

Weighted total least-squares with constraints: a universal formula for geodetic symmetrical transformations

Xing Fang

Received: 8 June 2014 / Accepted: 12 January 2015 / Published online: 5 February 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract Although the analytical solutions for total least-squares with multiple linear and single quadratic constraints were developed quite recently in different geodetic publications, these methods are restricted in number and type of constraints, and currently their computational efficiency and applications are mostly unknown. In this contribution, it is shown how the weighted total least-squares (WTLS) problem with arbitrary applicable constraints can be solved based on a Newton type methodology. This iterative process with quadratic convergence is expanded upon to become a compact solution for the WTLS with or without constraints. This compact solution is then further interpreted as a universal formula for the symmetrical adjustment of the errors-in-variables model which represents affine, similarity and rigid transformations in two- and three-dimensional space. Furthermore, statistical analysis of the constrained WTLS including the first-order approximation of precision and the bias was investigated. In order to substantiate our proposed method's applicability, it was used to solve the affine, similarity and rigid transformation problem in two- and three-dimensional cases, where the structure of the coefficient matrix and multiple constraints were taken into account simultaneously.

Keywords Total least-squares · Errors-in-variables model · Nonlinear constraints · Newton iteration · Universal formula · Statistical analysis · Geodetic transformations

1 Introduction

The well-known least-squares (LS) has been widely used for adjusting the Gauss–Markov model (GMM) in geodesy and related fields. The usual prerequisite for applying this linear LS technique within the GMM is that the coefficient matrix is error free. However, this assumption is not always satisfied in many cases, for example, in regression problems (e.g., [Schaffrin and Wieser 2008](#); [Grafarend and Awange 2012](#); [Amiri-Simkooei et al. 2014](#)) as well as in geodetic transformations (e.g., [Teunissen 1988](#); [Schaffrin and Felus 2008](#); [Cai and Grafarend 2009](#)). In mathematics, if the coefficient matrix is affected by random errors, this nonlinear LS estimation within the GMM containing such a coefficient matrix is called total least-squares (TLS) within the errors-in-variables (EIV) model (see [Golub and Van Loan 1980](#); [Van Huffel and Vandewalle 1991](#)). It is acknowledged that in [Teunissen \(1988\)](#) it is the first time that an EIV or TLS problem was formulated and solved in the geodetic literature. Moreover, the solution given was an exact, non-iterative, analytical solution with an outstanding example of the planar similarity transformation. From a geodetic point of view, this is in fact ([Teunissen 1988](#)) as well as [Teunissen \(1985\)](#) reformulated the EIV model to the standard nonlinear Gauss–Markov model, and therefore they have the advantage that all available knowledge (numerical and statistical) of solving nonlinear observation equations by least-squares can be directly applied. Of course, the EIV model can be also handled as a nonlinear Gauss–Helmert model (GHM) (see [Neitzel 2010](#)), and equivalently transformed to condition equations containing an uncertain coefficient matrix (see [Schaffrin and Wieser 2011](#)).

Recently a large number of iterative algorithms were investigated which attempted to solve the weighted TLS (WTLS) problem (see [Schaffrin and Wieser 2008](#); [Neitzel](#)

X. Fang (✉)
School of Geodesy and Geomatics, Wuhan University,
Wuhan, China
e-mail: wflsfang1981@hotmail.com

2010; Amiri-Simkooei and Jazaeri 2012, 2013; Mahboub 2012; Snow 2012; Xu et al. 2012; Amiri-Simkooei 2013; Jazaeri et al. 2014; Xu and Liu 2014). These algorithms were individually developed using different approaches, but all of them were based on the first derivative of the WTLS objective function. Fang (2011) and Fang (2013) proposed an iterative Newton method based on the second-order approximation of the WTLS objective function. This Newton approach presented its local quadratic convergence rate (e.g., Teunissen 1990) and can treat the WTLS problem with an arbitrary positive definite cofactor matrix.

It is quite common to introduce constraints of unknown parameters in a system of equations. As far as we know, incorporating equality constraints frequently avoid a trivial solution, introduce the prior knowledge, guarantee the stability of estimates and mitigate bias. The biases are namely due to the nonlinearities involved. By making use of differential geometric concepts, the bias of the nonlinear least-squares estimators of the similarity transformation has been determined in Teunissen (1989a). The bias results in Teunissen (1989a) are also given in the textbook by Borre and Strang (2012), p. 186. The same consideration to introduce the constraints arises in the EIV model. The constrained TLS (CTLS) problem has been discussed primarily in mathematics and geodesy. Van Huffel and Vandewalle (1991, p. 275) and Dowling et al. (1992) proposed a closed form solution for the TLS problem with linear constraints. Schaffrin and Felus (2005) and Schaffrin (2006) presented an iterative CTLS solution with the fixed and stochastic right-hand-side vector in linear constraint equations, respectively. In quadratic constraints, Golub et al. (1999) and Sima et al. (2004) and Beck and Ben-Tal (2006) established the CTLS algorithms that dealt with a quadratic constraint in the context of regularization. In this case, where linear and quadratic constraints are considered simultaneously, Schaffrin and Felus (2009) proposed a solution to treat the mixed constrained TLS problem, which also could be extended to the fairly weighted case, according to Schaffrin and Wieser (2008). The mixed constrained TLS solution is obtained iteratively by solving the Lagrange multiplier associated with the quadratic constraint within a deduced quadratic equation. Recently, Mahboub and Sharifi (2013) extended the mixed constrained TLS solution to the mixed constrained WTLS (CWTLS) solution and presented its regularization effects. This estimator is capable of solving the constrained WTLS problem in a weighted case, where the structure of the weighted matrix of the coefficient matrix is not limited. Quite recently, Fang (2014a) generalized the CWTLS to a fully populated covariance matrix, including cross-correlations between the coefficient matrix and the traditional observation vector. One should not underestimate the significance in the differences of weights, as they represent different levels of generality in admission of weights into the EIV model. Regarding the importance and progres-

sion of weighting schemes within the EIV model, we refer to Snow (2012) discussion for the WTLS, which is also valid in the constrained WTLS environment. Furthermore, De Moor (1990), Zhang et al. (2013) and Fang (2014b) investigated the TLS or WTLS problem with linear inequality constraints.

The types and the number of the constraints, however, have still been limited until now. For example, existing methods only allow linear constraints and a maximum of one quadratic constraint to be considered. In addition, Schaffrin and Felus (2009) admitted that the exact convergence behavior of their algorithm is still to be determined, which is also the case for Mahboub and Sharifi (2013) and Fang (2014a). These algorithms might be inefficient, especially in a case with large perturbations, see Example 2 in Schaffrin and Felus (2009). It is also mentioned in Schaffrin and Felus that the CTLS solution should be examined in real geodetic applications. Furthermore, the statistical analysis of the constrained WTLS solution have not received due attention.

In this contribution, a WTLS problem with arbitrary types and numbers of applicable constraints (defined in the next section) is investigated. Based on sequential quadratic programming (SQP), the algorithm is designed, which is then presented to be directly applicable to affine, similarity and rigid transformation in the two- and three-dimensional space. The quality description of the CWTLS solution is given. Applications are then employed to demonstrate the use of the algorithm in the geodetic problems.

2 CWTLS problem and its nonlinear normal equations

2.1 CWTLS problem

Let the constrained EIV model be defined by the functional model and constraints:

$$\begin{aligned} \mathbf{y} - \mathbf{e}_y &= (\mathbf{A} - \mathbf{E}_A)\xi \\ \text{subject to } \mathbf{c}(\xi) &= \mathbf{0} \end{aligned} \quad (1)$$

and the stochastic model representing properties of random errors

$$\mathbf{e} := \begin{bmatrix} \text{vec}(\mathbf{E}_A) \\ \mathbf{e}_y \end{bmatrix} = \begin{bmatrix} \mathbf{e}_A \\ \mathbf{e}_y \end{bmatrix} \sim \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \sigma_0^2 \mathbf{Q}_{\Pi} = \sigma_0^2 \mathbf{P}^{-1} \right) \quad (2)$$

In the above constrained EIV model \mathbf{y} and \mathbf{e}_y are the $n \times 1$ observation vector and the corresponding random error vector, respectively. Matrices \mathbf{A} and \mathbf{E}_A are the full column-rank $n \times u$ stochastic model matrix and the corresponding random error matrix, respectively. Vector ξ is the unknown parameter vector with dimension $u \times 1$. \mathbf{e} is the extended random error vector, where $\mathbf{e}_A = \text{vec}(\mathbf{E}_A)$ ('vec' denotes the operator that stacks one column of a matrix underneath the previous one, see Schaffrin and Felus 2008). The symbol σ_0^2 denotes the

unknown variance component. Matrices \mathbf{Q}_{Π} and \mathbf{P} are the cofactor matrix and the weight matrix of the extended observation vector $\mathbf{l} := \text{vec}([\mathbf{A}, \mathbf{y}])$. The applicable constraints $\mathbf{c}(\xi) = \mathbf{0}$ with dimension $m \times 1$ denote a compact formulation of constraints. ‘Applicable’ refers to 1) that the constraints $\mathbf{c}(\xi)$ are assumed to be twice differentiable; 2) that the number of constraints must be smaller than the number of unknown parameters, i.e., $m < u$; and 3) that the Jacobian matrix of the constraints with respect to the unknown parameter vector has full row-rank. Except for the three constraint properties mentioned above, the types and the number of the constraint functions can be arbitrary.

The cofactor matrix is segmented in further detail (see Fang 2011) as

$$\mathbf{Q}_{\Pi} = \begin{bmatrix} \mathbf{Q}_{AA} & \mathbf{Q}_{Ay} \\ \mathbf{Q}_{yA} & \mathbf{Q}_{yy} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_1 \\ \vdots \\ \mathbf{Q}_k \\ \vdots \\ \mathbf{Q}_{u+1} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{Q}_{11} & \cdots & \mathbf{Q}_{1j} & \cdots & \mathbf{Q}_{1(u+1)} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \mathbf{Q}_{k1} & \cdots & \mathbf{Q}_{kj} & \cdots & \mathbf{Q}_{k(u+1)} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \mathbf{Q}_{(u+1)1} & \cdots & \mathbf{Q}_{(u+1)j} & \cdots & \mathbf{Q}_{(u+1)(u+1)} \end{bmatrix}, \quad (3)$$

where \mathbf{Q}_{AA} , \mathbf{Q}_{Ay} , \mathbf{Q}_{yA} and \mathbf{Q}_{yy} are the block cofactor matrices with dimension $nu \times nu$, $nu \times n$, $n \times nu$ and $n \times n$, respectively. Matrix \mathbf{Q}_k is an $n \times n(u+1)$ cofactor matrix and \mathbf{Q}_{kj} is an $n \times n$ cofactor matrix, both being certain blocks of whole cofactor matrix \mathbf{Q}_{Π} .

2.2 Constrained normal equation through using the Lagrange approach

A CWTLS problem can be formulated by a WTLS target function subject to Eq. (1) as

$$\begin{aligned} \min \mathbf{e}^T \mathbf{P} \mathbf{e} \\ \text{subject to } \mathbf{y} - \mathbf{A}\xi + \mathbf{E}_A \xi - \mathbf{e}_y = \mathbf{0} \\ \text{and } \mathbf{c}(\xi) = \mathbf{0}. \end{aligned} \quad (4)$$

According to the traditional Lagrange approach, the CWTLS target function is formed as follows:

$$\begin{aligned} \Phi(\mathbf{e}, \lambda, \xi, \mu) &= \mathbf{e}^T \mathbf{P} \mathbf{e} + 2\lambda^T (\mathbf{y} - \mathbf{A}\xi + \mathbf{E}_A \xi - \mathbf{e}_y) + 2\mu^T \mathbf{c}(\xi) \\ &= \mathbf{e}^T \mathbf{P} \mathbf{e} + 2\lambda^T (\mathbf{y} - \mathbf{A}\xi + \mathbf{B}\mathbf{e}) + 2\mu^T \mathbf{c}(\xi), \end{aligned} \quad (5)$$

where the vectors λ and μ are the Lagrange multipliers vectors associated with the functional model and the constraint functions, respectively. The full row-rank matrix $\mathbf{B}_{n \times n(u+1)}$ is defined by $[\xi^T \otimes \mathbf{I}_n, -\mathbf{I}_n]$ where \otimes is the Kronecker product. \mathbf{I}_n is identity matrix with dimension $n \times n$.

Setting the partial derivatives of the target function with respect to vectors ξ , \mathbf{e} , λ and μ each equal to zero, gives the necessary conditions as

$$\frac{1}{2} \frac{\partial \Phi}{\partial \xi} \Big|_{\hat{\xi}, \hat{\mathbf{e}}, \hat{\lambda}, \hat{\mu}} = -\mathbf{A}^T \hat{\lambda} + \tilde{\mathbf{E}}_A^T \hat{\lambda} + \hat{\mathbf{C}}^T \hat{\mu} = \mathbf{0}, \quad (6)$$

$$\frac{1}{2} \frac{\partial \Phi}{\partial \mathbf{e}} \Big|_{\hat{\xi}, \hat{\mathbf{e}}, \hat{\lambda}} = \mathbf{P} \hat{\mathbf{e}} + \hat{\mathbf{B}}^T \hat{\lambda} = \mathbf{0}, \quad (7)$$

$$\frac{1}{2} \frac{\partial \Phi}{\partial \lambda} \Big|_{\hat{\xi}, \hat{\mathbf{e}}} = \mathbf{y} - \mathbf{A} \hat{\xi} + \hat{\mathbf{B}} \hat{\mathbf{e}} = \mathbf{0}, \quad (8)$$

$$\frac{1}{2} \frac{\partial \Phi}{\partial \mu} \Big|_{\hat{\xi}} = \mathbf{c}(\hat{\xi}) = \mathbf{0} \quad (9)$$

with $\mathbf{C} = \mathbf{C}(\xi) := \partial \mathbf{c}(\xi) / \partial \xi^T$. From Eq. (7) the random error vector is predicted as

$$\tilde{\mathbf{e}} = -\mathbf{Q}_{\Pi} \hat{\mathbf{B}}^T \hat{\lambda}. \quad (10)$$

Using Eq. (10) in Eq. (8) $\hat{\lambda}$ can be estimated by

$$\hat{\lambda} = (\hat{\mathbf{B}} \mathbf{Q}_{\Pi} \hat{\mathbf{B}}^T)^{-1} (\mathbf{y} - \mathbf{A} \hat{\xi}). \quad (11)$$

Then by reinserting Eq. (11) into Eq. (10), the predicted error vector is directly formed by $\mathbf{e}(\hat{\xi})$, a function of the estimated parameter vector:

$$\tilde{\mathbf{e}} = \mathbf{e}(\hat{\xi}) = -\mathbf{Q}_{\Pi} \hat{\mathbf{B}}^T (\hat{\mathbf{B}} \mathbf{Q}_{\Pi} \hat{\mathbf{B}}^T)^{-1} (\mathbf{y} - \mathbf{A} \hat{\xi}). \quad (12)$$

The above predicted error vector leads to the predicted random error matrix $\tilde{\mathbf{E}}_A$ by $\mathbf{E}_A(\hat{\xi})$ as:

$$\begin{aligned} \tilde{\mathbf{E}}_A &= \mathbf{E}_A(\hat{\xi}) \\ &= - \left[\mathbf{Q}_1^B T (\hat{\mathbf{B}} \mathbf{Q}_{\Pi} \hat{\mathbf{B}}^T)^{-1} (\mathbf{y} - \mathbf{A} \hat{\xi}), \dots, \right. \\ &\quad \left. \mathbf{Q}_u^B T (\hat{\mathbf{B}} \mathbf{Q}_{\Pi} \hat{\mathbf{B}}^T)^{-1} (\mathbf{y} - \mathbf{A} \hat{\xi}) \right]. \end{aligned} \quad (13)$$

Combining Eqs. (6), (9) and (11), the constrained normal equations can be formulated by

$$\begin{aligned} (\tilde{\mathbf{E}}_A - \mathbf{A})^T (\hat{\mathbf{B}} \mathbf{Q}_{\Pi} \hat{\mathbf{B}}^T)^{-1} (\mathbf{y} - \mathbf{A} \hat{\xi}) + \hat{\mathbf{C}}^T \hat{\mu} &= \mathbf{0} \\ \mathbf{c}(\hat{\xi}) &= \mathbf{0}. \end{aligned} \quad (14)$$

The solution of the CWTLS problem is now converted to solve the constrained nonlinear normal equations (CNNE). It is well known that nonlinear equations can be solved by a Newton type method. Forming an equivalent CWTLS problem based on the knowledge presented in Fang (2013) will facilitate directly applying the Newton method.

2.3 Equivalent approach in forming the constrained normal equations

Another formulation of the CWTLS problem can be written based on the alternative WTLS objective function by Fang (2011) and Fang (2013) with the same constraint functions as follows:

$$\begin{aligned} \min f(\xi) &= \frac{1}{2}(\mathbf{y} - \mathbf{A}\xi)^T (\mathbf{BQ}_{\parallel}\mathbf{B}^T)^{-1}(\mathbf{y} - \mathbf{A}\xi) \\ \text{subject to } \mathbf{c}(\xi) &= \mathbf{0}, \end{aligned} \quad (15)$$

where $\mathbf{B} = [\xi^T \otimes \mathbf{I}_n, -\mathbf{I}_n]$.

The analytical form of the gradient vector $\mathbf{g}(\xi)$ and the Hessian matrix $\mathbf{H}(\xi)$ of the function $f(\xi)$ was presented in Fang (2013), Fang (2014b) and repeated here as

$$\begin{aligned} \mathbf{g}(\xi) &= \frac{\partial f(\xi)}{\partial \xi} = -(\mathbf{A} + \mathbf{A}^*)^T (\mathbf{BQ}_{\parallel}\mathbf{B}^T)^{-1}(\mathbf{y} - \mathbf{A}\xi) \\ &= (\mathbf{E}_A(\xi) - \mathbf{A})^T (\mathbf{BQ}_{\parallel}\mathbf{B}^T)^{-1}(\mathbf{y} - \mathbf{A}\xi), \end{aligned} \quad (16)$$

$$\begin{aligned} \mathbf{H}(\xi) &= \frac{\partial^2 f(\xi)}{\partial \xi \partial \xi^T} = (\mathbf{A} + \mathbf{A}^* + \mathbf{A}^{**})^T (\mathbf{BQ}_{\parallel}\mathbf{B}^T)^{-1} \\ &\quad \times (\mathbf{A} + \mathbf{A}^* + \mathbf{A}^{**}) - [\varpi_{kj}], \end{aligned} \quad (17)$$

with the auxiliary matrices \mathbf{A}^* , \mathbf{A}^{**} and $[\varpi_{kj}]$ defined by:

$$\begin{aligned} \mathbf{A}^* &= \begin{bmatrix} \mathbf{Q}_1 \mathbf{B}^T (\mathbf{BQ}_{\parallel}\mathbf{B}^T)^{-1} (\mathbf{y} - \mathbf{A}\xi), \dots, \\ \mathbf{Q}_u \mathbf{B}^T (\mathbf{BQ}_{\parallel}\mathbf{B}^T)^{-1} (\mathbf{y} - \mathbf{A}\xi) \end{bmatrix}, \\ \mathbf{A}^{**} &= \begin{bmatrix} \mathbf{BQ}_1^T (\mathbf{BQ}_{\parallel}\mathbf{B}^T)^{-1} (\mathbf{y} - \mathbf{A}\xi), \dots, \\ \mathbf{BQ}_u^T (\mathbf{BQ}_{\parallel}\mathbf{B}^T)^{-1} (\mathbf{y} - \mathbf{A}\xi) \end{bmatrix}, \\ \varpi_{kj} &= (\mathbf{y} - \mathbf{A}\xi)^T (\mathbf{BQ}_{\parallel}\mathbf{B}^T)^{-1} \mathbf{Q}_{kj} (\mathbf{BQ}_{\parallel}\mathbf{B}^T)^{-1} (\mathbf{y} - \mathbf{A}\xi). \end{aligned} \quad (18)$$

In order to obtain the minimum of (15), the Lagrange function is again used for this context as

$$L(\xi, \mu) = f(\xi) + \mu^T \mathbf{c}(\xi). \quad (19)$$

The first order necessary condition of this problem, minimization of the function $L(\xi, \mu)$, can be written as a system of $u + m$ equations including the $u + m$ unknowns ξ and μ :

$$\mathbf{f}(\hat{\xi}, \hat{\mu}) = \frac{\partial L(\xi, \mu)}{\partial [\xi^T, \mu^T]^T} \bigg|_{\hat{\xi}, \hat{\mu}} = \begin{bmatrix} \mathbf{g} + \mathbf{C}^T \hat{\mu} \\ \mathbf{c} \end{bmatrix} = \mathbf{0}. \quad (20)$$

Note that the identity of (14) and (20) can be recognized by the formulation of the vector \mathbf{g} . For brevity of notation, we denote the estimates ξ and μ by $\hat{\xi}$ and $\hat{\mu}$ but keep the notation of \mathbf{c} , \mathbf{g} , \mathbf{H} and \mathbf{C} to denote their estimates $\mathbf{c}(\hat{\xi})$, $\mathbf{g}(\hat{\xi})$, $\mathbf{H}(\hat{\xi})$ and $\mathbf{C}(\hat{\xi})$ in the following part, respectively.

Equation (20), the first-order derivatives $\mathbf{f}(\hat{\xi}, \hat{\mu}) = \mathbf{0}$, are identical to the CNNE (14). This equivalence indicates that the alternative CWTLs objective function (15) yields an identical solution to the solution based on the optimization problem (4). The strict CTLS solution is further verified by the positive quadratic form of the Hessian matrix \mathbf{H} and

any feasible active direction which is determined by the constraints.

3 WTLS and CWTLs solution based on Newton iteration

As far as we know, Eq. (15) is a standard nonlinear program. Thus, the CWTLs problem (15) can be categorized as a local SQP problem (see Nocedal and Wright 2006, p. 530). One approach to solving Eq. (20) is here suggested by Newton's method. The total Hessian matrix with respect to the augmented vector $[\xi^T, \mu^T]^T$ is given by

$$\begin{aligned} \frac{\partial^2 L(\xi, \mu)}{\partial [\xi^T, \mu^T]^T \partial [\xi^T, \mu^T]^T} \bigg|_{\hat{\xi}, \hat{\mu}} &= \frac{\partial \mathbf{f}(\hat{\xi}, \hat{\mu})}{\partial [\xi^T, \mu^T]^T} \bigg|_{\hat{\xi}, \hat{\mu}} \\ &= \begin{bmatrix} \mathbf{H} + \sum_{t=1}^m \hat{\mu}_t \mathbf{H}_t & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix}, \end{aligned} \quad (21)$$

where the Hessian matrix of the t th constraint is defined as $\mathbf{H}_t = \mathbf{H}_t(\hat{\xi}) = \partial^2 c_t(\hat{\xi}) / \partial \xi \partial \xi^T$, and $\hat{\mu}_t$ is the t th component of the vector $\hat{\mu}$ ($t = 1, \dots, m$).

By combining Eqs. (20) and (21), the Newton step to compute increments of the parameter vector and the Lagrange multiplier vector is presented as

$$\begin{bmatrix} \mathbf{H} + \sum_{t=1}^m \hat{\mu}_t \mathbf{H}_t & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} d\hat{\xi} \\ d\hat{\mu} \end{bmatrix} = - \begin{bmatrix} \mathbf{g} + \mathbf{C}^T \hat{\mu} \\ \mathbf{c} \end{bmatrix}. \quad (22)$$

By elimination of the increment of the Lagrange multipliers vector $d\hat{\mu}$, the estimates of the parameter increment $d\hat{\xi}$ and the Lagrange multipliers vector $\hat{\mu}$ can be iteratively derived based on Eq. (22) by

$$\begin{bmatrix} d\hat{\xi} \\ \hat{\mu} \end{bmatrix} = - \begin{bmatrix} \mathbf{H} + \sum_{t=1}^m \hat{\mu}_t \mathbf{H}_t & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{g} \\ \mathbf{c} \end{bmatrix}. \quad (23)$$

This Newton iteration is well defined when the augmented Hessian matrix is nonsingular. If the regular inversion of the augmented matrix is not available, we can establish a positive definite matrix as the approximation of the original Hessian matrix.

If these constraints are not considered, the unconstrained WTLS solution can be simplified based on Eq. (23). The increment of the parameter $d\hat{\xi}$ can be obtained by multiplying the gradient vector by the negative inverted Hessian matrix (Fang 2013):

$$d\hat{\xi} = -\mathbf{H}^{-1} \mathbf{g}. \quad (24)$$

In summation, the algorithm of this WTLS solution with or without arbitrary applicable constraints is designed as:

Algorithm 1: An SQP algorithm to solve the WTLS problem with or without constraints
INPUT:

- data matrix \mathbf{A} ($n \times u$) and \mathbf{y} ($n \times 1$)
- Cofactor matrix \mathbf{Q}_{\parallel} with dimension $(u+1)n \times (u+1)n$
- Constraints $\mathbf{c}(\xi) = \mathbf{0}$ if constraints are available
- small positive value ε , for example $\varepsilon = 10^{-10}$ (dependent on application)

BEGIN

- Initialize a parameter vector, for example $\hat{\xi}^0 = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}$
- Initialize a Lagrange vector $\mu^0 = \mathbf{0}$

Begin loop

- ☆ Create the Hessian matrix $\mathbf{H}^{i-1} = \mathbf{H}(\hat{\xi}^{i-1})$ and the gradient vector

$$\mathbf{g}^{i-1} = \mathbf{g}(\hat{\xi}^{i-1})$$

- ☆ If constraints are available,

Compute Hessian matrix of the constraints $\mathbf{H}_r^{i-1} = \mathbf{H}_r(\hat{\xi}^{i-1})$ and the

Jacobian of the constraints $\mathbf{C}^{i-1} = \mathbf{C}(\hat{\xi}^{i-1})$ by the parameter vector $\hat{\xi}^{i-1}$

Calculate the increment of the parameter $d\hat{\xi}^i$ and the Lagrange vector

$\hat{\mu}^i$ according to (23)

- ☆ Else

Compute the increment of the parameter $d\hat{\xi}^i$ according to (24);

- ☆ Set the parameter vector $\hat{\xi}^i = \hat{\xi}^{i-1} + d\hat{\xi}^i$

- **While** $\left\| \begin{matrix} \mathbf{g} + \mathbf{C}^T \hat{\mu} \\ \mathbf{c} \end{matrix} \right\| > \varepsilon$ **repeat**;

End loop**END****Output:**

- **The estimate of the parameter vector:** $\hat{\xi} := \hat{\xi}^i$
-

Being a Newton's type, this compact algorithm for solving the WTLS with or without constraints works efficiently, and provides a local quadratic convergence at the desired local minimum (Nocedal and Wright 2006)—bearing in mind

that the convergence behavior of other CTLS algorithms is unknown. It is noted that this local behavior implies (Teunissen 1990): (1) that convergence is not guaranteed and (2) that approximate values still need to be good enough. The

Table 1 The common functional model and stochastic model, and the different constraints represent the three transformations in the 2D and 3D cases

2D	Functional model	$\mathbf{y} - \mathbf{e}_y = (\mathbf{A} - \mathbf{E}_A)\xi$ $\xi = [\text{vec}^T(\Xi_{2 \times 2}), [\Delta x \ \Delta y]]^T$ $\mathbf{A} = [\mathbf{I}_2 \otimes [\mathbf{x}_s \ \mathbf{y}_s], \mathbf{I}_2 \otimes \mathbf{1}]$ $\mathbf{y} = [\mathbf{x}_t^T, \mathbf{y}_t^T]^T$		
	Stochastic model	$D(\text{vec}[\mathbf{A} \ \mathbf{y}]) = \sigma_0^2 \mathbf{Q}_\Pi$ $= \begin{bmatrix} \mathbf{F}_{2D} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} D([\mathbf{x}_s^T \ \mathbf{y}_s^T \ \mathbf{x}_t^T \ \mathbf{y}_t^T]^T) \begin{bmatrix} \mathbf{F}_{2D} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^T$		
	Constraints $\mathbf{c}(\xi) = \mathbf{0}$ for	Affine	Similarity	Rigid
		No	$\xi_{11}^2 + \xi_{12}^2 - \xi_{21}^2 - \xi_{22}^2 = 0$ $\xi_{11}\xi_{21} + \xi_{12}\xi_{22} = 0$	$\xi_{11}^2 + \xi_{12}^2 - \xi_{21}^2 - \xi_{22}^2 = 0$ $\xi_{11}\xi_{21} + \xi_{12}\xi_{22} = 0$ $\xi_{11}^2 + \xi_{12}^2 - 1 = 0$
3D	Functional model	$\mathbf{y} - \mathbf{e}_y = (\mathbf{A} - \mathbf{E}_A)\xi$ $\xi = [\text{vec}^T(\Xi_{3 \times 3}), [\Delta x \ \Delta y \ \Delta z]]^T$ $\mathbf{A} = [\mathbf{I}_3 \otimes [\mathbf{x}_s \ \mathbf{y}_s \ \mathbf{z}_s], \mathbf{I}_3 \otimes \mathbf{1}]$ $\mathbf{y} = [\mathbf{x}_t^T \ \mathbf{y}_t^T \ \mathbf{z}_t^T]^T$		
	Stochastic model	$D(\text{vec}[\mathbf{A} \ \mathbf{y}]) = \sigma_0^2 \mathbf{Q}_\Pi$ $= \begin{bmatrix} \mathbf{F}_{3D} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} D([\mathbf{x}_s^T \ \mathbf{y}_s^T \ \mathbf{z}_s^T \ \mathbf{x}_t^T \ \mathbf{y}_t^T \ \mathbf{z}_t^T]^T) \begin{bmatrix} \mathbf{F}_{3D} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^T$		
	Constraints $\mathbf{c}(\xi) = \mathbf{0}$ for	Affine	Similarity	Rigid
		No	$\xi_{11}^2 + \xi_{12}^2 + \xi_{13}^2 - \xi_{21}^2 - \xi_{22}^2 - \xi_{23}^2 = 0$ $\xi_{21}^2 + \xi_{22}^2 + \xi_{23}^2 - \xi_{31}^2 - \xi_{32}^2 - \xi_{33}^2 = 0$ $\xi_{11}\xi_{21} + \xi_{12}\xi_{22} + \xi_{13}\xi_{23} = 0$ $\xi_{11}\xi_{31} + \xi_{12}\xi_{32} + \xi_{13}\xi_{33} = 0$ $\xi_{31}\xi_{21} + \xi_{32}\xi_{22} + \xi_{33}\xi_{23} = 0$	$\xi_{11}^2 + \xi_{12}^2 + \xi_{13}^2 - \xi_{21}^2 - \xi_{22}^2 - \xi_{23}^2 = 0$ $\xi_{21}^2 + \xi_{22}^2 + \xi_{23}^2 - \xi_{31}^2 - \xi_{32}^2 - \xi_{33}^2 = 0$ $\xi_{11}^2 + \xi_{12}^2 + \xi_{13}^2 - 1 = 0$ $\xi_{11}\xi_{21} + \xi_{12}\xi_{22} + \xi_{13}\xi_{23} = 0$ $\xi_{11}\xi_{31} + \xi_{12}\xi_{32} + \xi_{13}\xi_{33} = 0$ $\xi_{31}\xi_{21} + \xi_{32}\xi_{22} + \xi_{33}\xi_{23} = 0$

linear search strategy to find the step length, which satisfies the (strong) Wolfe condition, can be introduced to improve convergence behavior of the iterative TLS solution (e.g., Teunissen 1990; Fang 2014c). Furthermore, the compact form presented will be a promising tool for estimating the transformation parameters in different types of geodetic transformations.

4 Universal formula for geodetic symmetrical transformations

It is well known that attaining an estimation of the transformation parameters in the presence of two sets of (2D or 3D) coordinates for a given geodetic network is a long-standing problem in geodetic science (see Schaffrin et al. 2012; Fang 2014c). In Table 1, the adjustment of the different geodetic transformations, including affine, similarity and rigid transformations, can be interpreted as the WTLS or CWTLS problem if one perfectly describes the constraints for the corresponding 2D or 3D transformation models.

In 2D case, the matrix $\Xi_{2 \times 2} = \begin{bmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{bmatrix}$ contains the four synthetic parameters, and its transpose is computed

using the product of the rotation matrix, the affinity matrix and the scaling matrix containing two scales for the x and y coordinates (see Mahboub 2012). It is noted that if one deals with a rotation matrix, then not only the columns are of unit length and mutually orthogonal, but also the determinant of the matrix must be +1. \mathbf{x}_t and \mathbf{y}_t are the random vectors of the x and y coordinates of the target system while \mathbf{x}_s and \mathbf{y}_s are the random vectors of the x and y coordinates of the source system. $[\Delta x \ \Delta y]^T$ are the translations in the x and y orientations. $\mathbf{1}$ is the vector of ones with the corresponding size (number of common points). The matrix $\mathbf{F}_{2D} = \frac{\partial \text{vec}(\mathbf{A})}{\partial [\mathbf{x}_t^T \ \mathbf{y}_t^T]}$ represents the relationship between the vectorized coefficient matrix and the coordinates of the target system.

In 3D case, the parameter matrix $\Xi_{3 \times 3} = \begin{bmatrix} \xi_{11} & \xi_{12} & \xi_{13} \\ \xi_{21} & \xi_{22} & \xi_{23} \\ \xi_{31} & \xi_{32} & \xi_{33} \end{bmatrix}$ is enlarged with a dimension of 3×3 and the translation Δz in the z axis is considered. The covariance matrix calculated by $\mathbf{F}_{3D} = \frac{\partial \text{vec}(\mathbf{A})}{\partial [\mathbf{x}_t^T \ \mathbf{y}_t^T \ \mathbf{z}_t^T]}$, which can be obtained as it is computed in the 2D case, is also valid for all three types of 3D transformations.

Although the matrix \mathbf{A} , the observation vector \mathbf{y} and the parameter vector ξ in a 3D case are different from those in

a 2D case, we can still use the same symbols \mathbf{A} , \mathbf{y} and ξ to hold the formulation of the EIV model. Furthermore, the functional model (the matrix \mathbf{A} , the observation vector \mathbf{y} and the parameter vector ξ) and the stochastic model (the cofactor matrix \mathbf{Q}_{II}) are common for the three kinds of transformations in the corresponding dimension.

It has been presented that affine, similarity and rigid transformations in 2D and 3D space can be interpreted as the WTLS problem with or without constraints. Hence, Algorithm 1 can be directly applicable to all of the transformations and can be referred to as a universal formula for geodetic symmetrical transformations. When some extra constraints (e.g., fixed baselines) need to be taken into consideration simultaneously (see Shen et al. 2011), our proposed algorithm is also applicable.

5 Statistical analysis of CWTLS solution

Although our algorithm has been developed for the CWTLS problem, statistical aspects of the CWTLS estimation were not investigated. Teunissen (1989b) provided general diagnostics of the first and second moments of the nonlinear LS estimate for the first time in the geodetic literature. Xu et al. (2012) applied the theory and methods by Box (1971) to work out the statistical aspects of parameter estimation in the unconstrained EIV model.

In order to obtain the quality description by using the existing knowledge, we can formulate the functional part of the TLS model as a nonlinear LS problem (Teunissen 1988):

$$\mathbf{y} - \mathbf{e}_y = \bar{\mathbf{A}}\xi = (\xi^T \otimes \mathbf{I})(\mathbf{F}_t \mathbf{a}_I + \mathbf{d})$$

$$\mathbf{a}_I - \mathbf{e}_{A_I} = \bar{\mathbf{a}}_I$$

The symbol $\bar{\mathbf{a}}_I$ with dimension $s \times 1$ and \mathbf{e}_{A_I} denotes the true vector of the independent elements within the true coefficient matrix $\bar{\mathbf{A}}$ and its error vector, respectively. \mathbf{F}_t is the transformation matrix (\mathbf{F}_{2D} or \mathbf{F}_{3D}) to convert the vector $\bar{\mathbf{a}}_I$ to the vectorized (true) coefficient matrix $\text{vec}(\bar{\mathbf{A}}) = \mathbf{F}_t \bar{\mathbf{a}}_I + \mathbf{d}$.

In the case of the CWTLS model, it is usually formulated that the constraints approximately equal pseudo-equations: $\mathbf{0} - \delta \approx \mathbf{c}(\xi)$ with $\delta \sim (\mathbf{0}, \sigma_0^2 \mathbf{Q}_{\delta\delta})$. However, we should note that it is numerically admissible to treat the constraints completely as the pseudo-equations ($\mathbf{0} - \hat{\delta} = \mathbf{c}(\hat{\xi})$), as the variance of the random vector δ is sufficiently small (see Fang 2011, p. 45).

Combining the above models, a nonlinear Gauss–Markov model can be written as follows:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{a}_I \\ \mathbf{0} \end{bmatrix} = \mathbf{f}(\mathbf{x}) + \varepsilon = \mathbf{f}\left(\begin{bmatrix} \xi \\ \bar{\mathbf{a}}_I \end{bmatrix}\right) + \begin{bmatrix} \mathbf{e}_y \\ \mathbf{e}_{A_I} \\ \delta \end{bmatrix}$$

Based on the existing knowledge, the first-order approximation of the dispersion matrix and the bias vector \mathbf{b}_x of the

CWTLS estimates of the extended unknown parameter vector \mathbf{x} are simply given as follows:

$$D(\mathbf{x}) = \sigma_0^2 (\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1} \quad (25)$$

and

$$\mathbf{b}_x = -\frac{\sigma_0^2}{2} (\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{W} \mathbf{m} \quad (26)$$

with

$$\mathbf{F} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}^T} = \begin{bmatrix} \bar{\mathbf{A}} & (\xi^T \otimes \mathbf{I}_n) \mathbf{F}_t \\ \mathbf{0} & \mathbf{I}_s \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \quad \mathbf{W} = \begin{bmatrix} \mathbf{P}_I & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{\delta\delta}^{-1} \end{bmatrix}$$

$$\mathbf{m} = \begin{bmatrix} \text{tr}((\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1} \mathbf{M}_1) \\ \text{tr}((\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1} \mathbf{M}_2) \\ \vdots \\ \text{tr}((\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1} \mathbf{M}_{s+n+t}) \end{bmatrix} \quad \mathbf{M}_i = \frac{\partial^2 f_i}{\partial \mathbf{x} \partial \mathbf{x}^T}$$

where f_i is the i th function of \mathbf{f} , and \mathbf{P}_I denotes the weight matrix of the independent errors.

In the case of the above nonlinear adjustment problem, we have the second partial derivatives \mathbf{M}_i for each f_i , which is given by

$$\mathbf{M}_i = \begin{cases} \begin{bmatrix} \mathbf{0} & (\mathbf{I}_u \otimes \mathbf{e}_i) \mathbf{F}_t \\ \mathbf{F}_t^T (\mathbf{I}_u \otimes \mathbf{e}_i^T) & \mathbf{0} \end{bmatrix} & i = 1, \dots, n \\ \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{H}_{i-n-s} & \mathbf{0} \end{bmatrix} & i = n+1, \dots, n+s \\ \begin{bmatrix} \mathbf{0} & \mathbf{0} \end{bmatrix} & i = n+s+1, \dots, n+s+t \end{cases}$$

The total sum of squared residual (TSSR) is easily determined from objective function $f(\hat{\xi}_{\text{CWTLS}})$ so that a estimate of the variance component can be obtained through

$$\hat{\sigma}_0^2 = \frac{f(\hat{\xi}_{\text{CWTLS}})}{n - u + t}. \quad (27)$$

6 Case study of 2D and 3D affine, similarity and rigid transformations

In this case study, we demonstrate Algorithm 1 as the universal formula for affine, similarity and rigid transformations in 2D and 3D space, and provide the quality description for the solutions. The data for the 2D transformations, as seen in Table 2, originate from Mikhail and Gracie (1981, p. 397–402), and was also presented in Neitzel (2010) and Schaffrin et al. (2012) for the WTLS problem.

For these problems of affine, similarity and rigid transformations in 2D, Algorithm 1 was implemented when the constraints were perfectly described, as proposed in Sect. 4. In Table 3, we present the WTLS solution for the affine transformation and the CWTLS solution for the similarity and rigid transformations. The results of the similarity transformation

Table 2 Target and source coordinates from Mikhail and Gracie (1981, p. 397–402)

Point no.	\mathbf{x}_t (target)	\mathbf{y}_t (target)	\mathbf{x}_s (source)	\mathbf{y}_s (source)
1	−117.478	0	17.856	144.794
2	117.472	0	252.637	154.448
3	0.015	−117.41	140.089	32.326
4	−0.014	117.451	130.40	267.027

Table 3 Estimates of the case study by Algorithm 1 for three types of 2D transformation

Estimates	Affine	Similarity	Rigid
$\hat{\xi}_{11}$	0.99902905	0.99900748	0.99915487
$\hat{\xi}_{12}$	0.04111867	0.04109806	0.04110413
$\hat{\xi}_{21}$	−0.04107747	−0.04109806	−0.04110413
$\hat{\xi}_{22}$	0.99898590	0.99900748	0.99915487
Δx	−141.26879	−141.26279	−141.28363
Δy	−143.93120	−143.93164	−143.95288
Objective function	0.00061868	0.00064325	0.00124378

Table 4 Standard deviations of the parameter estimates for three types of 2D transformation

	Standard deviation		
	Affine	Similarity	Rigid
$\hat{\xi}_{11}$	1.4968×10^{-4}	0.7632×10^{-4}	0.3902×10^{-4}
$\hat{\xi}_{12}$	1.4974×10^{-4}	0.7632×10^{-4}	0.9487×10^{-4}
$\hat{\xi}_{21}$	1.4968×10^{-4}	0.7632×10^{-4}	0.9487×10^{-4}
$\hat{\xi}_{22}$	1.4973×10^{-4}	0.7632×10^{-4}	0.3902×10^{-4}
Δx	3.2661×10^{-2}	1.7816×10^{-2}	1.7641×10^{-2}
Δy	3.2660×10^{-2}	1.7816×10^{-2}	1.7444×10^{-2}

coincide with the iterative GHM solution presented in Neitzel (2010). The results of the affine and rigid transformation are not much different from the results of the similarity transformation since the noises of the data set are small. It is clear that the last row of Table 3 shows that the values of the objective functions become larger from the affine to similarity to rigid transformation, as more constraints are considered for the affine to similarity to rigid transformation.

According to Eqs. (25) and (26), we computed the standard deviations and biases of the estimates of the 2D transformations and show them in Tables 4 and 5, respectively. The standard deviations of the parameter vector for affine, similarity and rigid are presented at the level of 10^{-4} for the elements of the matrix $\Xi_{2 \times 2}$ and at the level of 10^{-2} for the translations. Table 4 indicates the fact that the more constraints are considered, the smaller the standard deviation of each parameter is. The biases of the parameters are

Table 5 Biases of the parameter estimates for three types of 2D transformation

	Bias		
	Affine	Similarity	Rigid
$\hat{\xi}_{11}$	0.0111×10^{-6}	0×10^{-6}	-0.0045×10^{-6}
$\hat{\xi}_{12}$	0.0004×10^{-6}	0×10^{-6}	-0.0002×10^{-6}
$\hat{\xi}_{21}$	-0.0004×10^{-6}	0×10^{-6}	0.0002×10^{-6}
$\hat{\xi}_{22}$	0.0112×10^{-6}	0×10^{-6}	-0.0045×10^{-6}
Δx	-1.5831×10^{-6}	0×10^{-6}	0.6368×10^{-6}
Δy	-1.6141×10^{-6}	0×10^{-6}	0.6489×10^{-6}

presented in the magnitude of 10^{-6} . The biases of the similarity can be neglected in 10^{-6} , whereas the biases of the affine and the rigid have opposite signs for each parameter.

A recent survey by Felus and Burtch (2009) lists a large number of methods that have been developed to compute 3D geodetic transformations between two sets of corresponding points. Here, we will use the data set presented in this paper to demonstrate the proposed solutions. In this demonstration, six control points are identified and recorded in the two coordinate systems (WGS 84 and a local datum), and those x , y and z coordinate values are given in Table 6.

Algorithm 1 was used to compute the transformation parameters displayed in Tables 7, 8 and 9 for the affine, similarity and rigid transformations, respectively. The results for the similarity transformation are identical to the solution presented in Felus and Burtch. In these three transformations, the differences between the estimated translations for the similarity transformation and the estimated translations for the affine and rigid transformation are significant. This could be explained by the restricted rotation matrices and scale factor as well as the magnitude of the coordinates. Note that the estimated rotation angles for the three transformations are quite small (only at the level of seconds). The objective function for the three types of transformations demonstrates the fact that the similarity and rigid transformation includes more constraints, thus resulting in a larger objective function than the affine.

Now we calculate the standard deviations and biases of the estimated parameters, for the 3D affine, similarity and rigid transformations, and present them in Tables 10 and 11, respectively. All the standard deviations for the elements within the matrix $\Xi_{3 \times 3}$ are given in 10^{-4} while the standard deviations of the translations are relatively large, in the magnitude of 10^4 for the affine, and in the magnitude of 10^2 for similarity and rigid. It needs to be pointed out, however, that the computed standard deviations of the translations depends on the datum, i.e., if we treat the real data by multiplying the idempotent matrix (e.g., Felus and Burtch 2009), the stan-

Table 6 Coordinates of source system and target system from Felus and Burtch (2009)

Point no.	x (source)	y (source)	z (source)	X (target)	Y (target)	Z (target)
80601	5,234,251.25	905,003.2011	3,518,869.674	5,233,991.482	905,003.1064	3,519,305.459
32,127	5,218,851.932	919,148.9749	3,537,928.348	5,218,595.021	919,152.3244	3,538,363.627
80,600	5,220,818.669	772,128.3613	3,569,828.606	5,220,565.466	772,130.563	3,570,253.01
32,136	5,148,067.252	803,912.306	3,668,491.426	5,147,806.722	803,921.3223	3,668,928.371
80,598	5,081,676.23	771,786.8122	3,765,023.787	5,081,410.788	771,799.4256	3,765,460.689
80,597	5,022,479.06	955,283.5487	3,801,754.143	5,022,218.176	955,297.2538	3,802,185.975

Table 7 Estimates of the case study by Algorithm 1 for the 3D affine transformation

Estimated translates for the affine transformation	Estimated parameter matrix for the affine transformation			
Δx	4,274.5841	0.999438049	0.000622536	0.0021992977
Δy	-5,094.9028	-0.000101814	1.000112976	0.0004077426
Δz	-17,013.5995	-0.000425541	0.000493015	1.001581579
Objective function	58.5666			

Table 8 Estimates of the case study by Algorithm 1 for the 3D similarity transformation

Estimated translates for the similarity transformation	Estimated parameter matrix for the similarity transformation			
Δx	-293.3670	1.000010668	-0.000021228	0.000010763
Δy	40.7974	0.000021228	1.000010668	-0.000018196
Δz	354.7273	-0.000010763	0.000018196	1.000010668
Objective function	115.2651			

Table 9 Estimates of the case study by Algorithm 1 for the 3D rigid transformation

Estimated translates for the rigid transformation	Estimated parameter matrix for the rigid transformation			
Δx	-238.3801	1.000000000	-0.000021228	0.000010763
Δy	49.9133	0.000021228	1.000000000	-0.0000181963
Δz	393.5986	-0.000010763	0.0000181963	1.000000000
Objective function	123.4183			

dard deviations are largely reduced. The phenomenon that the standard deviations become smaller when more constraints are available is also valid for the 3D case.

Comparing the most insignificant biases of the estimated parameters for the similarity, the biases of the estimated parameters for the affine transformation are large, especially for the translations. These large biases may be explained by the significant differences between the translation estimates of the affine and similarity transformations. In full analogy in

Table 10 Standard deviations of the parameter estimates for three types of 3D transformation

	Standard deviation		
	Affine	Similarity	Rigid
$\hat{\xi}_{11}$	15.3817×10^{-4}	0.2090×10^{-4}	0.1267×10^{-4}
$\hat{\xi}_{12}$	2.7838×10^{-4}	0.2370×10^{-4}	0.2341×10^{-4}
$\hat{\xi}_{13}$	11.0183×10^{-4}	0.1890×10^{-4}	0.1692×10^{-4}
$\hat{\xi}_{21}$	15.3869×10^{-4}	0.3362×10^{-4}	0.3330×10^{-4}
$\hat{\xi}_{22}$	2.7848×10^{-4}	0.1843×10^{-4}	0.1712×10^{-4}
$\hat{\xi}_{23}$	11.0220×10^{-4}	0.2577×10^{-4}	0.2553×10^{-4}
$\hat{\xi}_{31}$	15.3982×10^{-4}	0.2452×10^{-4}	0.2369×10^{-4}
$\hat{\xi}_{32}$	2.7868×10^{-4}	0.2226×10^{-4}	0.2203×10^{-4}
$\hat{\xi}_{33}$	11.0301×10^{-4}	0.1732×10^{-4}	0.1460×10^{-4}
Δx	1.2180×10^4	1.5959×10^2	1.0836×10^2
Δy	1.2184×10^4	2.5654×10^2	2.5415×10^2
Δz	1.2193×10^4	1.7358×10^2	1.6071×10^2

Table 11 Biases of the parameter estimates for three types of 3D transformation

	Bias		
	Affine	Similarity	Rigid
$\hat{\xi}_{11}$	1.1826×10^{-6}	0	-0.0003×10^{-6}
$\hat{\xi}_{12}$	0.2132×10^{-6}	0	0
$\hat{\xi}_{13}$	0.8470×10^{-6}	0	-0.0001×10^{-6}
$\hat{\xi}_{21}$	0.2145×10^{-6}	0	0
$\hat{\xi}_{22}$	0.0389×10^{-6}	-0.0004×10^{-6}	-0.0005×10^{-6}
$\hat{\xi}_{23}$	0.1537×10^{-6}	0	0
$\hat{\xi}_{31}$	0.8518×10^{-6}	-0.0001×10^{-6}	-0.0001×10^{-6}
$\hat{\xi}_{32}$	0.1535×10^{-6}	0	0
$\hat{\xi}_{33}$	0.6102×10^{-6}	0	-0.0002×10^{-6}
Δx	-9.3641	0.5828×10^{-3}	1.9238×10^{-3}
Δy	-1.6994	0.6929×10^{-3}	0.8862×10^{-3}
Δz	-6.7453	0.8479×10^{-3}	1.7937×10^{-3}

the 2D case, the biases of the affine and the rigid have opposite signs for each parameter here (see Table 11 columns 2 and 4).

7 Conclusions

This article investigated the WTLS solution with constraints, which is intuitively very appealing for a symmetrical adjustment of various geodetic transformations. Unlike other CWTLS solutions, the proposed algorithm adapts itself to arbitrarily applicable constraints. Furthermore, our method is also highly reliable and efficient from a numerical point of view due to having known quadratic convergence behavior.

Our proposed algorithm, which is referred to as the universal formula for the adjustment of transformations, was implemented to deal with various transformations in 2D and 3D space by simply changing constraints. Furthermore, the statistical analysis of the CWTLS estimates was investigated.

Along with Fang (2013, 2014b) this paper completes the WTLS framework with varying types of prior information about unknown parameters (random parameters, linear and nonlinear, equality and inequality constraints).

Acknowledgments I would like to thank the President at the Federal Agency for Cartography and Geodesy (BKG) in Germany, Prof. Kutterer, for his guidance. The author thanks three reviewers for their constructive comments. The second reviewer is particularly appreciated for his broad knowledge on nonlinear LS estimation and suggestions for the quality description, which have been fully implemented to improve the paper. This research was supported by the National Natural Science Foundation of China (41404005; 41474006) and the Fundamental Research Funds for the Central Universities (2042014kf053).

References

- Amiri-Simkooei A, Jazaeri S (2012) Weighted total least squares formulated by standard least squares theory. *J Geod Sci* 2(2):113–124
- Amiri-Simkooei A (2013) Application of least squares variance component estimation to errors-in-variables models. *J Geod* 87(10–12):935–944
- Amiri-Simkooei A, Jazaeri S (2013) Data-snooping procedure applied to errors-in-variables models. *Studia Geophysica et Geodaetica* 57(3):426–441
- Amiri-Simkooei A, Zangeneh-Nejad F, Asgari J, Jazaeri S (2014) Estimation of straight line parameters with fully correlated coordinates. *Measurement* 48:378–386
- Beck A, Ben-Tal A (2006) On the solution of the Tikhonov regularization of the total least squares. *SIAM J Optim* 17:98–118
- Borre K, Strang G (2012) Algorithms for global positioning. Wellesley-Cambridge Press, Cambridge
- Box MJ (1971) Bias in nonlinear estimation (with discussions). *J Roy Stat Soc B* 33:171–201
- Cai J, Grafarend E (2009) Systematical analysis of the transformation between Gauss–Krueger–Coordinate/DHDN and UTM-coordinate/ETRS89 in Baden–Württemberg with different estimation methods. In: Geodetic reference frames, International Association of Geodesy Symposia, vol 134, pp 205–211
- Dowling EM, Degroat RD, Linebarger DA (1992) Total least-squares with linear constraints. In: Acoustics, speech, and signal processing (ICASSP-92), vol 5, pp 341–344. doi:10.1109/ICASSP.1992.226613
- De Moor B (1990) Total linear least squares with inequality constraints. ESAT-SISTA Report 1990-02, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium
- Fang X (2011) Weighted total least squares solution for application in geodesy. Dissertation, Leibniz University Hanover, Nr 294
- Fang X (2013) Weighted total least squares: necessary and sufficient conditions, fixed and random parameters. *J Geod* 87(8):733–749
- Fang X (2014a) A structured and constrained total least-squares solution with cross-covariances. *Studia Geophysica et Geodaetica* 58(1):1–16
- Fang X (2014b) On non-combinatorial weighted total least squares with inequality constraints. *J Geod* 88(8):805–816
- Fang X (2014c) A total least squares solution for geodetic datum transformations. *Acta Geodaetica et Geophysica* 49(2):189–207
- Felus F, Burtch R (2009) On symmetrical three-dimensional datum conversion. *GPS Sol* 13(1):65–74
- Golub G, Van Loan C (1980) An analysis of the total least-squares problem. *SIAM J Numer Anal* 17(6):883–893
- Golub GH, Hansen PC, O’Leary DP (1999) Tikhonov regularization and total least-squares. *SIAM J Matrix Anal Appl* 21:185–194
- Grafarend E, Awange JL (2012) Applications of linear and nonlinear models. Fixed effects, random effects, and total least squares. Springer, Berlin
- Jazaeri S, Amiri-Simkooei AR, Sharifi MA (2014) An iterative algorithm for weighted total least-squares adjustment. *Surv Rev* 46:16–27
- Mahboub V (2012) On weighted total least-squares for geodetic transformation. *J Geod* 86(5):359–367
- Mahboub V, Sharifi MA (2013) On weighted total least-squares with linear and quadratic constraints. *J Geod* 87(3):279–286
- Mikhail EM, Gracie G (1981) Analysis and adjustment of survey measurements. Van Nostrand Reinhold Company, New York
- Neitzel F (2010) Generalization of total least-squares on example of unweighted and weighted 2D similarity transformation. *J Geod* 84:751–762
- Nocedal J, Wright S (2006) Numerical optimization. Springer, Berlin
- Schaffrin B, Felus YA (2005) On total least-squares adjustment with constraints. In: Sansò F (ed) A window on the future of geodesy. International Association of Geodesy Symposia, vol 128. Springer, Berlin, pp 175–180
- Schaffrin B (2006) A note on constrained total least-squares estimation. *Linear Alg Appl* 417:245–258
- Schaffrin B, Felus Y (2008) On the multivariate total least-squares approach to empirical coordinate transformations. Three algorithms. *J Geod* 82:373–383
- Schaffrin B, Wieser A (2008) On weighted total least-squares adjustment for linear regression. *J Geod* 82:415–421
- Schaffrin B, Felus Y (2009) An algorithmic approach to the total least-squares problem with linear and quadratic constraints. *Studia Geophysica et Geodaetica* 53(1):1–16
- Schaffrin B, Wieser A (2011) Total least-squares adjustment of condition equation. *Studia Geophysica et Geodaetica* 55(3):529–536
- Schaffrin B, Neitzel F, Uzun S, Mahboub V (2012) Modifying Cadzow’s algorithm to generate the optimal TLS-solution for the structured EIV-model of a similarity transformation. *J Geod Sci* 2(2):98–106
- Shen Y, Li B, Chen Y (2011) An iterative solution to weighted total least squares adjustment. *J Geod* 85(4):229–238. doi:10.1007/s00190-010-0431-1
- Sima DM, van Huffel S, Golub GH (2004) Regularized total least-squares based on quadratic eigenvalue problem solver. *Bit* 44:793–812
- Snow K (2012) Topics in total least-squares adjustment within the errors-in-variables model: singular cofactor matrices and priori information. PhD Dissertation, report No. 502, Geodetic Science

- Program, School of Earth Sciences, The Ohio State University, Columbus Ohio, USA
- Teunissen PJG (1985) The geometry of geodetic inverse linear mapping and nonlinear adjustment. Netherlands Geodetic Commission, Publications on Geodesy, new series, vol. 8, no 1, pp 1–186
- Teunissen PJG (1988) The non-linear 2D symmetric Helmert transformation: an exact non-linear least-squares solution. *J Geod* 62(1):1–15
- Teunissen PJG (1989a) A note on the bias in the Symmetric Helmert Transformation, *Festschrift Torben Krarup*, pp 1–8
- Teunissen PJG (1989b) First and second moments of nonlinear least-squares estimators. *J Geod* 63:253–262
- Teunissen PJG (1990) Nonlinear least squares. *Manuscripta Geodetica* 15(3):137–150
- Van Huffel S, Vandewalle J (1991) The total least -squares problem. Computational aspects and analysis. Society for Industrial and Applied Mathematics, Philadelphia
- Xu PL, Liu JN, Shi C (2012) Total least squares adjustment in partial errors-in-variables models: algorithm and statistical analysis. *J Geod* 86(8):661–675
- Xu PL, Liu JN (2014) Variance components in errors-in-variables models: estimability, stability and bias analysis. *J Geod* 88(8):719–734
- Zhang S, Tong X, Zhang K (2013) A solution to EIV model with inequality constraints and its geodetic applications. *J Geod* 87(1):23–28