

A New Simple, Fast and Robust Total Least Square Error Computation in E2: Experimental Comparison

Michal Smolik, Vaclav Skala, and Zuzana Majdisova

Faculty of Applied Sciences, University of West Bohemia, Plzen, Czech Republic,
{smolik, skala, majdisz}@kiv.zcu.cz

Abstract. Many problems, not only in signal processing, image processing, digital imaging, computer vision and visualization, lead to the Least Square Error (LSE) problem or Total (Orthogonal) Least Square Error (TLSE) problem computation. Usually the standard least square error approximation method is used due to its simplicity, but it is not an optimal solution, as it does not optimize the orthogonal distances, but only the vertical distances. There are many problems for which the LSE is not convenient and the TLSE is to be used. Unfortunately, the TLSE is computationally much more expensive. This paper presents a new, simple, robust and fast algorithm for the total least square error computation in E^2 .

Keywords: Least squares, Total least squares, Orthogonal distance, Approximation

1 Introduction

The fitting of geometric features to given $2D$ or $3D$ points is desired in various fields of science and engineering. The least squares error approximation method (LSE) [7], [6] is one of the best known, and most often applied, mathematical tools in various disciplines of science and engineering, e.g. signal processing [4], optics [18], surface modeling [12], regression modelling [8], nonlinear systems [14]. In the past, fitting problems have usually been solved through the LSE method with respect to the effective implementation and acceptable computing cost. However the natural and best choice of the error distance is the shortest distance between the given point and the model feature. This error definition is used in total (orthogonal) least squares error approximation method (TLSE), which fits the model more accurately than the standard LSE method [2], [1].

2 Least Square Error Approximation

In the vast majority the Least Square Error (LSE) methods measuring vertical distances are used. This approach is acceptable in the case of explicit functions $f(x, y) = h$, resp. $f(x, y, z) = h$. However, it should be noted that a user should

keep in a mind, that smaller differences than 1.0, will have significantly smaller weight than higher differences than 1.0 as the differences are taken in a square resulting to dependence in scaling of the approximated data, i.e. the result will depend on physical units used, etc. The main advantage of the LSE method is its simplicity for fitting polynomial curves and it is easy to implement. The standard LSE method leads to an over determined system of linear equations. This approach is also known as polynomial regression.

Let us consider a data set $\Omega = \{(x_i, y_i, f_i)\}_{i=1}^N$, i.e. data set containing for x_i and y_i measured functional value f_i and we want to find parameters $\mathbf{a} = [a, b, c, d]^T$ for optimal fitting function, as an example:

$$f(x, y, \mathbf{a}) = a + bx + cy + dxy \quad (1)$$

by minimizing the vertical squared distance D , i.e.:

$$D = \min_{a,b,c,d} \sum_{i=1}^N (f_i - f(x_i, y_i, \mathbf{a}))^2. \quad (2)$$

Conditions for an extreme are given as a vector equation:

$$\frac{\partial D}{\partial \mathbf{a}} = \sum_{i=1}^N (f_i - (a + bx_i + cy_i + dx_i y_i)) \frac{\partial f(x, y, \mathbf{a})}{\partial \mathbf{a}} = \mathbf{0}. \quad (3)$$

All those conditions can be rewritten in a matrix form as $\mathbf{Ax} = \mathbf{b}$. The selection of bilinear form was used to show the LSE method application to a non-linear case, if the case of linear function, i.e. $f(x, y, \mathbf{a}) = a + bx + cy$, the 4th row and column of the matrix \mathbf{A} is to be removed.

Several methods for LSE have been derived [5], [11], however those methods are sensitive to the vector \mathbf{a} orientation and not robust in general as a value of $\sum_{i=1}^N x_i^2 y_i^2$ might be too high in comparison with the value N which has an influence to robustness of numerical solution. Also the LSE methods are sensitive to a rotation as they measure vertical distances. Rotational and translation invariance is fundamental requirement not only in geometrically oriented applications.

3 Total Least Square Error Approximation

The Total (Orthogonal) Least Square (TLSE) method takes another approach as it measures distances orthogonally and approximation by a line or plane is used nearly exclusively [19], [22]. One significant property of the TLSE method is its rotational and translational invariance [21], [20]. This approach leads to an approximation by an implicit function $F(x, y) = 0$ in the E^2 case, resp. $F(x, y, z) = 0$, in the E^3 case, i.e. dependence expressible as an implicit function.

There are several approaches how to solve TLSE problem and comprehensive analysis is given in [9]. Many algorithms are based on Singular Value Decomposition (SVD) or on a "simple" solution based on the explicit line representation

[13] in the form $y = bx + a$. This formulation leads to a simple formula for calculation of the a, b coefficients. However, it is not robust and it is sensitive to a rotation. If we rotate the input points and then compute the LSE approximation, the result is different than when we first compute the LSE approximation and then rotate the line. Although when using the TLSE approximation, there is no such problem with rotation and the results are consistent. Also when a line is close to a vertical one, there is a high numerical imprecision for the LSE approximation and an overflow can appear as well, etc. If TLSE method is to be used many times, it is reasonable to consider robust and fast method specialized for the E^2 case. In the E^2 case and the linear case a linear function $F(\mathbf{x}) = ax + by + c = 0$ is used, the orthogonal distance d of the given point \mathbf{x} and the line p is determined as:

$$d = \frac{|ax + by + c|}{\sqrt{a^2 + b^2}}, \quad (4)$$

where $\mathbf{x} = [x, y]^T$ is the given point and a line p is given as $ax + by + c = 0$. The computational problem is determination of coefficients a, b, c of a line $p \in E^2$.

There is a difference in approximation results for the LSE and the TLSE method as each of them use a different minimization criteria. The difference can be high for higher angles between x axis and the line, see Fig. 1 for an example.

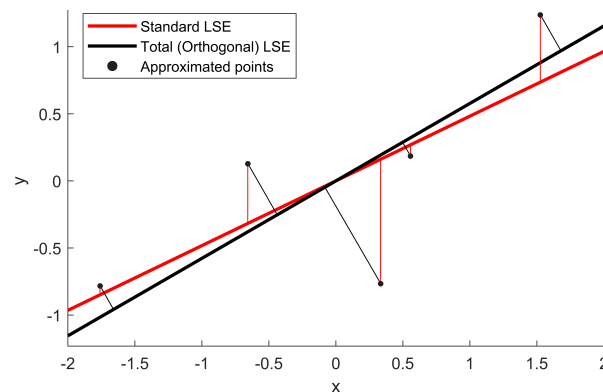


Fig. 1. Difference between standard least square error method and total (orthogonal) least square method.

In image processing, signal processing, digital imaging and computer graphics specialized algorithms should be used in the E^2 case. Such a solution for the E^2 case was published in [3] which is based on a line representation in the polar coordinates. Some specialized algorithms for a circle, resp. ellipse fitting were developed recently as well. The algorithm fully described in [3] is based on polar representation and leads to a formula which is stable. The derivation of the

algorithm is not simple and uses goniometric functions, i.e. $\sin(\theta)$ and $\cos(\theta)$. A special case for perfectly circular data is to be solved. The algorithm [3] is not extensible to the E^3 case.

3.1 Total Least Square Error - Goniometric Functions

The approach proposed by Alciatore [3] defines an implicit line as

$$x\sin(\theta) + y\cos(\theta) + \rho = 0, \quad (5)$$

where θ is chosen to be line orientation with respect to the x axis. The LSE method is computed as the sum of the squares of the related perpendicular distances is minimal. That is, minimize the value

$$\sum_{i=1}^N r_i^2(\theta, \rho), \quad (6)$$

where N is the number of points and r_i are perpendicular distances.

After lengthy derivation, the authors end up with the following formulas

$$\begin{aligned} A &= 2b' \\ B &= -\left(a' + \sqrt{(a')^2 + 4(b')^2}\right) \\ C &= A\bar{x} + B\bar{y}, \end{aligned} \quad (7)$$

where A , B and C are parameters of implicit line $Ax + By + C = 0$, $[\bar{x}, \bar{y}]^T$ is the centroid of the dataset $\{(x_i, y_i)\}$. Values of a' and b' are computed using

$$a' = \sum_{i=1}^N (x_i - \bar{x})^2 - \sum_{i=1}^N (y_i - \bar{y})^2 \quad b' = \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}). \quad (8)$$

3.2 Total Least Square Error - Parametric Form

The approach proposed by Skala [16] defines a parametric line as

$$y = \mathbf{x}_T + t\mathbf{s}, \quad (9)$$

where \mathbf{x}_T is the centroid of the dataset $\{(x_i, y_i)\}$, \mathbf{s} is a directional vector and t a parameter $t \in \mathfrak{R}$. The TLSE method is computed as the sum of the squares of the related perpendicular distances is minimal. That is, minimize the value

$$\sum_{i=1}^N r_i^2 = \sum_{i=1}^N \frac{((\mathbf{x}_i - \mathbf{x}_T) \times \mathbf{s})^T ((\mathbf{x}_i - \mathbf{x}_T) \times \mathbf{s})}{\mathbf{s}^T \mathbf{s}}, \quad (10)$$

where N is the number of points, r_i are perpendicular distances and vector \mathbf{s} is to be calculated. After derivation, the author end up with the following formula

$$(\mathbf{\Omega} - (\mathbf{s}^T \mathbf{\Omega} \mathbf{s}) \mathbf{I}) \mathbf{s} = \mathbf{0}, \quad (11)$$

where Ω is

$$\Omega = \sum_{i=1}^N \mathbf{x}_i \otimes \mathbf{x}_i^T, \quad (12)$$

where \otimes means a tensor product (result is a matrix).

4 Proposed Approach

In the following a new approach to TLSE computation will be described with experimental verification of the proposed method.

Fundamental requirement for any algorithm is its robustness. It should be fast and simple to implement as well. The proposed TLSE algorithm [17] is based on a squared orthogonal distance computation. As the TSLE method has to be translationally and rotationally invariant, the centroid of the given point set is to be $\mathbf{x}_0 = 0$, this was shown also in [3]. As it is not a general case, the first step is a data set transformation:

$$\mathbf{x}_i = \mathbf{x}_i - \mathbf{x}_0 \quad \mathbf{x}_0 = \frac{1}{2} \sum_{i=1}^N \mathbf{x}_i, \quad (13)$$

where N is the number of the given points, $\mathbf{x}_i = [x_i, y_i]^T$ are the given points, $i = 1, \dots, N$. This step has two consequences, the line p ($ax + by + c = 0$) passes the origin of the coordinate system and therefore the coefficient c of the line p is set $c = 0$ by definition, now.

There is a simple formulation of the TLSE problem using optimization and Lagrange multipliers, i.e.

$$\min_{a,b,\lambda} D(a, b, \lambda) = \min_{a,b,\lambda} \left((ax_i + by_i)^2 + \lambda g(a, b) \right) \quad \& \quad g(a, b) = 0, \quad (14)$$

where

$$g(a, b) = a^2 + b^2 - 1. \quad (15)$$

Unfortunately, this approach does not lead to a simple solution.

The proposed algorithm is based on direct minimization of a distance given as:

$$D(a, b) = \sum_{i=1}^N d_i^2 = \sum_{i=1}^N \frac{(ax_i + by_i)^2}{a^2 + b^2}. \quad (16)$$

For a minimum the following conditions must be fulfilled

$$\frac{\partial D(a, b)}{\partial a} = 0 \quad \& \quad \frac{\partial D(a, b)}{\partial b} = 0. \quad (17)$$

Using (16) and two conditions (17) we get two conditions for an extreme

$$\begin{aligned} \frac{\partial D(a, b)}{\partial a} &= ab \sum_{i=1}^N (x_i^2 - y_i^2) + (b^2 - a^2) \sum_{i=1}^N x_i y_i = 0 \\ \frac{\partial D(a, b)}{\partial b} &= ab \sum_{i=1}^N (x_i^2 - y_i^2) + (b^2 - a^2) \sum_{i=1}^N x_i y_i = 0. \end{aligned} \quad (18)$$

6 Michal Smolik et al.

It can be seen that both equations above are equivalent and actually we have got just one equation

$$ab \sum_{i=1}^N (x_i^2 - y_i^2) + (b^2 - a^2) \sum_{i=1}^N x_i y_i = 0, \quad (19)$$

using substitutions

$$\xi = \sum_{i=1}^N (x_i^2 - y_i^2) \quad \& \quad \eta = \sum_{i=1}^N x_i y_i, \quad (20)$$

(19) can be rewritten as

$$ab\xi + (b^2 - a^2)\eta = 0, \quad (21)$$

Now, we need to determine values a, b . As we keep the normalization condition for coefficients a, b during the extreme conditions, if

$$\sum_{i=1}^N x_i^2 \geq \sum_{i=1}^N y_i^2 \quad (22)$$

we can select the value a , e.g. $a = 1$, and solve the equation for b or vice versa. This leads to a quadratic equation

$$b\xi + (b^2 - 1)\eta = 0 \quad i.e. \quad \eta b^2 + \xi b - \eta = 0 \quad (23)$$

and therefore

$$b_1 = \frac{-\xi + \sqrt{\xi^2 + 4\eta^2}}{2\eta} \quad i.e. \quad b_2 = \frac{-\xi - \sqrt{\xi^2 + 4\eta^2}}{2\eta}. \quad (24)$$

The minimum distance is given by the b_2 value and

$$b = b_2 \quad i.e. \quad a = 1 \quad (25)$$

and the a, b values are of general values

Now, the computed line $p : ax + by = 0$, which is represented by the vector $\mathbf{p} = [a, b : 0]^T$, passes the origin of the coordinated system is to be "moved" back to the original coordinate system of the original data set using the standard geometric transformation represented by a matrix \mathbf{T} [15], i.e.

$$\begin{bmatrix} a' \\ b' \\ c' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -x_0 & -y_0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} \quad (26)$$

(26) can be rewritten as

$$\mathbf{p}' = \mathbf{T}\mathbf{p}. \quad (27)$$

Now, the line \mathbf{p}' represents the line which optimally fits data in the sense of the TLSE method.

The formula is simple, easy and robust. However, it should be noted, that the proposed method above is not directly extensible to the E^3 case as a line in E^3 cannot be represented by an implicit formula.

5 Experimental Results

In this section we will compare our proposed approach for total least squares approximation with other approaches. The first approach we will compare with, is the method presented in [3]. This method computes total orthogonal least squares approximation like our proposed method. It uses trigonometrical functions to derive the final formula for approximation. However the result of [3] is the same as result of our proposed approach, the derivation of final approximation formula in [3] is much more complicated. Derivation of our proposed approach is much more simpler. Next we will compare with the traditional least squares approximation method. This method minimizes the sum of square differences in y axis, i.e. approximated line formula is $y = f(x)$.

We used 100 uniformly generated points for testing purposes. The visualization of generated input points is in Fig. 2a. The generated input points were rotated by an angle $\varphi = \{0^\circ, 1^\circ, \dots, 45^\circ\}$.

We computed all the three tested least squares approximation methods and visualize the results in Fig. 2 and Fig. 3. It can be seen that the proposed approach for total least squares approximation has the same results as Alciatore's method presented in [3], i.e. the both approximated lines have the same mathematical formula. However the proposed approach is 21% faster than the Alciatore's method.

The results of traditional least squares approximation method differs from the results of the proposed approach. The reason for this is a different criteria for minimization, as the traditional method minimizes the sum of squared distances in y axis and the proposed approach minimizes the sum of squared orthogonal distances. For line that is parallel with the x axis both methods give the same

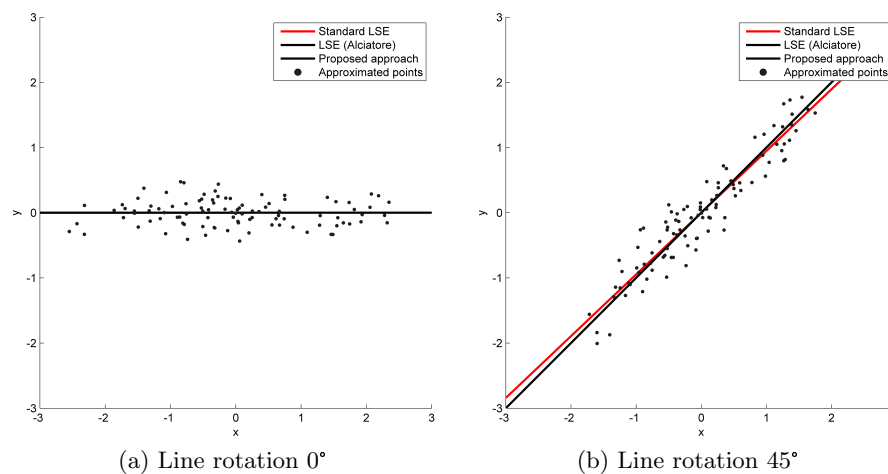


Fig. 2. Results of line approximation for different line rotation angles. Rotation angle is with respect to the x axis.

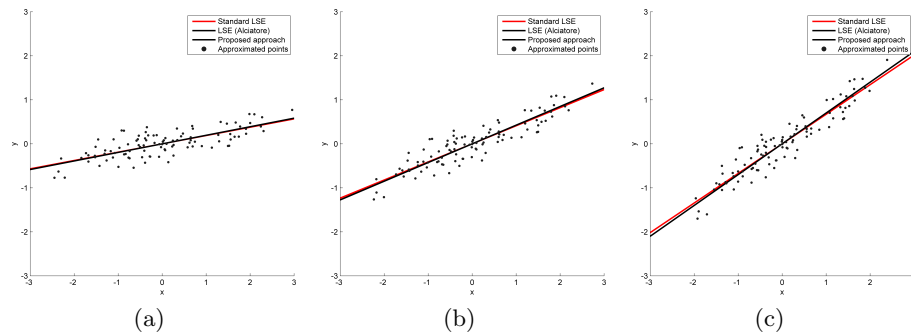


Fig. 3. Results of line approximation for different line rotation angles. Line rotation 11° (a), 23° (b) and 35° (c)

result as the orthogonal distances are identical to the distances in y axis. The more the angle of line with x axis increases, the more differs orthogonal distances from the distances in y axis and thus both approximated lines differs more and more, see Fig. 2 and Fig. 3. The additional comparison of TLSE approximation method with other approximation methods is in [10] and the TLSE method gives the most accurate results.

The more the angle of line with x axis increases, the more differs the standard approximation and the proposed approach. To compute how much the standard approximation differs from our proposed method for total least squares approximation we need to define the distance measurement

$$dist([a, b, c], \mathbf{x}) = \frac{\sum_{i=1}^N |ax_i + by_i + c|}{\sqrt{a^2 + b^2}}, \quad (28)$$

where $[a, b, c]$ are coefficients of approximated line $ax + by + c = 0$ and $\mathbf{x} = \{[x_1, y_1], \dots, [x_N, y_N]\}$ are approximated points. To compute how much the standard LSE approximation differs from our proposed method, which is the correct one, we define the following formula

$$error = \frac{(dist([a, b, c]_{StandardLSE}, \mathbf{x}) - dist([a, b, c]_{ProposedApproach}, \mathbf{x}))}{dist([a, b, c]_{ProposedApproach}, \mathbf{x})} \quad (29)$$

Using (29) we can compute the approximation line distance error, see Fig. 4. It can be seen that for line with angle 0° the approximation error is 0%. The error of the standard LSE increases with increasing the line angle. We performed the error computation for line angle up to 45°, as for angle from 45° to 90° we should use $x = g(y)$ for the standard LSE approximation, otherwise the approximation error will be too high for the standard LSE method ($y = f(x)$).

It should be noted that the TLSE approximation method by Alciatore [3] and the proposed TLSE method give the identical results.

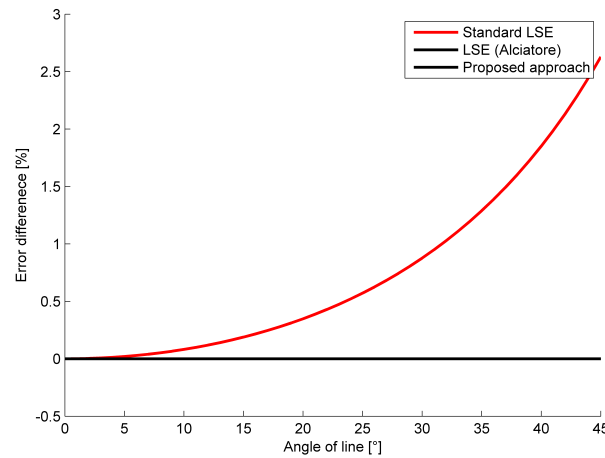


Fig. 4. Approximation line distance error in %. The error is computed using (29).

6 Conclusion

We presented a new approach for total least squares error approximation. The algorithm proved its simplicity and robustness and was compared with a LSE and a TLSE approaches. The proposed TLSE method has similar computational costs as standard LSE method and moreover, the approximation results are more correct.

In the future we plan to extend this approach for implicit plane approximation in E^3 as the extension is not straightforward due to non-linearity.

Acknowledgment

The authors would like to thank their colleagues at the University of West Bohemia, Plzen, for their discussions and suggestions, and anonymous reviewers for their valuable comments and hints provided. The research was supported by projects Czech Science Foundation (GACR) No. 17-05534S and SGS 2016-013.

References

1. S. J. Ahn. *Least squares orthogonal distance fitting of curves and surfaces in space*, volume 3151. Springer Science & Business Media, 2004.
2. S. J. Ahn, W. Rauh, and H.-J. Warnecke. Least-squares orthogonal distances fitting of circle, sphere, ellipse, hyperbola, and parabola. *Pattern Recognition*, 34(12):2283–2303, 2001.
3. D. Alciatore and R. Miranda. Iii.1 - the best least-squares line fit. In A. W. Paeth, editor, *Graphics Gems V*, pages 91 – 97. Academic Press, Boston, 1995.

4. J. A. Cadzow. Signal processing via least squares error modeling. *IEEE ASSP Magazine*, 7(4):12–31, 1990.
5. S. C. Chapra and R. P. Canale. *Numerical methods for engineers*, volume 2. McGraw-Hill New York, 1998.
6. N. Chernov. *Circular and linear regression: Fitting circles and lines by least squares*. CRC Press, 2010.
7. N. Chernov and C. Lesort. Least squares fitting of circles. *Journal of Mathematical Imaging and Vision*, 23(3):239–252, 2005.
8. T. B. A. Duong, J. Tsuchida, and H. Yadohisa. Multivariate multiple orthogonal linear regression. In *International Conference on Intelligent Decision Technologies*, pages 44–53. Springer, 2018.
9. G. H. Golub and C. F. Van Loan. An analysis of the total least squares problem. *SIAM journal on numerical analysis*, 17(6):883–893, 1980.
10. S. Gutta, M. Bhatt, S. K. Kalva, M. Pramanik, and P. K. Yalavarthy. Modeling errors compensation with total least squares for limited data photoacoustic tomography. *IEEE Journal of Selected Topics in Quantum Electronics*, 25(1):1–14, 2019.
11. E. Kreyszig. *Advanced engineering mathematics*. John Wiley & Sons, 2010.
12. P. Lancaster and K. Salkauskas. Surfaces generated by moving least squares methods. *Mathematics of computation*, 37(155):141–158, 1981.
13. S. L. Lee. A note on the total least square fit to coplanar points. Technical report, Tech. Rep. ORNL-TM-12852, Oak Ridge National Laboratory, 1994.
14. I. Rusnak and L. Peled-Eitan. Least squares error criterion based estimator of nonlinear systems. In *Control Conference (ASCC), 2017 11th Asian*, pages 2522–2527. IEEE, 2017.
15. V. Skala. Projective geometry and duality for graphics, games and visualization. In *SIGGRAPH Asia 2012 Courses*, SA '12, pages 10:1–10:47, New York, NY, USA, 2012. ACM.
16. V. Skala. A new formulation for total least square error method in d-dimensional space with mapping to a parametric line. In *AIP Conference Proceedings*, volume 1738, page 480106. AIP Publishing, 2016.
17. V. Skala. Total least square error computation in E2: A new simple, fast and robust algorithm. In *Proceedings of the 33rd Computer Graphics International*, pages 1–4. ACM, 2016.
18. J. Stutz and U. Platt. Numerical analysis and estimation of the statistical error of differential optical absorption spectroscopy measurements with least-squares methods. *Applied Optics*, 35(30):6041–6053, 1996.
19. S. Van Huffel, C.-L. Cheng, N. Mastronardi, C. Paige, and A. Kukush. Total least squares and errors-in-variables modeling, 2007.
20. S. Van Huffel and P. Lemmerling. *Total least squares and errors-in-variables modeling: analysis, algorithms and applications*. Springer Science & Business Media, 2013.
21. S. Van Huffel and J. Vandewalle. *The total least squares problem: computational aspects and analysis*, volume 9. Siam, 1991.
22. P. Xu, J. Liu, and C. Shi. Total least squares adjustment in partial errors-in-variables models: algorithm and statistical analysis. *Journal of Geodesy*, 86(8):661–675, 2012.