Acta Geod Geophys (2017) 52:479–496
DOI 10.1007/s40328-016-0184-2

ORIGINAL STUDY

Adjusting a 2D Helmert transformation within a Gauss–Helmert Model with a singular dispersion matrix where *BQ* is of smaller rank than *B*

Frank Neitzel[1,2] · **Burkhard Schaffrin**[2]

Received: 31 March 2016 / Accepted: 26 August 2016 / Published online: 31 October 2016

Abstract The case of a *singular* dispersion matrix within the Gauss–Helmert Model has been considered before, most recently even allowing the rank of *BQ* to be smaller than the rank of *B*. In this contribution the emphasis is shifted towards an illuminating example, the 2D Helmert transformation.

Keywords Gauss–Helmert Model · Singular dispersion matrix · Singular variance–covariance matrix · 2D Helmert transformation

1 Introduction

In a recent contribution, the Gauss–Helmert Model with *singular* dispersion matrix has been analyzed once more, but with the emphasis on *necessary and sufficient* conditions for the existence of a *unique solution* for both the residual vector as well as the estimated parameter vector. Unlike earlier work by Bjerhammar (1973), Wolf (1979) or Perelmuter (1981), and others, the contribution by Neitzel and Schaffrin (2016) no longer *assumed* that the rank deficiency was small enough to guarantee a unique solution, which is certainly the case if $rk\, BQ = rk\, B$. If $rk\, BQ < rk\, B$, however, the rank condition $rk[A|BQ] = rk\; B = r + q$ must be fulfilled in order for a unique solution of type BLUMBE (Best Linear Uniformly Minimum Bias Estimate) to exist according to Neitzel and Schaffrin (2016, Theorem 2.2).

In the following, after a short summary of the key results when $rk\, BQ < rk\, B$, the 2D Helmert transformation is being chosen as an application with some relevance, thereby illuminating the hidden relationships that ought to be fulfilled if meaningful results are

✉ Frank Neitzel
frank.neitzel@tu-berlin.de

[1] Institute of Geodesy and Geoinformation Science, Technische Universität, Berlin, Germany

[2] Geodetic Science Program, School of Earth Sciences, The Ohio State University, Columbus, OH, USA

expected. For earlier discussions of this application, see, e.g., Teunissen (1988), Bleich and Illner (1989), Koch et al. (2000), Fang (2014), or Chang (2015) among many others. For an alternative approach, see Schaffrin (2003), as well as Schaffrin et al. (2014).

2 The Gauss–Helmert Model with singular dispersion matrix: A short summary when rk BQ < rk B

In the following, key results from Neitzel and Schaffrin (2016) are summarized. Let us assume the (linearized) *Gauss–Helmert Model*

$$w = A\xi + Be, \quad e \sim \left(0,\ \sigma_0^2 Q\right), \tag{1.1a}$$

with
w as $(r + q) \times 1$ vector of so-called "mis-closures",
ξ as $m \times 1$ vector of (unknown) parameters,
e as $n \times 1$ vector of random observation errors,
A as $(r + q) \times m$ coefficient matrix with $q := rk\ A$,
B as $(r + q) \times n$ condition matrix with $r + q := rk\ B$
(not restricting the generality);
$r := rk\ B - rk\ A$ is called "redundancy".

Furthermore, the *expectation* of e is zero, $E\{e\} = 0$, and its *dispersion matrix* is given by $D\{e\} = \sigma_0^2 Q$; here, σ_0^2 is the (unknown) variance component, and Q denotes the $n \times n$ symmetric and *positive-semidefinite* cofactor matrix with $rk\ Q := t < n$. Since Q is *singular*, the theorem of Aitken (1935) is no longer applicable, according to which a weighted least-squares approach with the inverse cofactor matrix as weight matrix would provide the Best Linear Uniformly Unbiased Estimate (BLUUE) of the vector $A\xi$; for more details, see Grafarend and Schaffrin (1993, Chap. 3(a)).

To ensure the consistency of model (1.1a), it is further assumed that

$$w \in \mathcal{R}([A|BQ]) \text{ with probability } 1; \tag{1.1b}$$

here, $\mathcal{R}$ denotes the "*range space*" (or "column space") of a matrix.

For a *linear estimate* of type

$$\hat{\xi} = Lw + \kappa \tag{1.2}$$

with *unknown* $m \times (r + q + 1)$ matrix $[L, \kappa]$, the *bias vector* is defined as

$$\beta := E\left\{\hat{\xi} - \xi\right\} = (LA - I_m)\xi + \kappa, \tag{1.3}$$

which involves the unknown, but arbitrary, vector ξ. If ξ is known to belong to the range space of a certain (symmetric nonnegative-definite) matrix S,

$$\xi \in \mathcal{R}(S) \quad \text{with } rk(AS) = rk\ A = q, \tag{1.4}$$

it makes sense to minimize the expected bias vector (1.3) by setting

$$\kappa := 0 \tag{1.5a}$$

and replacing, in the MSE-matrix

$$MSE\{\hat{\xi}\} = \sigma_0^2\left[\left(LBQB^{\mathrm{T}}L^{\mathrm{T}}\right) + (I_m - LA)\xi\sigma_0^{-2}\xi^{\mathrm{T}}(I_m - LA)^{\mathrm{T}}\right], \tag{1.5b}$$

the unknown rank-1 matrix $(\xi\sigma_0^{-2}\xi^{\mathrm{T}})$ by the known matrix S itself, thereby minimizing

$$\sigma_0^2 \cdot tr(I_m - LA)S(I_m - LA)^{\mathrm{T}} = \min_{L^{\mathrm{T}}} \tag{1.5c}$$

uniformly over $\mathcal{R}(S)$; obviously, $S := S_{pd}$ could be *positive-definite* in which case it holds:

$$\mathcal{R}\left(S_{pd}\right) = \mathbb{R}^m. \tag{1.5d}$$

(Obviously, the case where $\kappa \neq 0$ deserves investigation, too.)

It is noted that the condition (1.4) does not permit the rank-deficiency of S to exceed $(m - q)$ since, otherwise, the rank of AS would fall below q automatically. Thus, if ξ can be restricted to an even lower-dimensional subspace, other techniques ought to be applied. Now, the variational principle (1.5c) readily leads to the (necessary) equation system

$$\left(ASA^{\mathrm{T}}\right) \cdot L^{\mathrm{T}} = AS, \tag{1.6}$$

which turns out to be sufficient as well, thanks to the *nonnegative-definite* matrix S. All the estimates of type $\hat{\xi} = Lw$ where L^{T} fulfills (1.6) constitute the class of *Linear S-Uniformly Minimum Biased Estimators* of ξ (i.e., *S-LUMBE*). In this class, the "Best" estimate (or S-BLUMBE) is formed by minimizing the S-modified Mean Squared Error of $\hat{\xi}$ on average, namely by solving the variational problem

$$tr\ MSE_S\left\{\hat{\xi}_{BLUMBE}\right\} := \sigma_0^2 \cdot tr\ \left(LBQB^{\mathrm{T}}L^{\mathrm{T}}\right) + \sigma_0^2 \cdot tr\ (I_m - LA)S(I_m - LA)^{\mathrm{T}} = \min_{L^{\mathrm{T}}} \tag{1.7}$$

or, equivalently, by making the Lagrange target function

$$\Phi\left(L^{\mathrm{T}},\ \Lambda\right) := tr\ \left(LBQB^{\mathrm{T}}L^{\mathrm{T}}\right) + 2tr\ \Lambda^{\mathrm{T}}\left(ASA^{\mathrm{T}}L^{\mathrm{T}} - AS\right) \tag{1.8}$$

stationary. Thus, the resulting *necessary conditions* read:

$$BQB^{\mathrm{T}} \cdot L^{\mathrm{T}} + ASA^{\mathrm{T}} \cdot \Lambda \doteq 0 \tag{1.9a}$$

$$ASA^{\mathrm{T}} \cdot L^{\mathrm{T}} - AS \doteq 0 \tag{1.9b}$$

while the *sufficient condition* holds true since the matrix $BQB^{\mathrm{T}} \otimes I_m$ is *positive-definite*; here, $\otimes$ denotes the "Kronecker-Zehfuss product" of matrices (Grafarend and Schaffrin 1993, p. 409). For more details, see, e.g., Schaffrin (1989).

In the following, Q might be an arbitrary *symmetric and positive-semidefinite* (thus singular) matrix. The key problem is then concerned with the *unique invertibility* of the system (1.9a–b) in which case unique estimates for ξ result. This does, however, not necessarily imply a unique residual vector unless an interpretation as weighted LEast-Squares Solution (LESS) is possible. The key results of Neitzel and Schaffrin (2016) are now summarized in:

Corollary 1.1:

(i) *In the Gauss–Helmert Model* (1.1) *under condition* (1.4) *the system* (1.9a–b) *has a unique solution for L if and only if*

$$rk[A|BQ] = r + q = rk\ B. \tag{1.10}$$

In this case, the S-BLUMBE of ξ *exists uniquely and is represented by*

$$\hat{\hat{\xi}}_{BLUMBE} = SA^{\mathrm{T}}\left[ASA^{\mathrm{T}}\left(ASA^{\mathrm{T}} + BQB^{\mathrm{T}}\right)^{-1}ASA^{\mathrm{T}}\right]^{-}ASA^{\mathrm{T}}\left(ASA^{\mathrm{T}} + BQB^{\mathrm{T}}\right)^{-1}w \tag{1.11}$$

for any g-inverse $\left[ASA^{\mathrm{T}}(ASA^{\mathrm{T}} + BQB^{\mathrm{T}})^{-1}ASA^{\mathrm{T}}\right]^{-}$ *with the dispersion matrix*

$$D\left\{\hat{\hat{\xi}}_{BLUMBE}\right\} = \sigma_0^2 \cdot SA^{\mathrm{T}}\left\{\left[ASA^{\mathrm{T}}\left(ASA^{\mathrm{T}} + BQB^{\mathrm{T}}\right)^{-1}ASA^{\mathrm{T}}\right]^{-} - \left(ASA^{\mathrm{T}}\right)^{+}\right\}AS, \tag{1.12}$$

and the minimized bias vector

$$\beta = -\left[I_m - SA^{\mathrm{T}}\left(ASA^{\mathrm{T}}\right)^{+}A\right] \cdot \xi \tag{1.13}$$

such that the S-modified Mean Squared Error matrix of $\hat{\hat{\xi}}_{BLUMBE}$ *results in*

$$MSE_S\left\{\hat{\hat{\xi}}_{BLUMBE}\right\} = D\left\{\hat{\hat{\xi}}_{BLUMBE}\right\} + \sigma_0^2\left[S - SA^{\mathrm{T}}\left(ASA^{\mathrm{T}}\right)^{+}AS\right]. \tag{1.14}$$

For the rank of the above matrices, it holds:

$$rk\ D\left\{\hat{\hat{\xi}}_{BLUMBE}\right\} = rk\ A + rk(BQ) - rk[A|BQ] = rk(BQ) - r, \tag{1.15}$$

$$\begin{aligned} rk\ MSE_s\left\{\hat{\hat{\xi}}_{BLUMBE}\right\} &= rk\ D\left\{\hat{\hat{\xi}}_{BLUMBE}\right\} + rk\left[I_m - SA^{\mathrm{T}}\left(ASA^{\mathrm{T}}\right)^{+}A\right]S \\ &= rk(BQ) + rk\ S - (r+q). \end{aligned} \tag{1.16}$$

(ii) *In the special case that* $q = rk\ A = m$, *the system* (1.9a–b) *turns into the system*

$$\begin{bmatrix} BQB^{\mathrm{T}} & A \\ A^{\mathrm{T}} & 0 \end{bmatrix}\begin{bmatrix} L^{\mathrm{T}} \\ \Lambda \end{bmatrix} = \begin{bmatrix} 0 \\ I_m \end{bmatrix}, \tag{1.17}$$

which has a unique solution if and only if the rank condition (1.10) *is fulfilled. In this case, the BLUUE of* ξ *exists uniquely and is represented by*

$$\hat{\hat{\xi}}_{BLUUE} = \left[A^{\mathrm{T}}\left(ASA^{\mathrm{T}} + BQB^{\mathrm{T}}\right)^{-1}A\right]^{-1}A^{\mathrm{T}}\left(ASA^{\mathrm{T}} + BQB^{\mathrm{T}}\right)^{-1}w \tag{1.18}$$

for any arbitrary symmetric and nonnegative-definite matrix S with $rk(AS) = rk\ A$ *as in* (1.4). *Its dispersion matrix is given by*

$$D\left\{\hat{\hat{\xi}}_{BLUUE}\right\} = \sigma_0^2 \cdot \left(\left[A^{\mathrm{T}}\left(ASA^{\mathrm{T}} + BQB^{\mathrm{T}}\right)^{-1}A\right]^{-1} - S\right) = MSE\left\{\hat{\hat{\xi}}_{BLUUE}\right\}, \tag{1.19}$$

which coincides with the Mean Squared Error matrix of $\hat{\hat{\xi}}_{BLUUE}$ *and has the rank*

$$rk\ D\left\{\hat{\hat{\xi}}_{BLUMBE}\right\} = rk(BQ) - r = rk\ MSE\left\{\hat{\hat{\xi}}_{BLUMBE}\right\}. \tag{1.20}$$

It is obvious that *Corollary 1.1*, in particular, establishes the *rank inequality*

$$q + r \geq rk(BQ) \geq rk[A|BQ] - rk\, A = r \tag{1.21}$$

as *necessary condition* for the *unique* existence of the matrix L for $\hat{\hat{\xi}}_{BLUMBE} = Lw$, as well as for $\hat{\hat{\xi}}_{BLUUE} = Lw$, in the general case of a singular dispersion matrix Q. Note that the uniqueness in *Corollary 1.1* has only been established "with probability 1", thanks to the consistency condition (1.1b).

Now, in order to recover the residual vector $\tilde{e}$ or, at least, the transformed residual vector $B\tilde{e} = w - A\hat{\hat{\xi}}$, along with the quadratic form Ω, an *equivalent interpretation* of the above BLUMBE/BLUUE approach by means of weighted least-squares adjustment is suggested. This proved possible along the lines of *Theorem 3.20* in Grafarend and Schaffrin (1993) in the case of a *positive-definite* choice for the matrix

$$S := S_{pd} \tag{1.22}$$

such that S_{pd}^{-1} exists. Again, the results of Neitzel and Schaffrin (2016) are summarized in:

Corollary 1.2:

(i) *In the Gauss–Helmert Model* (1.1) *under condition* (1.10), *any* $B^{\mathrm{T}}(BQB^{\mathrm{T}})^{-}B$-*weighted LESS of* ξ *fulfills the normal equation system*

$$\begin{bmatrix} BQB^{\mathrm{T}} & A \\ A^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} \hat{v} \\ \hat{\hat{\xi}}_{LESS} \end{bmatrix} = \begin{bmatrix} w \\ 0 \end{bmatrix} \tag{1.23}$$

independent of the g-inverse $(BQB^{\mathrm{T}})^{-}$. *If the residual vector* $\tilde{e}$ *is assumed to belong to the range space of* Q, *just like* e *itself belongs to* $\mathcal{R}(Q)$ *with probability 1, then the auxiliary* $(r + q) \times 1$ *vector* $\hat{v}$ *is obtained uniquely, and fulfills the formula*

$$\hat{v} = \left(AS_{pd}A^{\mathrm{T}} + BQB^{\mathrm{T}}\right)^{-1} B\tilde{e} \tag{1.24}$$

with

$$0 = A^{\mathrm{T}}\hat{v} = A^{\mathrm{T}}\left(AS_{pd}A^{\mathrm{T}} + BQB^{\mathrm{T}}\right)^{-1}\left(w - A\hat{\hat{\xi}}_{LESS}\right). \tag{1.25}$$

The corresponding residual vector $\tilde{e}$ *can now also be recovered uniquely for any* $\hat{\hat{\xi}}_{LESS}$ *under the further restriction* $\tilde{e} \in \mathcal{R}(QB^{\mathrm{T}})$ *as*

$$\tilde{e} = QB^{\mathrm{T}} \cdot \hat{v} = QB^{\mathrm{T}}\left(AS_{pd}A^{\mathrm{T}} + BQB^{\mathrm{T}}\right)^{-1}\left(w - A\hat{\hat{\xi}}_{LESS}\right), \tag{1.26}$$

and its weighted quadratic form as

$$\begin{aligned} \Omega &= \tilde{e}^{\mathrm{T}}B^{\mathrm{T}}\left(BQB^{\mathrm{T}}\right)^{-}B\tilde{e} = \tilde{e}^{\mathrm{T}}B^{\mathrm{T}}\hat{v} = \\ &= \left(w - A\hat{\hat{\xi}}\right)^{\mathrm{T}}\hat{v} = w^{\mathrm{T}}\hat{v} - \hat{\hat{\xi}}^{\mathrm{T}}\left(A^{\mathrm{T}}\hat{v}\right) = w^{\mathrm{T}}\hat{v} = \hat{\sigma}_0^2(rk\, B - rk\, A), \end{aligned} \tag{1.27}$$

thereby leading to a suitable estimate of σ_0^2.

(ii) *In the special case that* $q = rk\ A = m$, *the BLUUE of* ξ *can be interpreted equivalently as* $B^{\mathrm{T}}(BQB^{\mathrm{T}})^{-}B$-*weighted LESS as long as* $\tilde{e} \in \mathcal{R}(Q)$ *is assumed. If, moreover,* $\tilde{e} \in \mathcal{R}(QB^{\mathrm{T}})$ *can be assumed, then the residual vector is represented by* (1.26) *with*

$$\hat{\hat{\xi}}_{LESS} = \left[A^{\mathrm{T}}\left(AS_{pd}A^{\mathrm{T}} + BQB^{\mathrm{T}}\right)^{-1}A\right]^{-1}A^{\mathrm{T}}\left(AS_{pd}A^{\mathrm{T}} + BQB^{\mathrm{T}}\right)^{-1}w, \tag{1.28}$$

while its weighted quadratic form Ω *is obtained from* (1.27) *along with a suitable variance component estimate* $\hat{\sigma}_0^2$.
In addition, the respective dispersion matrices can be taken uniquely from

$$\begin{bmatrix} D\{\hat{v}\} & \times \\ \times & -D\{\hat{\hat{\xi}}_{LESS}\} \end{bmatrix} = \sigma_0^2 \begin{bmatrix} BQB^{\mathrm{T}} & A \\ A^{\mathrm{T}} & 0 \end{bmatrix}^{-1} \tag{1.29}$$

with covariance $C\{\hat{v},\ \hat{\hat{\xi}}_{LESS}\} = 0$ *and*

$$D\{\tilde{e}\} = QB^{\mathrm{T}} \cdot D\{\hat{v}\} \cdot BQ, \tag{1.30}$$

where σ_0^2 *may be replaced by its estimate* $\hat{\sigma}_0^2$ *in accordance with* (*1.27*).

Neitzel and Schaffrin (2016) already pointed out that it is not so easy to characterize all the other solutions for $\tilde{e}$ that solve the identity $B\tilde{e} = w - A\hat{\hat{\xi}}$, but may not belong to the range space $\mathcal{R}(QB^{\mathrm{T}}) \subset \mathcal{R}(Q)$. The answer to this question had to be left to a future publication.

After having summarized the extended analysis for the Gauss–Helmert Model with *positive-semidefinite* dispersion matrix Q, the various situations will be illustrated by applying the above results to the case of a 2D Helmert transformation.

3 Application to the 2D Helmert transformation

In the following, the over-determined 2D similarity transformation will be considered, commonly known as symmetric Helmert transformation. The functional model can be based on four parameters, namely:

ξ_0, ξ_1 for the translation of the origin of the frame,
α for the rotation angle, and
ω for the scale factor.

The transformation is then described approximately by

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} \approx \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \omega & 0 \\ 0 & \omega \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} \xi_0 \\ \xi_1 \end{bmatrix}, \tag{2.1}$$

where

(x_i, y_i) are the observed coordinates in the ("old") *source system*, and
(X_i, Y_i) are the observed coordinates in the ("new") *target system*;
i denotes the point number ($i = 1,\ldots,n/2$).

After executing the multiplications, (2.1) turns into

$$X_i \approx (\omega\cos\alpha)x_i - (\omega\sin\alpha)y_i + \xi_0, \tag{2.2a}$$

$$Y_i \approx (\omega\sin\alpha)x_i + (\omega\cos\alpha)y_i + \xi_1, \tag{2.2b}$$

and, with the substitutions

$$\xi_2 := \omega\cos\alpha, \quad \xi_3 := \omega\sin\alpha, \tag{2.3}$$

into the two approximate equations

$$X_i \approx x_i\xi_2 - y_i\xi_3 + \xi_0, \tag{2.4a}$$

$$Y_i \approx x_i\xi_3 + y_i\xi_2 + \xi_1. \tag{2.4b}$$

Taking the random errors of the observed quantities into account, the Eqs. (2.4a–b) more explicitly read:

$$X_i - e_{X_i} = (x_i - e_{x_i})\xi_2 - (y_i - e_{y_i})\xi_3 + \xi_0, \tag{2.5a}$$

$$Y_i - e_{Y_i} = (x_i - e_{x_i})\xi_3 + (y_i - e_{y_i})\xi_2 + \xi_1, \tag{2.5b}$$

thereby forming a *Structured Errors-In-Variables (EIV) Model* which could be either handled along the lines of Felus and Schaffrin (2005), resp. Schaffrin et al. (2012), directly, or by giving it the form of (*nonlinear*) *condition equations with unknowns* (i.e., Gauss–Helmert Model):

$$\underline{b}\left(\underset{2n\times 1}{\mu}, \underset{4\times 1}{\xi}\right) := \begin{bmatrix} \cdots \\ X_i \\ Y_i \\ \cdots \end{bmatrix} - \begin{bmatrix} \cdots \\ e_{X_i} \\ e_{Y_i} \\ \cdots \end{bmatrix} - \begin{bmatrix} 1 & 0 & x_i - e_{x_i} & -(y_i - e_{y_i}) \\ 0 & 1 & y_i - e_{y_i} & x_i - e_{x_i} \end{bmatrix} \begin{bmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = 0, \tag{2.6}$$

where $\underline{b} : \mathbb{R}^{2(n+2)} \to \mathbb{R}^n$ represents a nonlinear function with

$y := [\cdots, \quad X_i, \quad Y_i, \quad \cdots, \quad x_i, \quad y_i, \quad \cdots]^{\mathrm{T}}$ as $2n \times 1$ vector of observed coordinates,

$e := [\cdots, \quad e_{X_i}, \quad e_{Y_i}, \quad \cdots, \quad e_{x_i}, \quad e_{y_i}, \quad \cdots]^{\mathrm{T}}$ as $2n \times 1$ random error vector,

$\mu := y - e$ as $2n \times 1$ vector of actual ("true") coordinates, and

$\xi := [\,\xi_0 \quad \xi_1 \quad \xi_2 \quad \xi_3\,]^{\mathrm{T}}$ as the 4×1 (unknown) parameter vector.

Schaffrin (2015) has shown how the system (2.5) can be equivalently described by "direct observations with nonlinear constraints". On the other hand, it could as well be handled by an extension of the approach by Schaffrin and Wieser (2011) for *structured condition equations*, possibly after some sort of differencing to eliminate ξ_0 and ξ_1, or by the more traditional approach of *iterative linearization* in accordance with the provisions by Pope (1972); for more details, see also Neitzel (2010) and Schaffrin and Snow (2010), and particularly Lenzmann and Lenzmann (2004) who very clearly specify under which approximations *rather inaccurate* results may be produced.

Such insufficient approximations can, unfortunately, be found in a host of textbooks, including those by Mikhail and Gracie (1981), Wolf and Ghilani (1997), Benning (2007), and Niemeier (2008), which led to a situation where the provisions for their iterative algorithms may ensure convergence, but *not necessarily to the nonlinear least-squares solution*.

Here, an approach is chosen that resembles the procedure first proposed by Deming (1931, 1934) for a different example. Thus, for the linearization of (2.6), approximate values $\xi^0 := [\,\xi_0^0 \quad \xi_1^0 \quad \xi_2^0 \quad \xi_3^0\,]^{\mathrm{T}}$ as well as $\mu^0 := y - \underset{\sim}{0}$ are needed where $\underset{\sim}{0}$ indicates the "*random zero vector*" that strips y of its random nature without changing its values. This so-called "Helmert's knack" (or "Helmertscher Kunstgriff" in German) makes sure that the error propagation will turn out correctly. Consequently, the linearized form reads:

$$\begin{aligned} \underline{b}(\mu,\ \xi) &= \underline{b}(\mu^0,\ \xi^0) - A_0(\xi - \xi^0) + B_0 \cdot (\mu - \mu_0) \pm \cdots = \\ &= w_0 - A_0(\xi - \xi^0) - B_0 e \pm \cdots = 0, \end{aligned} \tag{2.7a}$$

with the $n \times 1$ vector of (initial) "misclosures"

$$w_0 := \underline{b}(\mu^0, \xi^0) - B_0 \cdot \underset{\sim}{0} \approx \underline{b}(y, \xi^0) \tag{2.7b}$$

and the combined coefficient matrices of size $n \times 2(n + 2)$

$$[-A_0|B_0] := \left.\frac{\partial \underline{b}(\mu, \xi)}{\partial[\xi^T|\mu^T]}\right|_{\xi=\xi^0, \mu=\mu^0} \tag{2.7c}$$

in the first iteration step and, after introducing the new approximate values $\xi^1 := \xi^0 + (\widehat{\xi - \xi^0}) - \underset{\sim}{0}$ as well as $\mu^1 := y - \tilde{e}^1 - \underset{\sim}{0}$,

$$\begin{aligned} \underline{b}(\mu,\ \xi) = \underline{b}(\mu^1,\ \xi^1) - A_1(\xi - \xi^1) + B_1(\mu - \mu^1) \pm \cdots = \\ = w_1 - A_1(\xi - \xi^1) - B_1 e \pm \cdots = 0, \end{aligned} \tag{2.8a}$$

with the updated vector of "misclosures"

$$w_1 := \underline{b}(\mu^1, \xi^1) + B_1\left(\underset{\sim}{0} + \tilde{e}^1\right) \approx \underline{b}(y,\ \xi^1), \tag{2.8b}$$

and the new combined coefficient matrices

$$[-A_1|B_1] := \left.\frac{\partial \underline{b}(\mu, \xi)}{\partial[\xi^T|\mu^T]}\right|_{\xi=\xi^1, \mu=\mu^1} \tag{2.8c}$$

It was Pope (1972) who had drawn attention to the fact that the update (2.8b) is oftentimes computed incorrectly, thereby potentially changing the convergence point during the iteration. However, the slight modification by Lenzmann and Lenzmann (2004) who replaced (2.8a) with

$$\begin{aligned} \underline{b}(\mu,\ \xi) = \underline{b}(\mu^1,\ \xi^1) - A_1(\xi - \xi^1) + B_1(\mu - \mu^1) \pm \cdots = \\ = [w_1 - B_1\tilde{e}^1] - A_1(\xi - \xi^1) - B_1(e - \tilde{e}^1) \pm \cdots = 0 \end{aligned} \tag{2.9}$$

is obviously equivalent and, therefore, represents another valid approach (although the error propagation becomes more complex). For the present case of the planar similarity transformation, the matrices involved are readily obtained in the first iteration as:

$$B_0 = [B_{1_0}|B_{2_0}] \tag{2.10a}$$

with

$$B_{1_0} = I_n\ (n \times n \text{ identity matrix}), \tag{2.10b}$$

$$B_{2_0} = \begin{bmatrix} -\xi_2^0 & \xi_3^0 & 0 & 0 & \cdots & 0 & 0 \\ -\xi_3^0 & -\xi_2^0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -\xi_2^0 & \xi_3^0 & \cdots & 0 & 0 \\ 0 & 0 & -\xi_3^0 & -\xi_2^0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\xi_2^0 & \xi_3^0 \\ 0 & 0 & 0 & 0 & 0 & -\xi_3^0 & -\xi_2^0 \end{bmatrix}, \tag{2.10c}$$

and

$$-A_0 = \begin{bmatrix} \cdots & \cdots & \cdots & \cdots \\ -1 & 0 & -x_i & y_i \\ 0 & -1 & -y_i & -x_i \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}, \tag{2.11}$$

while the initial "vector of misclosures" is taken from (2.7b) as:

$$w_0 := \underline{b}(y, \xi^0) = \begin{bmatrix} \cdots \\ X_i - x_i\xi_2^0 + y_i\xi_3^0 - \xi_0^0 \\ Y_i - y_i\xi_2^0 - x_i\xi_3^0 - \xi_1^0 \\ \cdots \end{bmatrix}. \tag{2.12}$$

Hence, with a suitably defined cofactor matrix Q of size $2n \times 2n$ for both the old and the new coordinates, that fulfills condition (1.10), the normal equations

$$\begin{bmatrix} B_0QB_0^{\mathrm{T}} & A_0 \\ A_0^{\mathrm{T}} & 0 \end{bmatrix} \cdot \begin{bmatrix} \hat{v}^1 \\ \widehat{\xi - \xi^0} \end{bmatrix} = \begin{bmatrix} w_0 \\ 0 \end{bmatrix} \tag{2.13}$$

ought to be solved from which the new approximation vector

$$\xi^1 := \xi^0 + \left(\widehat{\xi - \xi^0}\right) - \underset{\sim}{0} \tag{2.14a}$$

results as well as the (first) residual vector

$$\tilde{e}^1 := QB_0^{\mathrm{T}} \cdot \hat{v}^1. \tag{2.14b}$$

In the next iteration the matrices are updated as:

$$B_1 = [I_n | B_{2_1}] \tag{2.15a}$$

with

$$B_{2_1} = I_{n/2} \otimes \begin{bmatrix} -\xi_2^1 & \xi_3^1 \\ -\xi_3^1 & -\xi_2^1 \end{bmatrix} \tag{2.15b}$$

and

$$-A_1 = \begin{bmatrix} \cdots & \cdots & \cdots & \cdots \\ -1 & 0 & -(x_i - \tilde{e}_{x_i^1}) & (y_i - \tilde{e}_{y_i}^1) \\ 0 & -1 & -(y_i - \tilde{e}_{y_i}^1) & -(x_i - \tilde{e}_{x_i}^1) \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}, \tag{2.16}$$

and the "vector of misclosures" as:

$$w_1 = \underline{b}(\mu^1, \xi^1) + B_1\left(\underset{\sim}{0} + \tilde{e}^1\right) = \begin{bmatrix} \cdots \\ X_i - x_i\xi_2^1 + y_i\xi_3^1 - \xi_0^1 \\ Y_i - y_i\xi_2^1 - x_i\xi_3^1 - \xi_1^1 \\ \cdots \end{bmatrix} = \underline{b}(y,\ \xi^1), \tag{2.17}$$

which may be modified further in accordance with (2.9), eventually resulting in the normal equations

$$\begin{bmatrix} B_1QB_1^{\mathrm{T}} & A_1 \\ A_1^{\mathrm{T}} & 0 \end{bmatrix} \cdot \begin{bmatrix} \hat{v}^2 \\ \widehat{\xi - \xi^1} \end{bmatrix} = \begin{bmatrix} w_1 - B_1\tilde{e}^1 \\ 0 \end{bmatrix}, \tag{2.18}$$

and the new approximation vector

$$\xi^2 := \xi^1 + \left(\widehat{\xi - \xi^1}\right) - \underset{\sim}{0}, \tag{2.19a}$$

respectively the (second) residual vector

$$\tilde{e}^2 := \tilde{e}^1 + QB_1^{\mathrm{T}} \cdot \hat{v}^2 \approx QB_1^{\mathrm{T}} \cdot \left[\hat{v}^2 + \left(B_1QB_1^{\mathrm{T}}\right)^{-} B_1\tilde{e}^1\right]. \tag{2.19b}$$

After convergence, indicated by

$$\left\| \widehat{\xi - \xi^k} \right\| < \delta \tag{2.20}$$

for a chosen value of $\delta > 0$, the final estimate

$$\hat{\xi} = \xi^k + \left(\widehat{\xi - \xi^k}\right) \tag{2.21a}$$

and the final residual vector

$$\tilde{e} = \tilde{e}^k + QB_k^{\mathrm{T}} \cdot \hat{v}^{k+1} = QB_k^{\mathrm{T}}\left[\hat{v}^{k+1} + \left(B_kQB_k^{\mathrm{T}}\right)^{-} B_k\tilde{e}^k\right] \tag{2.21b}$$

will be *uncorrelated*, with their dispersion matrices stemming from the relationships

$$\begin{bmatrix} D\left\{\hat{v}^{k+1} + (B_kQB_k^{\mathrm{T}})^{-} B_k\tilde{e}^k\right\} & \times \\ \times & -D\{\hat{\xi}\} \end{bmatrix} = \sigma_0^2 \begin{bmatrix} B_kQB_k^{\mathrm{T}} & A_k \\ A_k^{\mathrm{T}} & 0 \end{bmatrix}^{-1} \tag{2.22a}$$

and

$$D\{\tilde{e}\} = QB_k^{\mathrm{T}} \cdot D\left\{\hat{v}^{k+1} + \left(B_kQB_k^{\mathrm{T}}\right)^{-} B_k\tilde{e}^k\right\} \cdot B_kQ, \tag{2.22b}$$

while the sum of weighted squared residuals is obtained from

$$\Omega = w_k^{\mathrm{T}}\left[\hat{v}^{k+1} + \left(B_kQB_k^{\mathrm{T}}\right)^{-} B_k\tilde{e}^k\right], \tag{2.23a}$$

resulting in the unique variance component estimate

$$\hat{\sigma}_0^2 = \Omega/r = \Omega/(rk\ B - rk\ A). \tag{2.23b}$$

4 Numerical example

In the following, a real-life example is presented that, thanks to its small size, allows to see the mechanics of the new approach rather clearly. For the trilateration network depicted in Fig. 1 the approximate values for the coordinates (X^0, Y^0) in the ("new") target system and (x^0, y^0) in the ("old") source system are listed in Table 1.

The horizontal distances s_{ij} are listed in Table 2. These distances are introduced as uncorrelated observations into a free net adjustment with a standard deviation of ±0.5 cm for the distances in the target system and ±1 cm for the distances in the source system.

Fig. 1 Trilateration network

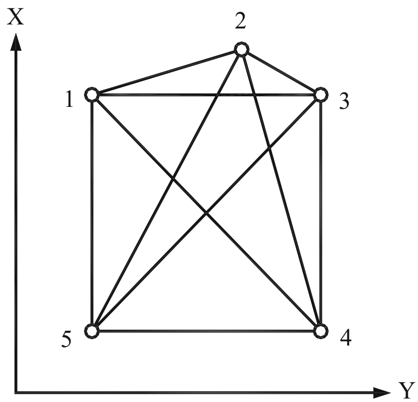

Table 1 Approximate coordinates in the target and the source system

Point No.	Y_i^0 [m]	X_i^0 [m]	y_i^0 [m]	x_i^0 [m]
1	100.000	400.000	137.612	453.800
2	300.000	500.000	350.795	521.282
3	400.000	400.000	433.921	406.869
4	400.000	100.000	386.991	110.559
5	100.000	100.000	90.681	157.490

Table 2 Horizontal distances in the target and the source system

	Target system	Source system
$s_{1,2}$ (m)	223.598	223.607
$s_{1,3}$ (m)	299.990	300.008
$s_{1,4}$ (m)	424.255	424.281
$s_{1,5}$ (m)	300.011	299.998
$s_{2,3}$ (m)	141.422	141.421
$s_{2,4}$ (m)	412.309	412.321
$s_{2,5}$ (m)	447.220	447.235
$s_{3,4}$ (m)	299.988	300.009
$s_{3,5}$ (m)	424.255	424.279
$s_{4,5}$ (m)	300.007	299.996

From a 2D free network adjustment of the trilateration network, the following coordinate estimates in the ("new") target system and in the ("old") source system have been obtained; they are listed in Table 3 respectively in Table 4.

Table 3 Coordinate estimates in the target system

Point No.	Y_i [m]	X_i [m]
1	100.0072	400.0040
2	299.9994	500.0019
3	399.9933	399.9925
4	400.0022	100.0059
5	99.9978	99.9956

Table 4 Coordinate estimates in the source system

Point No.	y_i [m]	x_i [m]
1	137.6099	453.8001
2	350.7972	521.2865
3	433.9247	406.8728
4	386.9880	110.5545
5	90.6802	157.4861

The corresponding cofactor matrix Q_{XY} has a rank deficiency of $d_1 = 3$ and reads

$$Q_{XY} = \begin{bmatrix} 6.79448757180059 & -1.24637379742034 & -2.06749265251556 & -0.87905006133625 & -1.75237198700040 & -1.16357954126859 & -0.45154752590708 & 0.97970173045535 & -2.52307540637754 & 2.30930166956982 \\ -1.24637379742034 & 6.09498927220848 & 1.09012834165340 & -2.49924282264974 & 0.93808137930146 & -1.20479964504946 & -0.10681237404477 & -0.69930007052657 & -0.67502354948975 & -1.69164673398272 \\ -2.06749265251556 & 1.09012834165340 & 6.42931803954814 & 0.55451061543664 & -1.97051230040660 & -0.91755859035999 & -1.40374219486592 & -0.95534188201874 & -0.98757089176006 & 0.22826151528869 \\ -0.87905006133625 & -2.49924282264975 & 0.55451061543664 & 5.91276039572392 & 0.36212974333293 & -3.44014208938923 & 1.21258998841997 & -0.11791264515374 & -1.25018028585329 & 0.14453716146879 \\ -1.75237198700040 & 0.93808137930146 & -1.97051230040660 & 0.36212974333293 & 6.22939970049753 & 1.44337480762821 & -2.05129094829068 & -2.01871586069104 & -0.45522446479986 & -0.72487006957155 \\ -1.16357954126860 & -1.20479964504946 & -0.91755859035999 & -3.44014208938923 & 1.44337480762820 & 7.20646286085327 & 0.58294810376109 & -1.84740759328547 & 0.05481522023931 & -0.71411353312912 \\ -0.45154752590708 & -0.10681237404477 & -1.40374219486592 & 1.21258998841998 & -2.05129094829069 & 0.58294810376109 & 5.07997251772178 & -0.04815330743166 & -1.17339184865810 & -1.64057241070464 \\ 0.97970173045535 & -0.69930007052657 & -0.95534188201873 & -0.11791264515374 & -2.01871586069103 & -1.84740759328547 & -0.04815330743165 & 6.36060256564705 & 2.04250931968606 & -3.69598225668127 \\ -2.52307540637755 & -0.67502354948976 & -0.98757089176006 & -1.25018028585330 & -0.45522446479986 & 0.05481522023930 & -1.17339184865810 & 2.04250931968607 & 5.13926261159557 & -0.17212070458233 \\ 2.30930166956983 & -1.69164673398272 & 0.22826151528869 & 0.14453716146878 & -0.72487006957155 & -0.71411353312913 & -1.64057241070463 & -3.69598225668126 & -0.17212070458233 & 5.95720536232432 \end{bmatrix} \cdot 10^{-6}\ (\mathrm{m}^2).$$

The corresponding cofactor matrix Q_{xy} shows a rank deficiency of $d_2 = 3$ and is given by

$$
Q_{xy} = \left[\begin{matrix}
36.37028145702680 & -5.47084753104938 & -10.71785609522750 & -4.15196839574972 & -8.65298041790832 \\
-5.47084753104937 & 29.08218656854860 & 5.84822137965519 & -12.47141347170350 & 4.36257321048772 \\
-10.71785609522750 & 5.84822137965518 & 31.71485018459150 & 3.08331019238332 & -9.75441280523014 \\
-4.15196839574972 & -12.47141347170350 & 3.08331019238333 & 30.95803801957830 & 3.06122736829926 \\
-8.65298041790830 & 4.36257321048773 & -9.75441280523014 & 3.06122736829926 & 29.49000356229180 \\
-6.30957538097353 & -6.36323276123894 & -3.43751427095770 & -17.72069907008300 & 6.20362896765228 \\
-3.00872247368822 & -0.45688827110577 & -7.17013351915318 & 5.11629815381483 & -9.26471495157075 \\
5.06154166169345 & -2.83561006289195 & -5.89199924913936 & -0.55682802371483 & -10.23264832811930 \\
-13.99072247020280 & -4.28305878798777 & -4.07244776498071 & -7.10886731874769 & -1.81789538758257 \\
10.87084964607920 & -7.41193027271422 & 0.39798194805854 & -0.20909745407698 & -3.39478121831995
\end{matrix}\right.
$$

$$
\left.\begin{matrix}
-6.30957538097352 & -3.00872247368821 & 5.06154166169344 & -13.99072247020280 & 10.87084964607920 \\
-6.36323276123892 & -0.45688827110575 & -2.83561006289193 & -4.28305878798774 & -7.41193027271419 \\
-3.43751427095771 & -7.17013351915317 & -5.89199924913935 & -4.07244776498073 & 0.39798194805855 \\
-17.72069907008300 & 5.11629815381484 & -0.55682802371480 & -7.10886731874765 & -0.20909745407695 \\
6.20362896765225 & -9.26471495157075 & -10.23264832811930 & -1.81789538758258 & -3.39478121831996 \\
38.73483275416460 & 2.97933036065107 & -10.53144746790980 & 0.56413032362793 & -4.11945345493275 \\
2.97933036065104 & 26.03149620685810 & -1.23726448444412 & -6.58792526244601 & -6.40147575891599 \\
-10.53144746790980 & -1.23726448444413 & 32.06314207358760 & 12.30037040000930 & -18.13925651907110 \\
0.56413032362793 & -6.58792526244599 & 12.30037040000930 & 26.46899088521210 & -1.47257461690178 \\
-4.11945345493281 & -6.40147575891599 & -18.13925651907110 & -1.47257461690180 & 29.87973770079510
\end{matrix}\right] \cdot 10^{-6}\ (\mathrm{m}^2).
$$

It is emphasized that all five points participated in the datum definition for both free adjustments. But, since a different scale factor may have been assumed for the two network adjustments, here a 2D similarity transformation will be investigated, not just a rigid one.

In the following, it is shown how the full singular cofactor matrices can be utilized to estimate the parameters of this 2D similarity transformation via weighted least-squares (Corollary 1.2), without resorting to the common practice to only use their diagonal elements and thereby circumventing the singularity issue, but at the cost of neglecting the existing covariances.

To start the process of iteratively linearizing a nonlinear Gauss–Helmert Model, suitable approximate values for the parameters of the 2D similarity transformation must be computed. This can be done by following the classical procedure of determining the parameters of a traditional "Helmert transformation" where Q_{xy} is replaced by 0 and Q_{XY} by I_n. The resulting initial approximate values are $\xi_0^0 = -69.73$, $\xi_1^0 = 35.08$, $\xi_2^0 = 0.988$, $\xi_3^0 = -0.156$.

Obviously, the initial choice for the random error vector e is the zero vector, consistent with (2.11) when compared with (2.16). This allows to compute the matrices B_0 and A_0 from (2.10a–c) and (2.11), as well as the "vector of misclosures" w_0 from (2.12). By defining the 20 × 20 cofactor matrix

$$
Q := \begin{bmatrix} Q_{XY} & 0 \\ 0 & Q_{xy} \end{bmatrix}, \quad d = d_1 + d_2 = 6, \tag{3.1}
$$

with zero covariances between estimated target and source coordinates, the normal equations (2.13) can be set up and solved uniquely whenever the criterion (1.10) is fulfilled which is necessary and sufficient. To establish non-uniqueness, the criterion (1.21) has to be violated which may be somewhat easier to show.

Disregarding some rather exceptional cases, which can easily be avoided in practice, the rank of the matrix A_0 should be equal to the number of parameters:

$$
q := rk\, A_0 = 4 = m. \tag{3.2}
$$

Moreover, the rank of the matrix B_0 turns out to be:

$$rk\,B_0 = 2 \cdot (n/2) = n = 10. \tag{3.3}$$

A numerical check of the matrices (B_0Q) and $[A_0|B_0Q]$ reveals their ranks to be:

$$r := rk[A_0|B_0Q] - rk\,A_0 = 10 - 4 = 6 \le 8 = rk(B_0Q) \le 10. \tag{3.4}$$

Clearly, the criterion (1.21) is *not violated*, which however, does not yet establish uniqueness of $\hat{\hat{\xi}}_{LESS}$. For this, the criterion (1.10) ought to be applied which indeed results in a *positive decision*, due to:

$$r := rk[A_0|B_0Q] = 10 = rk\,B_0. \tag{3.5}$$

After few iterations, the *unique solution* $\hat{\hat{\xi}}_{LESS}$ of the (originally nonlinear) Gauss–Helmert Model is obtained as listed in Table 5.

Finally, the residuals after convergence are listed in Table 6.

The respective dispersion matrices for both the estimated parameters and the residuals are given in the Appendix. They represent the "gain of efficiency" of the newly estimated coordinates over the original coordinate estimates.

5 Conclusions and outlook

In an earlier contribution by Neitzel and Schaffrin (2016) the treatment of the Gauss–Helmert Model with a *singular* covariance matrix had been generalized beyond the case where $rk(BQ) = rk\,B$. In particular, the criterion (1.10) was found to be *necessary and*

Table 5 Weighted least-squares solution on the basis of an iteratively linearized Gauss–Helmert Model with a singular cofactor matrix

Parameters	Its estimate	rmse
x-shift ξ_0	$\hat{\hat{\xi}}_0 = -69.726354$ m	± 4.090 mm
y-shift ξ_1	$\hat{\hat{\xi}}_1 = 35.078215$ m	± 2.488 mm
$\xi_2 = \omega \cdot \cos\alpha$	$\hat{\hat{\xi}}_2 = 0.98765502$	$\pm 1.093 \cdot 10^{-5}$
$\xi_3 = \omega \cdot \sin\alpha$	$\hat{\hat{\xi}}_3 = -0.15642921$	$\pm 1.730 \cdot 10^{-6}$
Scale factor ω	$\hat{\omega} = 0.99996626$	$\pm 1.106 \cdot 10^{-5}$
Rotation angle α	$\hat{\alpha} = -10.00000154$ gon	$\pm 1.446 \cdot 10^{-5}$ mgon
Variance component σ_0^2	$\hat{\sigma}_0^2 = 1.027339$	

Table 6 Residuals on the basis of an iteratively linearized Gauss–Helmert Model with a singular cofactor matrix

Point No.	Target system		Source system	
	$\tilde{e}_{Y_i}$ [mm]	$\tilde{e}_{X_i}$ [mm]	$\tilde{e}_{y_i}$ [mm]	$\tilde{e}_{x_i}$ [mm]
1	0.900	1.020	−5.323	−4.403
2	−0.163	0.345	0.545	−1.862
3	−0.992	−1.581	6.232	7.139
4	1.201	1.040	−6.849	−4.262
5	−0.945	−0.825	5.395	3.387

sufficient for a unique solution of type $\hat{\xi}_{LESS} = \hat{\xi}_{BLUMBE}$ to exist. To check the *non-uniqueness*, the inequality (1.21) could be used alternatively, which, however, would not guarantee uniqueness if satisfied since it is only a necessary but not sufficient condition.

Here, through an illuminating example, the theory as summarized in Chap. 1 was tested in the context of a *2D similarity transformation* with singular cofactor matrices for both the ("new") target and the ("old") source coordinate estimates. This is certainly a rather relevant extension as, more often than not, the estimated coordinates may indeed be taken from a *free network adjustment*. Consequently, the resulting covariance matrices will be singular, a fact that has frequently be circumvented in practice by only considering the variances on the diagonal while setting all the covariances to zero. This unwarranted procedure is no longer required; even the case where one set of the estimated coordinate data are replaced by fixed coordinates can simply be handled by setting either $Q_{XY} = 0$ or $Q_{xy} = 0$.

While this paper treats the 2D similarity transformation in the framework of a nonlinear Gauss–Helmert Model by iterative linearization, it will be of major interest as well how it can be handled within an *EIV-Model* ("Errors-In-Variables") by setting up nonlinear normal equations and solving them iteratively, all with singular covariance matrices for both vector and matrix observations. Two other papers on this subject have recently been published; see Schaffrin et al. (2014) and Jazaeri et al. (2014).

Acknowledgments The first author would like to gratefully acknowledge the support of a Feodor-Lynen Research Fellowship from the Alexander-von-Humboldt Foundation (Germany), and the School of Earth Sciences at the Ohio State University (Columbus/OH, USA), with Prof. Schaffrin as his host. This manuscript was actually completed when the second author visited Prof. Neitzel, again with funds of the AvH-Foundation, which is also gratefully acknowledged.

Appendix

Estimated dispersion matrix of the estimated parameters $\hat{\xi}_0$, $\hat{\xi}_1$, $\hat{\xi}_2$, $\hat{\xi}_3$, respectively for $\hat{\omega}$ and $\hat{\alpha}$, from inverting (2.3), including their covariances, for the numerical example:

$$\hat{D}\left\{\begin{bmatrix}\hat{\xi}_0\\ \hat{\xi}_1\\ \hat{\xi}_2\\ \hat{\xi}_3\end{bmatrix}\right\} = \begin{bmatrix} 1.673\text{E}{-05} & 1.018\text{E}{-05} & -4.469\text{E}{-08} & 7.078\text{E}{-09}\\ 1.018\text{E}{-05} & 6.191\text{E}{-06} & -2.718\text{E}{-08} & 4.306\text{E}{-09}\\ -4.469\text{E}{-08} & -2.718\text{E}{-08} & 1.194\text{E}{-10} & -1.891\text{E}{-11}\\ 7.078\text{E}{-09} & 4.306\text{E}{-09} & -1.891\text{E}{-11} & 2.994\text{E}{-12}\end{bmatrix},$$
$$\hat{D}\left\{\begin{bmatrix}\hat{\xi}_0\\ \hat{\xi}_1\\ \hat{\omega}\\ \hat{\alpha}\end{bmatrix}\right\} = \begin{bmatrix} 1.673\text{E}{-05} & 1.018\text{E}{-05} & -4.524\text{E}{-08} & 3.653\text{E}{-15}\\ 1.018\text{E}{-05} & 6.191\text{E}{-06} & -2.752\text{E}{-08} & 2.196\text{E}{-15}\\ -4.524\text{E}{-08} & -2.752\text{E}{-08} & 1.224\text{E}{-10} & -9.849\text{E}{-18}\\ 3.653\text{E}{-15} & 2.196\text{E}{-15} & -9.849\text{E}{-18} & 5.159\text{E}{-20}\end{bmatrix}. \tag{A.1}$$

Estimated dispersion matrices for the residuals and their cross-covariance matrix:

$$
\hat{D}\{\tilde{e}_{xy}\} = \left[\begin{matrix}
1.115E-06 & -1.577E-07 & -4.157E-07 & -1.618E-07 & -3.294E-07 \\
-1.577E-07 & 9.454E-07 & 2.903E-07 & -4.013E-07 & 2.116E-07 \\
-4.157E-07 & 2.903E - 07 & 9.542E-07 & 6.724E-08 & -3.993E-07 \\
-1.618E-07 & -4.013E-07 & 6.724E-08 & 9.942E-07 & 4.795E-08 \\
-3.294E-07 & 2.116E-07 & -3.993E-07 & 4.795E-08 & 1.020E-06 \\
-2.431E-07 & -1.294E-07 & -2.479E-07 & -6.001E-07 & 1.976E-07 \\
-1.007E-08 & -1.241E-07 & -1.048E-07 & 2.315E-07 & -2.805E-07 \\
1.192E-07 & -4.400E-08 & -2.542E-07 & -3.847E-08 & -3.876E-07 \\
-3.602E-07 & -2.201E-07 & -3.442E-08 & -1.848E-07 & -1.069E-08 \\
4.434E-07 & -3.707E-07 & 1.446E-07 & 4.564E-08 & -6.951E-08
\end{matrix}\right.
$$

$$
\left.\begin{matrix}
-2.431E-07 & -1.007E-08 & 1.192E-07 & -3.602E-07 & 4.434E-07 \\
-1.294E-07 & -1.241E - 07 & -4.400E-08 & -2.201E - 07 & -3.707E-07 \\
-2.479E - 07 & -1.048E-07 & -2.542E-07 & -3.442E-08 & 1.446E-07 \\
-6.001E-07 & 2.315E-07 & -3.847E-08 & -1.848E-07 & 4.564E-08 \\
1.976E-07 & -2.805E-07 & -3.876E-07 & -1.069E-08 & -6.951E-08 \\
1.153E-06 & 1.913E-07 & -3.772E-07 & 1.020E-07 & -4.651E-08 \\
1.913E-07 & 7.262E-07 & 8.462E-08 & -3.309E-07 & -3.833E-07 \\
-3.772E-07 & 8.462E-08 & 1.010E-06 & 4.380E-07 & -5.506E-07 \\
1.020E - 07 & -3.309E-07 & 4.380E-07 & 7.362E-07 & -1.351E-07 \\
-4.651E-08 & -3.833E-07 & -5.506E-07 & -1.351E-07 & 9.222E-07
\end{matrix}\right], \tag{A.2}
$$

$$
\hat{D}\{\tilde{e}_{xy}\} = \left[\begin{matrix}
2.991E-05 & -3.190E - 06 & -1.122E-05 & -4.312E-06 & -8.240E-06 \\
-3.190E-06 & 2.323E-05 & 7.346E-06 & -9.843E-06 & 4.679E-06 \\
-1.122E-05 & 7.346E-06 & 2.410E-05 & 1.490E-06 & -9.628E-06 \\
-4.312E-06 & -9.843E-06 & 1.490E-06 & 2.614E-05 & 2.163E-06 \\
-8.240E-06 & 4.679E-06 & -9.628E-06 & 2.163E-06 & 2.481E-05 \\
-7.043E-06 & -3.590E-06 & -5.465E-06 & -1.614E-05 & 4.314E-06 \\
-2.618E-07 & -3.061E-06 & -2.569E-06 & 5.718E-06 & -6.511E-06 \\
3.213E-06 & -1.133E-06 & -6.805E-06 & -1.125E-06 & -9.485E-06 \\
-1.018E-05 & -5.774E-06 & -6.766E-07 & -5.058E-06 & -4.276E-07 \\
1.133E-05 & -8.664E-06 & 3.435E-06 & 9.654E-07 & -1.671E-06
\end{matrix}\right. \tag{A.3}
$$

$$
\left.\begin{matrix}
-7.043E-06 & -2.618E-07 & 3.213E-06 & -1.018E-05 & 1.133E-05 \\
-3.590E-06 & -3.061E-06 & -1.133E-06 & -5.774E-06 & -8.664E-06 \\
-5.465E-06 & -2.569E-06 & -6.805E-06 & -6.766E-07 & 3.435E-06 \\
-1.614E-05 & 5.718E-06 & -1.125E-06 & -5.058E-06 & 9.654E-07 \\
-.314E-06 & -6.511E-06 & -9.485E-06 & -4.276E-07 & -1.671E-06 \\
3.122E-05 & 5.442E-06 & -1.045E-05 & 2.752E-06 & -1.047E-06 \\
5.442E-06 & 1.823E-05 & 9.436E-07 & -8.888E-06 & -9.043E-06 \\
-1.045E-05 & 9.436E-07 & 2.654E-05 & 1.213E-05 & -1.384E-05 \\
2.752E-06 & -8.888E-06 & 1.213E-05 & 2.018E-05 & -4.054E-06 \\
-1.047E-06 & -9.043E-06 & -1.384E-05 & -4.054E-06 & 2.258E-05
\end{matrix}\right],
$$

$$Cov\left\{\tilde{e}_{XY},\ \tilde{e}_{xy}\right\} = \left[\begin{array}{ccccc}
-5.719E-06 & -9.516E-08 & 1.956E-06 & 1.142E-06 & 1.459E-06 \\
1.542E-06 & -4.617E-06 & -1.775E-06 & 1.782E-06 & -1.164E-06 \\
2.316E-06 & -1.126E-06 & -4.733E-06 & -1.095E-06 & 1.806E-06 \\
4.929E-07 & 2.141E-06 & 4.525E-07 & -5.039E-06 & -7.172E-07 \\
1.820E-06 & -7.995E-07 & 2.041E-06 & 7.664E-08 & -4.958E-06 \\
1.116E-06 & 8.423E-07 & 7.664E-07 & 3.206E-06 & -7.501E-08 \\
-4.803E-08 & 6.302E-07 & 7.093E-07 & -1.078E-06 & 1.559E-06 \\
-6.329E-07 & 1.260E-07 & 1.245E-06 & 3.949E-07 & 1.644E-06 \\
1.632E-06 & 1.390E-06 & 2.581E-08 & 9.542E-07 & 1.347E-07 \\
-2.518E-06 & 1.507E-06 & -6.890E-07 & -3.437E-07 & 3.117E-07
\end{array}\right.$$

$$\left.\begin{array}{ccccc}
1.481E-06 & 1.452E-07 & -5.900E-07 & 2.159E-06 & -1.937E-06 \\
4.812E-07 & 5.873E-07 & 3.193E-07 & 8.094E-07 & 2.034E-06 \\
1.560E-06 & 3.235E-07 & 1.358E-06 & 2.875E-07 & -6.979E-07 \\
2.971E-06 & -1.191E-06 & 9.159E-09 & 9.632E-07 & -8.203E-08 \\
-1.801E-06 & 1.099E-06 & 2.167E-06 & -1.657E-09 & 3.572E-07 \\
-5.941E-06 & -1.259E-06 & 1.740E-06 & -5.486E-07 & 1.521E-07 \\
-7.366E-07 & -3.575E-06 & -1.001E-06 & 1.355E-06 & 2.185E-06 \\
2.200E-06 & 3.781E-07 & -5.134E-06 & -2.634E-06 & 2.413E-06 \\
-5.032E-07 & 2.007E-06 & -1.934E-06 & -3.800E-06 & 9.283E-08 \\
2.885E-07 & 1.485E-06 & 3.066E-06 & 1.410E-06 & -4.518E-06
\end{array}\right]. \tag{A.4}$$

References

Aitken AC (1935) On least squares and linear combinations of observations. Proc R Soc Edinb 55:42–48

Benning W (2007) Statistics in geodesy, geoinformation and civil engineering, 2nd edn. Herbert Wichmann Verlag, Heidelberg (in German)

Bjerhammar A (1973) Theory of errors and generalized matrix inverses. Elsevier Scientific Publishing Company, Amsterdam

Bleich U, Illner M (1989) Rigorous solution of the spatial coordinate transformation by iteration. Allgem Verm Nachr 96:133–144 (in German)

Chang G (2015) On least-squares solution to 3D similarity transformation problem under Gauss–Helmert model. J Geod 89:573–576

Deming WE (1931) XI. The application of least squares. Lond Edinb Dublin Philos Mag J Sci 11(68):146–158

Deming WE (1934) LXVII. On the application of least squares—II. Lond Edinb Dublin Philos Mag J Sci 17(114):804–829

Fang X (2014) A total least-squares solution for geodetic datum transformations. Acta Geodaet Geophys 49:189–207

Felus Y, Schaffrin B (2005) Performing similarity transformations using the Error-in-Variables Model. In: Proceedings of the ASPRS meeting, Washington, DC (on CD)

Grafarend E, Schaffrin B (1993) Adjustment computations in linear models. BI Wissenschaftsverlag, Mannheim (in German)

Jazaeri S, Schaffrin B, Snow K (2014) On Weighted Total Least-Squares adjustment with multiple constraints and singular dispersion matrices. Z Vermess 139:229–240

Koch KR, Fröhlich H, Bröker G (2000) Transformation of spatial variable coordinates. Allgem Verm Nachr 107:293–295 (in German)

Lenzmann L, Lenzmann E (2004) Rigorous adjustment of the nonlinear Gauss–Helmert Model. Allgem Verm Nachr 111:68–73 (in German)

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 similarity transformation. J Geod 84:751–762

Neitzel F, Schaffrin B (2016) On the Gauss–Helmert model with a singular dispersion matrix where BQ is of smaller rank than B. J Comput Appl Math 291:458–467

Niemeier W (2008) Adjustment computation, 2nd edn. Walter de Gruyter, Berlin (in German)

Perelmuter A (1981) Adjustment with a singular weight matrix. Allgem Verm Nachr 88:239–242

Pope AJ (1972) Some pitfalls to be avoided in the iterative adjustment of nonlinear problems. In: Proceedings of the 38th annual meeting of the American Society of Photogrammetry, Washington, DC, p 449–477

Schaffrin B (1989) Advanced adjustment computations, lecture notes. Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus

Schaffrin B (2003) Reproducing estimators via least squares: an optimal alternative to the Helmert transformation. In: Grafarend EW, Krumm FW, Schwarze VS (eds) Geodesy—the Challenge of the 3rd millennium. Springer, Berlin, pp 387–392

Schaffrin B (2015) Adjusting the errors-in-variables model: linearized least-squares vs. nonlinear total least-squares. In: Sneeuw N, Novak P, Crespi M, Sanso F (eds) VIII Hotine-Marussi Symposium on Mathematical Geodesy (Rome/Italy, June 2013). IAG Symposia, vol 142. Springer, Berlin, pp 301–307

Schaffrin B, Snow K (2010) Total Least-Squares regularization of Tykhonov type and an ancient race track in Corinth. Linear Algebra Appl 432:2061–2076

Schaffrin B, Wieser A (2011) Total Least-Squares adjustment for condition equations. Stud Geophys Geodaet 55:529–536

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 2:98–106

Schaffrin B, Snow K, Neitzel F (2014) On the errors-in-variables model with singular dispersion matrices. J Geod Sci 4:28–36

Teunissen P (1988) The nonlinear 2D symmetric Helmert transformation. An exact least-squares solution. J Geod 62:1–16

Wolf H (1979) Singular covariances in the Gauss–Helmert Model. Z Vermess 104:437–442 (in German)

Wolf PR, Ghilani CD (1997) Adjustment computations: statistics and least squares in surveying and GIS. Wiley, New York