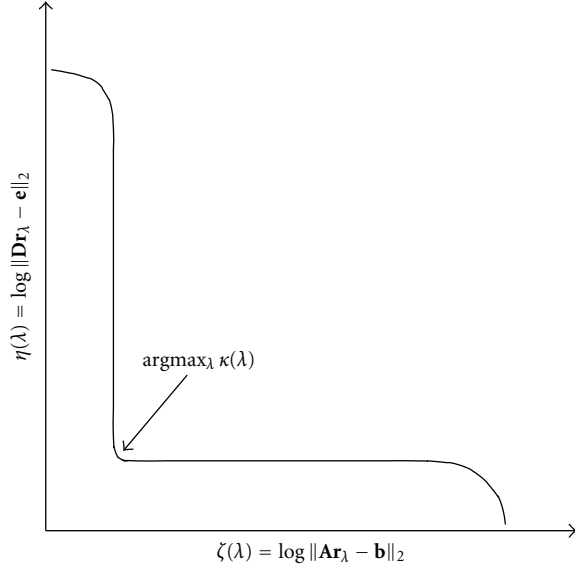


FIGURE 1: Generic form of the L-curve.

The reason for this choice is that the L-curve criterion is, in fact, accessible and a quite intuitive tool for the selection of λ . In the heart of this method is the shape of a certain curve (see Figure 1), which is parametrized as

$$[\zeta(\lambda), \eta(\lambda)] = [\log \|\mathbf{Ar}_\lambda - \mathbf{b}\|_2, \log \|\mathbf{Dr}_\lambda - \mathbf{e}\|_2]. \quad (16)$$

The idea of the L-curve criterion is to choose λ in such a way that there is a “good balance” between the values of the two norms in (5). In other words, balance between $\|\mathbf{e}\|_2$ and $\|\boldsymbol{\theta}\|_2$. Specifically, the L-curve criterion chooses the value of λ maximizing the curvature

$$\kappa(\lambda) = \frac{\dot{\zeta}\ddot{\eta} - \ddot{\zeta}\dot{\eta}}{\left[\left(\dot{\zeta} \right)^2 + \left(\dot{\eta} \right)^2 \right]^{3/2}} \quad (17)$$

of the curve (16). In (17), the dots represent first ($\dot{\cdot}$) and second ($\ddot{\cdot}$) derivatives of ζ and η with respect to λ . The location of the point with maximum curvature is demonstrated in Figure 1. In practice, a suitable one dimensional optimization routine, such as quasi-Newton, can be used to find this point. For a more complete discussion of the reasoning behind the L-curve method, see, for example, [8]. It would also be interesting to compare existing parameter-choice methods in navigation applications, but this does not fit the scope of this paper.

3. Practical Aspects of the Proposed Method

In this section, we will discuss some practical aspects of Tikhonov regularization with L-curve criterion, when it is applied to inertial navigation. First, we will give some insight on the choice of \mathbf{D} and then, discuss the numerical aspects of the problem.

3.1. The Choice of \mathbf{D} . As mentioned in the introduction, \mathbf{D} is determined by the application at hand. Indeed, an application with M position constraints, for example, results in different realization of \mathbf{D} than an application with M velocity constraints would. Both of these are examples of constraints that bound position estimates only locally. Average velocity over longer period of time, trajectory of the object, or possible symmetry in the computed result are examples of constraints bounding an arbitrary number of time instances.

We will now present some examples to demonstrate exploitation of possible constraints. For readability and notational convenience, these examples are one dimensional. It is, however, straightforward to generalize the presented examples to the three dimensions. In the examples, only the nonzero elements of the corresponding row of \mathbf{D} and \mathbf{e} are presented:

(I) (position constraint) $r_j \sim \mathcal{N}(E(r_j), \sigma_i^2)$:

$$D_{i,j} = \frac{1}{\sigma_i^2}, \quad (18)$$

$$e_i = \frac{E(r_j)}{\sigma_i^2},$$

(II) (velocity constraint) $\dot{r}_j \sim \mathcal{N}(E(\dot{r}_j), \sigma_i^2)$:

$$D_{i,j-1} = \frac{-1}{(2h\sigma_i^2)},$$

$$D_{i,j+1} = \frac{1}{(2h\sigma_i^2)}, \quad (19)$$

$$e_i = \frac{E(\dot{r}_j)}{\sigma_i^2},$$

(III) position vector \mathbf{r} is symmetric when mirrored around the middle element (where N is odd) with pointwise variance σ_i^2 (here, $j = N - i + 1$):

$$D_{i,i} = \frac{-1}{\sigma_i^2},$$

$$D_{i,j} = \frac{1}{\sigma_i^2}, \quad \forall i \in \left[1, \frac{N}{2} \right] \quad (20)$$

$$e_i = 0.$$

In these examples, variance σ_i^2 , $\forall i \leq M$, is used to give each constraint a suitable weight, similar to the weighted least squares method [31], assuming, of course, that θ_i are independent.

The fact that $\|\mathbf{e}\|_2$ is unknown and L-curve criterion is used to find λ gives rise to an interesting observation regarding the variance. In addition to its traditional interpretation, σ_i^2 can be treated as a measure of “relative” reliability of each constraint with respect to other constraints. The L-curve criterion is then used to find the “absolute” variances, given by $\lambda\sigma_i^2$.

In this section, we have only considered the situation where the M equations within (4) have different weights. Since (3) contains the measured accelerations and the required boundary conditions [1], the N equations within (3) can also have different significance. In this kind of a situation, one can introduce a diagonal weighting matrix \mathbf{W} , $W_{i,i} > 0$ and multiply both sides of (3) with it. In this paper, however, it is assumed that $\mathbf{W} = \mathbf{I}$.

3.2. Computational Aspects. The overall performance of the proposed method is determined by whether λ has to be solved or not. Namely, if we do not know it, it must be estimated as well, which requires that (5) must be solved for multiple different values of λ . The worse our initial guess for λ is, the harder it gets to find a reasonable value for it. While this “brute force” strategy might be a reasonable choice if the computations are made on a modern desktop computer, it is probably not a practical thing to do on a platform with limited resources. This is why we suggest that when possible, a suitable value of λ should be estimated beforehand, using either simulations or preferably, test measurements. After this, we can expect reasonable accuracy in similar situations, at least where the same type of measurement equipment is used. In example two, we use this strategy to avoid the re-evaluation of λ for each measurement run.

Independently of λ , problem (5) must be solved at least once. In order to avoid possible accuracy and/or performance issues, some attention should be paid. Although normal equations (12) and the augmented system (13) are mathematically equivalent, they are quite different from the numerical solution point of view. Namely, solving the system with normal equations is known to be inaccurate due to round-off errors when forming matrices $\mathbf{A}^T\mathbf{A}$ and $\lambda\mathbf{D}^T\mathbf{D}$. The augmented system, on the other hand, can be accurately solved for with QR or SVD decomposition. There are specific solvers available also for sparse augmented systems [32].

Considering the computational complexity, the normal equations seem quite attractive: with a tridiagonal and symmetric matrix \mathbf{A} [1], it is easy to see that $\mathbf{A}^T\mathbf{A} = \mathbf{A}^2$ will be a symmetric matrix with a bandwidth of five. The structure of the regularization matrix \mathbf{D} is, however, arbitrary in general. Yet in many situations, it will also be a banded matrix, like in the case of constraints (18) and (19). In case of (20), the situation looks worse, but significantly smaller bandwidth can be obtained by reordering the equations. In case of the normal equations, the resulting system has dimensions $3N \times 3N$ (for general three-dimensional problem) *independently* of M .

Consider a problem leading to a reasonably small bandwidth $p \in \mathbb{N}$ in case of the normal equations. For these situations, we propose an iterative refinement-based solution method similar to the *corrected seminormal equations* (CSNE) method [32]. Traditionally, CSNE is based on the use of a numerical approximation of the matrix $\tilde{\mathbf{A}} = \mathbf{A}^T\mathbf{A} + \lambda\mathbf{L}^T\mathbf{L}$, computed with QR decomposition. In this case ($p \ll 3N$), we can, however, compute elements of $\tilde{\mathbf{A}}$ analytically with negligible round-off errors. Thus, no QR decomposition is required and the corresponding algorithm

for finding \mathbf{r} is

```

 $\mathbf{r}_{\text{new}} \leftarrow \mathbf{0}, \mathbf{r}_{\text{old}} \leftarrow \mathbf{1}$ 
While  $\|\mathbf{r}_{\text{new}} - \mathbf{r}_{\text{old}}\| > \delta$  do
   $\mathbf{r}_{\text{old}} \leftarrow \mathbf{r}_{\text{new}}$ 
   $\delta_{\mathbf{r}} \leftarrow \begin{bmatrix} \mathbf{b} \\ \sqrt{\lambda}\mathbf{e} \end{bmatrix} - \begin{bmatrix} \mathbf{A} \\ \sqrt{\lambda}\mathbf{D} \end{bmatrix} \mathbf{r}_{\text{old}}$ 
   $\mathbf{r}_{\text{update}} \leftarrow \tilde{\mathbf{A}}^{-1}(\begin{bmatrix} \mathbf{A}^T & \sqrt{\lambda}\mathbf{D}^T \end{bmatrix} \delta_{\mathbf{r}})$ 
   $\mathbf{r}_{\text{new}} \leftarrow \mathbf{r}_{\text{old}} + \mathbf{r}_{\text{update}}$ 
end While

```

with stopping criterion δ [32]. Notice that the solution of the involved matrix equation can be carried out with linear time complexity, as it requires $3Np^2 + 24Np + 3N$ flops [31]. In many occasions, one refinement step is enough, and more than three refinement steps are seldom required. For problems with “large” p , this approach is not practical, and solvers designed for augmented systems should be used instead.

4. Example 1

In this section, we will work out a simple one dimensional example demonstrating the use of the Tikhonov regularization method. The object of this example is to compare three different solution strategies, all of which are based on the finite element method:

- (I) tikhonov regularization based on finite element method with the position fixed at both ends,
- (II) finite element method with the position fixed at both ends,
- (III) finite element method with the position fixed at both ends, where measurement error is modeled as described in [1].

4.1. Setting Up the Example. Let us generate the reference position $p(t)$, $t \in [0, 10]$ s as a fifth-order polynomial satisfying the following six conditions:

$$\begin{aligned}
 p(0) &= 0 \text{ m}, & \dot{p}(0) &= 0 \text{ m/s}, \\
 p(5) &= -10 \text{ m}, & \dot{p}(5) &= 0 \text{ m/s}, \\
 p(10) &= 10 \text{ m}, & \dot{p}(10) &= 0 \text{ m/s}.
 \end{aligned} \tag{21}$$

From the resulting $p(t)$, it is an easy task to compute the reference velocity $v(t)$ and acceleration $a(t)$. Erroneous “measurement” $\hat{a}(t)$, containing scale factor error s and bias term μ with additional white noise, is generated as follows:

$$\hat{a}(t) = (1 + s)a(t) + \mathcal{N}(\mu, \sigma^2). \tag{22}$$

In Figure 2, one realization of $\hat{a}(t)$ ($s = 0.05$, $\mu = -0.50$ and $\sigma^2 = 1$), sampled with a frequency of 50 Hz ($N = 501$), is presented with the reference acceleration $a(t)$.

Now, suppose that along with $\hat{a}(t)$, we know the position and velocity of the object at points $t = 0$ seconds and $t = 10$ seconds. Thus, we have the following problem: find $\hat{p}(t)$,

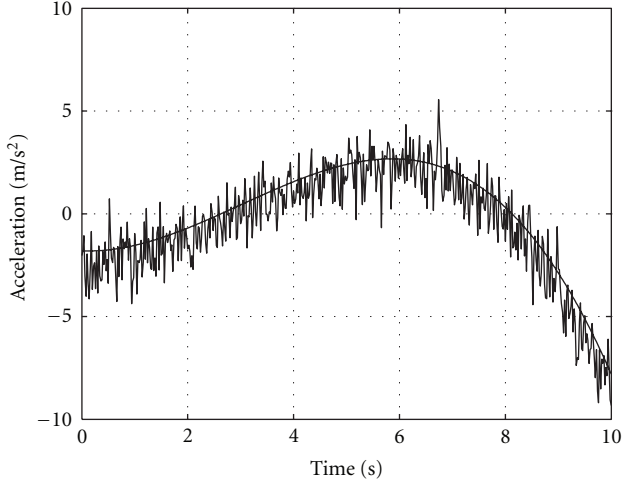


FIGURE 2: Actual acceleration (thick line) and “measured” acceleration (thin line) as a function of time.

$0 \leq t \leq 10$ such that

$$\begin{aligned}
 \hat{p}(t) &= \hat{a}(t), \\
 \hat{p}(0) &= 0 \text{ m}, \\
 \dot{\hat{p}}(0) &= 0 \text{ m/s}, \\
 \hat{p}(10) &= 10 \text{ m}, \\
 \dot{\hat{p}}(10) &= 0 \text{ m/s}.
 \end{aligned} \tag{23}$$

The problem here is, of course, that we do not know $a(t)$, but only the erroneous measurement $\hat{a}(t)$. Thus, ultimately, the goal is to minimize the effects of the measurement error using the given two additional boundary conditions.

To exploit method I, we must first generate a suitable prior, which can then be used to enhance the accuracy of the solution. For this, let us compute a polynomial of the lowest order that satisfies the given four conditions at the boundaries and use that as our prior. In this case, a suitable choice for the regularization matrix \mathbf{D} would be $N \times N$ identity matrix. However, to emphasize the fact that we have no trust in our prior between the two boundaries, we will zero out all but, say, 10 first and last rows of \mathbf{D} . Thus, in this example, $M = 20$. Method II follows from the method I by setting $\lambda = 0$. For method III, we use the error model presented in [1], modeling constant scale factor and bias errors using the velocity of the object at $t = 0$ and $t = 10$. Knowing the form of the generated measurements, this model will yield the exact solution when no noise is applied. The idea here is, however, to test the performance of the method III in case of noisy measurements.

4.2. Results. Table 1 summarizes the test results, which were run with a total of nine different combinations of error parameters s and μ . Each test was run several times with different realizations of random noise in order to show the average performance of the tested methods. As an overall comment, the tested methods are not particularly sensitive to

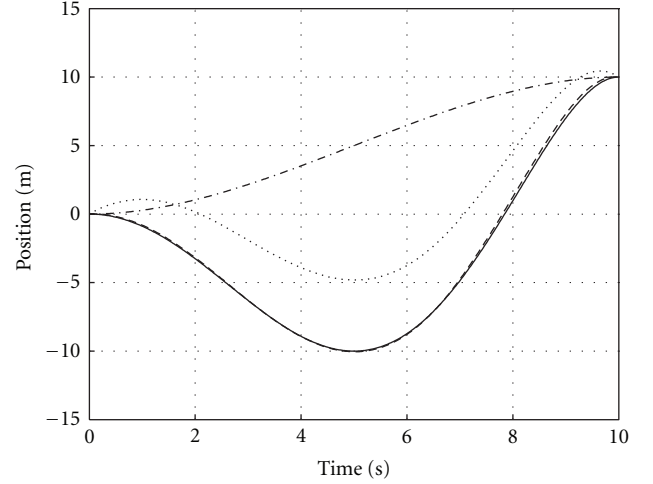


FIGURE 3: Actual position (solid line), estimated position computed with method I (dashed line, $\lambda = 3.611$), estimated position computed with method II (dotted line) and used prior (dash-dotted line) as a function of time.

noise. The accuracy of the method III is, however, dependent on the value of σ^2 . This was expected, since method III uses only two measurements to determine the two error parameters, which gets more inaccurate as the noise level increases. The value of λ minimizing the position error is presented only for reference, as there is no practical way to find it in real situations. Value of λ_L , however, can be found for each measurement using the L-curve criterion. During the test, the value of λ minimizing the velocity error was also recorded, but the difference to the value minimizing the position error was not significant.

Based on the results seen in Table 1, the L-curve method seems to work reasonably well. Indeed, in some cases it is even able to predict the optimal value of λ (or at least obtain position error levels very close to the minimum). On average, Tikhonov regularization with λ_L works much (five to ten times) better than the method II. It can also fail, meaning that the resulting error is larger than the one given by method II. This happened only when the optimal value of λ was zero, and the results of method II were optimal. The L-curve method, on the other hand, performed consistently among the test situations. When compared with method III, the difference is not as clear, although on average, Tikhonov regularization does slightly better. As mentioned above, this is due to the large value of σ^2 . In situations with lower noise and known main error sources, a combination of methods I and III is also a reasonable strategy. This is due to the fact that the error modeling can reduce such components of error (such as bias) with which Tikhonov regularization is not intended to work with.

Figures 3 and 4 illustrate the performance of methods I and II in terms of position and velocity. The measurement was the one seen in Figure 2 with parameters $s = 0.05$ and $\mu = -0.5$. Value $\lambda = \lambda_L$ was used to compute the regularized solution. The figures also demonstrate the used prior, only first and last 0.2 seconds of which were actually

TABLE 1: Simulation results for various values of parameters s and μ . Terms σ_I , σ_{II} and σ_{III} refer to the standard deviation of the position error (in meters) in case of solution methods I, II, and III. The value of the regularization parameter obtained via the L-curve method is represented as λ_L and λ represents the value of the regularization parameter minimizing the position error.

s	μ	σ^2	$\operatorname{argmin}_\lambda \sigma_I(\lambda)$	λ	$\sigma_I(\lambda_L)$	λ_L	σ_{II}	σ_{III}
-0.05	-0.5	1	1.137	4.200	1.137	4.501	5.310	1.236
-0.05	0.0	1	0.453	0.000	1.024	1.438	0.453	0.568
-0.05	0.5	1	0.109	0.149	0.995	5.663	4.040	1.413
0.00	-0.5	1	0.813	4.073	0.813	4.255	4.781	0.723
0.00	0.0	1	0.282	0.000	0.737	1.330	0.282	0.693
0.00	0.5	1	0.101	0.221	0.729	5.185	4.353	1.015
0.05	-0.5	1	0.286	3.407	0.286	3.611	3.823	0.757
0.05	0.0	1	0.139	1.230	0.260	1.581	0.666	1.634
0.05	0.5	1	0.132	0.678	0.336	5.243	4.881	0.758

$$\hat{a}(t) = (1 + s)a(t) + \mathcal{N}(\mu, \sigma^2).$$

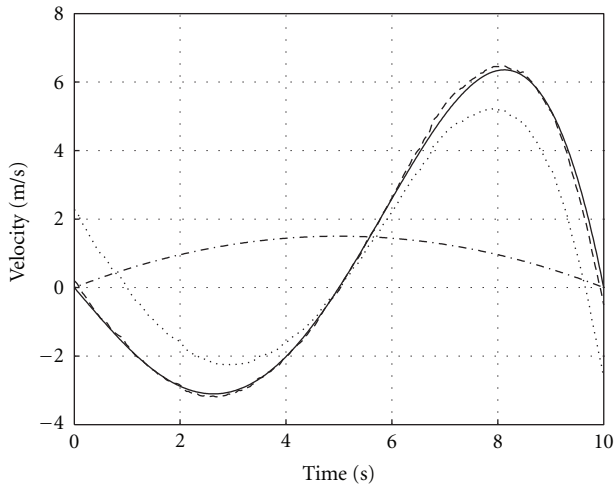


FIGURE 4: Actual velocity (solid line), estimated velocity computed with method I (dashed line, $\lambda = 3.611$), estimated velocity computed with method II (dotted line) and used prior (dash-dotted line) as a function of time.

exploited. Clearly, comparing to method II, the accuracy of the proposed method is remarkably better. In terms of the standard deviation, the proposed method is about 13 times more accurate than method II. When comparing to method III, the proposed method is about 2.5 times more accurate, which is also a significant improvement.

In Table 2, the computational complexity of the proposed method is demonstrated. The number of samples (N) and constraints (M) was increased by increasing the sampling frequency, keeping the problem otherwise unchanged. The value of λ was kept fixed. As the solution methods were not optimized for maximum performance, the differences in the solution times may not be fully comparable. Especially in case of the LSQR method, default tolerance (relative accuracy of $1e-6$) was used. Despite the “large” number of unknowns, the solution times are small for any solution method. Based on the results, one can expect linear time complexity with respect to $N + M$ for similar problems. In the last case with

$N = 128000$, the efficiency of the CSNE method drops as the number of iterations increases to seven. This is due to the fact that the conditioning of the system is comparable to N^2 . Otherwise, the CSNE method seems to provide good efficiency with a reasonably small cost in the accuracy.

The main purpose of this example was to demonstrate the performance of the proposed method in situations where the measurement error is biased (due to error term μ) and not normally distributed (due to error term s). That is, in situations where the presented optimality results are not valid. The results indicate that the proposed method is useful also in these situations, commonly encountered in practice. Moreover, the tests indicate that the proposed method can be very efficient with negligible computing times even for large-scale problems with over 100000 unknowns. This also guarantees that when required, also the value of λ can be determined with reasonably small computational cost.

5. Example 2

In this section, we will demonstrate the use of the regularization method with actual measurement data obtained in a real-world situation. The data was obtained for use with a television program of NHK (Japan Broadcasting Corporation) and presented here with due permission. The primary goal was to measure the accelerations to which the passengers were subjected in a car of a roller coaster (The roller coaster in question is called Insane, located at Gröna Lund, Stockholm, Sweden.) In this example, we present a way to estimate the velocity and trajectory of the car with knowledge about the trajectory of the car. It is important to notice that only the trajectory is known, not its time parametrization.

5.1. Measurement Setup. The car moved on a planar trajectory of length ≈ 250 m and rotated freely around its rotation axis (orthogonal to the plane) during the ride. In other words, the problem has three degrees of freedom: one for the rotation and two for the displacements. The trajectory is not known accurately, but roughly estimated using the

TABLE 2: Computing times as N and M increase. All times are average computing times over 100 cycles and include only the solution phase. Tests were performed with Matlab 7.9.0 running on a 32-bit operating system. The used computer has a 2 GHz processor and 2 GB of memory. CSNE stands for the method described in Section 3.2, “\” for the “backslash” operator, and LSQR for the iterative least-squares problems solution routine implemented in Matlab. “It” indicates the number of iterations made by CSNE and $\delta/N^{1/2}$ indicates the “size” of the last correction step.

N	M	It	CSNE		“\”	LSQR
			T (ms)	$\delta/N^{1/2}$	T (ms)	T (ms)
500	20	1	1	$6.4e-9$	2	5
1000	40	1	2	$3.4e-8$	4	7
2000	80	1	4	$3.1e-7$	8	12
4000	160	1	8	$5.5e-6$	16	20
8000	320	1	16	$8.7e-5$	34	36
16000	640	1	34	$4.1e-4$	76	63
32000	1280	2	103	$1.7e-5$	148	117
64000	2560	2	212	$4.0e-4$	302	271
128000	5120	7	1106	$3.1e-4$	635	618

length and an image of the trajectory. Because of this, the estimated trajectory is known to contain significant errors, which makes it a good example of situations with a rather vague prior.

The data was collected using an inertial measurement unit (IMU) consisting of the following sensors:

- (i) $3 \times \pm 75^\circ/\text{sec}$ Silicon Sensing CRS10 [33],
- (ii) $3 \times \pm 300^\circ/\text{sec}$ AD ADXRS300ABG [34],
- (iii) $3 \times \pm 3 \text{ g}$ VTI SCA610-CC5H1A [35],
- (iv) $3 \times \pm 12 \text{ g}$ VTI SCA620-CHCV1A [36].

The IMU is calibrated as described in [37]. Only the sensors of the lower dynamic range were exploited unless the input exceeded this range. In this case, only the sensors with the higher dynamical range were exploited. The sample rate was 1000 Hz, and the total number of unknown displacements was about 200000 in each measurement. Based on the construction of the roller coaster, the initial and the final angle were known to be identical. This allowed the angle to be computed with high accuracy. Thus, the accelerometers were the major source of error.

5.2. Exploitation of the Estimated Trajectory. In this section, the methods used to exploit the estimated trajectory are presented. Notice that only the trajectory is estimated, not the time when the car is at a certain point. It is assumed that the horizontal coordinates of the trajectory are stored to vector $\mathbf{X} \in \mathbb{R}^P$ and the vertical coordinates, respectively, to vector $\mathbf{Y} \in \mathbb{R}^P$. The number of the points ($P \in \mathbb{N}$) is assumed to be high enough and the trajectory data “smooth” enough such that the local tangent of the trajectory can be reliably estimated using two adjacent points of the trajectory.

Given a horizontal coordinate r_j , $j \in [1, N]$ such that $X_{i-1} \leq r_j < X_i$, $i \in [1, P]$, the corresponding vertical coordinate r_{N+j} satisfies the equation

$$(Y_i - Y_{i-1})r_j - (X_i - X_{i-1})r_{N+j} = X_{i-1}Y_i - X_iY_{i-1}. \quad (24)$$

Depending on the trajectory, it is possible that the index i is ambiguous, meaning that there can be more than one vertical coordinate corresponding to a single horizontal coordinate. This is the case also here, as can be verified from Figure 5. Thus, it is also required to choose correct i . It is also possible that no suitable index i is found, meaning that the value of r_j is out of range. In such a situation, the trajectory is not exploited for j in question. With these notions, the corresponding elements of the regularization term are

$$\begin{aligned} D_{j,j} &= Y_i - Y_{i-1}, \\ D_{j,N+j} &= X_i - X_{i-1}, \quad \forall j \in [1, N]. \\ e_j &= X_{i-1}Y_i - X_iY_{i-1}, \end{aligned} \quad (25)$$

5.3. Solution Process. The problem was solved in two phases. In the first phase, only few points of the trajectory were exploited. These were the points corresponding to the local extrema of the horizontal coordinates of the stored trajectory. The goal of this phase was to produce a reasonable initial guess for \mathbf{r} , which could then be used to obtain better results when linearizing the trajectory. For this, valid estimates of the indices j when the car crosses these points were required. In this phase, we used a rather high value $\lambda = 100$ to minimize the amount of values r_j that were out of range in the next phase. After this phase, the measurement vector \mathbf{b} was updated by setting $\mathbf{b} \leftarrow \mathbf{A}\mathbf{r}_\lambda$.

In phase two, the trajectory of the car was exploited throughout the track. This was done by linearizing the estimated trajectory at each time instance in the neighborhood of the horizontal position obtained in the first phase, as explained in Section 5.2. Variance of each constraint was set to one, as we had no knowledge of the possible differences in the reliability of these estimates. For the first run of the car, the value of the regularization parameter was obtained using the L-curve criterion. The same value of λ was then used also for the second and third run.

5.4. Results. In Figure 5, the computed three trajectories are compared to the used prior. In each case, the car starts

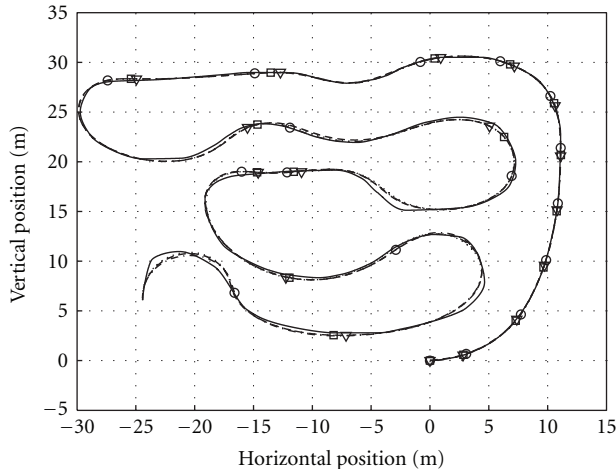


FIGURE 5: Actual trajectory (solid line), estimated trajectory of run 1 (dashed line), run 2 (dash-dotted line), and run 3 (dotted line). Small squares, circles, and triangles indicate the position of the car with three second increments, respectively, for runs 1, 2, and 3. Regularization parameter $\lambda = 0.352$ for all three runs.

at point $[0, 0]$ leaving to the upright direction. The end point was set to the last local extremum of the horizontal coordinates of the estimated trajectory, because the time instance corresponding to the actual stopping point could not be specified. Notice how the three computed trajectories are almost identical, while they follow the provided prior only approximately. Given that no information about the reliability of different points of the trajectory is provided, this indicates that the L-curve method provides useful estimates of the regularization parameter. Since this happens for all three runs, determining the value of λ in advance, as discussed before, is a plausible technique.

The small squares, circles, and triangles seen in Figure 5 indicate the position of the car with three second increments, respectively, for runs 1, 2, and 3. As seen in the figure, the distances between these marks change significantly, indicating that the used method “allows” the velocity of the car to change between different runs. For the first ≈ 50 meters of the trajectory, however, the distance between the marks does not change much. This is due to the fact that during these first meters, the cars were lifted up essentially the same way for each run. After the highest point of the trajectory, the car moved freely along the track “driven” only by gravity and affected by the moment of inertia, determined by the placements of the passengers. As seen in Figure 5, the run represented with the circles, for example, attained significantly higher velocities than the other two runs.

6. Conclusion

A BVP formulation of inertial navigation problems is further investigated using [1] as a starting point. It is suggested that the possible additional information can be taken into account by exploiting Tikhonov regularization, a basic tool of inverse problems. In addition to typical additional

information, such as a momentary position or velocity, it allows one to exploit more general forms of information. These include, but are not limited to, average velocity over longer period of time, trajectory of the object, and possible symmetry in the computed result. In the provided examples, significant accuracy improvements—up to an order of magnitude—over the basic FEM solution without any additional information are obtained.

It is also demonstrated that the proposed method can be viewed as a Bayesian estimator, yielding the maximum a posteriori likelihood estimator in case of unbiased and uncorrelated measurement errors. In addition to this, the provided examples show that the obtained accuracy improvement is significant, even in cases where these assumptions are not met.

A method for choosing the value of the regularization parameter is provided and demonstrated to work in a real-world example. This method, called the L-curve criterion, does not require any prior knowledge of the measurement errors. Thus, it can be used in many real-world situations, where measurement error is always present, but seldom reliably characterized. For performance-critical situations, it is possible to determine a suitable value for the regularization parameter in advance and use this value to obtain good results.

In some applications, it is possible and advantageous to combine the method based on sensor error modeling [1] with the method proposed herein. Sensor error modeling is used first to eliminate the modeled deterministic measurement errors, and Tikhonov regularization is then used to minimize the effects of the remaining stochastic components of the measurement error.

Acknowledgments

The authors would like to thank D.Tech. Saku Suuriniemi and Professor Stefan Kurz for proofreading and improving the clarity of the contents of this paper. They would also like to thank Olli Särkkä (M.S. degree) for his invaluable assistance while obtaining the measurements of Example 2.

References

- [1] T. Nieminen, J. Kangas, S. Suuriniemi, and L. Kettunen, “Accuracy improvements by boundary conditions for inertial navigation,” *International Journal of Navigation and Observation*, vol. 2010, Article ID 869127, 10 pages, 2010.
- [2] L. F. Shampine, I. Gladwell, and S. Thompson, *Solving ODEs with Matlab*, Cambridge University Press, Cambridge, UK, 2003.
- [3] U. M. Asher and L. R. Petzold, *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*, SIAM, Philadelphia, Pa, USA, 1998.
- [4] A. N. Tikhonov, “On the stability of inverse problems,” *Doklady Akademii Nauk SSSR*, vol. 39, no. 5, pp. 195–198, 1943.
- [5] A. N. Tikhonov, “Solution of incorrectly formulated problems and the regularization method,” *Doklady Akademii Nauk SSSR*, vol. 151, pp. 501–504, 1963.
- [6] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill Posed Problems*, VH Winston, Washington, Wash, USA, 1977.

- [7] D. L. Phillips, "A technique for the numerical solution of certain integral equations of the first kind," *Journal of the ACM*, vol. 9, no. 1, pp. 84–97, 1962.
- [8] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, SIAM, Philadelphia, Pa, USA, 1998.
- [9] A. E. Hoerl, "Application of ridge analysis to regression problems," *Chemical Engineering Progress*, vol. 58, pp. 54–59, 1962.
- [10] D. W. Marquardt, "Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation," *Technometrics*, vol. 12, no. 3, pp. 591–612, 1970.
- [11] A. Gelb, *Applied Optimal Estimation*, The M.I.T. Press, Cambridge, Mass, USA, 1974.
- [12] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, John Wiley & Sons, New York, NY, USA, 1st edition, 2001.
- [13] J. L. Crassidis and J. L. Junkins, *Optimal Estimation of Dynamic Systems*, Chapman & Hall/CRC, Boca Raton, Fla, USA, 2004.
- [14] P. D. Groves, *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*, Artech House, Boston, Mass, USA, 2008.
- [15] D. C. Fraser and J. E. Potter, "The optimum linear smoother as a combination of two optimum linear filters," *IEEE Transactions on Automatic Control*, vol. 14, no. 4, pp. 387–390, 1969.
- [16] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA Journal*, vol. 3, no. 8, pp. 1445–1450, 1965.
- [17] K. Gade, "NavLab, a generic simulation and post-processing tool for navigation," *European Journal of Navigation*, vol. 2, no. 4, pp. 51–59, 2004.
- [18] Y. Yang, Z. Jin, W. Tian, and F. Qian, "Application of fixed interval smoothing to gps/dr integrated navigation system," in *Proceedings of Intelligent Transportation Systems*, vol. 2, pp. 1027–1031, October 2003.
- [19] A. B. Willumsen and Ø. Hegrenæs, "The joys of smoothing," in *IEEE Bremen: Balancing Technology with Future Needs (OCEANS '09)*, pp. 1–7, May 2009.
- [20] T. Roths, M. Marth, J. Weese, and J. Honerkamp, "A generalized regularization method for nonlinear ill-posed problems enhanced for nonlinear regularization terms," *Computer Physics Communications*, vol. 139, no. 3, pp. 279–296, 2001.
- [21] G. H. Golub, P. C. Hansen, and D. P. O'Leary, "Tikhonov regularization and total least squares," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 1, pp. 185–194, 2000.
- [22] M. Vauhkonen, D. Vadász, P. A. Karjalainen, E. Somersalo, and J. P. Kaipio, "Tikhonov regularization and prior information in electrical impedance tomography," *IEEE Transactions on Medical Imaging*, vol. 17, no. 2, pp. 285–293, 1998.
- [23] E. Demidenko, *Mixed Models: Theory and Applications*, John Wiley & Sons, Hoboken, NJ, USA, 2004.
- [24] A. Tarantola, *Inverse Problem Theory*, SIAM, Philadelphia, Pa, USA, 2005.
- [25] J. N. Franklin, "Well-posed stochastic extensions of ill-posed linear problems," *Journal of Mathematical Analysis and Applications*, vol. 31, no. 3, pp. 682–716, 1970.
- [26] D. D. Jackson, "The use of a priori data to resolve non-uniqueness in linear inversion," *Geophysical Journal of the Royal Astronomical Society*, vol. 57, no. 1, pp. 137–157, 1979.
- [27] M. Hanke and T. Raus, "A general heuristic for choosing the regularization parameter in ill-posed problems," *SIAM Journal of Scientific Computing*, vol. 17, no. 4, pp. 956–972, 1996.
- [28] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, no. 2, pp. 215–223, 1979.
- [29] H. W. Engl and W. Grever, "Using the l-curve for determining optimal regularization parameters," *Numerische Mathematik*, vol. 69, no. 1, pp. 25–31, 1994.
- [30] P. C. Hansen, "Analysis of discrete illposed problems by means of the l-curve," *SIAM Review*, vol. 34, no. 4, pp. 561–580, 1992.
- [31] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Md, USA, 3rd edition, 1996.
- [32] Å. Björck, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, Pa, USA, 1996.
- [33] "Silicon Sensing CRS10 Datasheet," 2007. Revision 2, Checked on 23.09.2010.
- [34] "Analog Devices ADXRS300ABG Datasheet," 2004. Revision B, Checked on 22.12.2009.
- [35] "VTI SCA610-CC5H1A Datasheet," 2005. Revision 3, Checked on 23.09.2010.
- [36] "VTI SCA620-CHCV1A Datasheet," 2006. Revision 2/2, Checked on 22.12.2009.
- [37] T. Nieminen, J. Kangas, S. Suuriniemi, and L. Kettunen, "An enhanced multi-position calibration method for consumer-grade inertial measurement units applied and tested," *Measurement Science and Technology*, vol. 21, no. 10, Article ID 105204, 2010.