Acta Geod Geophys (2017) 52:389–406
DOI 10.1007/s40328-016-0178-0

ORIGINAL STUDY

L_1 norm minimization in partial errors-in-variables model

Jun Zhao[1] · Qingming Gui[2] · Feixiao Guo[1,3]

Received: 21 April 2016 / Accepted: 6 July 2016 / Published online: 28 July 2016

Abstract The weighted total least-squares (WTLS) estimate is sensitive to outliers and will be strongly disturbed if there are outliers in the observations and coefficient matrix of the partial errors-in-variables (EIV) model. The L_1 norm minimization method is a robust technique to resist the bad effect of outliers. Therefore, the computational formula of the L_1 norm minimization for the partial EIV model is developed by employing the linear programming theory. However, the closed-form solution cannot be directly obtained since there are some unknown parameters in constrained condition equation of the presented optimization problem. The iterated procedure is recommended and the proper condition for stopping iteration is suggested. At the same time, by treating the partial EIV model as the special case of the non-linear Gauss–Helmert (G–H) model, another iterated method for the L_1 norm minimization problem is also developed. At last, two simulated examples and a real data of 2D affine transformation are conducted. It is illustrated that the results derived by the proposed L_1 norm minimization methods are more accurate than those by the WTLS method while the observations and elements of the coefficient matrix are contaminated with outliers. And the two methods for the L_1 norm minimization problem are identical in the sense of robustness. By comparing with the data-snooping method, the L_1 norm minimization method may be more reliable for detecting multiple outliers due to masking. But it leads to great computation burden.

Keywords Partial errors-in-variables model · Weighted total least-squares · Outlier · L_1 norm minimization · Linear programming · 2D affine transformation

✉ Jun Zhao
zhaojun4368@163.com

[1] Institute of Surveying and Mapping, Information Engineering University, Zhengzhou 450001, China

[2] Institute of Science, Information Engineering University, Zhengzhou 450001, China

[3] Xi'an Research Institute of Surveying and Mapping, Xi'an 710054, China

1 Introduction

In recent years, the total least-squares (TLS) or weighted TLS (WTLS) as a method of parameter estimation for the errors-in-variables (EIV) model has been researched intensively in geodetic field (Schaffrin and Wieser 2008; Shen et al. 2011; Amiri-Simkooei and Jazaeri 2012; Mahboub 2012; Fang 2013; Li et al. 2013; Jazaeri et al. 2014). Unfortunately, like LS estimate, the WTLS estimate is also very vulnerable to outliers in data (Schaffrin and Uzun 2011; Amiri-Simkooei and Jazaeri 2013), and even a single outlying observation can result in an entirely wrong conclusion. Therefore, a new method to adapt the problem may be an important issue.

Robust method and outlier detection are two essential ways to deal with outliers. There are many publications for Gauss-Markov (G–M) model from researchers and scholars in geodesy. One can refer to Baarda (1968), Pope (1976), Hekimoglu (1997, 1999), Gui and Liu (1999), Gui et al. (2005, 2007), Guo et al. (2007, 2010), Yang (1999), Yang et al. (2002), Xu (1989, 1993, 2005), Baselga (2007), Koch (2013), Yang et al. (2013). Although these methods may, in principle, be applied to the EIV models, bearing in mind that these models are special in structure, more efficient methods may be highly desirable.

To overcome those obstacles, Schaffrin and Uzun (2011) generalized the mean-shift method to adapt to the EIV models for detecting a single outlier located in the observations or coefficient matrix, and a test statistic following F distribution was constructed. Based on the WTLS method (Amiri-Simkooei and Jazaeri 2012), Amiri-Simkooei and Jazaeri (2013) applied the data-snooping procedure to identify outlier. In view of the masking and smearing (Hadi and Imon 2009; Gui et al. 2011), above-mentioned two approaches may be unreliable for multiple outliers in the EIV model. Although the robust methods based on M-estimation for the EIV model have been partly investigated in statistical literature (Brown 1982; Zamar 1989). Mahboub et al. (2013) pointed out that those methods would only be applied to linear regression. For this reason, the iteratively reweighted total least-squares (IRTLS) (Mahboub et al. 2013) as a robust method was proposed by making use of the WTLS method (Mahboub 2012), and an improved weight function was introduced. The robustness of resisting outlier was proved to be superior to the traditional strategies. In addition, the IRTLS method to the linearized Gauss–Helmert (G–H) model was also developed (Tao et al. 2014; Lu et al. 2014). These methods were applied in GPS height fitting and three-dimensional similarity coordinate transformation so that more reliable estimates of the unknown parameters were obtained.

Actually, the WTLS method requires that the weighted sum of squared residuals should be minimized as a L_2 norm minimization method. In contrast to the L_2 norm minimization method, the L_1 norm minimization method which has been thoroughly discussed in G–M model (Marshall and Bethel 1996; Amiri-Simkooei 2003; Yetkin and Inal 2011), is more immune to outlier as a robust technique. The L_1 norm minimization problem in the EIV model solved by a trust region method (Watson and Yiu 1991) was also studied preliminary. Sincerely, this method is not easy to be understandable and can only be effective for the independent observations with equal weight. Nevertheless, Xu (2005) proved that the L_1 norm minimization method would not be robust under certain conditions of weights. Additionally, it cannot deal with the case where there are the fixed elements in the coefficient matrix and the repeated random elements in different locations. To circumvent these difficulties, a new formula of the L_1 norm minimization problem for the partial EIV model is proposed by taking advantage of the linear programming theory in this paper.

The remaining of this paper is organized as follows. In Sect. 2, known as a generalized EIV model, the partial EIV model (Xu et al. 2012) is briefly introduced. In Sect. 3, two formulae of the L_1 norm minimization problem for the partial EIV model based on the linear programming theory are presented and the iterated algorithms are proposed. In Sect. 4, the simulated examples and the real data of a 2D affine transformation are utilized to demonstrate the availability of the proposed methods. Finally, some conclusions and further discussions are given.

2 Partial EIV model and WTLS method

Firstly, let us consider the EIV model (Schaffrin and Wieser 2008; Shen et al. 2011) as follows:

$$\boldsymbol{L} = (\boldsymbol{A} - \boldsymbol{E}_A)\boldsymbol{X} + \boldsymbol{\Delta} \tag{1}$$

where $\boldsymbol{L}$ is the $n \times 1$ vector of observations, $\boldsymbol{A}$ is the $n \times t$ coefficient matrix with full column rank affected by the random errors $\boldsymbol{E}_A$, $\boldsymbol{X}$ is the $t \times 1$ vector of unknown parameters and $\boldsymbol{\Delta}$ is the $n \times 1$ vector of random errors. The stochastic character is described as

$$\begin{bmatrix} \boldsymbol{\Delta} \\ \boldsymbol{e} = vec(\boldsymbol{E}_A) \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} \boldsymbol{Q}_\Delta & 0 \\ 0 & \boldsymbol{Q}_e \end{bmatrix} \right) \tag{2}$$

where "*vec*" denotes an operator that transforms a matrix into a vector by stacking the columns of the matrix one underneath the other, σ^2 is the unknown variance component, $\boldsymbol{Q}_\Delta$ and $\boldsymbol{Q}_e$ are the $n \times n$ and $nt \times nt$ known cofactor matrices of $\boldsymbol{\Delta}$ and $\boldsymbol{e}$, respectively.

As a matter of fact, not all of the elements of the coefficient matrix $\boldsymbol{A}$ are random and there are some repeated random elements in different locations such as the coordinate transformation. Taking above into consideration and to eliminate the correlations of the repeated random elements in the coefficient matrix, the partial EIV model (Xu et al. 2012) is more appropriate to be used, and it is represented as follows:

$$\begin{cases} \boldsymbol{L} = \left(\boldsymbol{X}^T \otimes \boldsymbol{I}_n\right)(\boldsymbol{h} + \boldsymbol{B}\bar{\boldsymbol{a}}) + \boldsymbol{\Delta} \\ \boldsymbol{a} = \bar{\boldsymbol{a}} + \boldsymbol{\gamma} \end{cases} \tag{3}$$

where $\boldsymbol{h}$ is the $nt \times 1$ vector of deterministic constants constituted of zero and the non-random elements of $\boldsymbol{A}$, $\boldsymbol{I}_n$ is the $n \times n$ identity matrix, $\boldsymbol{B}$ is the $nt \times s$ structured matrix, s is the number of different random elements of $\boldsymbol{A}$ which are stored in the vector $\boldsymbol{a}$, $\bar{\boldsymbol{a}}$ is the true value of $\boldsymbol{a}$, the cofactor matrix of $\boldsymbol{a}$ is denoted by $\boldsymbol{Q}_\gamma$. It is assumed that $\boldsymbol{\Delta}$ is statistically independent with $\boldsymbol{\gamma}$.

The estimates of $\bar{\boldsymbol{a}}$ and $\boldsymbol{X}$ can be obtained by solving the following minimization problem (Xu et al. 2012):

$$\begin{aligned} \min : \Omega_1 \\ &= (\bar{\boldsymbol{a}} - \boldsymbol{a})^T \boldsymbol{Q}_\gamma^{-1} (\bar{\boldsymbol{a}} - \boldsymbol{a}) \\ &\quad + \left[\left(\boldsymbol{X}^T \otimes \boldsymbol{I}_n\right) \times (\boldsymbol{h} + \boldsymbol{B}\bar{\boldsymbol{a}}) - \boldsymbol{L}\right]^T \boldsymbol{Q}_\Delta^{-1} \left[\left(\boldsymbol{X}^T \otimes \boldsymbol{I}_n\right) \times (\boldsymbol{h} + \boldsymbol{B}\bar{\boldsymbol{a}}) - \boldsymbol{L}\right] \end{aligned} \tag{4}$$

which is equivalent to

$$\min : \Omega_1 = \gamma^T \boldsymbol{Q}_\gamma^{-1} \gamma + \boldsymbol{\Delta}^T \boldsymbol{Q}_\Delta^{-1} \boldsymbol{\Delta} \tag{5}$$

Obviously, the minimization problem (5) can be regarded as a L_2 norm minimization one. The WTLS solution is achieved by conducting first partial derivatives of Ω_1 with respect to the variables $\bar{\boldsymbol{a}}$ and $\boldsymbol{X}$, and letting them equal to zero.

3 Formulation of L_1 norm minimization in partial EIV model

When there are outliers in $\boldsymbol{L}$ and $\boldsymbol{a}$ simultaneously, the objective function Ω_1 will be larger than in the case without outliers. A method to control the bad effect of outliers for Ω_1 is proposed, which minimizes the weighted sum of the absolute values of errors in the observations and the different random elements of the coefficient matrix. Namely, the L_1 minimization problem of (3) is regarded as

$$\min: \Omega_2 = \boldsymbol{p}_\gamma^T |\gamma| + \boldsymbol{p}_\Delta^T |\boldsymbol{\Delta}| \tag{6}$$

where $\boldsymbol{p}_\gamma$ and $\boldsymbol{p}_\Delta$ are the $s \times 1$ and $n \times 1$ vectors consisting of the diagonal elements of the weight matrix $\boldsymbol{Q}_\gamma^{-1}$ and $\boldsymbol{Q}_\Delta^{-1}$, respectively, $|\,|$ is a mathematical operator that derives the absolute value of variable. That is to say, for a vector, one will obtain a new vector whose each component is the absolute value of the original component.

Unfortunately, the above optimization solution cannot be realized by taking partial derivative of Ω_2 with respect to variables $\bar{\boldsymbol{a}}$ and $\boldsymbol{X}$ and equating these derivatives to zero because there are absolute notations. An efficient method to solve this optimization problem is to bring in the slack variables.

Firstly, in order to remove the correlations between observations, the Cholesky factorizations

$$\boldsymbol{Q}_\Delta = \boldsymbol{W}_\Delta^T \boldsymbol{W}_\Delta, \quad \boldsymbol{Q}_\gamma = \boldsymbol{W}_\gamma^T \boldsymbol{W}_\gamma \tag{7}$$

are conducted and after some transformations, the model (3) is expressed as follows

$$\begin{cases} \tilde{\boldsymbol{L}} = \left(\boldsymbol{W}_\Delta^T\right)^{-1} \left(\boldsymbol{X}^T \otimes \boldsymbol{I}_n\right) \left(\boldsymbol{h} + \boldsymbol{B} \boldsymbol{W}_\gamma^T \tilde{\bar{\boldsymbol{a}}}\right) + \tilde{\boldsymbol{\Delta}} \\ \tilde{\boldsymbol{a}} = \tilde{\bar{\boldsymbol{a}}} + \tilde{\gamma} \end{cases} \tag{8}$$

where

$$\tilde{\boldsymbol{L}} = \left(\boldsymbol{W}_\Delta^T\right)^{-1} \boldsymbol{L},\ \tilde{\boldsymbol{\Delta}} = \left(\boldsymbol{W}_\Delta^T\right)^{-1} \boldsymbol{\Delta},\ \tilde{\boldsymbol{a}} = \left(\boldsymbol{W}_\gamma^T\right)^{-1} \boldsymbol{a},\ \tilde{\bar{\boldsymbol{a}}} = \left(\boldsymbol{W}_\gamma^T\right)^{-1} \bar{\boldsymbol{a}},\ \tilde{\gamma} = \left(\boldsymbol{W}_\gamma^T\right)^{-1} \gamma.$$

However, as mentioned in Xu (1989), outliers will be spread to all the observations no matter whether the original observations contain outliers or not.

To eliminate the absolute value notations, some slack vectors $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ for $\boldsymbol{X}$, $\boldsymbol{u}$ and $\boldsymbol{w}$ for $\tilde{\gamma}$, and $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ for $\tilde{\boldsymbol{\Delta}}$ are introduced. If one uses these slack variables to replace the corresponding variables, that is,

$$\begin{aligned} \tilde{\gamma} &= \boldsymbol{u} - \boldsymbol{w}, \quad \boldsymbol{u}, \boldsymbol{w} \geq 0 \\ \tilde{\boldsymbol{\Delta}} &= \boldsymbol{\alpha} - \boldsymbol{\beta}, \quad \boldsymbol{\alpha}, \boldsymbol{\beta} \geq 0 \\ \boldsymbol{X} &= \boldsymbol{\eta} - \boldsymbol{\xi}, \quad \boldsymbol{\eta}, \boldsymbol{\xi} \geq 0 \end{aligned} \tag{9}$$

the nonnegativity of parameters will be satisfied. Then, the L_1 minimization problem (6) can be rewritten as

$$\min\colon \Omega_3 = \begin{bmatrix} 0^T & 0^T & \bar{\boldsymbol{h}}^T & \bar{\boldsymbol{h}}^T & \bar{\boldsymbol{h}}^T & \bar{\boldsymbol{h}}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta} \\ \boldsymbol{\xi} \\ \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{w} \\ \boldsymbol{u} \end{bmatrix} \tag{10}$$

subject to

$$\begin{bmatrix} \tilde{\boldsymbol{A}} & -\tilde{\boldsymbol{A}} & \boldsymbol{I}_n & -\boldsymbol{I}_n & \left(\boldsymbol{W}_\Delta^T\right)^{-1}\left(\boldsymbol{X}^T \otimes \boldsymbol{I}_n\right)\boldsymbol{B}\boldsymbol{W}_\gamma^T & -\left(\boldsymbol{W}_\Delta^T\right)^{-1}\left(\boldsymbol{X}^T \otimes \boldsymbol{I}_n\right)\boldsymbol{B}\boldsymbol{W}_\gamma^T \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta} \\ \boldsymbol{\xi} \\ \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{w} \\ \boldsymbol{u} \end{bmatrix} = \tilde{\boldsymbol{L}} \tag{11}$$

where $\boldsymbol{\eta}, \boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{u} \geq 0, \tilde{\boldsymbol{A}} = \left(\boldsymbol{W}_\Delta^T\right)^{-1}\boldsymbol{A}, \bar{\boldsymbol{h}}^T = [1, 1, \ldots, 1]$.

Let us define

$$\bar{\boldsymbol{X}} = \begin{bmatrix} \boldsymbol{\eta} \\ \boldsymbol{\xi} \\ \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{w} \\ \boldsymbol{u} \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \bar{\boldsymbol{h}} \\ \bar{\boldsymbol{h}} \\ \bar{\boldsymbol{h}} \\ \bar{\boldsymbol{h}} \end{bmatrix} \tag{12}$$

And

$$\bar{\boldsymbol{A}} = \begin{bmatrix} \tilde{\boldsymbol{A}} & -\tilde{\boldsymbol{A}} & \boldsymbol{I}_n & -\boldsymbol{I}_n & \left(\boldsymbol{W}_\Delta^T\right)^{-1}\left(\boldsymbol{X}^T \otimes \boldsymbol{I}_n\right)\boldsymbol{B}\boldsymbol{W}_\gamma^T & -\left(\boldsymbol{W}_\Delta^T\right)^{-1}\left(\boldsymbol{X}^T \otimes \boldsymbol{I}_n\right)\boldsymbol{B}\boldsymbol{W}_\gamma^T \end{bmatrix} \tag{13}$$

Then, the L_1 norm minimization problem with (10) and (11) can be taken compactly as

$$\min\colon \Omega_4 = \boldsymbol{C}^T\bar{\boldsymbol{X}} \tag{14}$$

subject to

$$\bar{\boldsymbol{A}}\bar{\boldsymbol{X}} = \tilde{\boldsymbol{L}}, \quad \bar{\boldsymbol{X}} \geq \boldsymbol{0} \tag{15}$$

As we know, the unknown parameter $\boldsymbol{X}$ is located in $\bar{\boldsymbol{A}}$, which limits the usefulness of the linear programming theory (Vanderbei 2014). Therefore, an iterated method named as algorithm 1 is proposed to solve the optimization problem with (14) and (15). The implemented procedure of algorithm 1 is summarized as follows:

Step 1 Give the initial value

$$\boldsymbol{X}^{(0)} = \left(\boldsymbol{A}^T\boldsymbol{Q}_\Delta^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{Q}_\Delta^{-1}\boldsymbol{L};$$

Step 2 Compute $\bar{\boldsymbol{A}}$ by replacing $\boldsymbol{X}$ with $\boldsymbol{X}^{(0)}$;

Step 3 For any i, compute $\hat{\boldsymbol{X}}^{(i)}$ by solving the optimization problem with (14) and (15).

Step 4 The implemented procedure will be stopped when the

$$\left\|\hat{X}^{(i)}-\hat{X}^{(i-1)}\right\| \text{ or } \left|\left\|\hat{X}^{(i)}-\hat{X}^{(i-1)}\right\|-\left\|\hat{X}^{(i-1)}-\hat{X}^{(i-2)}\right\|\right|$$

is less than 0.00001; otherwise, one returns to Step 2 by taking $\hat{X}^{(i)}$ as the initial value for next iteration, where $\hat{X}^{(i)}$ is the ith iterated solution.

As we see, the above algorithm requires iterations so that the accurate solution cannot be guaranteed to be derived all the time, which is reported in next section with the numerical results.

As a matter of fact, the partial EIV model is a non-linear model. An efficient method for obtaining the WTLS solution is to replace the original model by a sequence of linearized G–H model through the LS adjustment. Therefore, making a transformation

$$\tilde{L}=\left(W_{\Delta}^{T}\right)^{-1}\left(X^{T}\otimes I_{n}\right)\left(h+BW_{\gamma}^{T}(\tilde{a}-\tilde{\gamma})\right)+\tilde{\Delta} \tag{16}$$

for model (8), one has

$$\tilde{L}=\tilde{A}X-\left(W_{\Delta}^{T}\right)^{-1}\left(X^{T}\otimes I_{n}\right)BW_{\gamma}^{T}\tilde{\gamma}+\tilde{\Delta} \tag{17}$$

If the approximate value $X^{(0)}$ is given,the partial EIV model is transformed to a linearized G–H model as follows:

$$\tilde{L}-\tilde{A}X^{(0)}=\tilde{A}x+\begin{bmatrix} I_n & -\left(W_{\Delta}^{T}\right)^{-1}\left(\left(X^{(0)}\right)^{T}\otimes I_n\right)BW_{\gamma}^{T}\end{bmatrix}\begin{bmatrix}\tilde{\Delta}\\ \tilde{\gamma}\end{bmatrix} \tag{18}$$

By introducing the above slack variables (9), one can form an optimization problem as follows:

$$\min\colon \Omega_5=\begin{bmatrix} 0^T & 0^T & \bar{h}^T & \bar{h}^T & \bar{h}^T & \bar{h}^T \end{bmatrix}\begin{bmatrix}\eta\\ \xi\\ \alpha\\ \beta\\ w\\ u\end{bmatrix} \tag{19}$$

subject to

$$\begin{bmatrix}\tilde{A} & -\tilde{A} & I_n & -I_n & \left(W_{\Delta}^{T}\right)^{-1}\left(\left(X^{(0)}\right)^{T}\otimes I_n\right)BW_{\gamma}^{T} & -\left(W_{\Delta}^{T}\right)^{-1}\left(\left(X^{(0)}\right)^{T}\otimes I_n\right)BW_{\gamma}^{T}\end{bmatrix}\begin{bmatrix}\eta\\ \xi\\ \alpha\\ \beta\\ w\\ u\end{bmatrix}=\tilde{L}-\tilde{A}X^{(0)} \tag{20}$$

Here $x=\eta-\xi$. Certainly, one should also take iterations to solve the minimization problem with (19) and (20) for estimating unknown parameters. Above method is taken as Algorithm 2.

With the basic linear programming theory (Vanderbei 2014), the computation complexity of two proposed algorithms in this paper is proportional to $\mathrm{O}(2(n+t+s)^2 n)$. But the computation complexity of structured EIV approaches is proportional to $\mathrm{O}((n+t)^3)$ (Abatzoglou et al. 1991). As a result, the two L_1 norm minimization methods require more computation burden than the WTLS method because $2n$ is greater than $n+t$. In addition, it should be pointed out that the two optimization problems are valid in the absence of correlation between observations and coefficient matrix.

4 Numerical results and discussions

The linear regression and 2D affine transformation examples are chosen to demonstrate the effectiveness of the L_1 norm minimization method in presence of outliers.

4.1 Example 1: linear regression

Consider a simple linear regression model as follows:

$$y_i - e_{y_i} = \xi_2 \cdot (x_i - e_{x_i}) + \xi_1 \tag{21}$$

where y_i and x_i are the observations containing the errors, ξ_1 and ξ_2 represent intercept and slope, respectively. The major data are taken from Schaffrin and Wieser (2008). The random errors with nominal standard deviation 0.01 are added to y_i and x_i based on the G–M model and the reference values for producing the simulated observations is

$$[\xi_1, \xi_2] = [5.4799 - 0.4805] \tag{22}$$

Considering the case with single outlier, we add the gross error of size 0.1 to the x component of point 5. In this example, the criterion of stopping iteration for the WTLS method and L_1 norm minimization method is that the Euclidean distances between the two estimates of the unknown parameters for consecutive iterated steps are less than 0.00001 or the number of iterations is beyond 200. The unknown parameters are estimated by the WTLS method and L_1 norm minimization method with two algorithms, respectively. Then the Euclidean distances between the estimates of the transformation parameters and the reference values are computed, which is displayed in Table 1. The results show that the unknown parameters obtained by the L_1 norm minimization method are closer to the reference values than those by the WTLS method, which means that the L_1 norm

Table 1 Unknown parameters estimated by the WTLS method (without outlier and with single outlier), L_1 norm minimization method (with single outlier) and the WTLS method after deleting the outlier with the data-snooping method and L_1 norm minimization method

Parameter	WTLS (without outlier)	WTLS (with outlier)	L_1 norm (algorithm 1)	L_1 norm (algorithm 2)	WTLS (After deleting outlier by data-snooping)	WTLS (After deleting outlier by L_1 norm)
ξ_1	5.4820	5.5126	5.4869	5.4869	5.4830	5.4830
ξ_2	−0.4811	−0.4861	−0.4817	−0.4817	−0.4813	−0.4813
$\Vert\hat{\boldsymbol{X}} - \boldsymbol{X}_{ref}\Vert$	0.0022	0.0332	0.0071	0.0071	0.0032	0.0032

minimization method is not sensitive to outliers, and two algorithms for the L_1 norm minimization problem are identical in terms of robustness. The goodness-of-fit test is a global test to evaluate whether the underlying method may be distorted by outliers (Amiri-Simkooei and Jazaeri 2013). The global test statistic is computed as $\chi^2 = 320.58$, but the threshold value is $\chi^2_{(0.975,8)} = 17.53$, which shows that at least one outlier exists in the observations or coefficient matrix. To keep the same with the data-snooping method proposed by Amiri-Simkooei and Jazaeri 2013) and make the comparative analysis for detecting outliers, the residuals of the L_1 norm minimization method for detecting outlier is adopted as

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{I}_n & -\left(\boldsymbol{W}_\Delta^T\right)^{-1}\left(\left(\boldsymbol{X}^{(0)}\right)^T \otimes \boldsymbol{I}_n\right)\boldsymbol{B}\boldsymbol{W}_\gamma^T \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\Delta}} \\ \tilde{\boldsymbol{\gamma}} \end{bmatrix}. \tag{23}$$

From the w test statistics and the residuals displayed in Tables 2 and 3, one can clearly judge that one outlier is located in point 5. After deleting point 5, the new residuals and w test statistics are derived. At the same time, the corresponding global test statistic is 15.22, which now is smaller than the threshold value 16.01, and there are no larger residuals for the L_1 norm minimization method. Therefore, the only one outlier is identified correctly for the simulated case. After deleting all suspicious outlying observations, the unknown parameters are estimated again by the WTLS method, which is given in rank 6 and 7 of Table 1. The Euclidean distance between the estimates after deleting the outlier and the reference values is just 0.003, which is superior to the result given by the WTLS method while there is an outlier in coefficient matrix.

To give an accurate evaluation for the proposed method with multiple outliers, two new gross errors of size 0.1 are added to the y components in point 1 and point 7 simultaneously. Table 4 presents the unknown parameters estimated by the L_1 norm minimization method and WTLS method, respectively. Those results show that the L_1 norm minimization method is more reliable than the WTLS method in terms of resisting outliers. And the proposed algorithms are convergent for only 2 iterations. To demonstrate the superiority of the proposed method for detecting multiples outliers, a comparison between the proposed method and the data-snooping method proposed by Amiri-Simkooei and Jazaeri (2013) is performed. For the data-snooping method, the global test is rejected because the statistic $\chi^2 = 3352.17$ is greater than the threshold value $\chi^2_{(0.975,8)} = 17.53$, which indicates

Table 2 Residuals and w-test statistics of the data-snooping procedure proposed by Amiri-Simkooei and Jazaeri (2013) (with single outlier)

Eq. no.	With outlier		Deleting one outlier	
	$\hat{e}$	w	$\hat{e}$	w
1	−0.0029	−0.0487	0.0266	1.9231
2	−0.0174	−0.3960	0.0076	0.7505
3	−0.0178	−0.6150	0.0030	0.4611
4	−0.0208	−1.0402	−0.0039	−0.8687
5	0.0335	2.7603	–	–
6	−0.0122	−0.9009	−0.0039	−1.2708
7	−0.0049	−0.7239	−0.0005	−0.3453
8	0.0009	0.1129	0.0009	0.5115
9	−0.0049	0.3214	0.0053	1.0247
10	0.0081	0.2779	0.0018	0.2770

Table 3 Residuals obtained by the WTLS method (without outlier, with single outlier and after deleting outlier) and L_1 norm minimization method (with single outlier and after deleting outlier)

Eq. no.	With outlier		After deleting outlier	
	WTLS $\hat{e}$	L_1 norm V	WTLS $\hat{e}$	L_1 norm V
1	−0.0029	0.0227	0.0266	0.0328
2	−0.0176	0.0056	0.0076	0.0172
3	−0.0178	0	0.0030	0.0145
4	−0.0208	−0.0183	−0.0038	−0.0014
5	0.0335	0.1985	–	–
6	−0.0122	−0.0253	−0.0039	−0.0112
7	−0.0049	−0.0158	−0.0005	0
8	0.0009	0	0.0009	0.0039
9	−0.0049	0.0459	0.0053	0.0443
10	0.0081	0.0350	0.0018	0

Table 4 Transformation parameters estimated by the WTLS method (without outlier, with multiple outliers and after deleting the outlier) and L_1 norm minimization method (with multiple outliers and after deleting the outlier)

Parameter	WTLS (without outlier)	WTLS (with outlier)	L_1 norm (algorithm 1)	L_1 norm (algorithm 2)	WTLS (After deleting outlier by data-snooping)	WTLS (After deleting outlier by L_1 norm)
ξ_1	5.4820	5.5328	5.4914	5.4914	5.5223	5.4796
ξ_2	−0.4811	−0.4830	−0.4821	−0.4821	−0.4874	−0.4806
$\lVert\hat{X} - X_{ref}\rVert$	0.0022	0.0530	0.0116	0.0116	0.0430	0.0003

that the observations are contaminated with outliers. Table 5 presents the residuals and w-test statistics obtained by the data-snooping procedure. From rank 3 and 4 of Table 5, the maximum of the absolute values of the w-test statistics is 2.66 while the threshold value is 2.31. As a result, the point 7 is considered as an outlying one under the criterion for identifying outlier (Amiri-Simkooei and Jazaeri 2013). For the case with multiple outliers, the data-snooping procedure needs to be implemented step by step. In order to detect next outlier, the point 7 should be deleted. But the global test statistic fulfills the condition

$$\chi^2 = 383.32 > \chi^2_{(0.975,7)} = 16.01 \tag{24}$$

after deleting the point 7, which indicates that there are additional outliers in the observations. Then the new residuals and w-test statistics are obtained, which is displayed in rank 5 and 6 of Table 5. However, the maximum value of the absolute values of the w-test statistics is 2.02 which is smaller than the threshold value 2.36. Obviously, the masking is emerged in accordance with the set simulated case. Therefore, we can make a conclusion that the data-snooping procedure is not reliable for detecting multiple outliers.

Following above discussions, we will employ the L_1 norm minimization method to detect multiple outliers. The residuals for the purpose of detecting outliers are presented in Table 6. Because the global test is not accepted, the corresponding point 7 is judged as the observations contaminated with outliers on the base of the residuals. After deleting the point 7, the global test is still rejected due to the statistic

Table 5 Residuals and w-test statistics of the data-snooping procedure proposed by Amiri-Simkooei and Jazaeri (2013) (with multiple outliers)

Eq. no.	With outlier		Deleting one outlier	
	$\hat{e}$	w	$\hat{e}$	w
1	0.0768	0.3930	0.0873	1.2380
2	−0.0406	−0.2825	−0.0262	−0.5036
3	−0.0437	−0.4645	−0.0252	−0.7459
4	−0.0491	−0.7569	−0.0272	−1.1703
5	0.0028	0.0704	0.0282	2.0402
6	−0.0460	−1.0516	−0.0163	−1.0629
7	0.0588	2.6616	–	–
8	−0.0382	−1.5423	−0.0010	−0.1384
9	−0.0331	−0.4587	0.0059	0.2269
10	−0.0350	−0.3755	0.0078	0.2338

Table 6 Residuals obtained by the L_1 norm minimization method (with multiple outliers)

Eq. no.	With outlier V	After deleting one outlier V	After deleting two outliers V	After deleting three outliers V
1	0.118	0.1227	0.1227	–
2	0	0.0056	0.0056	0.0185
3	−0.0076	0	0	0.0159
4	−0.0282	−0.0183	−0.0183	0
5	0.1844	0.1985	–	–
6	−0.0376	−0.0252	−0.0253	−0.0112
7	0.8005	–	–	–
8	−0.0174	0	0	0
9	0.0267	0.0459	0.0459	0.0384
10	0	0.0350	0.0350	−0.0187

$$\chi^2 = 383.32 > \chi^2_{(0.975,7)} = 16.01. \tag{25}$$

Therefore, the outlier detection should be sustained for next steps. The new results from Table 6 show that the residuals of point 5 are obviously greater than others. The same with above, point 5 is considered as an outlier and should be removed. After that, the global statistics $\chi^2 = 155.39$ is smaller than the threshold value $\chi^2_{(0.975,6)} = 14.45$, which implies that there are contaminated observations with outliers. By contrasting with the residuals, the point 1 is verified as the outlier. After deleting point 1, the global test is accepted because

$$\chi^2 = 7.06 < \chi^2_{(0.975,5)} = 12.83. \tag{26}$$

According to above analysis, all outliers are accurately identified. After deleting all outliers suggested by the data-snooping procedure and the L_1 norm minimization method, the unknown parameters are estimated again by the WTLS method, which is presented in rank 6 and 7 of Table 4. The Euclidean distance between the estimates and reference values are 0.043 and 0.0003 for the data-snooping method and the L_1 norm minimization method, respectively, which demonstrate that the L_1 norm minimization method is superior to the data-snooping method in case of the masking with multiple outliers.

4.2 Example 2: 2D affine transformation

The mathematical model of the 2D affine transformation can be expressed as

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} x_s & y_s & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_s & y_s & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \\ c_1 \\ a_2 \\ b_2 \\ c_2 \end{bmatrix} \tag{27}$$

where (x_s, y_s) and (x_t, y_t) are the coordinates of the same point in the start system and target system, respectively; a_i, b_i and $c_i (i = 1, 2)$ are the transformation parameters to be estimated. The data are taken from Mahboub et al. (2013). Suppose that the reference values of the transformation parameters are

$$[a_1 \quad b_1 \quad c_1 \quad a_2 \quad b_2 \quad c_2] = [0.9 \quad -0.8 \quad 1 \quad 0.6 \quad 0.7 \quad 5]. \tag{28}$$

The new coordinates without any errors in the target system are generated by the coordinates of points in the start system with Eq. (27). To verify the effectiveness of the proposed method, the Monte Carlo simulation has been performed. The Gaussian noises whose variance–covariance matrixes are $\boldsymbol{Q}_{start}$ and $\boldsymbol{Q}_{target}$ are introduced to the error-free coordinates in the start system and target system, respectively,

$$\boldsymbol{Q}_{start} = \boldsymbol{I}_2 \otimes \boldsymbol{Q}, \quad \boldsymbol{Q}_{target} = \boldsymbol{I}_2 \otimes \boldsymbol{Q}_T \tag{29}$$

$$\begin{aligned} \boldsymbol{Q}_S &= 0.005 Diag(1, 2, 3, 1, 5, 4, 2, 7, 2, 1, 8, 3, 6), \boldsymbol{Q}_T \\ &= 0.005 Diag(1, 3, 6, 1, 1, 8, 4, 3, 6, 5, 4, 5, 2). \end{aligned}$$

And the gross error of magnitude of 2 is added to the x_s component of point 4 in the start system. The outlier with size 2 is greater than some observations and simulated random noises, which may influence the robustness of the L_1 norm minimization method (Xu 2005).

In order to analyse the convergence of the proposed algorithm, the simulation is implemented for 500 replications. Firstly, the Euclidean distances between the estimated transformation parameters and reference values are computed, which is plotted in Fig. 1. As expected, the WTLS method without outliers achieves the best estimate of the transformation parameter among three schemes. The proposed L_1 norm minimization method generally produces more accurate and reliable transformation parameter than the WTLS method for 494 of 500 simulations in total. For the several invalid resutls, we abandon adding the outlier, and in this case the L_1 norm minimization algorithm is converged after several iterations. The reason is that there are some large outliers in the coefficient matrix so that the L_1 norm minimization method is not efficient in the case of robustness (See Xu 2005). Some statistical results are presented in Tables 7 and 8. As for the case with multiple outliers, the 4 gross errors of size 2 are put simultaneously in both components of point 4 in the start system and both components of point 7 in the target system. The

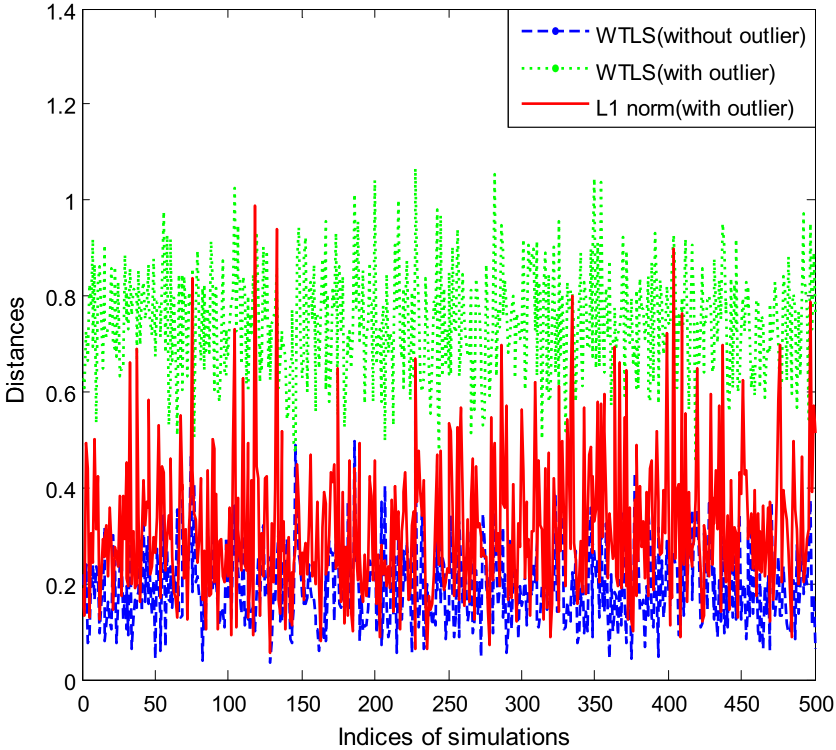

Fig. 1 Euclid distances between transformation parameters estimated by different methods and reference value for 500 replications with single outlier

Table 7 Statistical results of the Euclidean distances between the estimate of the transformation parameters estimated by the different methods and reference values for 500 replications with single outlier

Distance	WTLS (without outlier)		WTLS (with outlier)		L_1 norm (with outlier)	
	Mean	Max	Mean	Max	Mean	Max
$\|\|\hat{\boldsymbol{X}} - \boldsymbol{X}_{real}\|\|$	0.1941	0.5389	0.7462	1.0682	0.3150	0.9897

Euclidean distances between the estimated transformation parameters and the reference values are displayed in Fig. 2. As expected, the WTLS estimates are disturbed more seriously when the observations are contaminated with multiple outliers than the case with single outlier. And the transformation parameters estimated by the L_1 norm minimization method outperform those by the WTLS method. The statistical results in Tables 9 and 10 further make clear that the proposed L_1 norm minimization algorithms are effective and feasible to resist the bad effect of outlier.

Table 8 Euclidean distance between the estimate of the transformation parameters estimated by the different methods and reference values for 500 replications with single outlier

Differences	WTLS (without outlier)		WTLS (with outliers)		L_1 norm (with outliers)	
	Mean	Max	Mean	Max	Mean	Max
$\lvert da_1 \rvert$	0.0557	0.2385	0.1310	0.4465	0.0909	0.3956
$\lvert db_1 \rvert$	0.0306	0.1006	0.3041	0.4979	0.0556	0.2218
$\lvert dc_1 \rvert$	0.1130	0.4533	0.5241	0.9197	0.1780	0.8245
$\lvert da_2 \rvert$	0.0473	0.1831	0.0869	0.2573	0.0705	0.3584
$\lvert db_2 \rvert$	0.0272	0.1079	0.1881	0.3316	0.0490	0.2260
$\lvert dc_2 \rvert$	0.0959	0.3604	0.3171	0.5992	0.1584	0.7458

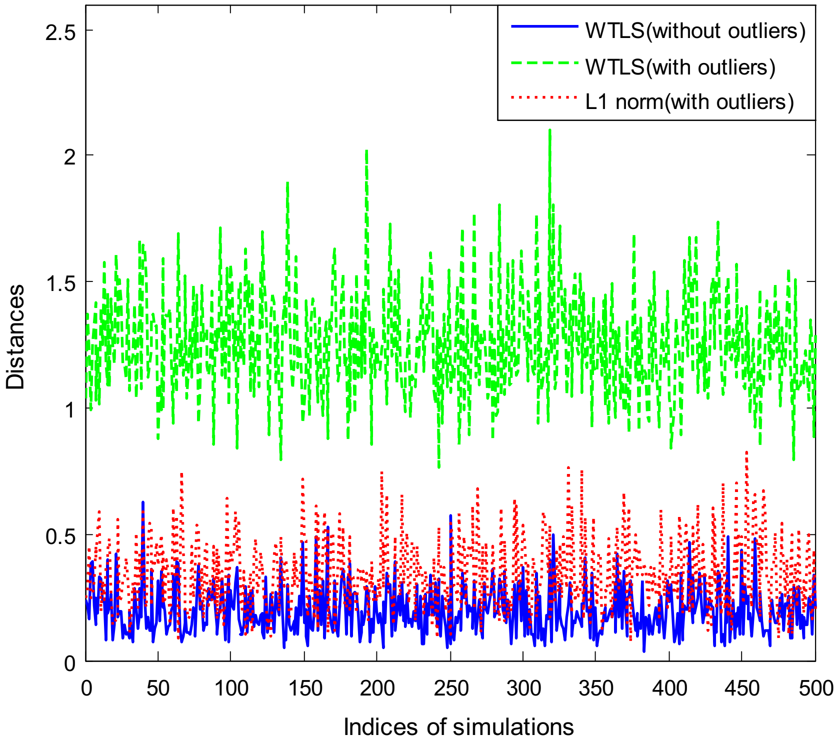

Fig. 2 Euclid distances between transformation parameters estimated by different methods and reference value for 500 replications with multiple outliers

4.3 Example 3: Real data about map rectification

The example is about the map rectification. The 2D affine transformation is used to rectify the map. The scale of the map is 1:500 for Fig. 3. The theoretical coordinates of the 10 common points and 15 non-common points are previously known. Then we sample the coordinates on the distorted map. The sampled coordinates and the theoretical coordinates

Table 9 Statistical results of the Euclidean distances between the estimated transformation parameters and reference values for 500 replications with multiple outliers

Distance	WTLS (without outlier)		WTLS (with outliers)		L_1 norm (with outliers)	
	Mean	Max	Mean	Max	Mean	Max
$\|\hat{X} - X_{real}\|$	0.1969	0.6244	1.2524	2.1046	0.3447	0.8250

Table 10 Euclidean distance between the transformation parameters estimated by the different methods and reference values for 500 replications with multiple outliers

Differences	WTLS (without outlier)		WTLS (with outliers)		L_1 norm (with outliers)	
	Mean	Max	Mean	Max	Mean	Max
$\|da_1\|$	0.0535	0.2115	0.5221	0.8288	0.0690	0.2904
$\|db_1\|$	0.0313	0.1752	0.1748	0.3344	0.0420	0.1452
$\|dc_1\|$	0.1124	0.5503	0.6051	1.2009	0.1283	0.5398
$\|da_2\|$	0.0484	0.2492	0.6422	1.2197	0.0866	0.3146
$\|db_2\|$	0.0279	0.1116	0.6163	0.8170	0.0910	0.3150
$\|dc_2\|$	0.0979	0.4322	0.2494	0.8770	0.2405	0.7396

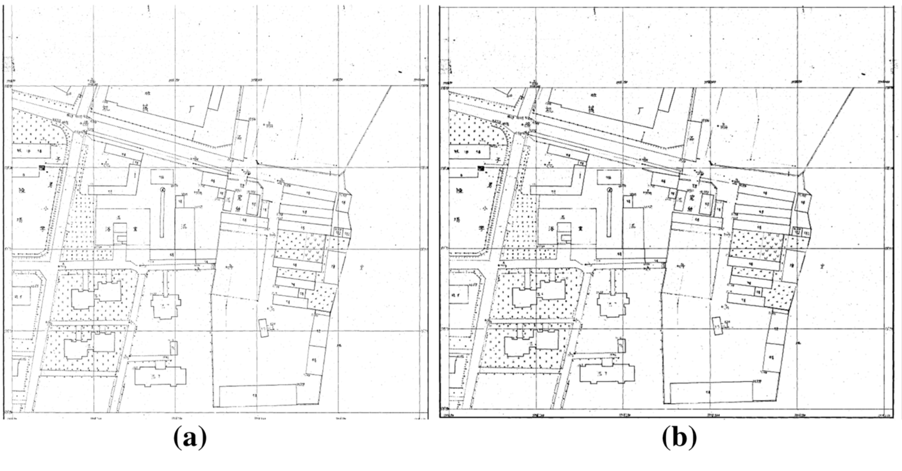

Fig. 3 The distorted map (**a**) and its rectified map (**b**) using affine transformation

can be treated as the coordinates of the start coordinate system and target system. The transformation parameters can be estimated by using the common points with the 2D affine transformation. Then the coordinates of non-common points in the target system can be derived by the coordinates of the non-common points in the start system and the estimated transformation parameters. The estimates of the transformation parameters with different methods are presented in Table 11.

To judge whether the observations are contaminated with outliers, the data-snooping method and the proposed L_1 norm minimization method are employed to detect the

Table 11 Transformation parameters estimated by the WTLS method and L_1 norm minimization method before and after deleting outliers

Before deleting outlier			After deleting outlier	
WTLS	L_1 norm (algorithm 1)	L_1 norm (algorithm 2)	Data-snooping	L_1 norm
0.30309	0.30305	0.30305	0.30309	0.30311
0.00003	0.00005	0.00005	0.00003	0.00003
10.47529	10.47486	10.47486	10.47529	10.47511
0.00140	0.00008	0.00008	0.00001	0.00001
0.30313	0.30377	0.30377	0.30382	0.303821
58.46941	58.49029	58.49029	58.48958	58.48958

outliers, respectively. From Table 12, the maximum of the absolute values of w-test statistics is 3.73 which is greater than the threshold value 2.14. So the no. 14 equation is distinguished as an outlying one. Then the new residuals and w-test statistics are obtained after deleting the outlying observation. Apparently, right now, there is no outlier in the observations and coefficient matrix because all the absolute values of w-test statistics are less than the threshold value 2.16. On the other hand, we know that the y_t component of the point 7 is contaminated with outlier by comparing with the residuals in Table 13 given by

Table 12 Residuals and w-test statistics of the data-snooping procedure by Amiri-Simkooei and Jazaeri (2013)

Eq. no.	Point no.	Coord.	$\hat{e}$	w	$\hat{e}$	w
1	1	x	0.0060	0.1137	0.0060	1.5793
2		y	−0.0252	−0.4801	0.0029	0.7775
3	2	x	−0.0052	−0.1003	−0.0052	−1.39
4		y	0.0498	0.9511	−0.0021	−0.5918
5	3	x	0.0047	0.0798	0.0047	1.1082
6		y	−0.0184	−0.3155	−0.0016	−0.3773
7	4	x	−0.0044	−0.0779	−0.0044	−1.0774
8		y	0.0607	1.0679	−0.0024	−0.6210
9	5	x	−0.0011	−0.0175	−0.0011	−0.2438
10		y	−0.0023	−0.0377	0.0033	0.7586
11	6	x	−0.0033	−0.0537	−0.0033	−0.7462
12		y	0.0071	0.1169	0.0015	0.3521
13	7	x	0.0019	0.0354	0.0019	0.4721
14		y	−0.2119	−3.7327	–	–
15	8	x	−0.0063	−0.1085	−0.0063	−1.5083
16		y	0.0114	0.1951	−0.0055	−1.3026
17	9	x	0.0077	0.1458	0.0077	2.0386
18		y	0.1015	1.9339	0.0046	1.3923
19	10	x	0.0001	0.0014	0.0001	0.0228
20		y	0.0273	0.5213	−0.0007	−0.2001

Table 13 Residuals obtained by the L_1 norm minimization method and WTLS method

Point no.	Eq. no.	L_1 norm (with outlier) $\boldsymbol{V}$	L_1 norm (deleting outlier) $\boldsymbol{V}$	WTLS (before deleting outlier) $\hat{\boldsymbol{e}}$	WTLS (after deleting outlier) $\hat{\boldsymbol{e}}$
1	1	0.0078	0.0078	0.0060	0.0057
	2	0	0	−0.0252	0.0029
2	3	−0.0060	−0.0004	−0.0052	−0.0048
	4	−0.0005	−0.0005	0.0498	−0.002
3	5	0.0061	0.0061	0.0047	0.0045
	6	−0.0038	−0.0038	−0.0184	−0.0016
4	7	−0.0055	0	−0.0044	−0.0038
	8	0	0	0.0607	−0.0024
5	9	0	0	−0.0011	−0.0011
	10	0.0019	0.0019	−0.0023	0.0033
6	11	−0.0026	−0.0026	−0.0033	−0.0032
	12	0.0008	0.0008	0.0071	0.0015
7	13	0	–	0.0019	–
	14	−0.2943	–	−0.2119	–
8	15	−0.0060	−0.0060	−0.0063	−0.0062
	16	−0.0054	−0.0054	0.0114	−0.0055
9	17	0.0054	0.01094	0.0077	−0.0062
	18	0.0092	0.0092	0.1015	−0.0055
10	19	0	0	8.671e−005	0.0003
	20	0	0	0.0273	−0.0007

the proposed L_1 norm minimization method. After rejecting the outlying observation, the more reliable transformation parameters are obtained, which can be found in Table 11.

By comparing with the reliability of the proposed method, the existing non-common points in the target system are treated as the check points. The RMSE (root mean square error) can be used to evaluate the reliability of the proposed algorithm. Therefore, the RMSE is computed as 0.033, 0.01, 0.0089, 0.0091 for WTLS method with outliers, L_1 norm minimization method, WTLS method after deleting outlier identified by L_1 norm minimization method and WTLS method after deleting outlier identified by data-snooping procedure, respectively, which show that the coordinates obtained by the L_1 norm minimization method are more accurate than those obtained by the WTLS method. The reason is that the transformation parameters estimated by the WTLS method are disturbed with the outlying observations. In this case, the proposed L_1 norm minimization method is more reliable to be employed for resisting the bad effect of outlier.

5 Conclusions

In this paper, the L_1 norm minimization method for the partial EIV model based on the linear programming theory is developed. However, the close-form solutions would not be exploited due to the unknown constrained condition equations for the optimization problem and the iterative method is proposed. At the same time, by treating the partial EIV model as a non-linear G–H one, another iterative algorithm is also proposed. The results of some numerical experiments show that the two proposed methods are superior to the WTLS method when the observations and coefficient matrix are contaminated with outliers simultaneously. And the two algorithms are equivalent in the case of robustness.

It is found through the Monte Carlo simulation that the L_1 norm minimization method may not efficient if there are the larger outliers in the coefficient matrix (Xu 2005) that can lead to divergence. If one wants to detect a single outlier, the data-snooping procedure proposed by Amiri-Simkooei and Jazaeri (2013) and the L_1 norm minimization method can do it all. But the L_1 norm minimization method maybe more reliable for detecting multiple outliers due to masking. Unfortunately, the computation burden is significantly increased.

Acknowledgments The authors would like to show the appreciations to the anonymous reviewers for their constructive comments so that the original paper has been substantial improved. The research works are sponsored by the National Natural Science Foundation of China (Project No. 41174005, 41474009).

References

Abatzoglou TJ, Mendel JM, Harada GA (1991) The constrained total least squares technique and its applications to harmonic superresolution. IEEE Trans Signal Proc 39(5):1070–1087

Amiri-Simkooei AR (2003) Formulation of L_1 norm minimization in Gauss-Markov models. J Surv Eng 129(1):37–43

Amiri-Simkooei AR, Jazaeri S (2013) Data-snooping procedure applied to errors-in-variables models. Stud Geophys Geod 57:426–441

Amiri-Simkooei AR, Jazaeri S (2012) Weighted total least squares formulated by standard least squares theory. J Geod Sci 2(2):113–124

Baarda W (1968) A testing procedure for use in geodetic networks. Netherlands Geod Comm Publ Geod 2(5):1–97

Baselga S (2007) Global optimization solution of robust estimation. J Surv Eng 133(3):123–128

Brown M (1982) Robust line estimation with errors in both variables. J Am Stat Assoc 77:71–79

Fang X (2013) Weighted total least squares: necessary and sufficient conditions, fixed and random parameters. J Geod 87:733–749

Gui QM, Liu JS (1999) Generalized shrunken-type robust estimation. J Surv Eng 125(4):177–184

Gui QM, Li GZ, Ou JK (2005) Robust-biased estimation based on quasi-accurate detection. J Surv Eng 131(3):67–72

Gui Q, Gong Y, Li G, Li B (2007) A Bayesian approach to the detection of gross errors based on posterior probability. J Geod 81:651–659

Gui Q, Li X, Gong Y, Li B, Li G (2011) A Bayesian unmasking method for locating multiple gross errors based on posterior probabilities of classification variables. J Geod 85:191–203

Guo JF, Ou JK, Wang HT (2007) Quasi-accurate detection of outliers for correlated observations. J Surv Eng 133(3):129–133

Guo JF, Ou JK, Wang HT (2010) Robust estimation for correlated observations: two local sensitivity-based downweighting strategies. J Geod 84:243–250

Hadi AS, Imon A (2009) Detection of outliers. Wiley Interdiscip Rev Comput Stat 1:57–70

Hekimoglu S (1997) Finite sample breakdown points of outlier detection procedures. J Surv Eng 125(1):15–31

Hekimoglu S (1999) Robustifying conventional outlier detection procedures. J Surv Eng 76:706–713

Jazaeri S, Amiri-Simkooei AR, Sharifi MA (2014) Iterative algorithm for weighted total least squares adjustment. Surv Rev 46(334):19–27

Koch KR (2013) Robust estimation by expectation maximization algorithm. J Geod 87:107–116

Li B, Shen Y, Zhang X, Li C, Lou L (2013) Seamless multivariate affine error-in-variables transformation and its application to map rectification. J GIS 27(8):1572–1592

Lu J, Chen Y, Li BF, Fang X (2014) Robust total least squares with reweighting iteration for three-dimensional similarity transformation. Surv Rev 46(334):28–36

Mahboub V (2012) On weighted total least-squares for geodetic transformations. J Geod 86:359–367

Mahboub V, Amiri-Simkooei AR, Sharifi MA (2013) Iteratively reweighted total least squares: a robust estimation in errors-in-variables models. Surv Rev 45(329):92–99
Marshall J, Bethel J (1996) Basic concept of L_1 norm minimization for surveying applications. J Surv Eng 122(4):168–179
Pope AJ (1976) The statistics of residuals and the detection of outliers. NOAA Technique Report Nos 65 GNS 1, Rockville
Schaffrin B, Uzun S (2011) Errors-in-variables for mobile mapping algorithms in the presence of outliers. Arch Photogr Cartogr Remote Sens 22:337–387
Schaffrin B, Wieser A (2008) On weighted total least-squares adjustment for linear regression. J Geod 82:415–421
Shen YZ, Li BF, Chen Y (2011) An iterative solution of weighted total least-squares adjustment. J Geod 85:229–238
Tao YQ, Gao JX, Yao YF (2014) TLS Algorithm for GPS height fitting based on robust estimation. Surv Rev 46(336):184–188
Vanderbei RJ (2014) Linear programming foundations and extensions, 4th edn. Springer, New York
Watson GA, Yiu KFC (1991) On the solution of the errors invariables problem using the L_1 Norm. BIT 31:697–710
Xu PL (1989) On robust estimation with correlated observation. Bull Geod 63:237–252
Xu PL (1993) Consequences of constant parameters and confidence intervals of robust estimation. Boll Geod Sci Aff 52(3):231–249
Xu PL (2005) Sign-constrained robust least squares, subjective break-down point and the effect of weights of observations on robustness. J Geod 79(1–3):146–159
Xu PL, Liu JN, Shi C (2012) Total least squares adjustment in partial errors-in-variables models: algorithm and statistical analysis. J Geod 86:661–675
Yang YX (1999) Robust estimation of geodetic datum transformation. J Geod 73(5):268–274
Yang YX, Song LJ, Xu TH (2002) Robust estimator for correlated observations based on bifactor equivalent weights. J Geod 76(6–7):353–358
Yang L, Wang JL, Knight NL, Shen YZ (2013) Outlier separability analysis with a multiple alternative hypotheses test. J Geod 87:591–604
Yetkin M, Inal C (2011) L_1 norm minimization in GPS networks. Surv Rev 43(322):523–532
Zamar RH (1989) Robust estimation in the errors-in-variables model. Biometrika 76(1):149–160