J. Geod. Sci. 2014; 4:28–36 — DOI 10.2478/jogs-2014-0004 DE GRUYTER OPEN

Research Article Open Access

B. Schaffrin, K. Snow*, and F. Neitzel

On the Errors-In-Variables model with singular dispersion matrices

Abstract: While the Errors-In-Variables (EIV) Model has been treated as a special case of the nonlinear Gauss-Helmert Model (GHM) for more than a century, it was only in 1980 that Golub and Van Loan showed how the Total Least-Squares (TLS) solution can be obtained from a certain minimum eigenvalue problem, assuming a particular relationship between the diagonal dispersion matrices for the observations involved in both the data vector and the data matrix. More general, but always nonsingular, dispersion matrices to generate the "properly weighted" TLS solution were only recently introduced by Schaffrin and Wieser, Fang, and Mahboub, among others. Here, the case of singular dispersion matrices is investigated, and algorithms are presented under a rank condition that indicates the existence of a unique TLS solution, thereby adding a new method to the existing literature on TLS adjustment. In contrast to more general "measurement error models," the restriction to the EIV-Model still allows the derivation of (nonlinear) closed formulas for the weighted TLS solution. The practicality will be evidenced by an example from geodetic science, namely the over-determined similarity transformation between different coordinate estimates for a set of identical points.

Keywords: Errors-In-Variables; singular dispersion matrices; Total Least-Squares

B. Schaffrin: Geodetic Science Program, School of Earth Sciences, The Ohio State University, Columbus, Ohio, U. S. A.
***Corresponding Author: K. Snow:** Geodetic Science Program, School of Earth Sciences, The Ohio State University, Columbus, Ohio, U. S. A., E-mail: ksnow6378137@gmail.com
F. Neitzel: Department of Geodesy and Geoinformation, Berlin University of Technology, Berlin, Germany

1 Introduction

Models equivalent to the Errors-In-Variables (EIV) Model have been used (at least) since Pearson (1901) for the fitting of straight lines and planes when *all* variables have been observed. For the longest time, the so-called nonlinear Gauss-Helmert Model (GHM) would have been applied, in conjunction with iterative linearization, in order to generate the standard LEast-Squares Solution (LESS); see, e.g., Helmert (1907). But, in 1980, Golub and Van Loan proved that, under the assumption of diagonal dispersion matrices for the data vector as well as for the vectorized data matrix, their Total Least-Squares (TLS) solution may equivalently be obtained from a certain minimum eigenvalue problem, where the weight matrices are chosen to be proportional to the inverse dispersion matrices in the manner shown in Eq. (3) below.

Since the seminal paper by Golub and Van Loan (1980), various attempts have been made to generalize their form of "element-wise weighting"; see, e.g., Markovsky et al. (2006). But only Schaffrin and Wieser (2008) succeeded in overcoming the "element-wise weighting" limitation, allowing for fairly general dispersion matrices, one of which would be positive-definite and the other the Kronecker-Zehfuss product of two nonnegative-definite matrices. The latter restriction was eventually overcome by Fang (2011) and Mahboub (2012) whose TLS algorithms allowed both dispersion matrices to be of a general structure, but positive-definite.

Meanwhile, however, Neitzel and Schaffrin (2013) had found a criterion for the uniqueness of the standard LESS within a linearized GHM even in the case of singular dispersion matrices. After adapting this criterion to the EIV-Model, two novel algorithms will be presented in the following that will allow computation of the TLS solution within an EIV-Model even if both dispersion matrices are singular, i.e., positive-semidefinite, thereby enriching the existing literature on methods for TLS adjustment.

In Section 2, a brief review of the TLS solution within EIV-Models with positive-definite dispersion matrices will be presented that covers both Fang's and Mahboub's algorithms. Then, in Section 3, novel algorithms will be presented that are able to generate the TLS solution even for singular dispersion matrices as long as the suitably adapted uniqueness criterion of Neitzel and Schaffrin (2013) is fulfilled. The practicality of the new algorithm will then be shown, in Section 4, by handling the case of an over-determined similarity transformation. Conclusions and an outlook will conclude this contribution.

2 The Errors-In-Variables model with general positive-definite dispersion matrices: a Lagrangian approach

The model, commonly known as the Errors-In-Variables (EIV) Model, can be defined as

$$y = \underbrace{(A - E_A)}_{n\times m} \xi + e_y, \ \text{rk} A = m < n, \tag{1a}$$

$$\begin{bmatrix} e_y \\ e_A := \text{vec}\, E_A \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma_0^2 \begin{bmatrix} Q_y & 0 \\ 0 & Q_A \end{bmatrix} = \sigma_0^2 Q = \sigma_0^2 P^{-1} \right), \tag{1b}$$

where

y	is the $n \times 1$ observation vector,
A	the $n \times m$ (random) data matrix of coefficients,
E_A	the associated $n \times m$ (unknown) random error matrix,
ξ	the $m \times 1$ (unknown) parameter vector,
e_y	the $n \times 1$ (unknown) random error vector associated with y,
e_A	the $nm \times 1$ vectorized form of the error matrix E_A, and
σ_0^2	the (unknown) variance component.

Moreover,

Q_y	is the $n \times n$ symmetric positive-definite cofactor matrix of e_y, with $P_y := Q_y^{-1}$ as the corresponding weight matrix, and
Q_A	is the $nm \times nm$ symmetric positive-definite cofactor matrix of e_A, with $P_A := Q_A^{-1}$ as the corresponding weight matrix.

Note that this model is slightly different from the one used by Schaffrin and Wieser (2008) who assumed

$$Q_A = \underset{m\times m}{Q_0} \otimes \underset{n\times n}{Q_x}, \tag{2a}$$

where Q_0 and Q_x are symmetric nonnegative-definite matrices such that the inverse

$$\left[Q_y + (\hat{\xi}^T Q_0 \hat{\xi}) \cdot Q_x\right]^{-1} \text{ exists} \tag{2b}$$

for the TLS solution $\hat{\xi}$. It is certainly more general than the model by Golub and Van Loan (1980) who allowed only "element-wise weighting" in the sense that

$$P_y := Q_y^{-1} = t_{m+1}^2 \cdot D^2 \text{ and } P_A := Q_A^{-1} = T_m^2 \otimes D^2 \tag{3}$$

with $\begin{bmatrix} T_m & 0 \\ 0 & t_{m+1} \end{bmatrix} := T$ and D as nonsingular diagonal matrices (cf. Snow (2012, pg. 22)). Here, $\otimes$ denotes the Kronecker-Zehfuss product of matrices, which is defined by

$$G \otimes H := [g_{ij} \cdot H] \quad \text{if} \quad G = [g_{ij}]. \tag{4}$$

Some of the basic rules include

$$(G \otimes H)^T = G^T \otimes H^T, \tag{5a}$$

$$(G \otimes H)^{-1} = G^{-1} \otimes H^{-1} \text{ if both inverses exist,} \tag{5b}$$

$$(A \otimes B) \cdot (G \otimes H) = (AG) \otimes (BH) \text{ if these products exist,} \tag{5c}$$

$$\text{vec}(ABC^T) = (C \otimes A)\, \text{vec}\, B, \tag{5d}$$

$$\text{tr}(ABC^T D^T) = \text{tr}(D^T ABC^T) = (\text{vec}\, D)^T (C \otimes A)\, \text{vec}\, B, \tag{5e}$$

$$\text{tr}(G \otimes H) = (\text{tr}\, G) \cdot (\text{tr}\, H) \text{ if both matrices are symmetric.} \tag{5f}$$

Moreover, the factors of the Kronecker-Zehfuss product can be switched by applying suitable "commutation matrices," in accordance with Magnus and Neudecker (2007), leading to

$$H \otimes G = K(G \otimes H)K, \tag{6a}$$

and specifically to

$$H \otimes g = K(g \otimes H), \tag{6b}$$

if g is a vector. The "commutation matrix" can also be interpreted as a "vec-permutation matrix" in the sense of

$$\text{vec}\left(G^T\right) = K \cdot \text{vec}\, G. \tag{6c}$$

Now, the task consists in finding the Total Least-Squares (TLS) solution for the EIV-Model in Eqs. (1a) and (1b). For this purpose, a Lagrangian approach will be applied that is based on the following Lagrange target function

$$\Phi(e_y, e_A, \xi, \lambda) := e_y^T P_y e_y + e_A^T P_A e_A + \\ + 2\lambda^T \left(y - A\xi - e_y + \left(\xi^T \otimes I_n\right) e_A\right), \tag{7}$$

which depends on the $n \times 1$ vector λ of Lagrange multipliers and needs to be made stationary while maintaining a minimum for the first two terms. Then, the Euler-Lagrange necessary conditions are obtained as:

$$\frac{1}{2}\frac{\partial \Phi}{\partial e_y} = P_y \tilde{e}_y - \hat{\lambda} \doteq 0 \quad \Rightarrow \quad \tilde{e}_y = Q_y \hat{\lambda}, \tag{8a}$$

$$\frac{1}{2}\frac{\partial \Phi}{\partial e_A} = P_A \tilde{e}_A + (\hat{\xi} \otimes I_n)\hat{\lambda} \doteq 0 \quad \Rightarrow \quad \tilde{e}_A = -Q_A(\hat{\xi} \otimes \hat{\lambda}), \tag{8b}$$

$$\frac{1}{2}\frac{\partial \Phi}{\partial \xi} = -\left(A - \tilde{E}_A\right)^T \hat{\lambda} \doteq 0 \quad \Rightarrow \quad A^T \hat{\lambda} = \left(\hat{\lambda}^T \otimes I_m\right) \text{vec}\left(\tilde{E}_A^T\right), \tag{8c}$$

$$\frac{1}{2}\frac{\partial \Phi}{\partial \lambda} = y - A\hat{\xi} - \tilde{e}_y + \left(\hat{\xi}^T \otimes I_n\right) \tilde{e}_A \doteq 0 \tag{8d}$$

$$\Rightarrow y - A\hat{\xi} = \left[Q_y + \left(\hat{\xi} \otimes I_n\right)^T Q_A \left(\hat{\xi} \otimes I_n\right)\right] \cdot \hat{\lambda} =: Q_1 \cdot \hat{\lambda} \tag{8e}$$

$$\Rightarrow \hat{\boldsymbol{\lambda}} = \boldsymbol{Q}_1^{-1}\left(\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\xi}}\right), \tag{9}$$

since $\boldsymbol{Q}_1$ is nonsingular if $\boldsymbol{Q}_y$ is positive-definite. Also, the sufficient condition holds true, as both $\boldsymbol{P}_y$ and $\boldsymbol{P}_A$ are positive-definite.

The corresponding "nonlinear normal equations" now follow from Eq. (8c) and Eq. (9), in combination with Eq. (8b), as:

$$\begin{bmatrix} \boldsymbol{Q}_1 & (\boldsymbol{A} - \tilde{\boldsymbol{E}}_A) \\ (\boldsymbol{A} - \tilde{\boldsymbol{E}}_A)^T & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\lambda}} \\ \hat{\boldsymbol{\xi}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{y} - \tilde{\boldsymbol{E}}_A\hat{\boldsymbol{\xi}} \\ \boldsymbol{0} \end{bmatrix} \tag{10a}$$

with the update

$$\operatorname{vec}\tilde{\boldsymbol{E}}_A = \tilde{\boldsymbol{e}}_A = -\boldsymbol{Q}_A\left(\hat{\boldsymbol{\xi}} \otimes \hat{\boldsymbol{\lambda}}\right) \tag{10b}$$

for the next iteration. This results in the algorithm suggested by Fang (2011), here labeled Algorithm 1; see also Xu et al. (2012) for a similar scheme to Eq. (10a), which could already be found in Fang (2011).

Algorithm 1: (according to Fang):

Step 1: Set $\tilde{\boldsymbol{E}}_A^{(0)} := \boldsymbol{0}$ and $\hat{\boldsymbol{\xi}}^{(0)} := (\boldsymbol{A}^T\boldsymbol{Q}_y^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Q}_y^{-1}\boldsymbol{y}$;

Step 2: For $i \in \mathbb{N}$, compute

$$\begin{aligned}
\boldsymbol{Q}_1^{(i)} &:= \left[\boldsymbol{Q}_y + \left(\hat{\boldsymbol{\xi}}^{(i-1)} \otimes \boldsymbol{I}_n\right)^T\boldsymbol{Q}_A\left(\hat{\boldsymbol{\xi}}^{(i-1)} \otimes \boldsymbol{I}_n\right)\right],\\
\hat{\boldsymbol{\xi}}^{(i)} &:= \left[\left(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)}\right)^T\left(\boldsymbol{Q}_1^{(i)}\right)^{-1}\left(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)}\right)\right]^{-1}\cdot\\
&\quad\cdot\left[\left(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)}\right)^T\left(\boldsymbol{Q}_1^{(i)}\right)^{-1}\left(\boldsymbol{y} - \tilde{\boldsymbol{E}}_A^{(i-1)}\hat{\boldsymbol{\xi}}^{(i-1)}\right)\right],\\
\hat{\boldsymbol{\lambda}}^{(i)} &= \left(\boldsymbol{Q}_1^{(i)}\right)^{-1}\left[\left(\boldsymbol{y} - \tilde{\boldsymbol{E}}_A^{(i-1)}\hat{\boldsymbol{\xi}}^{(i-1)}\right) - \left(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)}\right)\hat{\boldsymbol{\xi}}^{(i)}\right],\\
\tilde{\boldsymbol{e}}_A^{(i)} &= -\boldsymbol{Q}_A\left(\hat{\boldsymbol{\xi}}^{(i)} \otimes \hat{\boldsymbol{\lambda}}^{(i)}\right) \quad \text{and} \quad \tilde{\boldsymbol{E}}_A^{(i)} = \operatorname{Invec}\tilde{\boldsymbol{e}}_A^{(i)};
\end{aligned}$$

Step 3: Stop when $\|\hat{\boldsymbol{\lambda}}^{(i)} - \hat{\boldsymbol{\lambda}}^{(i-1)}\| < \delta$ and $\|\hat{\boldsymbol{\xi}}^{(i)} - \hat{\boldsymbol{\xi}}^{(i-1)}\| < \delta$ for a chosen threshold δ.

In contrast to Algorithm 1, Mahboub (2012) started with applying the "vec-permutation matrix" to the right-hand side of equation Eq. (8c) so that Eq. (9), in conjunction with Eq. (8b), leads to the following sequence of identities:

$$\begin{aligned}
\boldsymbol{A}^T\boldsymbol{Q}_1^{-1}\left(\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\xi}}\right) &= \boldsymbol{A}^T\hat{\boldsymbol{\lambda}} = \left(\hat{\boldsymbol{\lambda}}^T \otimes \boldsymbol{I}_m\right)\boldsymbol{K}\cdot\left(\operatorname{vec}\tilde{\boldsymbol{E}}_A\right) =\\
&= \left(\boldsymbol{I}_m \otimes \hat{\boldsymbol{\lambda}}^T\right)\tilde{\boldsymbol{e}}_A = -\left[\left(\boldsymbol{I}_m \otimes \hat{\boldsymbol{\lambda}}^T\right)\boldsymbol{Q}_A\left(\hat{\boldsymbol{\xi}} \otimes \boldsymbol{I}_n\right)\right]\cdot\\
&\quad\cdot\boldsymbol{Q}_1^{-1}\left(\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\xi}}\right) =: -\boldsymbol{R}_1\boldsymbol{Q}_1^{-1}\left(\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\xi}}\right)
\end{aligned} \tag{11}$$

and consequently to the next algorithm, here labeled Algorithm 2.

Algorithm 2: (according to Mahboub):

Step 1: Set $\hat{\boldsymbol{\xi}}^{(0)} := (\boldsymbol{A}^T\boldsymbol{Q}_y^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Q}_y^{-1}\boldsymbol{y}$;

Step 2: For $i \in \mathbb{N}$, compute

$$\begin{aligned}
\boldsymbol{Q}_1^{(i)} &:= \left[\boldsymbol{Q}_y + \left(\hat{\boldsymbol{\xi}}^{(i-1)} \otimes \boldsymbol{I}_n\right)^T\boldsymbol{Q}_A\left(\hat{\boldsymbol{\xi}}^{(i-1)} \otimes \boldsymbol{I}_n\right)\right],\\
\hat{\boldsymbol{\lambda}}^{(i)} &= \left(\boldsymbol{Q}_1^{(i)}\right)^{-1}\left(\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\xi}}^{(i-1)}\right),\\
\boldsymbol{R}_1^{(i)} &:= \left(\boldsymbol{I}_m \otimes \hat{\boldsymbol{\lambda}}^{(i)}\right)^T\boldsymbol{Q}_A\left(\hat{\boldsymbol{\xi}}^{(i-1)} \otimes \boldsymbol{I}_n\right),\\
\hat{\boldsymbol{\xi}}^{(i)} &:= \left[\left(\boldsymbol{A}^T + \boldsymbol{R}_1^{(i)}\right)\left(\boldsymbol{Q}_1^{(i)}\right)^{-1}\boldsymbol{A}\right]^{-1}\cdot\left(\boldsymbol{A}^T + \boldsymbol{R}_1^{(i)}\right)\left(\boldsymbol{Q}_1^{(i)}\right)^{-1}\boldsymbol{y};
\end{aligned}$$

Step 3: Stop when $\|\hat{\boldsymbol{\lambda}}^{(i)} - \hat{\boldsymbol{\lambda}}^{(i-1)}\| < \delta$ and $\|\hat{\boldsymbol{\xi}}^{(i)} - \hat{\boldsymbol{\xi}}^{(i-1)}\| < \delta$ for a chosen threshold δ.

Note that "Mahboub's algorithm" does not make explicit use of $\tilde{\boldsymbol{e}}_A$ or $\tilde{\boldsymbol{E}}_A$, but requires the inversion of a non-symmetric matrix such as $(\boldsymbol{A}^T + \boldsymbol{R}_1)\boldsymbol{Q}^{-1}\boldsymbol{A}$ in every iteration, the invertibility of which should be ensured since $\boldsymbol{A}$ has full column rank.

On the other hand, in case of "Fang's algorithm," there might be a slim chance for the matrix $\boldsymbol{A} - \tilde{\boldsymbol{E}}_A$ to have a lower rank than $\boldsymbol{A}$ itself, in which case the regular inverse of $(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A)^T\boldsymbol{Q}_1^{-1}(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A)$ might be replaced by the "pseudo-inverse," for instance.

3 The case of singular dispersion matrices $\mathbf{Q}_y$ and $\mathbf{Q}_A$

Interestingly, the TLS solution within an EIV-Model may still be unique even in the case of singular dispersion matrices $\boldsymbol{Q}_y$ and $\boldsymbol{Q}_A$. This possibility had already been established by Schaffrin and Wieser (2008) who, however, had to assume a Kronecker product structure for $\boldsymbol{Q}_A$. Meanwhile, Neitzel and Schaffrin (2013) succeeded in deriving a criterion that indicates such uniqueness of the LESS within the (more general) GH-Model, defined by

$$\boldsymbol{b}(\underbrace{\boldsymbol{Y} - \boldsymbol{e}}_{k\times 1}, \underbrace{\Xi_0 + \boldsymbol{\xi}}_{m\times 1}) = \underset{n\times 1}{\boldsymbol{0}}, \quad \boldsymbol{e} \sim (\boldsymbol{0}, \sigma_0^2\boldsymbol{Q}), \tag{12}$$

where the function $\boldsymbol{b}$ maps $\mathbb{R}^{m+k}$ into $\mathbb{R}^n$. Using $\boldsymbol{\mu}_Y^0$ as an approximation for $\boldsymbol{\mu}_Y := \boldsymbol{Y} - \boldsymbol{e}$, the nonlinear GH-Model Eq. (12) can be linearized into

$$\mathbf{0} \approx \boldsymbol{b}(\boldsymbol{\mu}_Y^0, \Xi_0) + \left.\frac{\partial \boldsymbol{b}(\boldsymbol{\mu}_Y, \Xi)}{\partial \boldsymbol{\mu}_Y^T}\right|_{\boldsymbol{\mu}_Y^0, \Xi_0} \cdot (\boldsymbol{Y} - \boldsymbol{\mu}_Y^0 - \boldsymbol{e}) + \left.\frac{\partial \boldsymbol{b}(\boldsymbol{\mu}_Y, \Xi)}{\partial \Xi^T}\right|_{\boldsymbol{\mu}_Y^0, \Xi_0} \cdot (\Xi - \Xi_0) =:$$
$$=: \boldsymbol{b}(\boldsymbol{\mu}_Y^0, \Xi_0) + \boldsymbol{B} \cdot (\boldsymbol{Y} - \boldsymbol{\mu}_Y^0) - \boldsymbol{B} \cdot \boldsymbol{e} - \boldsymbol{A} \cdot \boldsymbol{\xi}, \tag{13}$$

or, even more compactly, into

$$\boldsymbol{w} := \boldsymbol{b}(\boldsymbol{Y}, \Xi_0) \approx \boldsymbol{b}(\boldsymbol{\mu}_Y^0, \Xi_0) + \boldsymbol{B} \cdot (\boldsymbol{Y} - \boldsymbol{\mu}_Y^0) = \underset{n \times m}{\boldsymbol{A}} \boldsymbol{\xi} + \underset{n \times k}{\boldsymbol{B}} \boldsymbol{e}, \tag{14a}$$
$$\text{with} \quad \boldsymbol{e} \sim (\mathbf{0}, \sigma_0^2 \boldsymbol{Q}). \tag{14b}$$

According to Neitzel and Schaffrin (2013), the least-squares residual vector $\tilde{\boldsymbol{e}}$ within the GH-Model Eqs. (14a) and (14b) will now be unique if and only if the rank condition

$$\operatorname{rk}\left[\boldsymbol{BQ} \mid \boldsymbol{A}\right] = n \tag{15}$$

is fulfilled. In addition, the LESS $\hat{\boldsymbol{\xi}}$ will also be unique if and only if $\operatorname{rk} \boldsymbol{A} = m < n$. The criterion Eq. (15) nicely generalizes a well known result for the Gauss-Markov Model where $\boldsymbol{B}$ is replaced by $\boldsymbol{I}_n$.

The translation of the criterion Eq. (15) to the EIV-Models requires the knowledge of the matrices involved, particularly

$$\underset{n \times k}{\boldsymbol{B}} = \left[-\boldsymbol{I}_n \mid \boldsymbol{\xi}^T \otimes \boldsymbol{I}_n\right], \quad k = n(m+1), \tag{16}$$

for $\boldsymbol{Y} := \operatorname{vec}[\boldsymbol{y} \mid \boldsymbol{A}]$ and $\boldsymbol{e} := \operatorname{vec}[\boldsymbol{e}_y \mid \boldsymbol{E}_A]$, whereas $\boldsymbol{A}$ and $\boldsymbol{Q}$ remain as defined in Eqs. (1a) and (1b), with $\operatorname{rk} \boldsymbol{A} = m$. Hence, for the EIV-Model, the criterion Eq. (15) after adaptation reads

$$n = \operatorname{rk}\left[\boldsymbol{BQ} \mid \boldsymbol{A}\right] = \operatorname{rk}\left[-\boldsymbol{Q}_y \mid (\boldsymbol{\xi} \otimes \boldsymbol{I}_n)^T \boldsymbol{Q}_A \mid \boldsymbol{A}\right] = \operatorname{rk}(\boldsymbol{Q}_1 + \boldsymbol{ASA}^T) \tag{17a}$$

for any symmetric positive-definite matrix $\boldsymbol{S}$ and

$$\boldsymbol{Q}_1 := \boldsymbol{Q}_y + (\boldsymbol{\xi} \otimes \boldsymbol{I}_n)^T \boldsymbol{Q}_A (\boldsymbol{\xi} \otimes \boldsymbol{I}_n) \tag{17b}$$

as in Eq. (8e). Consequently, by introducing the vector

$$\hat{\gamma} := (\hat{\boldsymbol{\xi}} \otimes \hat{\boldsymbol{\lambda}}) = (\boldsymbol{I}_m \otimes \hat{\boldsymbol{\lambda}})\hat{\boldsymbol{\xi}} \quad \text{with} \quad \tilde{\boldsymbol{e}}_A = -\boldsymbol{Q}_A \cdot \hat{\gamma}, \tag{18a-b}$$

the system Eq. (10) can be given the extended form

$$\begin{bmatrix} \boldsymbol{Q}_1 & \mathbf{0} & \boldsymbol{A} \\ \mathbf{0} & -\boldsymbol{Q}_A & \boldsymbol{Q}_A(\boldsymbol{I}_m \otimes \hat{\boldsymbol{\lambda}}) \\ \boldsymbol{A}^T & (\boldsymbol{I}_m \otimes \hat{\boldsymbol{\lambda}})^T \boldsymbol{Q}_A & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\lambda}} \\ \hat{\gamma} \\ \hat{\boldsymbol{\xi}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{y} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tag{19}$$

and finally, after adding

$$\boldsymbol{ASA}^T \cdot \hat{\boldsymbol{\lambda}} + (\boldsymbol{AS} \otimes \hat{\boldsymbol{\lambda}}^T)\boldsymbol{Q}_A \cdot \hat{\gamma} = \boldsymbol{AS}(\boldsymbol{A}^T + \boldsymbol{R}_1)\hat{\boldsymbol{\lambda}} = \mathbf{0} \tag{20a}$$

to the upper part, with $\boldsymbol{R}_1$ from Eq. (11), and introducing

$$\boldsymbol{Q}_2 := \boldsymbol{Q}_1 + \boldsymbol{ASA}^T, \tag{20b}$$

while preserving its symmetry, the equation system reads

$$\begin{bmatrix} \boldsymbol{Q}_2 & (\boldsymbol{AS} \otimes \hat{\boldsymbol{\lambda}}^T)\boldsymbol{Q}_A & \boldsymbol{A} \\ \boldsymbol{Q}_A(\boldsymbol{SA}^T \otimes \hat{\boldsymbol{\lambda}}) & \boldsymbol{Q}_A(\boldsymbol{S} \otimes \hat{\boldsymbol{\lambda}}\hat{\boldsymbol{\lambda}}^T)\boldsymbol{Q}_A - \boldsymbol{Q}_A & \boldsymbol{Q}_A(\boldsymbol{I}_m \otimes \hat{\boldsymbol{\lambda}}) \\ \boldsymbol{A}^T & (\boldsymbol{I}_m \otimes \hat{\boldsymbol{\lambda}}^T)\boldsymbol{Q}_A & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\lambda}} \\ \hat{\gamma} \\ \hat{\boldsymbol{\xi}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{y} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{21}$$

Since $\boldsymbol{Q}_2$ is invertible under the rank condition Eq. (17a), the system Eq. (21) can be solved successively as follows:

$$\hat{\boldsymbol{\lambda}} = \boldsymbol{Q}_2^{-1} \cdot \left[(\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\xi}}) - (\boldsymbol{AS} \otimes \hat{\boldsymbol{\lambda}}^T)\boldsymbol{Q}_A \cdot \hat{\gamma}\right] \Rightarrow \tag{22}$$

$$\Rightarrow \begin{cases} \left[\boldsymbol{Q}_A - \boldsymbol{Q}_A(\boldsymbol{S}\otimes\hat{\boldsymbol{\lambda}}\hat{\boldsymbol{\lambda}}^T)\boldsymbol{Q}_A\right]\cdot\hat{\gamma} + \left[\boldsymbol{Q}_A(\boldsymbol{S}\boldsymbol{A}^T\otimes\hat{\boldsymbol{\lambda}})\boldsymbol{Q}_2^{-1}(\boldsymbol{A}\boldsymbol{S}\otimes\hat{\boldsymbol{\lambda}}^T)\boldsymbol{Q}_A\right]\cdot\hat{\gamma} = \\ \quad = \boldsymbol{Q}_A(\boldsymbol{I}_m\otimes\hat{\boldsymbol{\lambda}})\hat{\boldsymbol{\xi}} + \boldsymbol{Q}_A(\boldsymbol{S}\boldsymbol{A}^T\otimes\hat{\boldsymbol{\lambda}})\boldsymbol{Q}_2^{-1}(\boldsymbol{y}-\boldsymbol{A}\hat{\boldsymbol{\xi}}), \quad \text{and:} \\ \boldsymbol{A}^T\boldsymbol{Q}_2^{-1}(\boldsymbol{y}-\boldsymbol{A}\hat{\boldsymbol{\xi}}) = -(\boldsymbol{I}_m\otimes\hat{\boldsymbol{\lambda}}^T)\boldsymbol{Q}_A\cdot\hat{\gamma} + \boldsymbol{A}^T\boldsymbol{Q}_2^{-1}(\boldsymbol{A}\boldsymbol{S}\otimes\hat{\boldsymbol{\lambda}}^T)\boldsymbol{Q}_A\cdot\hat{\gamma}, \end{cases} \tag{23a, 23b}$$

resulting in the reduced system of equations

$$\left[\begin{array}{c|c} \boldsymbol{A}^T\boldsymbol{Q}_2^{-1}\boldsymbol{A} & -\left[(\boldsymbol{I}_m\otimes\hat{\boldsymbol{\lambda}}) - (\boldsymbol{S}\boldsymbol{A}^T\boldsymbol{Q}_2^{-1}\boldsymbol{A}\otimes\hat{\boldsymbol{\lambda}})\right]^T\boldsymbol{Q}_A \\ -\boldsymbol{Q}_A\left[(\boldsymbol{I}_m\otimes\hat{\boldsymbol{\lambda}}) - (\boldsymbol{S}\boldsymbol{A}^T\boldsymbol{Q}_2^{-1}\boldsymbol{A}\otimes\hat{\boldsymbol{\lambda}})\right] & \boldsymbol{Q}_A\left[(\boldsymbol{S}\boldsymbol{A}^T\boldsymbol{Q}_2^{-1}\boldsymbol{A}\boldsymbol{S}-\boldsymbol{S})\otimes\hat{\boldsymbol{\lambda}}\hat{\boldsymbol{\lambda}}^T\right]\boldsymbol{Q}_A + \boldsymbol{Q}_A \end{array}\right]\begin{bmatrix}\hat{\boldsymbol{\xi}}\\ \hat{\gamma}\end{bmatrix} = \begin{bmatrix}\boldsymbol{A}^T\boldsymbol{Q}_2^{-1}\boldsymbol{y} \\ \boldsymbol{Q}_A(\boldsymbol{S}\boldsymbol{A}^T\boldsymbol{Q}_2^{-1}\boldsymbol{y}\otimes\hat{\boldsymbol{\lambda}})\end{bmatrix}, \tag{24}$$

from which the final identities are obtained as

$$\boldsymbol{A}^T\boldsymbol{Q}_2^{-1}(\boldsymbol{y}-\boldsymbol{A}\hat{\boldsymbol{\xi}}) = -\left[(\boldsymbol{I}_m - \boldsymbol{A}^T\boldsymbol{Q}_2^{-1}\boldsymbol{A}\boldsymbol{S})\otimes\hat{\boldsymbol{\lambda}}^T\right]\boldsymbol{Q}_A\cdot\hat{\gamma} = (\boldsymbol{A}^T\boldsymbol{Q}_2^{-1}\boldsymbol{A}\boldsymbol{S}-\boldsymbol{I}_m)\cdot\left[(\boldsymbol{I}_m\otimes\hat{\boldsymbol{\lambda}})^T\boldsymbol{Q}_A(\boldsymbol{I}_m\otimes\hat{\boldsymbol{\lambda}})\right]\cdot\hat{\boldsymbol{\xi}} =: \boldsymbol{R}_2\cdot\hat{\boldsymbol{\xi}}. \tag{25a}$$

This and, from Eqs. (20a) and (20b), the identity

$$\boldsymbol{y}-\boldsymbol{A}\hat{\boldsymbol{\xi}} = \boldsymbol{Q}_1\hat{\boldsymbol{\lambda}} = (\boldsymbol{Q}_2 + \boldsymbol{A}\boldsymbol{S}\boldsymbol{R}_1)\hat{\boldsymbol{\lambda}} \tag{25b}$$

lead to the novel algorithm of Mahboub type for singular dispersion matrices that fulfill the rank criterion Eq. (17a).

Algorithm 3: (of Mahboub type):

Step 1: Set $\hat{\boldsymbol{\xi}}^{(0)} := (\boldsymbol{A}^T\boldsymbol{Q}_y^+\boldsymbol{A})^+\cdot\boldsymbol{A}^T\boldsymbol{Q}_y^+\boldsymbol{y}$ and $\hat{\boldsymbol{\lambda}}^{(0)} := \boldsymbol{Q}_y^+\cdot\left(\boldsymbol{y}-\boldsymbol{A}\hat{\boldsymbol{\xi}}^{(0)}\right)$;

Step 2: For $i\in\mathbb{N}$ and a chosen matrix $\boldsymbol{S}$, compute

$$\begin{aligned} \boldsymbol{Q}_1^{(i)} &:= \boldsymbol{Q}_y + \left(\hat{\boldsymbol{\xi}}^{(i-1)}\otimes\boldsymbol{I}_n\right)^T\boldsymbol{Q}_A\left(\hat{\boldsymbol{\xi}}^{(i-1)}\otimes\boldsymbol{I}_n\right), \\ \boldsymbol{Q}_2^{(i)} &:= \boldsymbol{Q}_1^{(i)} + \boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^T, \\ \boldsymbol{R}_1^{(i)} &:= \left(\boldsymbol{I}_m\otimes\hat{\boldsymbol{\lambda}}^{(i-1)}\right)^T\boldsymbol{Q}_A\left(\hat{\boldsymbol{\xi}}^{(i-1)}\otimes\boldsymbol{I}_n\right), \\ \hat{\boldsymbol{\lambda}}^{(i)} &:= \left[\boldsymbol{Q}_2^{(i)} + \boldsymbol{A}\boldsymbol{S}\boldsymbol{R}_1^{(i)}\right]^{-1}\left(\boldsymbol{y}-\boldsymbol{A}\hat{\boldsymbol{\xi}}^{(i-1)}\right), \\ \boldsymbol{R}_2^{(i)} &:= \left[\boldsymbol{A}^T\left(\boldsymbol{Q}_2^{(i)}\right)^{-1}\boldsymbol{A}\boldsymbol{S}-\boldsymbol{I}_m\right]\cdot\left[\left(\boldsymbol{I}_m\otimes\hat{\boldsymbol{\lambda}}^{(i)}\right)^T\boldsymbol{Q}_A\left(\boldsymbol{I}_m\otimes\hat{\boldsymbol{\lambda}}^{(i)}\right)\right], \\ \hat{\boldsymbol{\xi}}^{(i)} &:= \left[\boldsymbol{A}^T\left(\boldsymbol{Q}_2^{(i)}\right)^{-1}\boldsymbol{A} + \boldsymbol{R}_2^{(i)}\right]^{-1}\boldsymbol{A}^T\left(\boldsymbol{Q}_2^{(i)}\right)^{-1}\boldsymbol{y}; \end{aligned}$$

Step 3: Stop when $\left\|\hat{\boldsymbol{\lambda}}^{(i)} - \hat{\boldsymbol{\lambda}}^{(i-1)}\right\| < \delta$ and $\left\|\hat{\boldsymbol{\xi}}^{(i)} - \hat{\boldsymbol{\xi}}^{(i-1)}\right\| < \delta$ for a chosen threshold δ.

An alternative algorithm of Fang type may be based on a different modification of the system Eq. (10); by adding

$$\left(\boldsymbol{A}-\tilde{\boldsymbol{E}}_A\right)\boldsymbol{S}\left(\boldsymbol{A}-\tilde{\boldsymbol{E}}_A\right)^T\cdot\hat{\boldsymbol{\lambda}} = \boldsymbol{0}, \tag{26}$$

for any chosen symmetric, positive-definite matrix $\boldsymbol{S}$, to the upper part, first the identities

$$\boldsymbol{Q}_3\cdot\hat{\boldsymbol{\lambda}} := \left[\boldsymbol{Q}_1 + \left(\boldsymbol{A}-\tilde{\boldsymbol{E}}_A\right)\boldsymbol{S}\left(\boldsymbol{A}-\tilde{\boldsymbol{E}}_A\right)^T\right]\cdot\hat{\boldsymbol{\lambda}} = \boldsymbol{y}-\boldsymbol{A}\hat{\boldsymbol{\xi}}, \tag{27a}$$

$$\boldsymbol{0} = \left(\boldsymbol{A}-\tilde{\boldsymbol{E}}_A\right)^T\hat{\boldsymbol{\lambda}} = \left(\boldsymbol{A}-\tilde{\boldsymbol{E}}_A\right)^T\boldsymbol{Q}_3^{-1}\left(\boldsymbol{y}-\boldsymbol{A}\hat{\boldsymbol{\xi}}\right), \tag{27b}$$

result and finally the formal TLS solution

$$\hat{\boldsymbol{\xi}} = \left[\left(\boldsymbol{A}-\tilde{\boldsymbol{E}}_A\right)^T\boldsymbol{Q}_3^{-1}\left(\boldsymbol{A}-\tilde{\boldsymbol{E}}_A\right)\right]^{-1}\left(\boldsymbol{A}-\tilde{\boldsymbol{E}}_A\right)^T\boldsymbol{Q}_3^{-1}\left(\boldsymbol{y}-\tilde{\boldsymbol{E}}_A\hat{\boldsymbol{\xi}}\right), \tag{28}$$

with

$$\hat{\boldsymbol{\lambda}} = \boldsymbol{Q}_3^{-1}\left(\boldsymbol{y}-\boldsymbol{A}\hat{\boldsymbol{\xi}}\right) \tag{29a}$$

and

$$\operatorname{vec}\tilde{\boldsymbol{E}}_A = \tilde{\boldsymbol{e}}_A = -\boldsymbol{Q}_A(\hat{\boldsymbol{\xi}}\otimes\hat{\boldsymbol{\lambda}}). \tag{29b}$$

In the following, the new algorithms 3 and 4 will be applied to a typical problem in geodetic science where indeed $\boldsymbol{Q}_1$ turns out to be singular, but $\boldsymbol{Q}_2$ (and $\boldsymbol{Q}_3$) are nonsingular.

Algorithm 4: (of Fang type):

Step 1: Set $\tilde{\boldsymbol{E}}_A^{(0)} := \boldsymbol{0}$ and $\hat{\boldsymbol{\xi}}^{(0)} := (\boldsymbol{A}^T\boldsymbol{Q}_y^+\boldsymbol{A})^+\boldsymbol{A}^T\boldsymbol{Q}_y^+\boldsymbol{y}$;

Step 2: For $i \in \mathbb{N}$ and a chosen matrix $\boldsymbol{S}$, compute

$$\begin{aligned}
\boldsymbol{Q}_1^{(i)} &:= \boldsymbol{Q}_y + \left(\hat{\boldsymbol{\xi}}^{(i-1)} \otimes \boldsymbol{I}_n\right)^T \boldsymbol{Q}_A \left(\hat{\boldsymbol{\xi}}^{(i-1)} \otimes \boldsymbol{I}_n\right),\\
\boldsymbol{Q}_3^{(i)} &:= \boldsymbol{Q}_1^{(i)} + \left(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)}\right)\boldsymbol{S}\left(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)}\right)^T,\\
\hat{\boldsymbol{\xi}}^{(i)} &:= \left[\left(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)}\right)^T \left(\boldsymbol{Q}_3^{(i)}\right)^{-1} \left(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)}\right)\right]^{-1} \cdot\\
&\quad \cdot \left[\left(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)}\right)^T \left(\boldsymbol{Q}_3^{(i)}\right)^{-1} \left(\boldsymbol{y} - \tilde{\boldsymbol{E}}_A^{(i-1)}\hat{\boldsymbol{\xi}}^{(i-1)}\right)\right],\\
\hat{\boldsymbol{\lambda}}^{(i)} &= \left(\boldsymbol{Q}_3^{(i)}\right)^{-1}\left[\left(\boldsymbol{y} - \tilde{\boldsymbol{E}}_A^{(i-1)}\hat{\boldsymbol{\xi}}^{(i-1)}\right) - \left(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A^{(i-1)}\right)\hat{\boldsymbol{\xi}}^{(i)}\right],\\
\tilde{\boldsymbol{e}}_A^{(i)} &= -\boldsymbol{Q}_A\left(\hat{\boldsymbol{\xi}}^{(i)} \otimes \boldsymbol{I}_n\right)\hat{\boldsymbol{\lambda}}^{(i)} \quad \text{and} \quad \tilde{\boldsymbol{E}}_A^{(i)} = \operatorname{Invec} \tilde{\boldsymbol{e}}_A^{(i)};
\end{aligned}$$

Step 3: Stop when $\left\|\hat{\boldsymbol{\lambda}}^{(i)} - \hat{\boldsymbol{\lambda}}^{(i-1)}\right\| < \delta$ and $\left\|\hat{\boldsymbol{\xi}}^{(i)} - \hat{\boldsymbol{\xi}}^{(i-1)}\right\| < \delta$ for a chosen threshold δ.

4 Example: 2-D similarity transformation

In the 2-D similarity transformation problem, four parameters are estimated for the purpose of transforming estimated coordinates from a source system (here labeled xy-coordinate system) to a target system (here labeled XY-coordinate system). The estimation requires redundant data consisting of observed (or previously estimated) coordinates in both systems at common reference points, together with the associated cofactor (scaled dispersion) matrices $\boldsymbol{Q}_{xy}$ and $\boldsymbol{Q}_{XY}$ for the source and target systems, respectively. Since the observed coordinates and their cofactor matrices come from different sources, it is assumed that there is no cross-correlation between them.

The four parameters for the 2-D similarity transformation are

ξ_0, ξ_1	**for the translation of the coordinate frame,**
α	**for the rotation angle,**
ω	**for the scale factor.**

To transform the problem into a (quasi)linear one, two additional intermediary parameters are defined as $\xi_2 := \omega \cos\alpha$ and $\xi_3 := \omega \sin\alpha$. The vector of unknown parameters to be estimated is then $\boldsymbol{\xi} = \left[\xi_0, \xi_1, \xi_2, \xi_3\right]^T$.

The EIV model for the 2-D similarity transformation with $n/2$ pairs of observed points in both source and target coordinate systems is given by

$$\underset{n\times 1}{\boldsymbol{y}} := \begin{bmatrix} X_1 \\ Y_1 \\ \dots \\ X_{n/2} \\ Y_{n/2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & x_1 & -y_1 \\ 0 & 1 & y_1 & x_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & x_{n/2} & -y_{n/2} \\ 0 & 1 & y_{n/2} & x_{n/2} \end{bmatrix} \begin{bmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} - \begin{bmatrix} 0 & 0 & e_{x_1} & -e_{y_1} \\ 0 & 0 & e_{y_1} & e_{x_1} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & e_{x_{n/2}} & -e_{y_{n/2}} \\ 0 & 0 & e_{y_{n/2}} & e_{x_{n/2}} \end{bmatrix} \begin{bmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} + \begin{bmatrix} e_{X_1} \\ e_{Y_1} \\ \vdots \\ e_{X_{n/2}} \\ e_{Y_{n/2}} \end{bmatrix} = \tag{30a}$$

$$= (\underset{n\times m}{\boldsymbol{A}} - \boldsymbol{E}_A) \cdot \boldsymbol{\xi} + \boldsymbol{e}_y, \quad \text{with} \quad \operatorname{rk} \boldsymbol{A} = m = 4. \tag{30b}$$

The random errors are distributed as

$$\begin{bmatrix} \boldsymbol{e}_y \\ \boldsymbol{e}_A \end{bmatrix} := \begin{bmatrix} \boldsymbol{e}_y \\ \operatorname{vec} \boldsymbol{E}_A \end{bmatrix} \sim \left(\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}, \sigma_0^2 \begin{bmatrix} \boldsymbol{Q}_y & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_A \end{bmatrix} \right), \tag{30c}$$

with

$$\underset{n\times n}{\boldsymbol{Q}_y} := \boldsymbol{Q}_{XY} \quad \text{and} \quad \underset{nm\times nm}{\boldsymbol{Q}_A} := \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{Q}_{A_{33}} & \boldsymbol{Q}_{A_{34}} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{Q}_{A_{43}} & \boldsymbol{Q}_{A_{44}} \end{bmatrix}. \tag{30d}$$

The relationship between the nonzero blocks of $\boldsymbol{Q}_A$ and the cofactor matrix $\boldsymbol{Q}_{xy}$ from the source coordinate-system is determined as follows: Define a 2 × 2 block-diagonal transformation matrix $\boldsymbol{T}$ of dimension $n \times n$ such that $\boldsymbol{a}_4 = \boldsymbol{T}\boldsymbol{a}_3$, where $\boldsymbol{a}_3$ and $\boldsymbol{a}_4$ are the third and fourth columns, respectively, of the data matrix $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4]$. The matrix $\boldsymbol{T}$ is then given by

$$\underset{n\times n}{\boldsymbol{T}} := \operatorname{Diag}(\boldsymbol{T}', \dots, \boldsymbol{T}'), \quad \text{with} \quad \boldsymbol{T}' := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \tag{31}$$

where the matrix $\boldsymbol{T}'$ obviously occurs $n/2$ times in the Diag argument. Note that $\boldsymbol{T}$ is orthogonal, and thus $\boldsymbol{T}^T = \boldsymbol{T}^{-1}$ and $(\boldsymbol{T}^T)^{-1} = \boldsymbol{T}$.

Applying the law of variance propagation leads to the following expressions for the non-zero blocks of $\boldsymbol{Q}_A$ in terms of $\boldsymbol{Q}_{xy}$:

$$\underset{n\times n}{\boldsymbol{Q}_{A_{33}}} = \boldsymbol{Q}_{xy} = \boldsymbol{Q}_{xy}^T, \ \underset{n\times n}{\boldsymbol{Q}_{A_{34}}} = \boldsymbol{Q}_{xy}\boldsymbol{T}^T,$$
$$\underset{n\times n}{\boldsymbol{Q}_{A_{43}}} = \boldsymbol{T}\boldsymbol{Q}_{xy}, \ \underset{n\times n}{\boldsymbol{Q}_{A_{44}}} = \boldsymbol{T}\boldsymbol{Q}_{xy}\boldsymbol{T}^T. \tag{32}$$

With the help of Eq. (32), the matrix of combined cofactors

$$\boldsymbol{Q}_1 = \boldsymbol{Q}_y + (\hat{\boldsymbol{\xi}} \otimes \boldsymbol{I}_n)^T \boldsymbol{Q}_A (\hat{\boldsymbol{\xi}} \otimes \boldsymbol{I}_n) \tag{33}$$

can be readily expressed in terms of the cofactor matrices $\boldsymbol{Q}_{XY}$ and $\boldsymbol{Q}_{xy}$ as

$$\boldsymbol{Q}_1 = \boldsymbol{Q}_{XY} + \hat{\xi}_2^2 \cdot \boldsymbol{Q}_{xy} + \hat{\xi}_2\hat{\xi}_3 \cdot (\boldsymbol{T}\boldsymbol{Q}_{xy} + \boldsymbol{Q}_{xy}\boldsymbol{T}^T) + \hat{\xi}_3^2 \cdot \boldsymbol{T}\boldsymbol{Q}_{xy}\boldsymbol{T}^T, \tag{34}$$

where $\hat{\xi}_2$ and $\hat{\xi}_3$ are the third and fourth elements, respectively, of the (weighted) TLS estimator $\hat{\boldsymbol{\xi}}$.

As shown in Snow (2012), the total sum of squared residuals (TSSR) is computed by

$$\Omega = \left\| \begin{bmatrix} \tilde{\boldsymbol{e}}_y^T, & \tilde{\boldsymbol{e}}_A^T \end{bmatrix}^T \right\|_{\boldsymbol{P}}^2 = \begin{bmatrix} \tilde{\boldsymbol{e}}_y^T, & \tilde{\boldsymbol{e}}_A^T \end{bmatrix} \boldsymbol{P} \begin{bmatrix} \tilde{\boldsymbol{e}}_y \\ \tilde{\boldsymbol{e}}_A \end{bmatrix} = \tag{35a}$$
$$= \hat{\boldsymbol{\lambda}}^T \boldsymbol{Q}_1 \hat{\boldsymbol{\lambda}} = \tag{35b}$$
$$= \hat{\boldsymbol{\lambda}}^T (\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\xi}}). \tag{35c}$$

Then a suitable approximation for the estimated variance component $\hat{\sigma}_0^2$ is given by dividing the TSSR by the model degrees of freedom (or redundancy) as in

$$\hat{\sigma}_0^2 = \Omega / r, \tag{36}$$

where the redundancy r is defined as $r := n - \operatorname{rk}\boldsymbol{A}$, or as $r = n - m$ if $\boldsymbol{A}$ has full column rank m as assumed here (cf. Schaffrin et al., 2012).

The data for the 2-D similarity transformation are taken from Neitzel and Schaffrin (2013). They are comprised of 2-D coordinates of five stations from both the source and target systems, together with their associated cofactor (scaled dispersion) matrices, both of which are fully populated and singular. It is noted again that the source and target data are not correlated with each other. The coordinates are listed in Table 1. The cofactor matrices are listed in Snow (2012).

Both 10 × 10 source and target cofactor matrices $\boldsymbol{Q}_{xy}$, resp. $\boldsymbol{Q}_{XY}$ have the rank of seven, and, when incorporated into the 10 × 10 matrix $\boldsymbol{Q}_1$ of Eq. (34), the resulting matrix $\boldsymbol{Q}_1$ is found to have rank eight. However, the adapted Neitzel/Schaffrin rank condition Eq. (17a) is still satisfied. Thus the problem can be solved with either Algorithm 3 or Algorithm 4.

Using a convergence tolerance $\delta = 1.0 \times 10^{-12}$ for Step 3, Algorithm 3 converged in five iterations, whereas Algorithm 4 took four iterations to converge. Both algorithms generated the same solution, at least to the precision shown in Tables 2 and 3, which list the estimated parameters and predicted residuals, respectively. A TSSR value of $\Omega = 6.1640345$ was computed using Eq. (35b). Considering the system redundancy of $r = 6$, the estimated variance component is then computed using Eq. (36) as $\hat{\sigma}_0^2 = \Omega/r = (1.0135774)^2 = 1.027339$.

In addition to the tabulated residuals in Table 3, the total predicted error matrix

$$\begin{bmatrix} \tilde{\boldsymbol{e}}_y \mid \tilde{\boldsymbol{E}}_A \end{bmatrix} = \left[\begin{array}{r|rrrr} 1.0204 & 0 & 0 & -4.4026 & 5.3231 \\ 0.8998 & 0 & 0 & -5.3231 & -4.4026 \\ 0.3453 & 0 & 0 & -1.8617 & -0.5454 \\ -0.1634 & 0 & 0 & 0.5454 & -1.8617 \\ -1.5805 & 0 & 0 & 7.1386 & -6.2318 \\ -0.9923 & 0 & 0 & 6.2318 & 7.1386 \\ 1.0399 & 0 & 0 & -4.2616 & 6.8490 \\ 1.2009 & 0 & 0 & -6.8490 & -4.2616 \\ -0.8250 & 0 & 0 & 3.3873 & -5.3948 \\ -0.9450 & 0 & 0 & 5.3948 & 3.3873 \end{array}\right] \text{mm}$$

reveals interesting features of the weighted TLS algorithms. A comparison between this matrix and equation

Table 1. Coordinate estimates in source and target systems

Point No.	x_i[m]	y_i[m]	X_i[m]	Y_i[m]
1	453.8001	137.6099	400.0040	100.0072
2	521.2865	350.7972	500.0019	299.9994
3	406.8728	433.9247	399.9925	399.9933
4	110.5545	386.9880	100.0059	400.0022
5	157.4861	90.6802	99.9956	99.9978

Table 2. Weighted TLS estimates for the 2-D similarity transformation, generated by both Algorithms 3 and 4

Parameter	Estimated value
x-shift ξ_0	−69.726354 m
y-shift ξ_1	35.078215 m
$\xi_2 = \omega \cos \alpha$	0.98765502
$\xi_3 = \omega \sin \alpha$	−0.15642921
scale factor ω	0.99996626
rotation angle α	−10.00000154 gon
var. component σ_0^2	1.027339

Table 3. Weighted TLS residual predictions for the 2-D similarity transformation, generated by both Algorithms 3 and 4

Point	Target System		Source System	
	$\tilde{e}_Y$ [mm]	$\tilde{e}_X$ [mm]	$\tilde{e}_y$ [mm]	$\tilde{e}_x$ [mm]
1	0.8998	1.0204	−5.3231	−4.4026
2	−0.1634	0.3453	0.5454	−1.8617
3	−0.9923	−1.5805	6.2318	7.1386
4	1.2009	1.0399	−6.8490	−4.2616
5	−0.9450	−0.8250	5.3948	3.3873

Eq. (30a) shows that the structure of the data matrix $\boldsymbol{A}$ has been replicated exactly in the residual matrix $\tilde{\boldsymbol{E}}_A$. The first two columns of both matrices contain only zeros. The structure of the last two columns of $\tilde{\boldsymbol{E}}_A$ is highlighted by drawing a box around the first two rows. This replication of structure in the residual matrix had already been pointed out by Fang (2011), Mahboub (2012), and Schaffrin et al. (2012), for EIV-models with dispersion matrices having full rank. The property holds here as well in the new algorithms that handle rank-deficient dispersion matrices.

The points are plotted in a 2-D map in Fig. 1, where the dotted lines represent a grid for the adjusted coordinates $(\boldsymbol{x} - \tilde{\boldsymbol{e}}_x, \boldsymbol{y} - \tilde{\boldsymbol{e}}_y)$ in the source system, and the dash-dotted lines represent a (rotated) grid for the adjusted coordinates $(\boldsymbol{X} - \tilde{\boldsymbol{e}}_X, \boldsymbol{Y} - \tilde{\boldsymbol{e}}_Y)$ in the target system. The origin of the source system has coordinates $(\hat{\xi}_0, \hat{\xi}_1)$ in the target system.

5 Conclusions and outlook

In this contribution, we have presented two novel algorithms that can compute the TLS estimator within the EIV-Model with singular dispersion, resp. cofactor matrices, provided that the adapted Neitzel/Schaffrin criterion Eq. (17a) for uniqueness is satisfied. The new TLS algorithms, of type Fang and Mahboub, respectively, extend opportunities for usage of the EIV-model in the presence of singular dispersion matrices, and the fact that the presented algorithms are valid for both cases of singular and nonsingular dispersion matrices makes them quite flexible.

The addition of a second variance component to the EIV-Model should be explored in future work. Such an extension to the model would be particularly useful in the case where the relative precision between the dispersion matrices for the observation vector and the data matrix is not well known.

Whether the extension of the methodology, as presented here, to more general “measurement error models” turns out promising (or rather not), may also be decided after further investigations.

Acknowledgement: The first author is particularly grateful for the repeated support that he received through the Alexander-von-Humboldt Foundation while working on this contribution. The third author also thanks the German Science Foundation (DFG) for their financial support.

References

Fang, X., 2011, *Weighted Total Least-Squares Solutions for Applications in Geodesy*, PhD dissertation, Publ. No. 294, Dept. of Geodesy and Geoinformatics, Leibniz University Hannover, Germany.

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

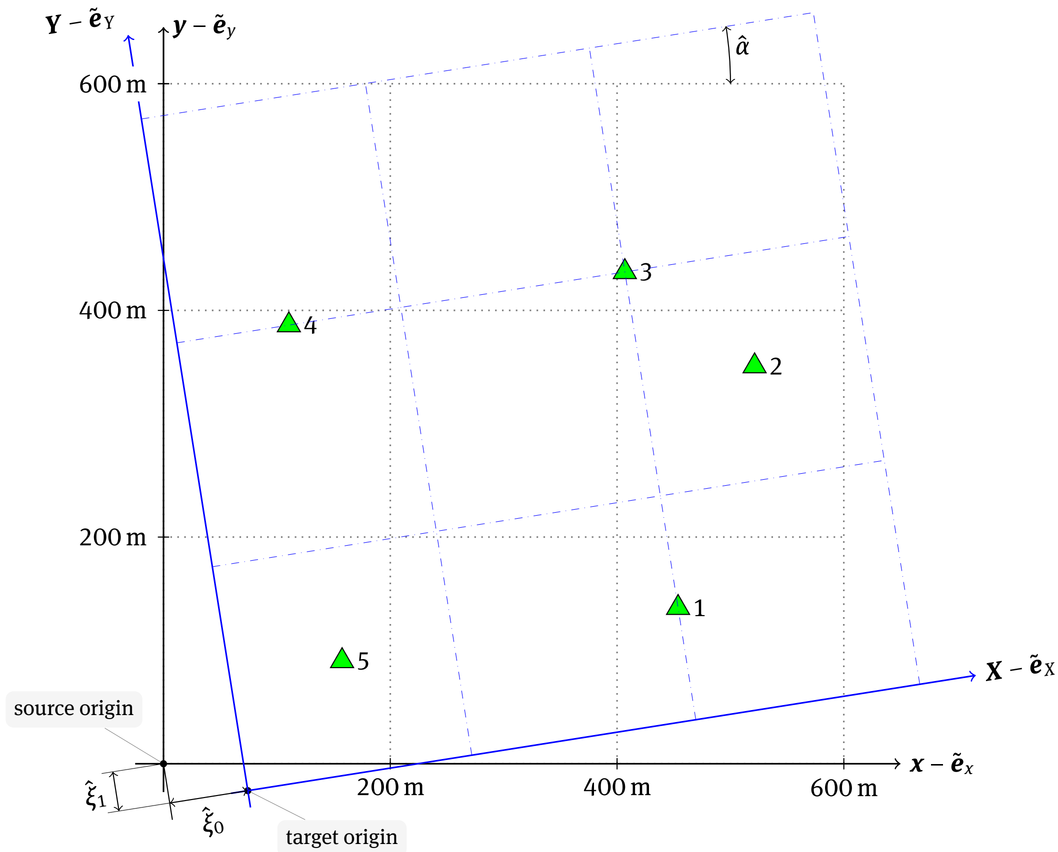

Fig. 1. Map view of the five data points before and after the 2-D similarity transformation. The dotted lines represent a grid for the adjusted coordinates ($\boldsymbol{x} - \tilde{\boldsymbol{e}}_x, \boldsymbol{y} - \tilde{\boldsymbol{e}}_y$) in the source system, and the dash-dotted lines represent a (rotated) grid for the adjusted coordinates ($\boldsymbol{X} - \tilde{\boldsymbol{e}}_X, \boldsymbol{Y} - \tilde{\boldsymbol{e}}_Y$) in the target system. The origin of the source system has coordinates ($\hat{\xi}_0, \hat{\xi}_1$) in the target system. The grid interval for both grids is 200 meters.

Helmert, F. R., 1907, *Adjustment Computations by the Method of Least Squares* (in German), Teubner: Leipzig, Germany, second edition.

Magnus, J. R. and Neudecker, H., 2007, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley: Chichester, England, third edition.

Mahboub, V., 2012, On weighted total least-squares for geodetic transformations, *Journal of Geodesy*, 86(5):359–367.

Markovsky, I., Rastello, M. L., Premoli, A., Kukush, A., and Van Huffel, S., 2006, The element-wise weighted total least-squares problem, *Computational Statistics & Data Analysis*, 50(1):181–209.

Neitzel, F. and Schaffrin, B., 2013, On the Gauss-Helmert Model with a singular dispersion matrix where ***BQ*** is of lesser row rank than ***B***, *Studia Geophysica et Geodaetica,* revised in June 2013.

Pearson, K., 1901, LIII. On lines and planes of closest fit to systems of points in space, *Philosophical Magazine, Series 6*, 2(11):559–572.

Schaffrin, B., Neitzel, F., Uzun, S., and Mahboub, V., 2012, Modifying Cadzow's algorithm to generate the optimal TLS-solution for the structured EIV-model of a similarity transformation, *Journal of Geodetic Science*, 2(2):98–106.

Schaffrin, B. and Wieser, A., 2008, On weighted total least-squares adjustment for linear regression, *Journal of Geodesy*, 82(7):415–421.

Snow, K., 2012, *Topics in Total Least-Squares Adjustment within the Errors-In-Variables Model: Singular Cofactor Matrices and Prior Information*, PhD dissertation, Report No. 502, Div. of Geodetic Science, School of Earth Sciences, The Ohio State Univ., Columbus/OH.

Xu, P., Liu, J., and Shi, C., 2012, Total least squares adjustment in partial errors-in-variables models: algorithm and statistical analysis, *Journal of Geodesy*, 86(8):661–675.

Received July 25, 2013 ; accepted February 18, 2013.