On the importance of intra-frame and inter-frame covariances in frame transformation theory

C. Kotsakis · A. Vatalis · F. Sansò

Received: 22 February 2014 / Accepted: 4 August 2014 / Published online: 3 September 2014

C. Kotsakis (✉)
Department of Geodesy and Surveying, Aristotle University of Thessaloniki, University Box 440, 54124 Thessaloniki, Greece
e-mail: kotsaki@topo.auth.gr

A. Vatalis · F. Sansò
DICA - Politecnico di Milano, c/o Polo Regionale di Como via Valleggio 11, 22100 Como, Italy
e-mail: antonios.vatalis@polimi.it

F. Sansò
e-mail: fernando.sanso@polimi.it

1 Introduction

The transformation of a coordinate set from a reference frame to another is a fundamental task in geodetic work with pro-found importance for Earth science applications (e.g. Leick and van Gelder 1975; Soler 1998; Altamimi et al. 2002; Tre-goning and van Dam 2005; Bevis and Brown 2014). As in many problems encountered in geodesy, coordinate frame transformation (CFT) has a dual character and it can be viewed as a forward or inverse problem depending on the prior knowledge of the associated parameters of the frame transformation model. Considering the general case where the underlying model parameters are not known beforehand, the frame transformation problem is usually split into two steps by combining both an inverse and a forward treatment. In particular, a least squares (LS) adjustment is first imple-mented for estimating the transformation model parameters on the basis of common points given in both frames. The coordinate accuracy of the common points in each frame is taken into account, at this step, through an appropriate weight matrix within the LS adjustment algorithm. To complete the task at hand, a forward computation step is further required to transform the known coordinates from their initial frame to the desired target frame using the estimated parameters from the previous step. This is a rather straightforward pro-cedure which has provided the standard framework for CFT problems in geodetic practice.

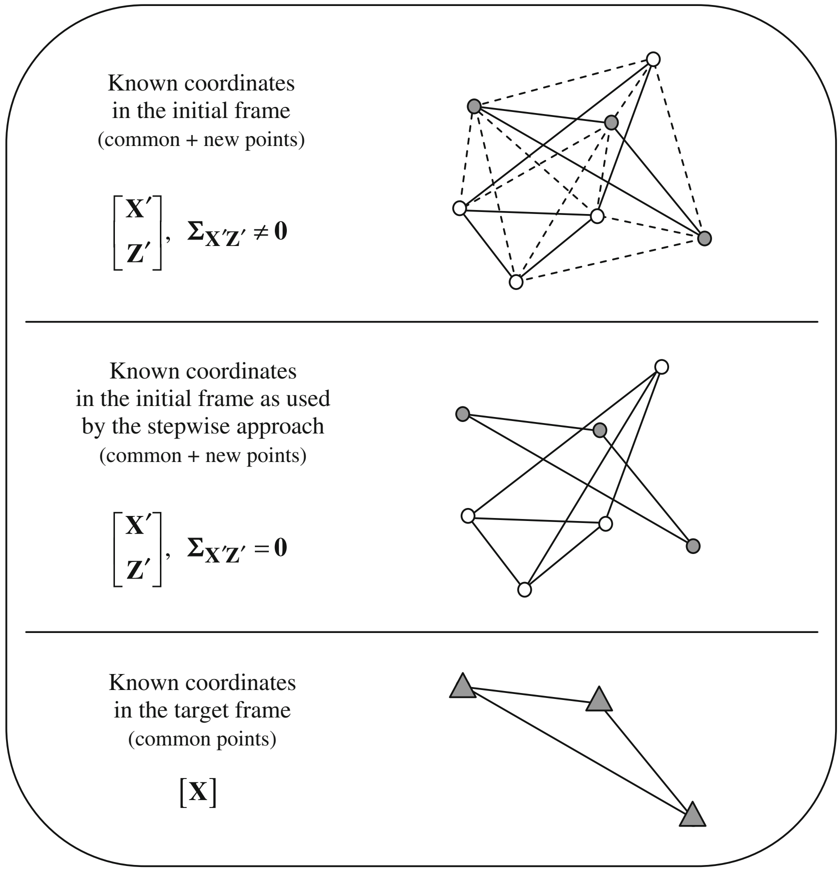


Fig. 1 Schematic representation of the typical data structure in geodetic CFT problems. The stepwise transformation approach treats the common and the new points independently by neglecting the cross-CV matrix $\boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{Z}'}$ (intra-frame covariances) of their initial coordinates. The inter-frame cross-CV matrices $\boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X}'}$ and $\boldsymbol{\Sigma}_{\mathbf{X}\mathbf{Z}'}$ are also ignored in the stepwise approach

The previous stepwise approach is not flawless as it overlooks (a crucial part of) the stochastic error model of the initial coordinates, a fact that can degrade the accuracy of their transformed values in the target frame. Typically, the group of points that needs to be analyzed consists of two subsets: a set of reference stations whose coordinates are known in both frames (hereafter called "common" points) and another set of stations whose coordinates are only known in the initial frame while the estimation of their spatial positions in the target frame is often the prime objective of the problem at hand (hereafter called "new" points). The second subset is left out of the LS inversion step since it does not contribute any information for the estimation of the transformation parameters, and its role is solely confined within the forward step of the transformation procedure. Despite its rational reasoning, such an approach is not based on any objective optimality criteria for the transformed coordinates of the new points and it cannot assure their best estimation accuracy from the available data.

A shortcoming in the standard CFT methodology arises if the coordinates of the common and the new points (in the initial frame) are stochastically correlated with each other due to their concurrent estimation from the same observation set. A common example is the transformation problem for an adjusted geodetic network that needs to be referred to another coordinate frame with the help of reference stations which participate in the network solution (e.g. Altamimi 2003). In such cases a reciprocal link exists between the common and the new points which is reflected in the cross-covariance part of the joint covariance (CV) matrix of their adjusted coordinates, yet it remains unexploited by the step-wise approach in CFT problems. This causes a negligence of the prior stochastic characteristics of the geodetic network and it implies a sloppy treatment of its spatial configura-tion under the transformation procedure. In fact, from a sta-tistical estimation perspective, the transformed coordinates through the stepwise approach remain unbiased but they will not have an optimal accuracy level as a result of mishandling the full stochastic model for the available data—this will be explained in more detail in the following sections of the paper. A schematic illustration of this situation is depicted in Fig. 1.

The aim of this paper is to solve the CFT problem by a single-step optimal inversion which employs all available data in the initial frame, that is the common and new station coordinates with their joint CV matrix, and thus exploit-ing the error covariance structure of the entire point set to be analyzed. To the authors' knowledge such an integrated approach has not been pursued in the geodetic literature and the present study is an attempt to reveal its significance for related applications, having particularly in mind the analysis of permanent GPS networks. Starting from the standard stepwise approach (Sect. 2), an enhanced data adjustment model is formulated to derive the optimal LS estimators for the geodetic CFT problem (Sect. 3). The characteristic of the new solution is that it leads to transformed coordinates with better statistical accuracy (i.e. smaller error variances) than the standard solution without changing the usual LS estimate of the frame transformation parameters. The considered *intra-frame covariances* between common and new points will be shown to affect the transformed coordinates through a correction term that is similar to a LS collocation predictor. The role of this term towards the transformation accuracy improvement is verified by the comparison of the error CV matrices of the final results obtained by the two approaches. Furthermore, its practical significance is evaluated in Sect. 4 through numerical experiments with the 3D Helmert transformation model and real coordinate sets obtained from weekly combined solutions of the EUREF Permanent Network (EPN). Our results show that the omission of this correction term causes a bias-like offset of several mm in the transformed coordinates of new points, whereas the corresponding influence for the common points is considerably larger but with a more random-like behaviour. Both of these effects are quite significant as their magnitude remains well above the statistical accuracy of the transformation results that are obtained by the standard (stepwise) solution.

In addition to the classic version of the geodetic CFT problem, the last part of our study (Sect. 5) presents a more comprehensive treatment with two crucial extensions from the usual setting of Fig. 1. Firstly, the prior realization of the target frame shall not be restricted on the so-called common points but it will also consider the additional information from other (non-common) reference points that belong to a larger network which provides the full realization of the desired target frame for the underlying transformation problem. Secondly, the *intra-frame and inter-frame covariances* among all point subsets will be taken into account, thus leading to the most general case for the stochastic model of the available coordinate data. Those theoretical generalizations may seem unnecessary in the current context of geodetic practice, however they provide a valuable all-inclusive view of the CFT problem that is missing from the geodetic literature.

2 Standard stepwise approach for geodetic frame transformation

2.1 Background

In this section we outline the main characteristics of the geodetic CFT problem and the related notation that will be used throughout the paper.

2.1.1 General mathematical model

The geodetic CFT problem is generally treated in terms of a linearized differential model

$$\mathbf{X} = \mathbf{X}' + \mathbf{G}\boldsymbol{\theta} \tag{1}$$

where $\mathbf{X}'$ and $\mathbf{X}$ denote the Cartesian coordinate vectors of a m-point set with respect to an initial frame and a target frame, respectively. The matrix $\mathbf{G}$ is dictated by the type of frame transformation and the vector $\boldsymbol{\theta}$ contains its associated parameters. The above model covers all variants of the Helmert transformation (e.g. similarity transformation, rigid transformation, shift-only transformation, rotational transformation) that are used in geodetic practice for static or time-dependent frames (Petit and Luzum 2010). For example, in the case of 3D similarity transformation, the parameter vector $\boldsymbol{\theta}$ consists of three translation components, three rotation angles and a scale factor, whereas the transformation matrix has the well known form (e.g. Sillard and Boucher 2001)

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & -Z_1 & Y_1 & X_1 \\ 0 & 1 & 0 & Z_1 & 0 & -X_1 & Y_1 \\ 0 & 0 & 1 & -Y_1 & X_1 & 0 & Z_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & -Z_m & Y_m & X_m \\ 0 & 1 & 0 & Z_m & 0 & -X_m & Y_m \\ 0 & 0 & 1 & -Y_m & X_m & 0 & Z_m \end{bmatrix} \tag{2}$$

In general, Eq. (1) represents a "close to the identity" formula which is valid for the coordinate transformation between reference frames with sufficiently small ($<10^{-5}$) orientation and scale differences. This allows us to discard any linearization errors associated with Eq. (1) or, at least, to safely assume that they are absorbed by the noise level of the available data.

Strictly speaking, in the context of the linearized Helmert transformation, the matrix $\mathbf{G}$ is formed by the initial frame coordinates and, thus, some of its elements are affected by the random errors of the coordinate vector $\mathbf{X}'$. Since the magnitude of the rotation and scale parameters multiplying those noisy elements does not exceed a few arcsec and a few ppm, respectively, we may easily neglect the stochastic character of the matrix $\mathbf{G}$ in geodetic CFT problems. In the rest of the paper, the design matrix of Eq. (1) is therefore treated as a fixed deterministic element without having the need to rely on total LS adjustment theory (e.g. Schaffrin and Wieser 2008; Mahboub 2012).

2.1.2 Coordinate datasets and true coordinate vectors

Given a set of points with known noisy coordinates in an initial frame, the problem at hand is to determine their coordinates, and their associated accuracy, with respect to a target frame which is realized by the known (and also noisy) coordinates in a subset of these points. The available data and their CV matrix in the initial frame are expressed in the partitioned form

$$\begin{bmatrix} \mathbf{X}' \\ \mathbf{Z}' \end{bmatrix}, \quad \begin{bmatrix} \mathbf{\Sigma}_{\mathbf{X}'} & \mathbf{\Sigma}_{\mathbf{X}'\mathbf{Z}'} \\ \mathbf{\Sigma}_{\mathbf{Z}'\mathbf{X}'} & \mathbf{\Sigma}_{\mathbf{Z}'} \end{bmatrix} \tag{3}$$

where $\mathbf{X}'$ and $\mathbf{Z}'$ denote the coordinate vectors of the common and new points, respectively. We consider a non-zero correlation between the two subsets ($\mathbf{\Sigma}_{\mathbf{X}'\mathbf{Z}'} \neq \mathbf{0}$) in line with our motivating discussion in Sect. 1. For the common points, an additional set of coordinates and their CV matrix, denoted as $\mathbf{X}$ and $\mathbf{\Sigma}_{\mathbf{X}}$, are also available with respect to the target frame.

It will be assumed that the coordinate vectors are uncorrelated between the two frames, that is

$$\mathbf{\Sigma}_{\mathbf{X}\mathbf{X}'} = \mathbf{0} \quad \text{and} \quad \mathbf{\Sigma}_{\mathbf{X}\mathbf{Z}'} = \mathbf{0} \tag{4}$$

The above simplification is justified if the known coordinates in the respective frames are obtained from different procedures and independent observation sets. For the sake of completeness and for the interested readers, the general treatment of the CFT problem without the presence of the last assumption is given later in the paper.

The general model of Eq. (1) provides the theoretical relationship between the two frames that are involved in our problem. This means that the systematic coordinate differences at the common points are presumably described as

$$E\{\mathbf{X} - \mathbf{X}'\} = \mathbf{G}\boldsymbol{\theta} \tag{5}$$

or, equivalently

$$\mathbf{x} = E\{\mathbf{X}'\} + \mathbf{G}\boldsymbol{\theta} \tag{6}$$

where $E\{\}$ denotes the expectation operator on a stochastic vector. A similar equation is also presumed for the new points

$$\mathbf{z} = E\{\mathbf{Z}'\} + \tilde{\mathbf{G}}\boldsymbol{\theta} \tag{7}$$

where $\tilde{\mathbf{G}}$ is the design matrix of the transformation model that refers to the new points. The lower case letters $\mathbf{x}$ and $\mathbf{z}$ correspond to the true coordinate vectors of the common and new points with respect to the target frame. Both of these vectors, together with the frame transformation parameters, constitute the unknowns of the geodetic CFT problem.

2.2 Standard CFT solution and its accuracy

The primary unknown in the CFT problem is the parameter vector $\boldsymbol{\theta}$ which can be directly estimated from the available data. By applying Eq. (1) over the common points and taking into account their coordinate noise in both frames, we get the system of observation equations

$$\mathbf{X} - \mathbf{X}' = \mathbf{G}\boldsymbol{\theta} + \mathbf{v}_{\mathbf{X}-\mathbf{X}'}, \quad \mathbf{v}_{\mathbf{X}-\mathbf{X}'} \sim (\mathbf{0}, \mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{X}'}) \tag{8}$$

which leads to the optimal LS estimate for the transformation parameters

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{G}^T(\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{X}'})^{-1}\mathbf{G}\right)^{-1} \mathbf{G}^T(\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{X}'})^{-1}(\mathbf{X} - \mathbf{X}') \tag{9}$$

The transformed coordinates in the target frame are then obtained via the forward implementation of Eq. (1), that is

$$\hat{\mathbf{x}}^{\mathrm{st}} = \mathbf{X}' + \mathbf{G}\hat{\boldsymbol{\theta}} \tag{10}$$

for the common points, and

$$\hat{\mathbf{z}}^{\mathrm{st}} = \mathbf{Z}' + \tilde{\mathbf{G}}\hat{\boldsymbol{\theta}} \tag{11}$$

for the new points. The above formulae yield unbiased estimates of the true coordinates in accordance to the modeling assumptions given in Eqs. (6) and (7). The superscript "st" indicates their association to the standard approach and it distinguishes them from the improved LS estimates that will be presented in Sect. 3.

The formal accuracy of the transformed coordinates in the target frame is determined by their CV matrices through straightforward error propagation to Eqs. (10) and (11). Hence, we have

Transformation accuracy at "common" points

$$\begin{aligned} \mathbf{\Sigma}_{\hat{\mathbf{x}}}^{\mathrm{st}} = \mathbf{\Sigma}_{\mathbf{X}'} + \mathbf{G}\mathbf{\Sigma}_{\hat{\boldsymbol{\theta}}}\mathbf{G}^T - \mathbf{\Sigma}_{\mathbf{X}'}(\mathbf{\Sigma}_{\mathbf{X}}+\mathbf{\Sigma}_{\mathbf{X}'})^{-1}\mathbf{G}\mathbf{\Sigma}_{\hat{\boldsymbol{\theta}}}\mathbf{G}^T \\ - \mathbf{G}\mathbf{\Sigma}_{\hat{\boldsymbol{\theta}}}\mathbf{G}^T(\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{X}'})^{-1}\mathbf{\Sigma}_{\mathbf{X}'} \end{aligned} \tag{12}$$

Transformation accuracy at "new" points

$$\begin{aligned} \mathbf{\Sigma}_{\hat{\mathbf{z}}}^{\mathrm{st}} = \mathbf{\Sigma}_{\mathbf{Z}'} + \tilde{\mathbf{G}}\mathbf{\Sigma}_{\hat{\boldsymbol{\theta}}}\tilde{\mathbf{G}}^T - \mathbf{\Sigma}_{\mathbf{Z}'\mathbf{X}'}(\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{X}'})^{-1}\mathbf{G}\mathbf{\Sigma}_{\hat{\boldsymbol{\theta}}}\tilde{\mathbf{G}}^T \\ - \tilde{\mathbf{G}}\mathbf{\Sigma}_{\hat{\boldsymbol{\theta}}}\mathbf{G}^T(\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{X}'})^{-1}\mathbf{\Sigma}_{\mathbf{X}'\mathbf{Z}'} \end{aligned} \tag{13}$$

where $\mathbf{\Sigma}_{\hat{\boldsymbol{\theta}}}$ is the CV matrix of the estimated transformation parameters

$$\mathbf{\Sigma}_{\hat{\boldsymbol{\theta}}} = \left(\mathbf{G}^T(\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{X}'})^{-1}\mathbf{G}\right)^{-1} \tag{14}$$

If the a priori coordinates of the common points in the target frame are considered errorless ($\mathbf{\Sigma}_{\mathbf{X}} \simeq \mathbf{0}$) then Eq. (12) is simplified as

$$\mathbf{\Sigma}_{\hat{\mathbf{x}}}^{\mathrm{st}} = \mathbf{\Sigma}_{\mathbf{X}'} - \mathbf{G}\mathbf{\Sigma}_{\hat{\boldsymbol{\theta}}}\mathbf{G}^T \tag{15}$$

and it reveals that the transformed coordinates at the common points will have better accuracy than their original values in the initial frame. This conditional improvement does not necessarily occur for the transformed coordinates at the new points. As a matter of fact, if the coordinate vectors of the common and new points are uncorrelated with each other then Eq. (13) takes the form

$$\mathbf{\Sigma}_{\hat{\mathbf{z}}}^{\mathrm{st}} = \mathbf{\Sigma}_{\mathbf{Z}'} + \tilde{\mathbf{G}}\mathbf{\Sigma}_{\hat{\boldsymbol{\theta}}}\tilde{\mathbf{G}}^T \tag{16}$$

and it shows that the transformed coordinates become more dispersed than their original values in the initial frame. The influence of the cross-CV matrix $\mathbf{\Sigma}_{\mathbf{X}'\mathbf{Z}'}$ may or may not alter this situation.

In any case, and regardless of the treatment of the a priori coordinates in the target frame ($\mathbf{\Sigma}_{\mathbf{X}} = \mathbf{0}$ or $\mathbf{\Sigma}_{\mathbf{X}} \neq \mathbf{0}$), the transformation results obtained from the standard approach are non-optimal as explained in the next section.

2.3 Remarks

In contrast to the LS solution for the frame transformation parameters, the determination of the transformed coordinates is not embedded in an optimal estimation framework. It is just based on the transformation model of Eq. (1) using the noisy initial coordinates and the estimated parameters in a forward manner, that is

$$\begin{bmatrix} \hat{\mathbf{x}}^{st} \\ \hat{\mathbf{z}}^{st} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \\ \mathbf{Z}' \end{bmatrix} + \begin{bmatrix} \mathbf{G} \\ \tilde{\mathbf{G}} \end{bmatrix} \hat{\boldsymbol{\theta}} \tag{17}$$

Such an approach is certainly not wrong, however it is not optimal either as it violates the theoretical requirements of the best linear unbiased estimation (BLUE, e.g., Koch 1999) due to the stochastic dependence between the original data and the frame transformation parameters, both of which appear at the right side of the last equation. The transformed coordinates are linearly related to a BLUE-type parameter estimator and they depend also on the additive terms $\mathbf{X}'$ and $\mathbf{Z}'$ which are directly correlated with that estimator. Indeed, $\mathbf{X}'$ is hidden in $\hat{\boldsymbol{\theta}}$ as per Eq. (9) while $\mathbf{Z}'$ is supposed to be correlated with $\mathbf{X}'$ and, thus, also with $\hat{\boldsymbol{\theta}}$. As a result, the generalized Gauss-Markov theorem does not hold in this case and therefore Eq. (17) will not provide the best (minimum variance) linear unbiased solution for the transformed coordinates in the target frame[1].

In simple words, the drawback of the standard stepwise approach is that the data noise is partially minimized in the result of Eq. (17) and only for the part contained in the frame transformation parameters. The existing noise in the additive vectors $\mathbf{X}'$ and $\mathbf{Z}'$ is not mitigated in the standard CFT solution and it is fully absorbed by the transformed coordinates.

[1] According to this theorem (Koch 1999, pp. 156–158) the BLUE of a linear function $\mathbf{q} = \mathbf{Q}\boldsymbol{\theta}+\mathbf{c}$ of a parameter vector is given by $\hat{\mathbf{q}} = \mathbf{Q}\hat{\boldsymbol{\theta}}+\mathbf{c}$, where $\hat{\boldsymbol{\theta}}$ is the BLUE of the parameter vector while $\mathbf{Q}$ and $\mathbf{c}$ are fixed (deterministic) quantities. The transformed coordinates in Eq. (17), on the other hand, have the form $\hat{\mathbf{q}} = \mathbf{Q}\hat{\boldsymbol{\theta}} + \mathbf{c}'$, where $\mathbf{c}'$ corresponds to an observed vector which is correlated with the parameter estimator.

3 Optimal LS approach for geodetic frame transformation

3.1 Enhanced adjustment model

Two straightforward augmentations in the classic adjustment model for CFT problems are employed in this section to derive the optimal LS estimators for the transformed coordinates in the target frame.

3.1.1 Inclusion of the new stations

A first step is the inclusion of the initial coordinates of the new points into the LS adjustment process. Following the formulation from the previous section, we may set up the extended system of observation equations

$$\mathbf{X} - \mathbf{X}' = \mathbf{G}\boldsymbol{\theta} + \mathbf{v}_{\mathbf{X}-\mathbf{X}'}, \quad \mathbf{v}_{\mathbf{X}-\mathbf{X}'} \sim (\mathbf{0}, \mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{X}'}) \tag{18}$$

and

$$\mathbf{Z}' = \mathbf{z} - \tilde{\mathbf{G}}\boldsymbol{\theta} + \mathbf{v}_{\mathbf{Z}'}, \quad \mathbf{v}_{\mathbf{Z}'} \sim (\mathbf{0}, \mathbf{\Sigma}_{\mathbf{Z}'}) \tag{19}$$

The random error vectors in the above system are correlated with each other in accordance with the adopted stochastic model for the coordinate datasets (see Sect. 2.1.2). Considering that $\mathbf{X}$ is uncorrelated with both $\mathbf{X}'$ and $\mathbf{Z}'$, we have

$$E\{\mathbf{v}_{\mathbf{X}-\mathbf{X}'}\mathbf{v}_{\mathbf{Z}'}^T\} = -\mathbf{\Sigma}_{\mathbf{X}'\mathbf{Z}'} \tag{20}$$

The system of Eqs. (18)–(19) is expressed in block-matrix notation as

$$\begin{bmatrix} \mathbf{X} - \mathbf{X}' \\ \mathbf{Z}' \end{bmatrix} = \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ -\tilde{\mathbf{G}} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta} \\ \mathbf{z} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_{\mathbf{X}-\mathbf{X}'} \\ \mathbf{v}_{\mathbf{Z}'} \end{bmatrix} \tag{21}$$

and it can be used for the joint LS estimation of (1) the frame transformation parameters and (2) the transformed coordinates of the new points. The weight matrix of the input data has the form

$$\mathbf{P} = \begin{bmatrix} \mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{X}'} & -\mathbf{\Sigma}_{\mathbf{X}'\mathbf{Z}'} \\ -\mathbf{\Sigma}_{\mathbf{Z}'\mathbf{X}'} & \mathbf{\Sigma}_{\mathbf{Z}'} \end{bmatrix}^{-1} \tag{22}$$

and it takes into account the full stochastic model of the entire point set to be analyzed.

From the weighted LS adjustment of Eq. (21) we obtain the solution (see Appendix)

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{G}^T(\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{X}'})^{-1}\mathbf{G}\right)^{-1} \mathbf{G}^T(\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{X}'})^{-1}(\mathbf{X} - \mathbf{X}') \tag{23}$$

$$\hat{\mathbf{z}} = \mathbf{Z}' + \tilde{\mathbf{G}}\hat{\boldsymbol{\theta}} + \mathbf{\Sigma}_{\mathbf{Z}'\mathbf{X}'}(\mathbf{\Sigma}_{\mathbf{X}} + \mathbf{\Sigma}_{\mathbf{X}'})^{-1}(\mathbf{X} - \mathbf{X}' - \mathbf{G}\hat{\boldsymbol{\theta}}) \tag{24}$$

The optimal estimate $\hat{\boldsymbol{\theta}}$ remains the same as the one obtained by the standard approach in Sect. 2.2. This was actually expected since the inclusion of the new points does not contribute any additional information for the estimation of the frame transformation parameters. On the other hand, the optimal estimate $\hat{\mathbf{z}}$ differs from the standard estimate $\hat{\mathbf{z}}^{\mathrm{st}}$ that was given in Eq. (11). The latter has lower statistical accuracy than the optimal estimate of Eq. (24) which is accompanied by the CV matrix (after straightforward covariance propagation)

$$\begin{aligned}\boldsymbol{\Sigma}_{\hat{\mathbf{z}}} &= \boldsymbol{\Sigma}_{\mathbf{Z}'} + \tilde{\mathbf{G}}\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\tilde{\mathbf{G}}^T - \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}\mathbf{G}\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\tilde{\mathbf{G}}^T \\ &\quad - \tilde{\mathbf{G}}\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\mathbf{G}^T(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}\boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{Z}'} \\ &\quad - \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'} - \mathbf{G}\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\mathbf{G}^T) \\ &\quad \times (\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}\boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{Z}'} \end{aligned} \quad (25)$$

Recalling the CV matrix of $\hat{\mathbf{z}}^{\mathrm{st}}$ from Eq. (13), we can express the last equation in the equivalent form

$$\boldsymbol{\Sigma}_{\hat{\mathbf{z}}} = \boldsymbol{\Sigma}_{\hat{\mathbf{z}}}^{\mathrm{st}} - \boldsymbol{\Delta\Sigma}_{\hat{\mathbf{z}}} \quad (26)$$

where

$$\begin{aligned}\boldsymbol{\Delta\Sigma}_{\hat{\mathbf{z}}} &= \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'} - \mathbf{G}\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\mathbf{G}^T) \\ &\quad \times (\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}\boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{Z}'} \end{aligned} \quad (27)$$

The matrix $\boldsymbol{\Delta\Sigma}_{\hat{\mathbf{z}}}$ is non-negative definite (the proof is trivial considering that the central term in the last equation is the CV matrix of the adjusted residuals at the common points) and therefore the transformed coordinates from Eq. (24) will have smaller error variances than the transformed coordinates by the standard stepwise approach.

3.1.2 Stacking of the common stations

The previous formulation does not provide directly an optimal LS estimate for the transformed coordinates at the common points. To obtain such an estimate we should use a more extended system of observation equations

$$\mathbf{X} = \mathbf{x} + \mathbf{v}_{\mathbf{X}}, \quad \mathbf{v}_{\mathbf{X}} \sim (\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{X}}) \quad (28)$$

$$\mathbf{X}' = \mathbf{x} - \mathbf{G}\boldsymbol{\theta} + \mathbf{v}_{\mathbf{X}'}, \quad \mathbf{v}_{\mathbf{X}'} \sim (\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{X}'}) \quad (29)$$

$$\mathbf{Z}' = \mathbf{z} - \tilde{\mathbf{G}}\boldsymbol{\theta} + \mathbf{v}_{\mathbf{Z}'}, \quad \mathbf{v}_{\mathbf{Z}'} \sim (\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{Z}'}) \quad (30)$$

or, equivalently, in block-matrix notation

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{X}' \\ \mathbf{Z}' \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} \\ -\mathbf{G} & \mathbf{I} & \mathbf{0} \\ -\tilde{\mathbf{G}} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta} \\ \mathbf{x} \\ \mathbf{z} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_{\mathbf{X}} \\ \mathbf{v}_{\mathbf{X}'} \\ \mathbf{v}_{\mathbf{Z}'} \end{bmatrix} \quad (31)$$

The above scheme incorporates explicitly the target frame coordinates of the common points to the unknown parameters of the CFT problem. Note that by subtracting Eqs. (28) and (29) we can eliminate the unknown term $\mathbf{x}$ and we obtain the same system of observation equations that was analyzed in the previous section.

The data weight matrix associated with Eq. (31) has the form

$$\mathbf{P} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{X}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\mathbf{X}'} & \boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{Z}'} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'} & \boldsymbol{\Sigma}_{\mathbf{Z}'} \end{bmatrix}^{-1} \quad (32)$$

and it (also) considers the full stochastic model for the available data. Note the sign difference compared to the weight matrix in Eq. (22) due to the new algebraic structure of the observation equations. The more general case, yet still unrealistic in geodetic practice, of a full weight matrix by assuming non-zero coordinate correlation *between the two frames* is treated later in Sect. 5.

The LS adjustment of Eq. (31) using the aforementioned weight matrix leads to the solution (see Appendix)

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{G}^T(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}\mathbf{G}\right)^{-1}\mathbf{G}^T(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}(\mathbf{X} - \mathbf{X}') \quad (33)$$

$$\hat{\mathbf{x}} = \mathbf{X}' + \mathbf{G}\hat{\boldsymbol{\theta}} + \boldsymbol{\Sigma}_{\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}(\mathbf{X} - \mathbf{X}' - \mathbf{G}\hat{\boldsymbol{\theta}}) \quad (34)$$

$$\hat{\mathbf{z}} = \mathbf{Z}' + \tilde{\mathbf{G}}\hat{\boldsymbol{\theta}} + \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}(\mathbf{X} - \mathbf{X}' - \mathbf{G}\hat{\boldsymbol{\theta}}) \quad (35)$$

The optimal estimates $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{z}}$ remain the same with the ones obtained by the weighted LS adjustment in Sect. 3.1.1—this is not surprising since they are in any case the unique BLUE of the problem at hand. The gain from the current scheme is the optimal estimate $\hat{\mathbf{x}}$ which differs from the standard estimate $\hat{\mathbf{x}}^{\mathrm{st}}$ given in Eq. (10). Their difference has the form of a prediction-like term of similar structure with the one that appears in the transformed coordinates of the new points. The presence of this term improves the accuracy of the transformed coordinates at the common points, a fact that can be inferred from their CV matrix

$$\begin{aligned}\boldsymbol{\Sigma}_{\hat{\mathbf{x}}} &= \boldsymbol{\Sigma}_{\mathbf{X}'} + \mathbf{G}\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\mathbf{G}^T - \boldsymbol{\Sigma}_{\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}\mathbf{G}\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\mathbf{G}^T \\ &\quad - \mathbf{G}\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\mathbf{G}^T(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}\boldsymbol{\Sigma}_{\mathbf{X}'} \\ &\quad - \boldsymbol{\Sigma}_{\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'} - \mathbf{G}\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\mathbf{G}^T) \\ &\quad \times (\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}\boldsymbol{\Sigma}_{\mathbf{X}'} \end{aligned} \quad (36)$$

Taking into account the CV matrix of $\hat{\mathbf{x}}^{\mathrm{st}}$ from Eq. (12), the last equation can be expressed as

$$\boldsymbol{\Sigma}_{\hat{\mathbf{x}}} = \boldsymbol{\Sigma}_{\hat{\mathbf{x}}}^{\mathrm{st}} - \boldsymbol{\Delta\Sigma}_{\hat{\mathbf{x}}} \quad (37)$$

where

$$\begin{aligned}\boldsymbol{\Delta\Sigma}_{\hat{\mathbf{x}}} &= \boldsymbol{\Sigma}_{\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'} - \mathbf{G}\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\mathbf{G}^T) \\ &\quad \times (\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}\boldsymbol{\Sigma}_{\mathbf{X}'} \end{aligned} \quad (38)$$

The matrix $\boldsymbol{\Delta\Sigma}_{\hat{\mathbf{x}}}$ is non-negative definite (again, the proof is trivial considering that the central term in the last equation is the CV matrix of the adjusted residuals at the common points) and therefore the transformed coordinates from Eq. (34) will have smaller error variances than the transformed coordinates by the standard stepwise approach.

3.2 Discussion

The traditional methodology for geodetic frame transformation does not provide an optimal solution for the computed coordinates in the target frame. Our previous analysis showed that the BLUE estimators for the transformed coordinates are related to the standard estimators of Eq. (17) via the expression

$$\begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{z}} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{x}}^{st} \\ \hat{\mathbf{z}}^{st} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{X}'} \\ \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'} \end{bmatrix} (\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1} \underbrace{(\mathbf{X} - \mathbf{X}' - \mathbf{G}\hat{\boldsymbol{\theta}})}_{\mathbf{X}-\hat{\mathbf{x}}^{st}} \tag{39}$$

or, in a more compact form

$$\begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{z}} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{x}}^{st} \\ \hat{\mathbf{z}}^{st} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\delta}\hat{\mathbf{x}} \\ \boldsymbol{\delta}\hat{\mathbf{z}} \end{bmatrix} \tag{40}$$

The prediction-like correction terms that appear in the above equations are similar to those found in least squares collocation theory, and they contain the required updating for the transformed coordinates obtained by the standard stepwise approach. From a computational viewpoint, these corrections do not require any additional matrix inversion other the one already used for the determination of the frame transformation parameters. It is worth noting that Eq. (39) is closely related to the so-called covariance adjustment technique for the approximation of a BLUE estimator as a sum of a linear unbiased estimator and a linear zero-mean estimator; for more details see Schwarz (1974).

At this point it is useful to clarify the main aspects of the terms $\boldsymbol{\delta}\hat{\mathbf{x}}$ and $\boldsymbol{\delta}\hat{\mathbf{z}}$ in view of their contribution to the increased accuracy of the transformation results. Both of them are obtained from a stochastic mapping of the zero-mean residuals at the common points $\mathbf{X} - \hat{\mathbf{x}}^{st}$ (which are often used for the statistical empirical assessment of the standard CFT solution) into appropriate corrections for the transformed coordinates. These corrections carry the statistical information for the available data that is either overlooked or misused in the context of the stepwise approach while their role is to minimize the propagated noise from the initial coordinates to their transformed values with respect to the target frame—see also the discussion in Sect. 2.3.

Interestingly enough, Eq. (39) resembles a Kalman-type filtering which improves the standard estimates of the transformed coordinates by exploiting the noise characteristics of the various datasets (Gibbs 2011). In the case of the common points, this improvement essentially stems from a weighted averaging of their available coordinates in the target frame. In fact from Eq. (39) we have

$$\hat{\mathbf{x}} = (\mathbf{I} - \boldsymbol{\Sigma}_{\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1})\hat{\mathbf{x}}^{st} + \boldsymbol{\Sigma}_{\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}\mathbf{X} \tag{41}$$

and by taking into account the well known matrix identities[2]

$$\begin{aligned} \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1} &= \mathbf{I} - (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1} \\ &= (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1} \end{aligned} \tag{42}$$

we get the equivalent expression

$$\hat{\mathbf{x}} = (\boldsymbol{\Sigma}_{\mathbf{X}}^{-1} + \boldsymbol{\Sigma}_{\mathbf{X}'}^{-1})^{-1} \left(\boldsymbol{\Sigma}_{\mathbf{X}'}^{-1}\hat{\mathbf{x}}^{st} + \boldsymbol{\Sigma}_{\mathbf{X}}^{-1}\mathbf{X}\right) \tag{43}$$

which represents a weighted average of the coordinate vectors $\hat{\mathbf{x}}^{st}$ and $\mathbf{X}$. This procedure reduces the noise that exists in the standard estimate of the transformed coordinates [actually the noise part caused by the additive vector $\mathbf{X}'$ in Eq. (17)] by taking advantage of the prior information for the common points in the target frame.

In the case of the new points, the improvement of the transformation solution is linked to the minimization of the propagated noise from the data vector $\mathbf{Z}'$ to their transformed coordinates (note that the standard estimate $\hat{\mathbf{z}}^{st}$ totally ignores this noise effect). For this purpose the following correction term is employed

$$\boldsymbol{\delta}\hat{\mathbf{z}} = \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}(\mathbf{X} - \hat{\mathbf{x}}^{st}) \tag{44}$$

which is predicted from the adjusted residuals at the common points by exploiting the cross-CV matrix $\boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'}$.

3.2.1 *Does the optimal CFT solution in Eq. (39) refer to the same frame with the standard CFT solution in Eq. (17)?*

In principle, the frame consistency between the BLUE and the standard estimates for the transformed coordinates is verified by the unbiasedness property

$$E\{\hat{\mathbf{x}}\} = E\{\hat{\mathbf{x}}^{st}\} = \mathbf{x} \;\text{ and }\; E\{\hat{\mathbf{z}}\} = E\{\hat{\mathbf{z}}^{st}\} = \mathbf{z} \tag{45}$$

which holds true provided that no systematic discrepancies exist between the initial and the target frame other than those already absorbed by the transformation parameters $\hat{\boldsymbol{\theta}}$; see also Sect. 2.1.2.

Alternatively, the nullification of the *posterior LS estimate* of the frame transformation parameters

$$\hat{\boldsymbol{\theta}}^{post} = \begin{cases} (\mathbf{G}^T\mathbf{W}\mathbf{G})^{-1}\mathbf{G}^T\mathbf{W}(\mathbf{X} - \hat{\mathbf{x}}^{st}) = \mathbf{0} \\ \qquad\qquad \textit{and} \\ (\mathbf{G}^T\mathbf{W}\mathbf{G})^{-1}\mathbf{G}^T\mathbf{W}(\mathbf{X} - \hat{\mathbf{x}}) = \mathbf{0} \end{cases} \tag{46}$$

can also confirm that the BLUE and the standard estimates of the transformed coordinates refer to the same frame as the prior coordinates $\mathbf{X}$ of the common points. The matrix $\mathbf{W}$

[2] These identities are directly obtained from the general matrix property (e.g. Blewitt 1998, p. 248) $(\boldsymbol{\Lambda}_1 \pm \boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_2^{-1}\boldsymbol{\Lambda}_{12}^T)^{-1} = \boldsymbol{\Lambda}_1^{-1} \mp \boldsymbol{\Lambda}_1^{-1}\boldsymbol{\Lambda}_{12}(\boldsymbol{\Lambda}_2 \pm \boldsymbol{\Lambda}_{12}^T\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\Lambda}_{12})^{-1}\boldsymbol{\Lambda}_{12}^T\boldsymbol{\Lambda}_1^{-1}$ by substituting $\boldsymbol{\Lambda}_1 = \mathbf{A}$, $\boldsymbol{\Lambda}_2^{-1} = \mathbf{B}$ and $\boldsymbol{\Lambda}_{12} = \boldsymbol{\Lambda}_{12}^T = \mathbf{I}$.

denotes a weight matrix which does not necessarily remain the same in both cases.

In the case of the standard solution ($\hat{\mathbf{x}}^{\mathrm{st}}$) is easily shown that

$$\mathbf{X} - \hat{\mathbf{x}}^{\mathrm{st}} = \left(\mathbf{I} - \mathbf{G}\left(\mathbf{G}^T(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}\mathbf{G}\right)^{-1} \times \mathbf{G}^T(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}\right)(\mathbf{X} - \mathbf{X}') \quad (47)$$

and by substituting the above expression into Eq. (46) we confirm that $\hat{\boldsymbol{\theta}}^{\mathrm{post}}=\mathbf{0}$ under the weight matrix choice $\mathbf{W} = (\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}$. On the other hand, the optimal solution ($\hat{\mathbf{x}}$) satisfies the filtering relationship

$$\mathbf{X} - \hat{\mathbf{x}} = \boldsymbol{\Sigma}_{\mathbf{X}}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1}(\mathbf{X} - \hat{\mathbf{x}}^{\mathrm{st}}) \quad (48)$$

which, in conjunction with Eq. (47), implies the nullification of $\hat{\boldsymbol{\theta}}^{\mathrm{post}}$ under the weight matrix choice $\mathbf{W} = \boldsymbol{\Sigma}_{\mathbf{X}}^{-1}$. The difference in the respective weight matrices corroborates the fact that $\hat{\mathbf{x}}$ has always higher statistical accuracy than $\hat{\mathbf{x}}^{\mathrm{st}}$ as already discussed in the previous sections.

3.2.2 What happens when the prior coordinates in the target frame are considered errorless?

This is a special case that should be understood in a relative sense without assigning necessarily the meaning of noise-free coordinates to the data vector $\mathbf{X}$. Practically, if the a priori coordinate accuracy of the common points is significantly better in the target frame compared to the initial frame, that is

$$\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'} \simeq \boldsymbol{\Sigma}_{\mathbf{X}'} \quad \text{and} \quad \boldsymbol{\Sigma}_{\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1} \simeq \mathbf{I} \quad (49)$$

then the optimally transformed coordinates at the common points become

$$\hat{\mathbf{x}} \simeq \hat{\mathbf{x}}^{\mathrm{st}} + (\mathbf{X} - \hat{\mathbf{x}}^{\mathrm{st}}) = \mathbf{X} \quad (50)$$

thus reproducing the prior high-quality coordinates in the target frame. The corresponding solution for the transformed coordinates at the new points takes the form

$$\hat{\mathbf{z}} \simeq \hat{\mathbf{z}}^{\mathrm{st}} + \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'}\boldsymbol{\Sigma}_{\mathbf{X}'}^{-1}(\mathbf{X} - \hat{\mathbf{x}}^{\mathrm{st}}) \quad (51)$$

Note that the magnitude of the correction vectors $\boldsymbol{\delta}\hat{\mathbf{x}}$ and $\boldsymbol{\delta}\hat{\mathbf{z}}$ will generally increase as the CV matrix $\boldsymbol{\Sigma}_{\mathbf{X}}$ tends to zero, a fact that can be inferred from the analytic expressions in Eq. (39); see also the numerical examples in the next section.

In the opposite scenario, that is if the a priori coordinate accuracy of the common points is significantly better in the initial frame compared to the target frame

$$\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'} \simeq \boldsymbol{\Sigma}_{\mathbf{X}}, \quad \boldsymbol{\Sigma}_{\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1} \simeq \mathbf{0},$$
$$\boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'}(\boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'})^{-1} \simeq \mathbf{0} \quad (52)$$

then the optimal estimates for the transformed coordinates become

$$\hat{\mathbf{x}} \simeq \hat{\mathbf{x}}^{\mathrm{st}} \quad \text{and} \quad \hat{\mathbf{z}} \simeq \hat{\mathbf{z}}^{\mathrm{st}} \quad (53)$$

for the common and the new points, respectively.

4 Numerical examples

The optimal CFT scheme that was described in the previous section will be evaluated herein using actual real-network data. As an example we consider the Helmert transformation of a weekly combined solution of the EUREF permanent network (EPN) to the International Terrestrial Reference Frame 2008 (ITRF08). The used EPN coordinates and their full CV matrix refer to GPS week 1762 and were obtained from the corresponding SINEX file produced by the BKG (Bundesamt für Kartographie und Geodäsie) data analysis center. This particular solution contains 244 stations, 95 of which are included in the ITRF08 realization of the global IGS network. The ITRF08 coordinates and their full CV matrix for these 95 stations have been extracted from the solution file ITRF2008-TRF-IGS.SNX which was obtained from the IGN/ITRF public ftp site, and they were subsequently reduced to the mean epoch of the weekly EPN solution.

Two different cases are considered for the treatment of the ITRF08 reference coordinates at the common points, namely (1) the "noise-free" case by setting their prior CV matrix to zero, and (2) the fully-weighted case by incorporating their actual CV matrix into the LS adjustment procedure. The estimated Helmert transformation parameters between the weekly EPN solution and ITRF08 have been computed from Eq. (33) and they are given in Table 1.

To assess the significance of the rigorous CFT approach we investigate the differences between the optimal and the standard estimates for the transformed coordinates in the underlying network. For this purpose, the correction vectors $\boldsymbol{\delta}\hat{\mathbf{x}}$ and $\boldsymbol{\delta}\hat{\mathbf{z}}$ have been respectively determined by the prediction-like expressions in Eq. (39) for each weighting choice of the ITRF08 reference coordinates (i.e. $\boldsymbol{\Sigma}_{\mathbf{X}} = \mathbf{0}$ and $\boldsymbol{\Sigma}_{\mathbf{X}} \neq \mathbf{0}$). The statistics of their Cartesian components are shown in Table 2 from which the following conclusions can be drawn.

The coordinate corrections at the new points ($\boldsymbol{\delta}\hat{\mathbf{z}}$) amount to several mm for all three Cartesian components. Comparing the respective *mean* and *rms* values in Table 2, it is evident that the major part of these corrections is related to an "apparent bias" that is hidden in the (suboptimal) transformed

Table 1 Estimated Helmert transformation parameters between the EPN weekly combined solution (GPS week 1762) and ITRF08

Treatment of the ITRF08 reference coordinates at the common points	t_x [cm]	t_y [cm]	t_z [cm]	ε_x [mas]	ε_y [mas]	ε_z [mas]	δs [ppb]
"Noise-free"	−2.1	−3.9	6.6	0.46	−6.97	5.48	−6.7
Fully-weighted	−4.8	−8.5	10.0	1.99	−8.27	4.83	−7.1

All 95 common points were used in the computations

Table 2 Statistics of the corrections vectors $\boldsymbol{\delta\hat{x}}$ and $\boldsymbol{\delta\hat{z}}$ for the transformed EPN weekly solution (GPS week 1762)

All values in mm

	$\boldsymbol{\delta\hat{x}}$ (at 95 common pts)				$\boldsymbol{\delta\hat{z}}$ (at 149 new pts)			
	max	*min*	*mean*	*rms*	*max*	*min*	*mean*	*rms*
"Noise-free" case: $\boldsymbol{\Sigma}_{\mathbf{X}} = \mathbf{0}$								
Cartesian component X	204.1	−115.1	2.1	36.1	9.6	−18.9	−3.3	6.4
Cartesian component Y	249.6	−578.0	−2.5	70.5	7.4	−8.3	−4.5	5.1
Cartesian component Z	377.1	−125.5	5.9	48.3	20.6	−13.4	8.3	9.2
Fully-weighted case: $\boldsymbol{\Sigma}_{\mathbf{X}} \neq \mathbf{0}$								
Cartesian component X	44.7	−45.1	−0.0	12.7	3.4	−9.9	−2.7	3.3
Cartesian component Y	63.6	−16.2	−0.4	10.5	1.5	−4.2	−2.3	2.4
Cartesian component Z	27.1	−95.2	1.2	15.1	8.8	−2.6	2.2	2.6

coordinates by the stepwise approach. In contrast, the coordinate corrections at the common points ($\boldsymbol{\delta\hat{x}}$) exhibit a more random-like pattern, a fact that can be confirmed from the larger discrepancy between the *mean* and *rms* values in the corresponding columns of Table 2; see also Figs. 1 and 2. The latter corrections reach up to several cm (or even dm in the case $\boldsymbol{\Sigma}_{\mathbf{X}} = \mathbf{0}$) and they are significantly larger from the respective corrections which are associated with the new points.

The actual differences between the optimal and the standard estimates for the transformed coordinates are plotted in Fig. 2 (noise-free case) and Fig. 3 (fully-weighted case). The magnitude of these differences is evaluated in terms of their signal-to-noise ratio (SNR) which is also depicted in the same figures. For the SNR computation at each point we have used the formula

$$SNR = 10\log_{10}\left\{\begin{array}{l}\left|\hat{X}_i - \hat{X}_i^{\rm st}\right| / \sigma_{\hat{X}_i^{\rm st}}\\ \left|\hat{Y}_i - \hat{Y}_i^{\rm st}\right| / \sigma_{\hat{Y}_i^{\rm st}}\\ \left|\hat{Z}_i - \hat{Z}_i^{\rm st}\right| / \sigma_{\hat{Z}_i^{\rm st}}\end{array}\right. \quad (54)$$

where the accuracy of the standard estimates is obtained from the rigorous CV matrices in Eqs. (12) and (13). The above equation quantifies (in dB) the significance of the Cartesian components of the coordinate corrections at each EPN station. In general, negative SNR values correspond to a statistically negligible effect while SNR values higher than ∼4.8 indicate a strong correction signal which exceeds the 99 % uncertainty level of the stepwise transformation approach.

The SNRs of the correction vector $\boldsymbol{\delta\hat{z}}$ lie well above 5 dB (noise-free case) and 15 dB (fully-weighted case) for most of the 149 new stations, thus indicating a significant difference between the optimal and the standard estimates of their transformed coordinates. On the other hand, the correction vector $\boldsymbol{\delta\hat{x}}$ shows a stronger variability over the 95 common stations and a smaller apparent bias in their transformed coordinates which are obtained by the stepwise approach. However the magnitudes of these corrections are considerably larger and even more significant as they have higher SNRs from the respective corrections that are associated with the new stations.

5 A (more) general CFT problem

The optimal treatment of the CFT problem, as presented in Sect. 3, can support all related applications in current geodetic practice. The improved estimators in Eq. (39) take into account the *intra-frame covariances* between common and new points which are commonly present, yet usually ignored, in network solutions that need to be aligned to another coordinate frame. On the other hand, the *inter-frame covariances* in the available datasets have been disregarded in the weight matrix of the LS adjustment model [see Eq. (32)], a fact that makes our previous estimators suboptimal in the general sense. Although the consideration of this effect seems difficult to be implemented in practical cases (due to the absence of knowledge of the corresponding cross-CV matrices), it is nevertheless essential for a complete and thorough treatment of the CFT problem.

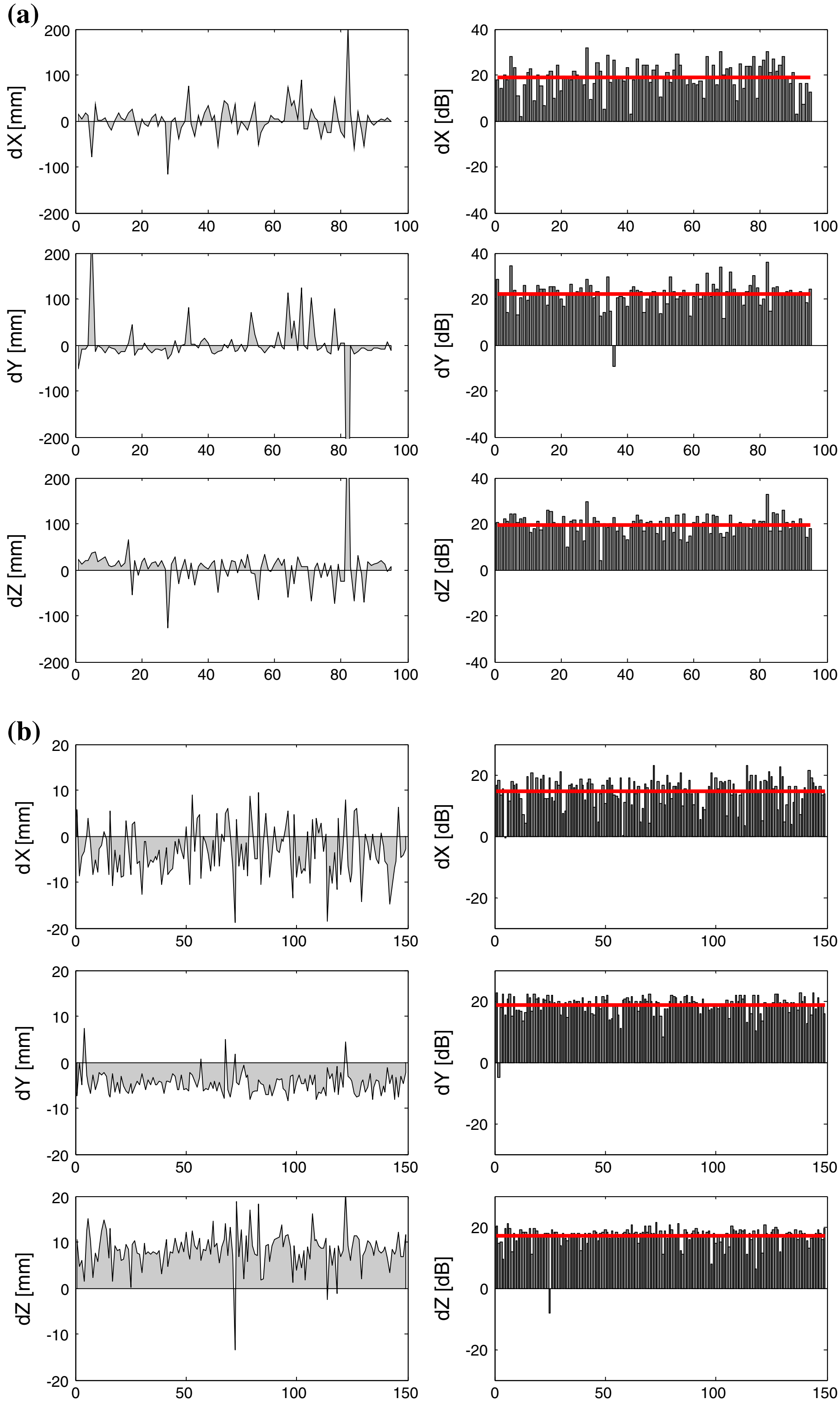
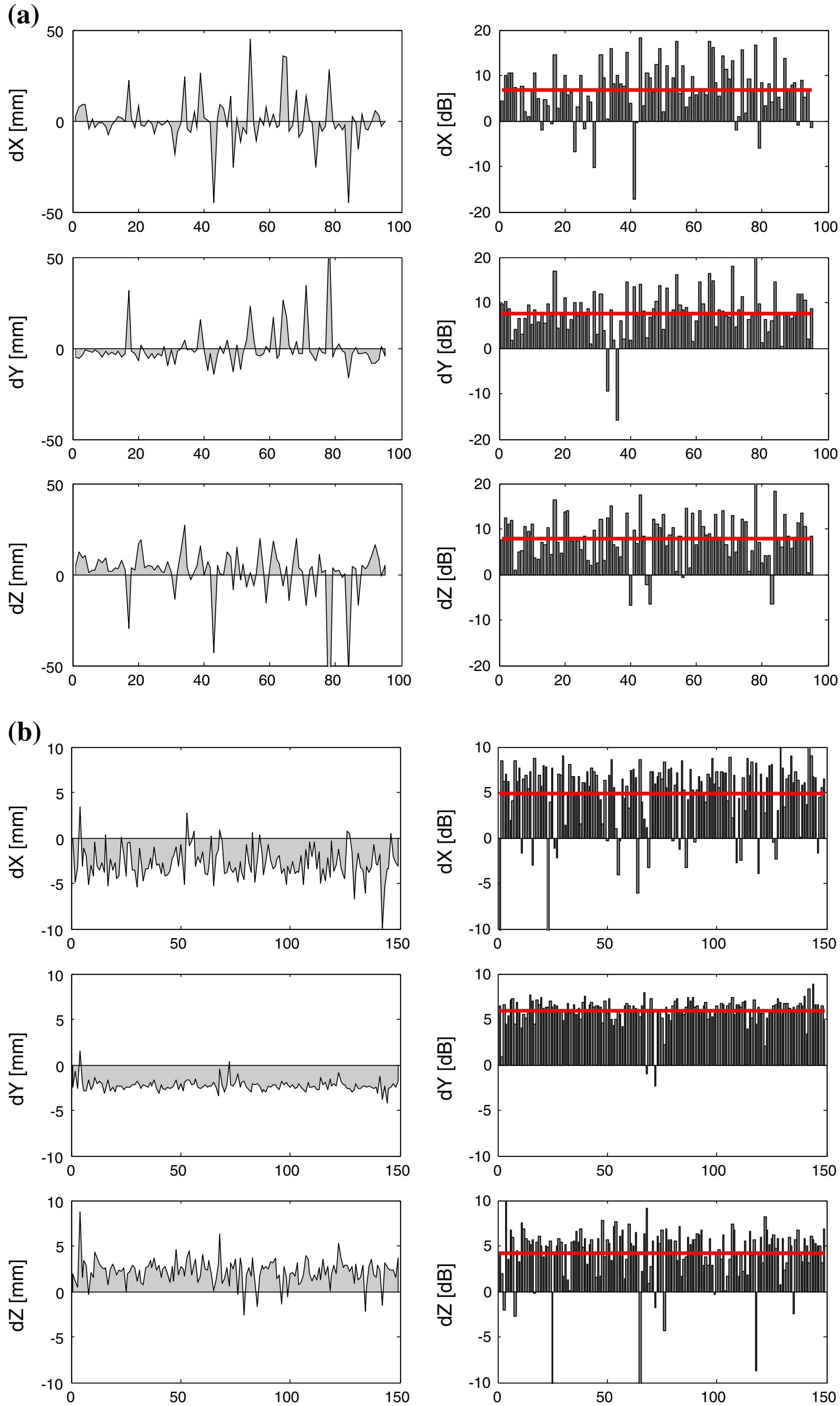


Fig. 2 Cartesian components and their corresponding SNRs of the correction vectors for the transformed coordinates at the 95 common points (**a**) and 149 new points (**b**) in the EPN weekly solution (GPS week 1,762). The results refer to the noise-free case ($\mathbf{\Sigma_X} = \mathbf{0}$). The straight red line in the right-column plots indicates the average SNR value

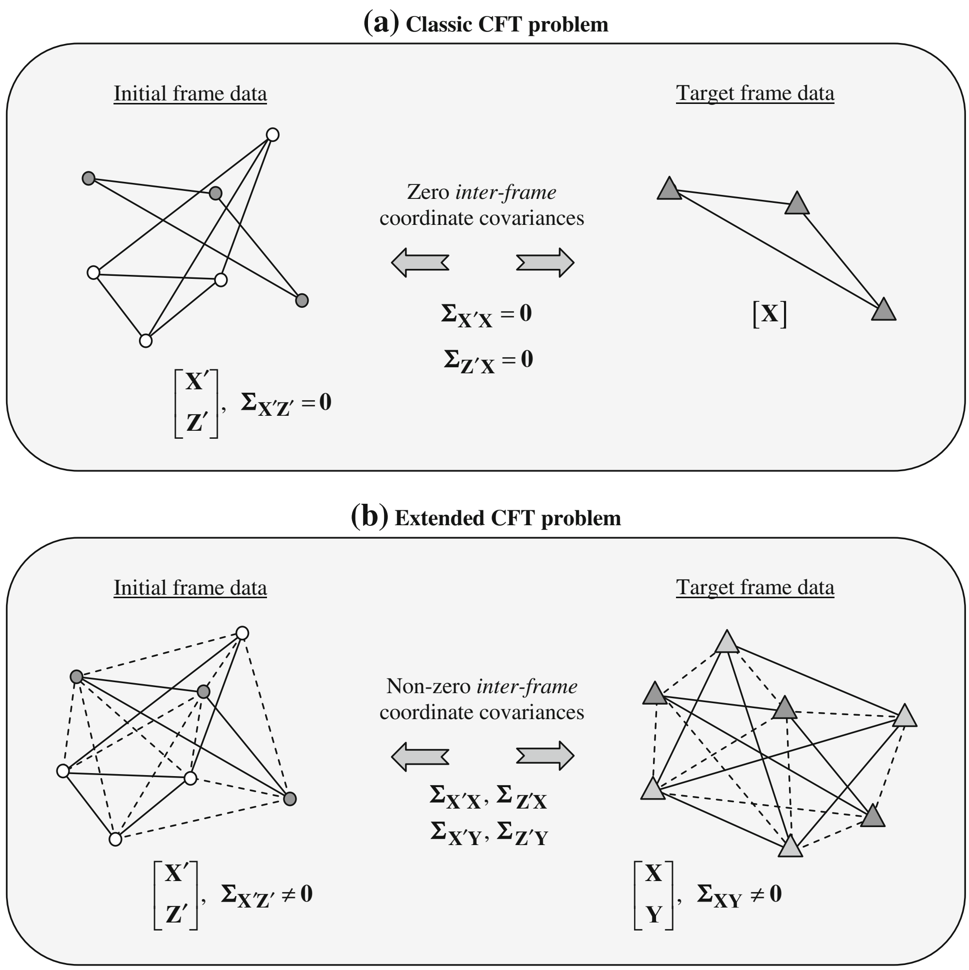
Another point of interest stems from the fact that, in most cases, the prior coordinates of the common points in the target frame (i.e. $\mathbf{X}$, $\mathbf{\Sigma_X}$) do not constitute a "stand-alone" dataset. In fact these reference coordinates are often estimated jointly with the coordinates of additional stations through a LS adjustment of a larger network

Fig. 3 Cartesian components and their corresponding SNRs of the correction vectors for the transformed coordinates at the 95 common points (**a**) and 149 new points (**b**) in the EPN weekly solution (GPS week 1,762). The results refer to the fully-weighted case ($\mathbf{\Sigma_X} \neq \mathbf{0}$). The straight red line in the right-column plots indicates the average SNR value



which provides the full realization of the desired target frame for our transformation problem. A legitimate concern is therefore the handling of the coordinate correlation between (1) the employed common stations and (2) the other (non-common) reference stations in the target frame, and its possible influence on the final results. A

Fig. 4 A schematic representation of the classic and the extended version of the geodetic CFT problem



schematic overview of the two aforementioned issues is given in Fig. 4.

In order to properly model the effects of the intra-frame and inter-frame covariances, and also to account for the information coming from other (non-common) reference stations in the target frame, we shall extend the system of observation Eqs. (28)–(30) as follows

$$\mathbf{Y} = \mathbf{y} + \mathbf{v_Y}, \qquad \mathbf{v_Y} \sim (\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{Y}) \tag{55}$$

$$\mathbf{X} = \mathbf{x} + \mathbf{v_X}, \qquad \mathbf{v_X} \sim (\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{X}) \tag{56}$$

$$\mathbf{X}' = \mathbf{x} - \mathbf{G}\boldsymbol{\theta} + \mathbf{v}_{\mathbf{X}'}, \qquad \mathbf{v}_{\mathbf{X}'} \sim (\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{X}'}) \tag{57}$$

$$\mathbf{Z}' = \mathbf{z} - \tilde{\mathbf{G}}\boldsymbol{\theta} + \mathbf{v}_{\mathbf{Z}'}, \qquad \mathbf{v}_{\mathbf{Z}'} \sim (\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{Z}'}) \tag{58}$$

or, in block-matrix form

$$\begin{bmatrix} \mathbf{Y} \\ \mathbf{X} \\ \mathbf{X}' \\ \mathbf{Z}' \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ -\mathbf{G} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ -\tilde{\mathbf{G}} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta} \\ \mathbf{y} \\ \mathbf{x} \\ \mathbf{z} \end{bmatrix} + \begin{bmatrix} \mathbf{v_Y} \\ \mathbf{v_X} \\ \mathbf{v}_{\mathbf{X}'} \\ \mathbf{v}_{\mathbf{Z}'} \end{bmatrix} \tag{59}$$

where the coordinate vector $\mathbf{Y}$ refers to a set of reference stations in the target frame which is correlated with the common and new points in the underlying network. The data weight matrix that is associated with the above system is considered to be a full matrix

$$\mathbf{P} = \left[\begin{array}{cc|cc} \boldsymbol{\Sigma}_\mathbf{Y} & \boldsymbol{\Sigma}_\mathbf{YX} & \boldsymbol{\Sigma}_{\mathbf{YX}'} & \boldsymbol{\Sigma}_{\mathbf{YZ}'} \\ \boldsymbol{\Sigma}_\mathbf{XY} & \boldsymbol{\Sigma}_\mathbf{X} & \boldsymbol{\Sigma}_{\mathbf{XX}'} & \boldsymbol{\Sigma}_{\mathbf{XZ}'} \\ \hline \boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{Y}} & \boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{X}} & \boldsymbol{\Sigma}_{\mathbf{X}'} & \boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{Z}'} \\ \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{Y}} & \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}} & \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'} & \boldsymbol{\Sigma}_{\mathbf{Z}'} \end{array}\right]^{-1} \tag{60}$$

where each CV submatrix is assumed known beforehand. The weighted LS adjustment of Eq. (59) leads to the following solution (see appendix for the proof)

Estimated transformation parameters

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{G}^T \boldsymbol{\Sigma}_{\mathbf{X}-\mathbf{X}'}^{-1} \mathbf{G}\right)^{-1} \mathbf{G}^T \boldsymbol{\Sigma}_{\mathbf{X}-\mathbf{X}'}^{-1} (\mathbf{X} - \mathbf{X}') \tag{61}$$

Estimated transformed coordinates at the common and new points

$$\begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{z}} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{x}}^{st} \\ \hat{\mathbf{z}}^{st} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{X}'} - \boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{X}} \\ \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'} - \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}} \end{bmatrix} \boldsymbol{\Sigma}_{\mathbf{X}-\mathbf{X}'}^{-1} (\mathbf{X} - \hat{\mathbf{x}}^{st}) \tag{62}$$

Estimated ("updated") coordinates at the non-common reference stations

$$\hat{\mathbf{y}} = \mathbf{Y} + (\boldsymbol{\Sigma}_{\mathbf{YX}'} - \boldsymbol{\Sigma}_{\mathbf{YX}})\boldsymbol{\Sigma}^{-1}_{\mathbf{X}-\mathbf{X}'}(\mathbf{X} - \hat{\mathbf{x}}^{st}) \quad (63)$$

The auxiliary terms $\hat{\mathbf{x}}^{st}$ and $\hat{\mathbf{z}}^{st}$ correspond to the transformed coordinates according to the stepwise approach (see Sect. 2) and the CV matrix $\boldsymbol{\Sigma}_{\mathbf{X}-\mathbf{X}'}$ is

$$\boldsymbol{\Sigma}_{\mathbf{X}-\mathbf{X}'} = \boldsymbol{\Sigma}_{\mathbf{X}} + \boldsymbol{\Sigma}_{\mathbf{X}'} - \boldsymbol{\Sigma}_{\mathbf{XX}'} - \boldsymbol{\Sigma}^{T}_{\mathbf{XX}'} \quad (64)$$

From Eqs. (61)–(63) we notice that:

- The general optimal estimators for the transformed coordinates retain the Kalman filtering structure of the corresponding optimal estimators given in Eq. (39);
- The inter-frame covariances affect the estimated transformation parameters and the corrections to the standard estimates of the transformed coordinates through the cross-CV matrices $\boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{X}}$ and $\boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}}$. In the absence of this effect, the optimal estimators $\hat{\boldsymbol{\theta}}$, $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ become identical to the (simpler) optimal estimators given in Sect. 3;
- The consideration of the prior information of the non-common reference stations ($\mathbf{Y}$) does not affect the results of the CFT problem even when these coordinates are correlated with the prior coordinates of the common and new points;
- The knowledge of the cross-CV matrices $\boldsymbol{\Sigma}_{\mathbf{YX}'}$ and $\boldsymbol{\Sigma}_{\mathbf{YX}}$ allows us to update the prior coordinate vector $\mathbf{Y}$ via Eq. (63). This is an important procedure, at least from a statistical perspective, if one wants to retain optimal coordinate consistency among all point subsets in the target frame.

The practical applicability of Eqs. (61)–(63) is limited by the lack of the cross-CV matrices between the "reference" network ($\mathbf{X}$, $\mathbf{Y}$) and the "new" network ($\mathbf{X}'$, $\mathbf{Z}'$); see Fig. 4. However, there are examples where such inter-frame covariances are inherently present. Be it enough to mention here the case that the reference and the new network are separately determined from GPS phase observations which are partly common in both networks due to overlapping permanent stations. The noise of the common observations will be spread by the adjustment to both networks, thus resulting to correlated errors in their estimated coordinates. Indeed, to come to know such error covariances might be as difficult as solving the problem by adjusting the two networks together in a batch solution. So the formulae derived in this section have primarily a theoretical interest, although they might be used for instance to get a priori bounds on the coordinate correction terms $\boldsymbol{\delta}\hat{\mathbf{x}}$ and $\boldsymbol{\delta}\hat{\mathbf{z}}$.

6 Conclusions

In this paper the CFT problem has been analyzed under different hypotheses on the stochastic model of the known coordinates which are involved in the frame transformation procedure. If we take into account only the internal CV matrices of the coordinates of the common points in the two frames, $\boldsymbol{\Sigma}_{\mathbf{X}}$ and $\boldsymbol{\Sigma}_{\mathbf{X}'}$, and we ignore that the coordinates of the new points $\mathbf{Z}'$ are usually correlated with $\mathbf{X}'$, we come out with a procedure that determines first the parameters $\boldsymbol{\theta}$ of the frame transformation model and then the transformed coordinates in the target frame. This is what we called a standard stepwise solution which represents the "orthodox" approach for CFT problems in geodetic practice. Of course, the standard solution for the transformed coordinates is unbiased but not optimal in the presence of internal covariances between $\mathbf{X}'$ and $\mathbf{Z}'$. The optimal LS solution, however, has a simple analytic expression [cf. Eq. (39)] that can be easily implemented in practical computations, and it leads to less dispersed estimates for the transformed coordinates, as shown in Sect. 3. This is by far the most interesting case from a practical point of view because all the required CV matrices, $\boldsymbol{\Sigma}_{\mathbf{X}}$, $\boldsymbol{\Sigma}_{\mathbf{X}'}$, $\boldsymbol{\Sigma}_{\mathbf{Z}'}$, $\boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{Z}'}$, are easily available in network solutions obtained from space geodetic techniques. A numerical example of the transformation of an EPN weekly solution to the ITRF08 frame shows that the corrections to the transformed coordinates due to the introduction of the full covariance matrix of the EPN weekly network can amount from several millimeters up to several centimeters

Also, in this study we have treated the most general CFT problem of fully correlated coordinate sets between two overlapping geodetic network solutions. Although this is more of theoretical interest since the cross-covariances between ($\mathbf{X}$, $\mathbf{Y}$) and ($\mathbf{X}'$, $\mathbf{Z}'$) are usually not available, yet the optimal LS solution looks interesting, displaying two general characteristics. First, all the related estimators depend on the inversion of $\boldsymbol{\Sigma}_{\mathbf{X}-\mathbf{X}'}$ which is always of manageable dimension in practical cases. Moreover, the transformed coordinates for the new network are independent of the stochastic model of the known coordinates of the (non-common) reference points ($\mathbf{Y}$) regardless of the inter-frame and intra-frame covariances that may be present in their values. The influence of the above considerations on the optimal analysis of coordinate time series in permanent geodetic networks is an interesting topic that will be the subject of future investigations.

7 Appendix

The BLUE estimators related to the solution of the geodetic CFT problem are analytically derived in this appendix. Our proof scheme considers the most general case of the problem by taking into account both the intra-frame and inter-frame covariances in the coordinate datasets. The optimal transformation formulae given in Sect. 3 [see e.g. Eq. (39)] stem directly as special cases of the following derivations.

Let us first express the system of observation equations from Eqs. (55)–(58) in the equivalent algebraic form

$$\mathbf{X}-\mathbf{X}' = \mathbf{G}\boldsymbol{\theta}+\mathbf{v}_{\mathbf{X}-\mathbf{X}'} \tag{65}$$

$$\mathbf{X} = \mathbf{x}+\mathbf{v}_{\mathbf{X}} \tag{66}$$

$$\mathbf{Y} = \mathbf{y}+\mathbf{v}_{\mathbf{Y}} \tag{67}$$

$$\mathbf{Z}' = \mathbf{z}-\tilde{\mathbf{G}}\boldsymbol{\theta}+\mathbf{v}_{\mathbf{Z}'} \tag{68}$$

Using block-matrix notation the above system can be written as

$$\begin{bmatrix} \mathbf{X}-\mathbf{X}' \\ \hline \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z}' \end{bmatrix} = \begin{bmatrix} \mathbf{G} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ -\tilde{\mathbf{G}} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta} \\ \hline \mathbf{x} \\ \mathbf{y} \\ \mathbf{z} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_{\mathbf{X}-\mathbf{X}'} \\ \hline \mathbf{v}_{\mathbf{X}} \\ \mathbf{v}_{\mathbf{Y}} \\ \mathbf{v}_{\mathbf{Z}'} \end{bmatrix} \tag{69}$$

or, in a more compact form

$$\begin{bmatrix} \delta\mathbf{X} \\ \boldsymbol{\Xi} \end{bmatrix} = \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{K} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta} \\ \boldsymbol{\xi} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_{\delta\mathbf{X}} \\ \mathbf{v}_{\boldsymbol{\Xi}} \end{bmatrix} \tag{70}$$

The meaning of all auxiliary terms in the last equation is deduced from (69). The data weight matrix that is associated with the above system has the general form

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_{12} \\ \mathbf{P}_{12}^T & \mathbf{P}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{\delta\mathbf{X}} & \boldsymbol{\Sigma}_{\delta\mathbf{X},\boldsymbol{\Xi}} \\ \boldsymbol{\Sigma}_{\delta\mathbf{X},\boldsymbol{\Xi}}^T & \boldsymbol{\Sigma}_{\boldsymbol{\Xi}} \end{bmatrix}^{-1} \tag{71}$$

where

$$\boldsymbol{\Sigma}_{\delta\mathbf{X}} = \boldsymbol{\Sigma}_{\mathbf{X}}+\boldsymbol{\Sigma}_{\mathbf{X}'}-\boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X}'}-\boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X}'}^T \tag{72}$$

$$\boldsymbol{\Sigma}_{\delta\mathbf{X},\boldsymbol{\Xi}} = \left[\, \boldsymbol{\Sigma}_{\delta\mathbf{X},\mathbf{X}} \quad \boldsymbol{\Sigma}_{\delta\mathbf{X},\mathbf{Y}} \quad \boldsymbol{\Sigma}_{\delta\mathbf{X},\mathbf{Z}'} \,\right] \tag{73}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\Xi}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{X}} & \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{Y}} & \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{Z}'} \\ \boldsymbol{\Sigma}_{\mathbf{Y}\mathbf{X}} & \boldsymbol{\Sigma}_{\mathbf{Y}} & \boldsymbol{\Sigma}_{\mathbf{Y}\mathbf{Z}'} \\ \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}} & \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{Y}} & \boldsymbol{\Sigma}_{\mathbf{Z}'} \end{bmatrix} \tag{74}$$

Using elementary covariance propagation rules, we may express the submatrices of the auxiliary cross-CV matrix $\boldsymbol{\Sigma}_{\delta\mathbf{X},\boldsymbol{\Xi}}$ in terms of the relationships

$$\boldsymbol{\Sigma}_{\delta\mathbf{X},\mathbf{X}} = \boldsymbol{\Sigma}_{\mathbf{X}}-\boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{X}} \tag{75}$$

$$\boldsymbol{\Sigma}_{\delta\mathbf{X},\mathbf{Y}} = \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{Y}}-\boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{Y}} \tag{76}$$

$$\boldsymbol{\Sigma}_{\delta\mathbf{X},\mathbf{Z}'} = \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{Z}'}-\boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{Z}'} \tag{77}$$

The weighted LS adjustment of (70) leads to the normal equations system

$$\begin{bmatrix} \mathbf{G}^T & \mathbf{K}^T \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_{12} \\ \mathbf{P}_{12}^T & \mathbf{P}_2 \end{bmatrix} \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{K} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\theta}} \\ \hat{\boldsymbol{\xi}} \end{bmatrix} = \begin{bmatrix} \mathbf{G}^T & \mathbf{K}^T \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_{12} \\ \mathbf{P}_{12}^T & \mathbf{P}_2 \end{bmatrix} \begin{bmatrix} \delta\mathbf{X} \\ \boldsymbol{\Xi} \end{bmatrix} \tag{78}$$

from which we obtain the following equations for the optimal estimators $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\xi}}$

$$\left(\mathbf{G}^T\mathbf{P}_1\mathbf{G}+\mathbf{K}^T\mathbf{P}_{12}^T\mathbf{G}+\mathbf{G}^T\mathbf{P}_{12}\mathbf{K}+\mathbf{K}^T\mathbf{P}_2\mathbf{K}\right)\hat{\boldsymbol{\theta}} + \left(\mathbf{G}^T\mathbf{P}_{12}+\mathbf{K}^T\mathbf{P}_2\right)\hat{\boldsymbol{\xi}} = \left(\mathbf{G}^T\mathbf{P}_1+\mathbf{K}^T\mathbf{P}_{12}^T\right)\delta\mathbf{X}+\left(\mathbf{G}^T\mathbf{P}_{12}+\mathbf{K}^T\mathbf{P}_2\right)\boldsymbol{\Xi} \tag{79}$$

and

$$\left(\mathbf{P}_{12}^T\mathbf{G}+\mathbf{P}_2\mathbf{K}\right)\hat{\boldsymbol{\theta}}+\mathbf{P}_2\hat{\boldsymbol{\xi}} = \mathbf{P}_{12}^T\delta\mathbf{X}+\mathbf{P}_2\boldsymbol{\Xi} \tag{80}$$

Solving the last equation for $\hat{\boldsymbol{\xi}}$ we get

$$\hat{\boldsymbol{\xi}} = \mathbf{P}_2^{-1}\mathbf{P}_{12}^T\delta\mathbf{X}+\boldsymbol{\Xi}-\left(\mathbf{P}_2^{-1}\mathbf{P}_{12}^T\mathbf{G}+\mathbf{K}\right)\hat{\boldsymbol{\theta}} \tag{81}$$

and by substituting back to (79), and after several cancelations of similar terms, we end up with the relationship

$$\mathbf{G}^T\left(\mathbf{P}_1-\mathbf{P}_{12}\mathbf{P}_2^{-1}\mathbf{P}_{12}^T\right)\mathbf{G}\hat{\boldsymbol{\theta}} = \mathbf{G}^T\left(\mathbf{P}_1-\mathbf{P}_{12}\mathbf{P}_2^{-1}\mathbf{P}_{12}^T\right)\delta\mathbf{X} \tag{82}$$

Taking into account from (71) that

$$\mathbf{P}_1-\mathbf{P}_{12}\mathbf{P}_2^{-1}\mathbf{P}_{12}^T = \boldsymbol{\Sigma}_{\delta\mathbf{X}}^{-1} = \boldsymbol{\Sigma}_{\mathbf{X}-\mathbf{X}'}^{-1} \tag{83}$$

we finally obtain the optimal estimate of the frame transformation parameters

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{G}^T\boldsymbol{\Sigma}_{\mathbf{X}-\mathbf{X}'}^{-1}\mathbf{G}\right)^{-1}\mathbf{G}^T\boldsymbol{\Sigma}_{\mathbf{X}-\mathbf{X}'}^{-1}(\mathbf{X}-\mathbf{X}') \tag{84}$$

Based again on (71) we have the useful formula

$$\mathbf{P}_2^{-1}\mathbf{P}_{12}^T = -\boldsymbol{\Sigma}_{\delta\mathbf{X},\boldsymbol{\Xi}}^T\boldsymbol{\Sigma}_{\delta\mathbf{X}}^{-1} \tag{85}$$

which can be substituted to (81), thus leading to the optimal estimate of the auxiliary vector $\boldsymbol{\xi}$ in terms of the expression

$$\hat{\boldsymbol{\xi}} = \boldsymbol{\Xi}-\boldsymbol{\Sigma}_{\delta\mathbf{X},\boldsymbol{\Xi}}^T\boldsymbol{\Sigma}_{\delta\mathbf{X}}^{-1}\left(\delta\mathbf{X}-\mathbf{G}\hat{\boldsymbol{\theta}}\right)-\mathbf{K}\hat{\boldsymbol{\theta}} \tag{86}$$

or, in the equivalent form

$$\begin{bmatrix}\hat{\mathbf{x}}\\ \hat{\mathbf{y}}\\ \hat{\mathbf{z}}\end{bmatrix}=\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\\ \mathbf{Z}'\end{bmatrix}-\begin{bmatrix}\boldsymbol{\Sigma}^T_{\delta\mathbf{X},\mathbf{X}}\\ \boldsymbol{\Sigma}^T_{\delta\mathbf{X},\mathbf{Y}}\\ \boldsymbol{\Sigma}^T_{\delta\mathbf{X},\mathbf{Z}'}\end{bmatrix}\boldsymbol{\Sigma}^{-1}_{\delta\mathbf{X}}\left(\delta\mathbf{X}-\mathbf{G}\hat{\boldsymbol{\theta}}\right)-\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\\ -\tilde{\mathbf{G}}\end{bmatrix}\hat{\boldsymbol{\theta}} \tag{87}$$

Using the covariance expressions from (75)–(77), the last equation is expressed as

$$\begin{bmatrix}\hat{\mathbf{x}}\\ \hat{\mathbf{y}}\\ \hat{\mathbf{z}}\end{bmatrix}=\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\\ \mathbf{Z}'+\tilde{\mathbf{G}}\hat{\boldsymbol{\theta}}\end{bmatrix}-\begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{X}}-\boldsymbol{\Sigma}_{\mathbf{XX}'}\\ \boldsymbol{\Sigma}_{\mathbf{YX}}-\boldsymbol{\Sigma}_{\mathbf{YX}'}\\ \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}}-\boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'}\end{bmatrix}\boldsymbol{\Sigma}^{-1}_{\delta\mathbf{X}}\left(\delta\mathbf{X}-\mathbf{G}\hat{\boldsymbol{\theta}}\right) \tag{88}$$

and, by taking into account (72), in the equivalent form

$$\begin{bmatrix}\hat{\mathbf{x}}\\ \hat{\mathbf{y}}\\ \hat{\mathbf{z}}\end{bmatrix}=\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\\ \mathbf{Z}'+\tilde{\mathbf{G}}\hat{\boldsymbol{\theta}}\end{bmatrix}-\begin{bmatrix}\boldsymbol{\Sigma}_{\delta\mathbf{X}}-\boldsymbol{\Sigma}_{\mathbf{X}'}+\boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{X}}\\ \boldsymbol{\Sigma}_{\mathbf{YX}}-\boldsymbol{\Sigma}_{\mathbf{YX}'}\\ \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}}-\boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'}\end{bmatrix}\boldsymbol{\Sigma}^{-1}_{\delta\mathbf{X}}\left(\delta\mathbf{X}-\mathbf{G}\hat{\boldsymbol{\theta}}\right) \tag{89}$$

Considering that $\boldsymbol{\delta}\mathbf{X}=\mathbf{X}-\mathbf{X}'$, equation (89) can be finally reduced to the form

$$\begin{bmatrix}\hat{\mathbf{x}}\\ \hat{\mathbf{y}}\\ \hat{\mathbf{z}}\end{bmatrix}=\begin{bmatrix}\mathbf{X}'+\mathbf{G}\hat{\boldsymbol{\theta}}\\ \mathbf{Y}\\ \mathbf{Z}'+\tilde{\mathbf{G}}\hat{\boldsymbol{\theta}}\end{bmatrix}+\begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{X}'}-\boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{X}}\\ \boldsymbol{\Sigma}_{\mathbf{YX}'}-\boldsymbol{\Sigma}_{\mathbf{YX}}\\ \boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'}-\boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}}\end{bmatrix}\boldsymbol{\Sigma}^{-1}_{\mathbf{X}-\mathbf{X}'}\left(\mathbf{X}-\mathbf{X}'-\mathbf{G}\hat{\boldsymbol{\theta}}\right) \tag{90}$$

From the last equation we explicitly obtain the optimal estimates for the transformed coordinates from the initial frame to the target frame at the common and new points, respectively

$$\hat{\mathbf{x}}=\hat{\mathbf{x}}^{\mathrm{st}}+(\boldsymbol{\Sigma}_{\mathbf{X}'}-\boldsymbol{\Sigma}_{\mathbf{X}'\mathbf{X}})\,\boldsymbol{\Sigma}^{-1}_{\mathbf{X}-\mathbf{X}'}(\mathbf{X}-\hat{\mathbf{x}}^{\mathrm{st}}) \tag{91}$$

$$\hat{\mathbf{z}}=\hat{\mathbf{z}}^{\mathrm{st}}+(\boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}'}-\boldsymbol{\Sigma}_{\mathbf{Z}'\mathbf{X}})\,\boldsymbol{\Sigma}^{-1}_{\mathbf{X}-\mathbf{X}'}(\mathbf{X}-\hat{\mathbf{x}}^{\mathrm{st}}) \tag{92}$$

and also the "updated" coordinates at the non-common reference points (in the target frame)

$$\hat{\mathbf{y}}=\mathbf{Y}+(\boldsymbol{\Sigma}_{\mathbf{YX}'}-\boldsymbol{\Sigma}_{\mathbf{YX}})\,\boldsymbol{\Sigma}^{-1}_{\mathbf{X}-\mathbf{X}'}(\mathbf{X}-\hat{\mathbf{x}}^{\mathrm{st}}) \tag{93}$$

Note that the terms $\hat{\mathbf{x}}^{\mathrm{st}}$ and $\hat{\mathbf{z}}^{\mathrm{st}}$ correspond to the transformed coordinates according to the standard stepwise CFT approach [see Eq. (17)].

References

Altamimi Z (2003) Discussion on how to express a regional GPS solution in the ITRF. EUREF Publication No. 12, Verlag des Bundesamtes für Kartographie und Geodäsie, Frankfurt am Main, pp 162–167

Altamimi Z, Boucher C, Sillard P (2002) New trends for the realization of the international terrestrial reference system. Adv Space Res 30:175–184

Bevis M, Brown A (2014) Trajectory models and reference frames for crustal motion geodesy. J Geod 88:283–311

Blewitt G (1998) GPS data processing methodology. In: Teunissen PJG, Kleusberg A (eds) GPS for Geodesy, 2nd edition. Springer, Berlin, pp 231–270

Gibbs B (2011) Advanced Kalman filtering, least-squares and modelling: a practical handbook. Wiley, New York

Koch K-R (1999) Parameter estimation and hypothesis testing in linear models, 2nd edition. Springer, Berlin

Leick A, van Gelder BHW (1975) On similarity transformations and geodetic network distortions based on Doppler satellite observations. Technical Report No. 235, Department of Geodetic Science, Ohio State University, Columbus, Ohio

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

Petit G, Luzum B (eds) (2010) Internation earth rotation and reference systems service (IERS) conventions. IERS Technical Note No. 36, Verlag des Bundesamts fur Kartographie und Geodäsie, Frankfurt am Main

Schaffrin B, Wieser A (2008) On weighted total least squares adjustment for linear regression. J Geod 82:415–421

Schwarz K-P (1974) Combination of spatial networks using an estimated covariance matrix. Bull Geod 112(1):171–186

Sillard P, Boucher C (2001) A review of algebraic constraints in terrestrial reference frame datum definition. J Geod 75:63–73

Soler T (1998) A compendium of transformation formulae useful for GPS work. J Geod 72:482–490

Tregoning P, van Dam T (2005) Effects of atmospheric pressure loading and seven-parameter transformations on estimates of geocenter motion and station heights from space geodetic observations. J Geoph Res 110:B03408. doi:10.1029/2004JB003334