J Geod
DOI 10.1007/s00190-013-0637-0

ORIGINAL ARTICLE

Generalized inner constraints for geodetic network densification problems

C. Kotsakis

Received: 8 September 2012 / Accepted: 15 April 2013

Abstract The estimated coordinates from a minimum-constrained (MC) network adjustment are generally influenced by two different error sources, that is the data noise from the available measurements and the so-called datum noise due to random errors in the fiducial coordinates that are used for the datum definition with regard to an external reference frame. Although the latter source does not affect the estimable characteristics of a MC solution, it still contributes a datum-related noise to the estimated positions which reflects the uncertainty of the coordinate system itself for the adjusted network. The aim of this paper is to develop a new type of MCs which minimizes both of the aforementioned effects in the final coordinates of an adjusted network. This particular problem has not been treated in the geodetic literature and the solution which is presented herein offers an elegant unification of the classic inner constraints into a more general framework for the datum choice problem of network optimization theory. Furthermore, the findings of our study provide a useful and rigorous tool for frame densification problems by means of an optimal MC adjustment in geodetic networks.

Keywords Minimum constraints · Inner constraints · Data noise effect · Datum noise effect · Network adjustment · Datum choice problem · Zero-order network optimization

C. Kotsakis (✉)
Department of Geodesy and Surveying,
Aristotle University of Thessaloniki, University Box 440,
541 24 Thessaloniki, Greece
e-mail: kotsaki@topo.auth.gr

1 Introduction

The determination of station positions from geodetic measurements through a least squares network adjustment is an 'ill-posed' estimation problem of profound importance for several applications. Typically, a unique solution of this problem is obtained from the generalized inversion of a rank-deficient system of normal equations with the aid of the so-called *minimum constraints* (MCs); see Blaha (1971). Depending on the network's observational model, the primary role of these constraints is to specify the non-estimable components of the coordinate system (i.e. origin, orientation, scale, and their temporal evolution) with respect to which the station positions will be computed without affecting the estimable characteristics of the adjusted network. Their use enables the hierarchical alignment of geodetic networks within a unified spatial framework and the densification of global reference frames over regional networks through the realization of high-quality terrestrial reference frames (TRFs) from space geodetic data. The formulation of minimum constraints according to a novel optimization scheme for the estimated coordinates in network adjustment problems is the main objective of the present paper.

The search for an optimal type of MCs had received considerable attention in the past decades mainly in tandem with the zero-order design (ZOD) or datum choice problem of network optimization theory (Grafarend 1974; Schmitt 1982; Papo and Perelmuter 1981; Blaha 1982; Dermanis 1985, 1998; Schaffrin 1985; Teunissen 1985). A solution to this problem is known to be given by the so-called *inner constraints* which lead to the minimum-trace error covariance (CV) matrix for the estimated coordinates at the network stations. This particular type of MCs was originally introduced by Meissl (1969) yet it was Blaha (1971) and Pope (1971)

Published online: 23 May 2013

who first established in detail their optimality with regard to the statistical properties of the estimated coordinates in MC networks.

It should be noted that the optimality of inner constraints is restricted to the minimization of the propagated data noise on the adjusted network coordinates (hereafter called *data noise effect*). Therefore it does not consider the additional effect of random errors in the known coordinates of the reference stations which are commonly used for the practical implementation of MCs in geodetic networks. The latter does not influence the geometrical form of a MC solution yet it contributes a datum-related noise to the estimated positions which reflects the uncertainty of the coordinate system itself for the adjusted network (hereafter called *datum noise effect*—more details to be given later in the paper). The minimization of the datum noise is an important issue for the MC adjustment of a geodetic network with respect to a given TRF, yet it has not been treated within the theoretical realm of the datum choice problem and the ZOD optimization framework.

From a strict perspective of statistical estimation theory, the station coordinates in network adjustment problems are not estimable quantities and, thus, the optimization of their CV matrix with respect to different datum choices can be considered as a mathematical artifice with no actual physical relevance. Specifically, in the context of MC networks, the estimated coordinates (and their CV matrix) act as a 'depository' of information for the computation of estimable parameters (and their corresponding accuracy) that remain invariant under any choice of MCs for the underlying network (Grafarend and Schaffrin 1974, 1976). In this way, when it comes to specifying the statistical accuracy of the estimated coordinates it may be prudent to speak about their 'apparent' accuracy which depends directly on the datum choice and, obviously, it does not provide an objective quality measure of the MC network from the available data (Dermanis 2012). The use of inner constraints yields an optimal solution which makes only the final coordinates to have a seemingly better accuracy under the influence of the data noise effect, without offering any actual improvement for the geometrical accuracy of the adjusted network. In fact, if one is interested only to the shape analysis of a MC network (or to the error analysis of the available measurements) there is nothing to be gained from an 'optimal' datum choice through a preferred set of MCs.

The above viewpoint overlooks an important aspect of MCs, namely their role as a datum transfer and frame densification tool in network adjustment problems. Minimum-constrained adjustment schemes are often used to tie a geodetic network to a reference frame through a set of a priori known coordinates (and velocities in case of dynamic networks) at a number of fiducial stations (Altamimi et al. 2002). From this perspective, the choice of MCs poses a meaningful concern for practical applications since it affects the quality of the datum implementation in the adjusted network (e.g. single-station MCs are known to have significantly worse 'frame alignment' accuracy for regional GNSS networks compared to the no-net-translation condition over a number of reference stations).[1] The fact is that different sets of MCs may offer different stability for the realized coordinate frame over the same network in the sense that the adjusted station positions may exhibit higher or lower sensitivity to small perturbations (=random errors) in the used fiducial coordinates—this corresponds, in loose terms, to what we previously identified as datum noise effect. Some general aspects and practical examples related to this issue have been recently discussed in Kotsakis (2012) while a heuristic treatment within a rather different context can be found in the study by Coulot et al. (2010). The crucial problem to be exposed herein refers to the optimal filtering of the datum noise within the minimum-constrained adjustment of a geodetic network in support to its connection to an existing reference frame.

The aim of this study is to devise a new type of MCs based on the joint minimization of the data and datum noise effects in geodetic network adjustment. Using a straightforward analytic approach we derive a set of datum constraints that minimizes the trace of the complete error CV matrix (i.e. the one including the combined contribution of the data noise and the random errors within the fiducial station coordinates) for the estimated positions in a MC network. Our treatment covers also the case of the separate minimization of the datum noise effect which leads to a weighted type of inner constraints for the reference frame realization in a geodetic network. It is noted that the aforementioned minimum-trace property has not been used as an independent criterion for determining an optimal set of MCs—to the author's knowledge such an approach has only appeared in (Blaha 1971, pp. 53–68)—but it is deduced as an intrinsic property of the inner constraints in the context of generalized inverse theory for symmetric matrices and singular normal equations systems (Teunissen 1985; Koch 1999). The present paper extends Blaha's original approach (without the use of generalized inverses) and it leads to an elegant unification of the classic inner constraints into a more general optimization framework for the datum choice problem.

[1] To facilitate the reader's view at this point we note that different choices of MCs, such as $\mathbf{H}(\mathbf{x}-\mathbf{x}^{\mathrm{ext}})=\mathbf{0}$ vs. $\tilde{\mathbf{H}}(\mathbf{x}-\mathbf{x}^{\mathrm{ext}})=\mathbf{0}$, includes also the case $\tilde{\mathbf{H}}=\mathbf{H}\mathbf{P}_x$ where $\mathbf{P}_x$ is a weight matrix which controls the contribution of the reference stations towards the datum definition in the underlying network.

2 Preliminaries

The adjustment of geodetic networks is generally formulated as a parameter estimation problem in a nonlinear and rank-deficient Gauss-Markov model

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) + \mathbf{v}, \quad \mathbf{v} \sim (\mathbf{0}, \sigma^2\mathbf{P}^{-1}) \tag{1}$$

whose weighted least squares inversion yields a *singular* consistent system of linearized normal equations (NEQs)

$$\underbrace{(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})}_{\mathbf{N}}(\mathbf{x} - \mathbf{x}_{\mathrm{o}}) = \underbrace{\mathbf{A}^{\mathrm{T}}\mathbf{P}\,(\mathbf{y} - \mathbf{f}(\mathbf{x}_{\mathrm{o}}))}_{\mathbf{u}} \tag{2}$$

where $\mathbf{y}$ is the measurement vector, $\mathbf{x}_{\mathrm{o}}$ is an initial approximation of the unknown parameters, $\mathbf{A} = \mathbf{f}_x(\mathbf{x}_{\mathrm{o}})$ is the Jacobian matrix of the observables and $\mathbf{P}$ is the data weight matrix. Depending on the network type, the parameters $\mathbf{x}$ stem from a static (coordinates only) or dynamic (coordinates and velocities) modeling of the geodetic observables with respect to an appropriate reference system. Without loss of generality we assume that any nuisance parameters have been eliminated beforehand from the NEQ system, so that the term $\mathbf{x} - \mathbf{x}_{\mathrm{o}}$ contains only the unknown corrections to the approximate positions of the network stations.

The determination of a single solution of Eq. (2) constitutes the so-called datum choice problem that is generally resolved with the help of a consistent set of linear (or linearized) MCs

$$\mathbf{H}(\mathbf{x} - \mathbf{x}_{\mathrm{o}}) = \mathbf{c} = \mathbf{H}(\mathbf{x}^{\mathrm{ext}} - \mathbf{x}_{\mathrm{o}}) \tag{3a}$$

or, in equivalent form

$$\mathbf{H}(\mathbf{x} - \mathbf{x}^{\mathrm{ext}}) = \mathbf{0} \tag{3b}$$

which lead to the extended full-rank NEQ system

$$(\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})(\mathbf{x} - \mathbf{x}_{\mathrm{o}}) = \mathbf{u} + \mathbf{H}^{\mathrm{T}}\mathbf{c} \tag{4}$$

The above system has a unique solution (termed here as MC solution)

$$\hat{\mathbf{x}} = \mathbf{x}_{\mathrm{o}} + (\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{u} + (\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{c} \tag{5}$$

which satisfies both the singular NEQs in Eq. (2) and the MCs in Eq. (3a, 3b) (the necessary algebraic conditions for the matrix $\mathbf{H}$ are omitted here for brevity, for more details see Schaffrin 1985; Koch 1999).

The classic inner constraints correspond to a special choice of the MC matrix, denoted as $\mathbf{H} = \mathbf{E}$, which fulfills the orthogonality property $\mathbf{A}\mathbf{E}^{\mathrm{T}} = \mathbf{0}$, and thus $\mathbf{N}\mathbf{E}^{\mathrm{T}} = \mathbf{0}$, and it minimizes the propagated data noise on the final solution of Eq. (5) (Blaha 1971). The structure of the inner constraint matrix $\mathbf{E}$ depends on the network's datum defect and it originates from the so-called Helmert matrix that is associated with the linearized similarity transformation for Cartesian coordinate frames (ibid.). Of particular importance for our study are the following formulae:

$$(\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N} = \mathbf{I} - \mathbf{E}^{\mathrm{T}}(\mathbf{H}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{H} \tag{6}$$

$$(\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}} = \mathbf{E}^{\mathrm{T}}(\mathbf{H}\mathbf{E}^{\mathrm{T}})^{-1} \tag{7}$$

where $\mathbf{H}$ denotes an arbitrary MC matrix for the underlying network. The analytic proof of the above equations can be found in Kotsakis (2012).

The vector $\mathbf{x}^{\mathrm{ext}}$ in Eq. (3a, 3b) refers to a set of known coordinates from an 'external' reference frame with respect to which the MC network is aligned. The preceding formulation implies a one-to-one correspondence between the vectors $\mathbf{x}$ and $\mathbf{x}^{\mathrm{ext}}$, and thus it implicitly assumes that the external frame information is available over all network stations. For the purpose of this paper we consider the (more realistic) case where the network is composed of two distinctive parts $\mathbf{x}_1$ and $\mathbf{x}_2$: the first contains the fiducial stations with a priori known coordinates $\mathbf{x}_1^{\mathrm{ext}}$ while the second includes the new stations. The rationale of our study relies on the premise that the datum definition involves only stations for which there exists prior information for their positions (and also their accuracy) with respect to an external reference frame. Hence, the MC formulation in Eq. (3a, 3b) and the corresponding solution from Eq. (5) will be considered in the context of the partition scheme

$$\mathbf{H} = [\,\mathbf{H}_1 \quad \mathbf{0}\,], \quad \mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \quad \mathbf{x}^{\mathrm{ext}} = \begin{bmatrix} \mathbf{x}_1^{\mathrm{ext}} \\ \mathbf{0} \end{bmatrix} \tag{8}$$

where the necessary algebraic conditions for the MC submatrix $\mathbf{H}_1$ are again omitted for brevity.

An important remark should be given regarding the use of a weight matrix for the incorporation of MCs in network adjustment problems. Theoretically such a matrix will not affect the estimated coordinates since the NEQ system of Eq. (4) is equivalent to the weighted NEQ system[2]:

$$(\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{W}\mathbf{H})(\mathbf{x} - \mathbf{x}_{\mathrm{o}}) = \mathbf{u} + \mathbf{H}^{\mathrm{T}}\mathbf{W}\mathbf{c} \tag{9}$$

for any symmetric positive-definite matrix $\mathbf{W}$ (e.g. Kotsakis 2012). Therefore, any prior weighting for the non-estimable part of the coordinate system (which is induced by a weight matrix for the datum related quantities $\mathbf{c}$) is a non-essential task in MC network adjustment, although for practical purposes it may be useful for improving the numerical condition of the final NEQ system. It is important to point out that prior weighting of the fiducial stations participating in the MCs can be an influential factor for the final solution, which nevertheless cannot be 'passed' into the adjusted coordinates (and their accuracy) through the aforementioned weight matrix $\mathbf{W}$. It is actually one of this paper's merits to reveal how an

[2] This algebraic equivalency is theoretically valid only for minimum-constrained NEQs and it does not generally hold in cases of over-constrained network adjustment schemes.

appropriate weighting for the fiducial station coordinates $\mathbf{x}_1^{\text{ext}}$ should be implemented to ensure certain optimal properties for the estimated positions in a MC network.

The objective of our study is to find the MC matrix $\mathbf{H}$, or rather its non-zero submatrix $\mathbf{H}_1$, that optimizes the accuracy of the estimated network coordinates with respect to the reference frame realized by the fiducial stations. For this purpose we shall consider the total noise effect in the MC solution, including the part caused by the random errors in the fiducial coordinates which is 'hidden' within the last term of Eq. (5). In practice this term is often set to zero by selecting the approximate coordinates of the fiducial stations to be equal to the a priori known values $\mathbf{x}_1^{\text{ext}}$ (i.e. $\mathbf{c} = \mathbf{0}$). This of course does not eliminate the presence of the datum noise effect which should be accounted for through the covariance propagation scheme that is described in the next section.

3 Datum-dependent error covariance matrices in MC networks

3.1 General formulae

The error CV matrix of the estimated coordinates from a minimum-constrained NEQ system can be expressed as a sum of two independent components

$$\boldsymbol{\Sigma}_{\hat{x}} = \boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}} + \boldsymbol{\Sigma}_{\hat{x}}^{\text{mc}} \tag{10}$$

according to the analytic forms

$$\begin{aligned}\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}} &= \sigma^2(\mathbf{N}+\mathbf{H}^{\text{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\text{T}}\mathbf{H})^{-1} \\ &= \sigma^2(\mathbf{N}+\mathbf{H}^{\text{T}}\mathbf{H})^{-1} - \sigma^2\mathbf{E}^{\text{T}}(\mathbf{H}\mathbf{E}^{\text{T}})^{-1}(\mathbf{E}\mathbf{H}^{\text{T}})^{-1}\mathbf{E} \qquad (11)\\ \boldsymbol{\Sigma}_{\hat{x}}^{\text{mc}} &= \mathbf{E}^{\text{T}}(\mathbf{H}\mathbf{E}^{\text{T}})^{-1}\boldsymbol{\Sigma}_c(\mathbf{E}\mathbf{H}^{\text{T}})^{-1}\mathbf{E} \qquad (12)\end{aligned}$$

The above formulae are obtained from a clear-cut error propagation to Eq. (5) by taking into account the fundamental relationships in Eqs. (6) and (7).

Considering that the matrix $\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}}$ is invariant with respect to any symmetric positive-definite weight matrix $\mathbf{W}$, that is

$$\begin{aligned}&(\mathbf{N}+\mathbf{H}^{\text{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\text{T}}\mathbf{H})^{-1} \\ &\quad = (\mathbf{N}+\mathbf{H}^{\text{T}}\mathbf{W}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\text{T}}\mathbf{W}\mathbf{H})^{-1}\end{aligned} \tag{13}$$

and subsequently

$$\begin{aligned}&(\mathbf{N}+\mathbf{H}^{\text{T}}\mathbf{H})^{-1} - \mathbf{E}^{\text{T}}(\mathbf{H}\mathbf{E}^{\text{T}})^{-1}(\mathbf{E}\mathbf{H}^{\text{T}})^{-1}\mathbf{E} \\ &\quad = (\mathbf{N}+\mathbf{H}^{\text{T}}\mathbf{W}\mathbf{H})^{-1} - \mathbf{E}^{\text{T}}(\mathbf{H}\mathbf{E}^{\text{T}})^{-1}\mathbf{W}^{-1}(\mathbf{E}\mathbf{H}^{\text{T}})^{-1}\mathbf{E}\end{aligned} \tag{14}$$

it is easily shown that the (total) error CV matrix of a MC solution takes the compact 'Bayesian form' (the proofs of the last two equations are given in the Appendix):

$$\boldsymbol{\Sigma}_{\hat{x}} = \boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}} + \boldsymbol{\Sigma}_{\hat{x}}^{\text{mc}} = \sigma^2(\mathbf{N}+\sigma^2\mathbf{H}^{\text{T}}\boldsymbol{\Sigma}_c^{-1}\mathbf{H})^{-1} \tag{15}$$

The CV matrices $\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}}$ and $\boldsymbol{\Sigma}_{\hat{x}}^{\text{mc}}$ contain the contributions from separate (and formally independent) error sources, namely the data and datum noise effects both of which are influenced by the choice of the MC matrix $\mathbf{H}$. Each of these matrices is rigorously singular, yet their sum yields a full-rank error CV matrix provided that the matrix $\boldsymbol{\Sigma}_c$ is invertible.

The CV matrix $\boldsymbol{\Sigma}_c$ specifies the statistical accuracy of the pseudo-observation vector $\mathbf{c}$ that appears in the MC formulation as per Eq. (3a, 3b). Its form is dictated by the selected type of MCs and the a priori accuracy of the fiducial station coordinates. In view of Eqs. (3a) and (8), we have

$$\boldsymbol{\Sigma}_c = \mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}} \tag{16}$$

where $\boldsymbol{\Sigma}_x^{\text{prior}}$ is the known CV matrix of the coordinate vector $\mathbf{x}_1^{\text{ext}}$.

In the context of our study the CV matrices $\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}}$ and $\boldsymbol{\Sigma}_{\hat{x}}^{\text{mc}}$ can be expressed in the equivalent forms

$$\begin{aligned}\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}} &= \sigma^2(\mathbf{N}+\mathbf{H}^{\text{T}}\mathbf{H})^{-1} - \sigma^2\mathbf{E}^{\text{T}}(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1} \\ &\quad \times(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\mathbf{E} \qquad (17)\\ \boldsymbol{\Sigma}_{\hat{x}}^{\text{mc}} &= \mathbf{E}^{\text{T}}(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\mathbf{E} \qquad (18)\end{aligned}$$

where $\mathbf{E}_1$ is the submatrix of the total inner-constraint matrix that corresponds to the fiducial stations, i.e. $\mathbf{E} = [\mathbf{E}_1\ \mathbf{E}_2]$, in accordance with the partitioning of the MC matrix $\mathbf{H} = [\mathbf{H}_1\ \mathbf{0}]$ that was introduced in Eq. (8).

3.2 The matrix $\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}}$

The CV matrix $\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}}$ is always singular and it contains the full effect of the data noise on the MC solution. Its rank defect is equal to the datum defect of the available observations and its elements do not carry any accuracy information for the (non-estimable part of the) coordinate system of the adjusted network. This matrix corresponds to a scaled (by σ^2) symmetric reflexive generalized inverse of the singular normal matrix $\mathbf{N}$, and it is always positive semi-definite (Koch 1999, pp. 52–53).

The minimization of the trace$\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}}$, or for a part of it, has been the main goal of the ZOD or datum choice problem in network optimization theory (Grafarend 1974; Schmitt 1982; Schaffrin 1985). It is well known that this problem is resolved through the classic inner constraints in full or partial form, i.e. $\mathbf{H} = \mathbf{E}$ or $\mathbf{H} = [\mathbf{E}_1\ \mathbf{0}]$, depending whether the minimization of the trace$\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}}$ is sought over all or some of the network stations (Koch 1999, pp. 62–64).

It is emphasized that the choice of classic inner constraints yields the optimal datum definition for a MC network only with regard to the data noise effect on the estimated coordinates. The trace of the second CV component in Eq. (10) is not necessarily minimized under this choice, and thus the

Table 1 Anatomy of the error CV matrices for the statistical accuracy assessment of a MC solution

NEQs and MCs	$\mathbf{N}(\mathbf{x}-\mathbf{x}_o)=\mathbf{u}$ $\mathbf{H}(\mathbf{x}-\mathbf{x}_o)=\mathbf{c}=\mathbf{H}(\mathbf{x}^{ext}-\mathbf{x}_o)$	CV matrix of the pseudo-observation vector of the MCs	$\mathbf{\Sigma}_c=\mathbf{H}_1\mathbf{\Sigma}_x^{prior}\mathbf{H}_1^T$
Network partition scheme	$\mathbf{x}=\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{bmatrix}$, $\mathbf{x}_1 \rightarrow$ fiducial stations, $\mathbf{x}_2 \rightarrow$ new stations	CV matrix of the non-estimable frame parameters in MC solution	$\mathbf{\Sigma}_\theta=(\mathbf{H}_1\mathbf{E}_1^T)^{-1}\mathbf{H}_1\mathbf{\Sigma}_x^{prior}\mathbf{H}_1^T(\mathbf{E}_1\mathbf{H}_1^T)^{-1}$
	$\mathbf{E}=\begin{bmatrix}\mathbf{E}_1 & \mathbf{E}_2\end{bmatrix}$ inner-constraint matrix $\mathbf{H}=\begin{bmatrix}\mathbf{H}_1 & \mathbf{0}\end{bmatrix}$ used MC matrix	**Data noise effect** ($\mathbf{\Sigma}_{\hat{x}}^{obs}$)	$\sigma^2(\mathbf{N}+\mathbf{H}^T\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^T\mathbf{H})^{-1}$ *or* $\sigma^2((\mathbf{N}+\mathbf{H}^T\mathbf{H})^{-1}-\mathbf{E}^T(\mathbf{H}_1\mathbf{E}_1^T)^{-1}(\mathbf{E}_1\mathbf{H}_1^T)^{-1}\mathbf{E})$
A priori datum information	$\mathbf{x}^{ext}=\begin{bmatrix}\mathbf{x}_1^{ext}\\ \mathbf{0}\end{bmatrix}$, $(\mathbf{x}_1^{ext}, \mathbf{\Sigma}_x^{prior})$	**Datum noise effect** ($\mathbf{\Sigma}_{\hat{x}}^{mc}$)	$\mathbf{E}^T(\mathbf{H}_1\mathbf{E}_1^T)^{-1}\mathbf{H}_1\mathbf{\Sigma}_x^{prior}\mathbf{H}_1^T(\mathbf{E}_1\mathbf{H}_1^T)^{-1}\mathbf{E}$
MC solution	$\hat{\mathbf{x}}=\mathbf{x}_o+(\mathbf{N}+\mathbf{H}^T\mathbf{H})^{-1}\mathbf{u}$ $+(\mathbf{N}+\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{c}$	**Total error CV matrix** ($\mathbf{\Sigma}_{\hat{x}}=\mathbf{\Sigma}_{\hat{x}}^{obs}+\mathbf{\Sigma}_{\hat{x}}^{mc}$)	$\sigma^2(\mathbf{N}+\sigma^2\mathbf{H}^T\mathbf{\Sigma}_c^{-1}\mathbf{H})^{-1}$

optimal treatment of the datum noise effect remains an open problem in geodetic network adjustment theory.

3.3 The matrix $\mathbf{\Sigma}_{\hat{x}}^{mc}$

The CV matrix $\mathbf{\Sigma}_{\hat{x}}^{mc}$ is always singular with its rank being equal to the datum defect of the geodetic observations in the underlying network. Its role is to specify the accuracy of the adjusted coordinates in a MC solution due to the uncertainty of the (non-estimable part of the) coordinate system that is caused from the erroneous fiducial coordinates. The essential role of this matrix is revealed by the following 'covariance mapping' expression

$$\mathbf{\Sigma}_{\hat{x}}^{mc}=\mathbf{E}^T\mathbf{\Sigma}_\theta\mathbf{E} \tag{19}$$

where the CV matrix $\mathbf{\Sigma}_\theta$ specifies the accuracy of the non-estimable frame parameters which are fixed by the chosen MCs (i.e. $\mathbf{H}_1(\mathbf{x}_1-\mathbf{x}_1^{ext})=\mathbf{0}$)

$$\mathbf{\Sigma}_\theta=(\mathbf{H}_1\mathbf{E}_1^T)^{-1}\mathbf{H}_1\mathbf{\Sigma}_x^{prior}\mathbf{H}_1^T(\mathbf{E}_1\mathbf{H}_1^T)^{-1} \tag{20}$$

The last two equations are deduced from the analytic expression in Eq. (18) and they were originally presented, within a more general context, in Kotsakis (2012).

The consideration of the matrix $\mathbf{\Sigma}_{\hat{x}}^{mc}$ is necessary for the accuracy assessment of TRF realizations by means of MCs over a number of fiducial stations in geodetic network adjustments. Loosely speaking, this matrix activates the stochasticity of the datum definition due to the noise presence in the fiducial coordinates, and it enables the dissemination of the total accuracy of the adjusted network positions in a straightforward statistical fashion (i.e. use of $\mathbf{\Sigma}_{\hat{x}}^{-1}=(\mathbf{\Sigma}_{\hat{x}}^{obs}+\mathbf{\Sigma}_{\hat{x}}^{mc})^{-1}$ as a weight matrix for future use in other estimation problems).

The minimization of the trace$\mathbf{\Sigma}_{\hat{x}}^{mc}$ signifies a zero-order optimization problem which has not been tackled in the geodetic literature. Its solution will lead to the datum definition scheme (i.e. the MC submatrix $\mathbf{H}_1$) that ensures the most accurate alignment of a geodetic network with an external reference frame which is characterized by a given CV matrix $\mathbf{\Sigma}_x^{prior}$ over the fiducial stations. This issue has been briefly discussed in Kotsakis (2012) and it will be analyzed in detail throughout the following sections of the paper.

Remark The matrix $\mathbf{\Sigma}_{\hat{x}}^{mc}$ does not contribute to the estimation accuracy of the adjusted observations (or of any other estimable quantity) in the underlying network, and it is thus irrelevant for the quality assessment of the estimable characteristics in a MC solution. This is true however only within the linearized treatment of the least squares network adjustment problem that we follow in this paper. Theoretically, the datum noise effect may cause a disturbance to the estimable parameters of a linearly adjusted geodetic network under a nonlinear observational model, see Kotsakis (2012).

A summary of the various types of error CV matrices that were presented in this section is given in Table 1.

4 Optimal datum definition for MC networks

4.1 Problem formulation

For the present study, the optimal datum definition in a MC solution is linked to the minimization of an objective functional that quantifies the total accuracy of the estimated coordinates with respect to the reference frame realized by a set of fiducial stations. A common choice for this functional is the trace of the error CV matrix $\mathbf{\Sigma}_{\hat{x}}$ which was analytically described in the previous section.

Considering the partition of the MC matrix $\mathbf{H}=[\mathbf{H}_1\ \mathbf{0}]$ as stated in Sect. 2 (i.e. the MCs are applied only to the fiducial stations of the network), the optimal datum choice problem can be expressed as

$$\min_{\mathbf{H}_1} \operatorname{trace} \boldsymbol{\Sigma}_{\hat{x}} \tag{21}$$

or, equivalently, in terms of the matrix equation

$$\frac{\partial \operatorname{trace} \boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{obs}}}{\partial \mathbf{H}_1} + \frac{\partial \operatorname{trace} \boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}}}{\partial \mathbf{H}_1} = \mathbf{0} \tag{22}$$

The previous formulation leads to a 'complete' datum optimization in the sense that the MC submatrix obtained from the solution of Eq. (22) will minimize the combined effect of the data noise and datum noise on the adjusted network's coordinates. It is noted that in the geodetic literature the datum choice problem has been tackled as a special case of the above optimization scheme by ignoring the presence of the error covariance matrix $\boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}}$. In order to provide a unified approach that covers both cases, a more general criterion will be adopted herein for the datum choice problem

$$\boxed{\min_{\mathbf{H}_1} (\operatorname{trace} \boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{obs}} + \lambda \operatorname{trace} \boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}})} \tag{23}$$

where λ is a non-negative factor that controls the relative significance of the data noise and the datum noise on the estimated network coordinates. Based on this formulation the following matrix equation is obtained

$$\boxed{\frac{\partial \operatorname{trace} \boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{obs}}}{\partial \mathbf{H}_1} + \lambda \frac{\partial \operatorname{trace} \boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}}}{\partial \mathbf{H}_1} = \mathbf{0}} \tag{24}$$

which can be explicitly solved for the desired submatrix $\mathbf{H}_1$ (see next section).

Note that the weight factor λ does not have a regularization role within the datum choice problem, but it is merely used as an auxiliary parameter allowing us to obtain a general result for the optimal MC submatrix which can be specialized to the pertinent cases: $\lambda = 0$ (optimization of the data noise effect only), $\lambda = \infty$ (optimization of the datum noise effect only) and $\lambda = 1$ (joint optimization of the data and datum noise effects).

4.2 Derivation of optimal MCs

Following the previous formulation and in order to determine the solution of Eq. (24), we need to obtain the partial derivatives of the traces of the CV matrices $\boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{obs}}$ and $\boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}}$ with respect to the MC submatrix $\mathbf{H}_1$. Their analytic expressions are (see Appendix for their proofs):

$$\begin{aligned}\frac{\partial \operatorname{trace} \boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{obs}}}{\partial \mathbf{H}_1} &= -2\sigma^2 (\mathbf{E}\mathbf{H}^{\mathrm{T}})^{-1} \mathbf{E}(\mathbf{I} - \mathbf{E}^{\mathrm{T}}(\mathbf{H}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{H}) \\ &\quad \times (\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}}\end{aligned} \tag{25}$$

$$\begin{aligned}\frac{\partial \operatorname{trace} \boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}}}{\partial \mathbf{H}_1} &= 2(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1} \\ &\quad \times \mathbf{H}_1 \boldsymbol{\Sigma}_x^{\mathrm{prior}} (\mathbf{I} - \mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}_1)\end{aligned} \tag{26}$$

where the auxiliary matrix $\mathbf{S}$ has the block structure

$$\mathbf{S} = [\,\mathbf{I} \quad \mathbf{0}\,] \tag{27}$$

with the dimensions of its submatrices being such that

$$\mathbf{H}_1 = [\,\mathbf{H}_1 \quad \mathbf{0}\,] \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix} = \mathbf{H}\mathbf{S}^{\mathrm{T}} \tag{28}$$

$$\mathbf{E}_1 = [\,\mathbf{E}_1 \quad \mathbf{E}_2\,] \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix} = \mathbf{E}\mathbf{S}^{\mathrm{T}} \tag{29}$$

Obviously the following relationships also hold

$$\mathbf{H}\mathbf{E}^{\mathrm{T}} = \mathbf{H}_1\mathbf{E}_1^{\mathrm{T}} \quad \text{and} \quad \mathbf{E}\mathbf{H}^{\mathrm{T}} = \mathbf{E}_1\mathbf{H}_1^{\mathrm{T}} \tag{30}$$

$$\mathbf{S}\mathbf{S}^{\mathrm{T}} = \mathbf{I} \tag{31}$$

Substituting Eqs. (25)–(26) into Eq. (24) and taking into account the previous formulae, we get

$$\begin{aligned}
&- 2\sigma^2 (\mathbf{E}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{E}(\mathbf{I} - \mathbf{E}^{\mathrm{T}}(\mathbf{H}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{H})(\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}} \\
&\quad + 2\lambda (\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1} \\
&\quad \times \mathbf{H}_1 \boldsymbol{\Sigma}_x^{\mathrm{prior}} (\mathbf{I} - \mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}_1) = \mathbf{0} \Leftrightarrow \\
&- \sigma^2 (\mathbf{E}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{E}(\mathbf{I} - \mathbf{E}^{\mathrm{T}}(\mathbf{H}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{H})(\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}} \\
&\quad + \lambda (\mathbf{E}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}\mathbf{E}^{\mathrm{T}})^{-1} \\
&\quad \times \mathbf{H}\mathbf{S}^{\mathrm{T}} \boldsymbol{\Sigma}_x^{\mathrm{prior}} (\mathbf{S}\mathbf{S}^{\mathrm{T}} - \mathbf{S}\mathbf{H}^{\mathrm{T}}(\mathbf{E}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{S}^{\mathrm{T}}) = \mathbf{0} \Leftrightarrow \\
&- \sigma^2 \mathbf{E}(\mathbf{I} - \mathbf{E}^{\mathrm{T}}(\mathbf{H}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{H})(\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}} \\
&\quad + \lambda \mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{H}\mathbf{S}^{\mathrm{T}} \boldsymbol{\Sigma}_x^{\mathrm{prior}} \mathbf{S}(\mathbf{I} - \mathbf{H}^{\mathrm{T}}(\mathbf{E}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{E})\mathbf{S}^{\mathrm{T}} = \mathbf{0}
\end{aligned} \tag{32}$$

By considering the fundamental relationship (see Eq. 6)

$$\mathbf{I} - \mathbf{H}^{\mathrm{T}}(\mathbf{E}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{E} = \mathbf{N}(\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} \tag{33}$$

then Eq. (32) is reduced to the form

$$\begin{aligned}&(-\sigma^2\mathbf{E} + \sigma^2\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{H} + \lambda\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{H}\mathbf{S}^{\mathrm{T}}\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{S}\mathbf{N}) \\ &\quad \times (\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}} = \mathbf{0}\end{aligned} \tag{34}$$

which is equivalent to

$$\begin{aligned}&(-\sigma^2\mathbf{H}\mathbf{E}^{\mathrm{T}}(\mathbf{E}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{E} + \sigma^2\mathbf{H} + \lambda\mathbf{H}\mathbf{S}^{\mathrm{T}}\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{S}\mathbf{N}) \\ &\quad \times (\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}} = \mathbf{0}\end{aligned} \tag{35}$$

The orthogonal projection matrix $\mathbf{E}^{\mathrm{T}}(\mathbf{E}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{E}$ can be expressed as (see Eq. 6)

$$\mathbf{E}^{\mathrm{T}}(\mathbf{E}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{E} = \mathbf{I} - (\mathbf{N} + \mathbf{E}^{\mathrm{T}}\mathbf{E})^{-1}\mathbf{N} \tag{36}$$

and therefore Eq. (35) becomes

$$\begin{aligned}&(-\sigma^2\mathbf{H}(\mathbf{I} - (\mathbf{N} + \mathbf{E}^{\mathrm{T}}\mathbf{E})^{-1}\mathbf{N}) + \sigma^2\mathbf{H} + \lambda\mathbf{H}\mathbf{S}^{\mathrm{T}}\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{S}\mathbf{N}) \\ &\quad \times (\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}} = \mathbf{0}\end{aligned} \tag{37}$$

or equivalently

$$\mathbf{H}(\lambda \mathbf{S}^{\mathrm{T}} \boldsymbol{\Sigma}_x^{\mathrm{prior}} \mathbf{S} + \sigma^2 (\mathbf{N} + \mathbf{E}^{\mathrm{T}}\mathbf{E})^{-1}) \times \mathbf{N}(\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}} = \mathbf{0} \tag{38}$$

Recalling Eq. (33) and applying straightforward block-matrix multiplications, we have

$$\mathbf{N}(\mathbf{N} + \mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}} = (\mathbf{I} - \mathbf{H}^{\mathrm{T}}(\mathbf{E}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{E})\mathbf{S}^{\mathrm{T}} = \begin{bmatrix} \mathbf{I} - \mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}_1 \\ \mathbf{0} \end{bmatrix} \tag{39}$$

and by representing the symmetric matrix $(\mathbf{N} + \mathbf{E}^{\mathrm{T}}\mathbf{E})^{-1}$ as

$$(\mathbf{N} + \mathbf{E}^{\mathrm{T}}\mathbf{E})^{-1} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{12}^{\mathrm{T}} & \mathbf{M}_{22} \end{bmatrix} \tag{40}$$

where the above partitioning conforms to the dimensions of the block matrix

$$\mathbf{E}^{\mathrm{T}}\mathbf{E} = \begin{bmatrix} \mathbf{E}_1^{\mathrm{T}} \\ \mathbf{E}_2^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{E}_1 & \mathbf{E}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{E}_1^{\mathrm{T}}\mathbf{E}_1 & \mathbf{E}_1^{\mathrm{T}}\mathbf{E}_2 \\ \mathbf{E}_2^{\mathrm{T}}\mathbf{E}_1 & \mathbf{E}_2^{\mathrm{T}}\mathbf{E}_2 \end{bmatrix}$$

it is then a matter of simple matrix operations to verify that Eq. (38) leads to the algebraic formula

$$\mathbf{H}_1(\lambda \boldsymbol{\Sigma}_x^{\mathrm{prior}} + \sigma^2 \mathbf{M}_{11})(\mathbf{I} - \mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}_1) = \mathbf{0} \tag{41}$$

The solution of the last equation is given by the form

$$\boxed{\mathbf{H}_1 = \mathbf{E}_1(\lambda \boldsymbol{\Sigma}_x^{\mathrm{prior}} + \sigma^2 \mathbf{M}_{11})^{-1}} \tag{42}$$

which corresponds to the MC submatrix for the datum definition according to the optimization principle of Eq. (23).

Since the network solution $\hat{\mathbf{x}}$ is not affected if the MC matrix $\mathbf{H}$ is multiplied by an arbitrary non-zero scalar (i.e. this effect is absorbed by the non-essential weight matrix $\mathbf{W}$ as explained in Sect. 2), the optimal MC submatrix of Eq. (42) can be also expressed in the alternate form:

$$\boxed{\mathbf{H}_1 = \mathbf{E}_1(\boldsymbol{\Sigma}_x^{\mathrm{prior}} + \frac{\sigma^2}{\lambda} \mathbf{M}_{11})^{-1}} \tag{43}$$

Regarding the value of the auxiliary parameter λ, three important cases can be identified as follows.

4.2.1 *Optimization of the data noise effect* (*trace* $\boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{obs}} = \min$)

In the case that $\lambda = 0$ the optimal MC submatrix from Eq. (42) becomes

$$\mathbf{H}_1 = \mathbf{E}_1 \, (\mathbf{M}_{11})^{-1} \tag{44}$$

(note that the presence of the multiplicative factor σ^2 can be omitted for the reason explained before). The matrix $\mathbf{M}_{11}$ is taken as per Eq. (40) and its inverse provides the required weight matrix for the fiducial stations within the MC formulation

$$\mathbf{E}_1 \, (\mathbf{M}_{11})^{-1} \, (\mathbf{x}_1 - \mathbf{x}_1^{\mathrm{ext}}) = \mathbf{0} \tag{45}$$

so that the data noise effect is minimized *over the entire network*.

The above result yields a generalization of the classic (unweighted) inner constraints

$$\mathbf{E}_1(\mathbf{x}_1 - \mathbf{x}_1^{\mathrm{ext}}) = \mathbf{0} \tag{46}$$

whose optimality is restricted to the minimization of the data noise effect *at the selected fiducial stations* and not on other parts of the adjusted network (e.g. Koch 1999, pp. 62–64).

4.2.2 *Optimization of the datum noise effect* (*trace* $\boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}} = \min$)

In the case that $\lambda = \infty$ the optimal MC submatrix from Eq. (43) becomes

$$\mathbf{H}_1 = \mathbf{E}_1(\boldsymbol{\Sigma}_x^{\mathrm{prior}})^{-1} \tag{47}$$

The weighting for the fiducial stations now employs the CV matrix of their a priori coordinates $\mathbf{x}_1^{\mathrm{ext}}$ according to the MC formulation

$$\mathbf{E}_1(\boldsymbol{\Sigma}_x^{\mathrm{prior}})^{-1}(\mathbf{x}_1 - \mathbf{x}_1^{\mathrm{ext}}) = \mathbf{0} \tag{48}$$

so that the datum noise effect is minimized over the entire network.

It is noted that the minimization of the trace$\boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}}$ is algebraically equivalent to the minimization of the trace$\boldsymbol{\Sigma}_{\theta}$ (see Appendix for a proof). This exemplifies the fact that the weighted inner constraints in Eq. (48) provide the optimal 'frame alignment' for network densification problems in the context of MC theory.

4.2.3 *Joint optimization of the data/datum noise effects* (*trace*$(\boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{obs}} + \boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}}) = \min$)

In the case that $\lambda = 1$ the optimal MC submatrix from Eq. (42) or (43) becomes

$$\mathbf{H}_1 = \mathbf{E}_1(\boldsymbol{\Sigma}_x^{\mathrm{prior}} + \sigma^2 \mathbf{M}_{11})^{-1} \tag{49}$$

and it leads to the most general form of weighted inner constraints for the optimal datum definition in a geodetic network. In particular, the MCs

$$\mathbf{E}_1(\boldsymbol{\Sigma}_x^{\mathrm{prior}} + \sigma^2 \mathbf{M}_{11})^{-1}(\mathbf{x}_1 - \mathbf{x}_1^{\mathrm{ext}}) = \mathbf{0} \tag{50}$$

provide the network solution with the minimum-trace CV matrix $\boldsymbol{\Sigma}_{\hat{x}}$ for the total random error in the adjusted coordinates.

Table 2 Weighting schemes for the fiducial coordinates in MC adjustment of a geodetic network according to different optimality criteria

Network optimization criterion	Weight matrix $\mathbf{P}_x$	Objective
trace $(\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}} + \boldsymbol{\Sigma}_{\hat{x}}^{\text{mc}}) = \min$	$(\boldsymbol{\Sigma}_x^{\text{prior}} + \sigma^2 \mathbf{M}_{11})^{-1}$	Joint minimization of the data/datum noise effect
trace $\boldsymbol{\Sigma}_{\hat{x}}^{\text{mc}} = \min$ (equivalent to trace $\boldsymbol{\Sigma}_\theta = \min$)	$(\boldsymbol{\Sigma}_x^{\text{prior}})^{-1}$	Minimization of the datum noise effect
trace $\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}} = \min$	$\mathbf{M}_{11}^{-1}$	Minimization of the data noise effect
trace $\mathbf{S}\Sigma_{\hat{x}}^{\text{obs}}\mathbf{S}^{\text{T}} = \min$	$\mathbf{I}$	Minimization of the data noise effect (over the fiducial stations only) (*)

(*) This is an already known case which has not been proved anew in the paper; see e.g. (Koch 1999, pp. 62-64)

The matrix $\mathbf{M}_{11}$ is defined according to Eq. (40) while the rest of the matrices have been described in Sects. 3 and 4

The matrix $\mathbf{H}_1$ in Eq. (49) depends on the data variance factor σ^2 and thus it can be computed, in a rigorous way, only if the accuracy of the network measurements is perfectly known beforehand. This theoretical obstacle suggests the use of an iterative procedure for the implementation of the optimal MCs in Eq. (50), where the unknown variance factor σ^2 is replaced at each step by its a posteriori estimate $\hat{\sigma}^2$ that was obtained in the previous step of the MC adjustment.

4.3 Recapitulation

The optimal MCs for the reference frame realization in geodetic network adjustment by means of known (and noisy) fiducial coordinates have the general form

$$\boxed{\mathbf{E}_1 \mathbf{P}_x (\mathbf{x}_1 - \mathbf{x}_1^{\text{ext}}) = \mathbf{0}} \tag{51}$$

where $\mathbf{E}_1$ is the usual inner-constraint submatrix and $\mathbf{P}_x$ is a weight matrix that depends on the error source (or the combination of the error sources) whose effect on the MC solution needs to be minimized. The possible cases with their respective weight matrices are summarized in Table 2.

The matrix $\mathbf{P}_x$ should be distinguished from the weight matrix $\mathbf{W}$ which may be used for incorporating any set of MCs into a singular NEQ system, see Eq. (9). The latter provides a non-essential weighting for the pseudo-observation vector of a general system of MCs $\mathbf{H}(\mathbf{x} - \mathbf{x}_\text{o}) = \mathbf{c}$, whereas the former assigns the required weights to the fiducial stations so that the weighted inner constraints in Eq. (51) lead to an optimal network solution in a certain statistical sense (see Table 2).

It is worth mentioning that Eq. (51) gives rise to a weighted form for the no-net-translation (NNT) and no-net-rotation (NNR) conditions in geodetic networks. Considering the analytic structure of the inner constraint matrix $\mathbf{E}$ (e.g. Sillard and Boucher 2001), it is easily deduced that Eq. (51) results in the datum-related conditions

generalized NNT condition

$$\sum_{i,k} \mathbf{P}_{ik}(\mathbf{x}_1^k - \mathbf{x}_1^{k,\text{ext}}) = \mathbf{0} \tag{52}$$

generalized NNR condition

$$\sum_i [\mathbf{x}_1^{i,\text{o}} \times] \sum_k \mathbf{P}_{ik}(\mathbf{x}_1^k - \mathbf{x}_1^{k,\text{ext}}) = \mathbf{0} \tag{53}$$

where $\mathbf{P}_{ik}$ is the submatrix of the total weight matrix $\mathbf{P}_x$ that refers to the pair of the fiducial stations $\mathbf{x}_1^k$ and $\mathbf{x}_1^i$. In the geodetic literature the NNT/NNR conditions have been usually considered in a weightless form or, at most, in a scalar-weighted form (e.g. Angermann et al. 2004) assuming that

$$\mathbf{P}_{ik} = \begin{cases} \mathbf{0}, & i \neq k \\ \mu_k \mathbf{I}, & i \equiv k \end{cases} \tag{54}$$

where μ_k is a weight factor for each fiducial station whose optimal selection has remained unclear. This paper has presented a rigorous framework for the formation of a full weight matrix $\mathbf{P}_x$ to be used within the NNT/NNR conditions in accordance to specific optimal criteria for the realized reference frame in a MC network (see Table 2).

5 Conclusions

A class of optimal MCs for geodetic network adjustment, with respect to a given TRF, has been developed in this study. Our approach is based on a dual-objective optimization principle of maximum accuracy for the estimated coordinates which takes into consideration both the data noise and datum noise in the MC solution. Each of these effects has a distinct contribution to the total accuracy of the estimated coordinates and their joint minimization is a desirable task for the establishment of high-quality TRFs through a MC network adjustment.

The analytic form of the optimal MCs, given by Eq. (51), entails an a priori weighting of the fiducial stations within the realm of the well-known inner constraints. Various options for this weighting have been identified in the paper, depending on the optimization scenario for the CV matrix of the adjusted network's coordinates (see Table 2). Note that simplified versions of such weighting schemes have occasionally appeared in the geodetic literature (e.g. Angermann et al. 2004, pp. 28–31), however this is the first time that their implementation is justified on the basis of specific optimality criteria for the estimated coordinates in the MC network. In that sense, the theoretical findings of our study can provide a useful and rigorous tool to be exploited for frame densification problems by means of an optimal MC adjustment in geodetic networks.

Acknowledgments Prof. A. Dermanis and two anonymous reviewers are acknowledged for their comments and suggestions which helped in the improvement of the final version of this paper.

Appendix

Proof of Eqs. (13) *and* (14)

Given a singular NEQ matrix $\mathbf{N} = \mathbf{A}^{\mathrm{T}}\mathbf{PA}$ and any MC matrix $\mathbf{H}$, then the following equation holds for every symmetric positive-definite matrix $\mathbf{W}$ (Kotsakis 2012)

$$(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1}\mathbf{A}^{\mathrm{T}} = (\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{A}^{\mathrm{T}} \tag{55}$$

Hence, it also holds that

$$\mathbf{PA}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1} = \mathbf{PA}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} \tag{56}$$

and, therefore, we have

$$\begin{aligned}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1}&\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1}\\ &= (\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{PA}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1}\\ &= (\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{PA}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\\ &= (\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\end{aligned} \tag{57}$$

Furthermore, if we consider the general relationships (Kotsakis 2012)

$$(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1}\mathbf{N} = \mathbf{I}-\mathbf{E}^{\mathrm{T}}(\mathbf{HE}^{\mathrm{T}})^{-1}\mathbf{H} \tag{58}$$

$$(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{W} = \mathbf{E}^{\mathrm{T}}(\mathbf{HE}^{\mathrm{T}})^{-1} \tag{59}$$

then we have

$$\begin{aligned}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1}&\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1}\\ &= \left(\mathbf{I}-\mathbf{E}^{\mathrm{T}}(\mathbf{HE}^{\mathrm{T}})^{-1}\mathbf{H}\right)(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1}\\ &= (\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1}\\ &\quad -\mathbf{E}^{\mathrm{T}}(\mathbf{HE}^{\mathrm{T}})^{-1}\mathbf{W}^{-1}\mathbf{WH}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1}\\ &= (\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1}\\ &\quad -\mathbf{E}^{\mathrm{T}}(\mathbf{HE}^{\mathrm{T}})^{-1}\mathbf{W}^{-1}(\mathbf{EH}^{\mathrm{T}})^{-1}\mathbf{E}\end{aligned} \tag{60}$$

which, in view of Eq. (57), leads to the relationship

$$\begin{aligned}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{WH})^{-1} &- \mathbf{E}^{\mathrm{T}}(\mathbf{HE}^{\mathrm{T}})^{-1}\mathbf{W}^{-1}(\mathbf{EH}^{\mathrm{T}})^{-1}\mathbf{E}\\ &= (\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} - \mathbf{E}^{\mathrm{T}}(\mathbf{HE}^{\mathrm{T}})^{-1}(\mathbf{EH}^{\mathrm{T}})^{-1}\mathbf{E}\end{aligned} \tag{61}$$

Partial derivatives of $\mathrm{trace}\boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}}$ *and* $\mathrm{trace}\boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{obs}}$ *with respect to the MC submatrix* $\mathbf{H}_1$

The following derivations make use of the well-known differentiation rules from matrix calculus (Harville 1997, ch. 15):

$$\frac{\partial \mathbf{A}^{-1}}{\partial t} = -\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial t}\mathbf{A}^{-1} \tag{62}$$

$$\frac{\partial \mathrm{trace}\mathbf{A}}{\partial t} = \mathrm{trace}\frac{\partial \mathbf{A}}{\partial t} \tag{63}$$

$$\frac{\partial \mathbf{AB}}{\partial t} = \frac{\partial \mathbf{A}}{\partial t}\mathbf{B} + \mathbf{A}\frac{\partial \mathbf{B}}{\partial t} \tag{64}$$

where $\mathbf{A}$ and $\mathbf{B}$ are any matrices whose elements depend on some variable t. If the latter corresponds to a particular element $A(i,k)$ of the matrix $\mathbf{A}$, then we have

$$\frac{\partial \mathbf{A}}{\partial A(i,k)} = \mathbf{e}_i\mathbf{e}_k^{\mathrm{T}}, \quad \frac{\partial \mathbf{A}^{\mathrm{T}}}{\partial A(i,k)} = \mathbf{e}_k\mathbf{e}_i^{\mathrm{T}} \tag{65}$$

and

$$\mathbf{e}_i^{\mathrm{T}}\mathbf{A}\mathbf{e}_k = A(i,k) \tag{66}$$

where $\mathbf{e}_i$, $\mathbf{e}_k$ denote column unit vectors of appropriate dimensions with their ith and kth element being respectively equal to one. For our derivations we also use the partitioned-matrix differentiation rule

$$\underset{k\times m}{\tilde{\mathbf{A}}} = \left[\underset{k\times m_1}{\mathbf{A}} \quad \underset{k\times m_2}{\mathbf{0}}\right] \tag{67}$$

$$\frac{\partial \tilde{\mathbf{A}}}{\partial A(i,k)} = \mathbf{e}_i\mathbf{e}_k^{\mathrm{T}}\mathbf{S}, \quad \frac{\partial \tilde{\mathbf{A}}^{\mathrm{T}}}{\partial A(i,k)} = \mathbf{S}^{\mathrm{T}}\mathbf{e}_k\mathbf{e}_i^{\mathrm{T}} \tag{68}$$

where the unit vectors $\mathbf{e}_i$, $\mathbf{e}_k$ have dimensions $k\times 1$ and $m_1\times 1$ respectively, and the auxiliary matrix $\mathbf{S}$ is

$$\underset{m_1\times m}{\mathbf{S}} = \left[\underset{m_1\times m_1}{\mathbf{I}} \quad \underset{m_1\times m_2}{\mathbf{0}}\right] \tag{69}$$

Based on the analytic form of $\boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}}$ given in Eq. (18), the partial derivative of its trace with respect to an arbitrary element of the MC submatrix $\mathbf{H}_1$ is:

$$
\begin{aligned}
\frac{\partial \mathrm{trace}\boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}}}{\partial H_1(i,k)} &= \mathrm{trace}\left\{\frac{\partial \boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}}}{\partial H_1(i,k)}\right\} \\
&= \mathrm{trace}\left\{\frac{\partial \mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}}{\partial H_1(i,k)}\right\} \\
&= \mathrm{trace}\left\{\frac{\partial \mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1}{\partial H_1(i,k)}\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E} + \mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\frac{\partial \boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}}{\partial H_1(i,k)}\right\} \\
&= \mathrm{trace}\left\{\begin{array}{l}\left(\frac{\partial \mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}}{\partial H_1(i,k)}\mathbf{H}_1 + \mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\frac{\partial \mathbf{H}_1}{\partial H_1(i,k)}\right)\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E} \\ +\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\left(\frac{\partial \mathbf{H}_1^{\mathrm{T}}}{\partial H_1(i,k)}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E} + \mathbf{H}_1^{\mathrm{T}}\frac{\partial(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}}{\partial H_1(i,k)}\right)\end{array}\right\} \\
&= \mathrm{trace}\left\{\begin{array}{l}(-\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\frac{\partial \mathbf{H}_1\mathbf{E}_1^{\mathrm{T}}}{\partial H_1(i,k)}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1 + \mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\frac{\partial \mathbf{H}_1}{\partial H_1(i,k)})\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E} \\ +\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\left(\frac{\partial \mathbf{H}_1^{\mathrm{T}}}{\partial H_1(i,k)}\left(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}}\right)^{-1}\mathbf{E} - \mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\frac{\partial \mathbf{E}_1\mathbf{H}_1^{\mathrm{T}}}{\partial H_1(i,k)}\left(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}}\right)^{-1}\mathbf{E}\right)\end{array}\right\} \\
&= \mathrm{trace}\left\{\begin{array}{l}(-\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{e}_i\mathbf{e}_k^{\mathrm{T}}\mathbf{E}_1^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1 + \mathbf{E}^{\mathrm{T}}\left(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}}\right)^{-1}\mathbf{e}_i\mathbf{e}_k^{\mathrm{T}})\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E} \\ +\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\left(\mathbf{e}_k\mathbf{e}_i^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E} - \mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}_1\mathbf{e}_k\mathbf{e}_i^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\right)\end{array}\right\} \\
&= \mathrm{trace}\left\{\begin{array}{l}-\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{e}_i\mathbf{e}_k^{\mathrm{T}}\mathbf{E}_1^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E} \\ -\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}_1\mathbf{e}_k\mathbf{e}_i^{\mathrm{T}}\left(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}}\right)^{-1}\mathbf{E}\end{array}\right\} \\
&\quad + \mathrm{trace}\left\{\begin{array}{l}\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{e}_i\mathbf{e}_k^{\mathrm{T}}\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E} \\ +\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{e}_k\mathbf{e}_i^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\end{array}\right\} \\
&= \mathrm{trace}\left\{\begin{array}{l}(-\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{e}_i\mathbf{e}_k^{\mathrm{T}}\mathbf{E}_1^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E})^{\mathrm{T}} \\ -\mathbf{E}^{\mathrm{T}}\left(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}}\right)^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}\left(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}}\right)^{-1}\mathbf{E}_1\mathbf{e}_k\mathbf{e}_i^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\end{array}\right\} \\
&\quad + \mathrm{trace}\left\{\begin{array}{l}(\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{e}_i\mathbf{e}_k^{\mathrm{T}}\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E})^{\mathrm{T}} \\ +\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{e}_k\mathbf{e}_i^{\mathrm{T}}\left(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}}\right)^{-1}\mathbf{E}\end{array}\right\} \\
&= -2\ \mathrm{trace}\{\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}_1\mathbf{e}_k\mathbf{e}_i^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\} \\
&\quad +2\ \mathrm{trace}\{\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{e}_k\mathbf{e}_i^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\}
\end{aligned} \tag{70}
$$

For the second last equality in the above derivations we have used the fact that $\mathrm{trace}\mathbf{A}^{\mathrm{T}} = \mathrm{trace}\mathbf{A}$. Taking into account that $\mathrm{trace}\mathbf{AB} = \mathrm{trace}\mathbf{BA}$, the last equality of the previous formula yields

$$
\begin{aligned}
\frac{\partial \mathrm{trace}\boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}}}{\partial H_1(i,k)} &= -2\ \mathrm{trace}\{\mathbf{e}_i^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1} \\
&\quad \times\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}_1\mathbf{e}_k\} \\
&\quad +2\ \mathrm{trace}\{\mathbf{e}_i^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{e}_k\} \\
&= -2\ \mathbf{e}_i^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}} \\
&\quad \times\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}_1\mathbf{e}_k \\
&\quad +2\ \mathbf{e}_i^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{e}_k \\
&= -2\ [(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1} \\
&\quad \times\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}_1]_{ik} \\
&\quad +2\ [(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}]_{ik}
\end{aligned} \tag{71}
$$

Generalizing the above result for all elements of the MC submatrix $\mathbf{H}_1$, we have

$$
\begin{aligned}
\frac{\partial \mathrm{trace}\boldsymbol{\Sigma}_{\hat{x}}^{\mathrm{mc}}}{\partial \mathbf{H}_1} &= -2\ (\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1} \\
&\quad \times\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}_1 \\
&\quad +2\ (\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\mathrm{prior}}
\end{aligned} \tag{72}
$$

or, equivalently

$$\boxed{\begin{aligned}\frac{\partial \text{trace}\boldsymbol{\Sigma}_{\hat{x}}^{\text{mc}}}{\partial \mathbf{H}_1} &= 2(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}\mathbf{E}^{\mathrm{T}}(\mathbf{H}_1\mathbf{E}_1^{\mathrm{T}})^{-1} \\ &\quad \times \mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}(\mathbf{I}-\mathbf{H}_1^{\mathrm{T}}(\mathbf{E}_1\mathbf{H}_1^{\mathrm{T}})^{-1}\mathbf{E}_1)\end{aligned}} \quad (73)$$

which is identical to Eq. (26) given in Sect. 4.2

Based on the general form of $\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}}$ given in Eq. (11), the partial derivative of its trace with respect to an arbitrary element of the non-zero submatrix $\mathbf{H}_1$ is (note that the total MC matrix is assumed to have the form $\mathbf{H} = [\mathbf{H}_1 \ \mathbf{0}]$):

$$\begin{aligned}
\frac{\partial \text{trace}\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}}}{\partial H_1(i,k)} &= \text{trace}\left\{\frac{\partial \boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}}}{\partial H_1(i,k)}\right\} \\
&= \text{trace}\left\{\frac{\partial \sigma^2\left(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{N}\left(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}}{\partial H_1(i,k)}\right\} \\
&= \sigma^2\text{trace}\left\{\frac{\partial\left(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{N}}{\partial H_1(i,k)}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} + (\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}\frac{\partial(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}}{\partial H_1(i,k)}\right\} \\
&= \sigma^2\text{trace}\left\{\begin{array}{l} -(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\frac{\partial(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})}{\partial H_1(i,k)}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} \\ -(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\frac{\partial(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})}{\partial H_1(i,k)}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\end{array}\right\} \\
&= -\sigma^2\text{trace}\left\{\begin{array}{l} (\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\left(\frac{\partial\mathbf{H}^{\mathrm{T}}}{\partial H_1(i,k)}\mathbf{H}+\mathbf{H}^{\mathrm{T}}\frac{\partial\mathbf{H}}{\partial H_1(i,k)}\right)(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} \\ +(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\left(\frac{\partial\mathbf{H}^{\mathrm{T}}}{\partial H_1(i,k)}\mathbf{H}+\mathbf{H}^{\mathrm{T}}\frac{\partial\mathbf{H}}{\partial H_1(i,k)}\right)(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\end{array}\right\} \\
&= -\sigma^2\text{trace}\left\{\begin{array}{l} (\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\left(\mathbf{S}^{\mathrm{T}}\mathbf{e}_k\mathbf{e}_i^{\mathrm{T}}\mathbf{H}+\mathbf{H}^{\mathrm{T}}\mathbf{e}_i\mathbf{e}_k^{\mathrm{T}}\mathbf{S}\right)(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} \\ +(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\left(\mathbf{S}^{\mathrm{T}}\mathbf{e}_k\mathbf{e}_i^{\mathrm{T}}\mathbf{H}+\mathbf{H}^{\mathrm{T}}\mathbf{e}_i\mathbf{e}_k^{\mathrm{T}}\mathbf{S}\right)(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\end{array}\right\} \\
&= -\sigma^2\text{trace}\left\{\begin{array}{l} (\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{e}_k\mathbf{e}_i^{\mathrm{T}}\mathbf{H}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} \\ +(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{e}_i\mathbf{e}_k^{\mathrm{T}}\mathbf{S}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} \\ +(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{e}_k\mathbf{e}_i^{\mathrm{T}}\mathbf{H}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} \\ +(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{e}_i\mathbf{e}_k^{\mathrm{T}}\mathbf{S}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\end{array}\right\} \\
&= -\sigma^2\left(\begin{array}{l} \mathbf{e}_i^{\mathrm{T}}\mathbf{H}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-2}\mathbf{S}^{\mathrm{T}}\mathbf{e}_k \\ +\mathbf{e}_k^{\mathrm{T}}\mathbf{S}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-2}\mathbf{H}^{\mathrm{T}}\mathbf{e}_i \\ +\mathbf{e}_i^{\mathrm{T}}\mathbf{H}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-2}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{e}_k \\ +\mathbf{e}_k^{\mathrm{T}}\mathbf{S}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-2}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{e}_i\end{array}\right) \\
&= -\sigma^2\left(\begin{array}{l} \mathbf{e}_i^{\mathrm{T}}\mathbf{H}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-2}\mathbf{S}^{\mathrm{T}}\mathbf{e}_k \\ +\mathbf{e}_i^{\mathrm{T}}\mathbf{H}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-2}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{e}_k \\ +\mathbf{e}_i^{\mathrm{T}}\mathbf{H}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-2}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{e}_k \\ +\mathbf{e}_i^{\mathrm{T}}\mathbf{H}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-2}\mathbf{S}^{\mathrm{T}}\mathbf{e}_k\end{array}\right) \\
&= -2\sigma^2\left(\begin{array}{l} \mathbf{e}_i^{\mathrm{T}}\mathbf{H}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-2}\mathbf{S}^{\mathrm{T}}\mathbf{e}_k \\ +\mathbf{e}_i^{\mathrm{T}}\mathbf{H}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-2}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{e}_k\end{array}\right)
\end{aligned} \quad (74)$$

Recalling from Eqs. (6) and (7) that

$$\mathbf{H}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} = (\mathbf{E}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{E} \quad (75)$$

$$\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} = \mathbf{I}-\mathbf{H}^{\mathrm{T}}(\mathbf{E}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{E} \quad (76)$$

then we have

$$\mathbf{H}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} = \mathbf{0} \quad (77)$$

$$\begin{aligned}&\mathbf{H}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{N} \\ &\quad = (\mathbf{E}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{E}(\mathbf{I}-\mathbf{E}^{\mathrm{T}}(\mathbf{H}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{H})\end{aligned} \quad (78)$$

and by substituting the last two equations into (74) we get

$$\begin{aligned}\frac{\partial \text{trace}\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}}}{\partial H_1(i,k)} &= -2\sigma^2\mathbf{e}_i^{\mathrm{T}}(\mathbf{E}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{E}(\mathbf{I}-\mathbf{E}^{\mathrm{T}}(\mathbf{H}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{H}) \\ &\quad \times(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{e}_k \\ &= -2\sigma^2[(\mathbf{E}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{E}(\mathbf{I}-\mathbf{E}^{\mathrm{T}}(\mathbf{H}\mathbf{E}^{\mathrm{T}})^{-1}\mathbf{H}) \\ &\quad \times(\mathbf{N}+\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{S}^{\mathrm{T}}]_{ik}\end{aligned} \quad (79)$$

Generalizing the above result for all elements of the MC submatrix $\mathbf{H}_1$, we finally have

$$\boxed{\begin{aligned}\frac{\partial \text{trace}\boldsymbol{\Sigma}_{\hat{x}}^{\text{obs}}}{\partial \mathbf{H}_1} &= -2\sigma^2(\mathbf{E}\mathbf{H}^{\text{T}})^{-1}\mathbf{E}(\mathbf{I}-\mathbf{E}^{\text{T}}(\mathbf{H}\mathbf{E}^{\text{T}})^{-1}\mathbf{H})\\ &\quad\times(\mathbf{N}+\mathbf{H}^{\text{T}}\mathbf{H})^{-1}\mathbf{S}^{\text{T}}\end{aligned}} \tag{80}$$

which is identical to Eq. (25) given in Sect. 4.2.

Equivalency of $\min\limits_{\mathbf{H}_1}\text{trace}\boldsymbol{\Sigma}_{\hat{x}}^{\text{mc}}$ *and* $\min\limits_{\mathbf{H}_1}\text{trace}\boldsymbol{\Sigma}_{\theta}$

Let us first determine the partial derivative of the $\text{trace}\boldsymbol{\Sigma}_{\theta}$ with respect to the MC submatrix $\mathbf{H}_1$. Using Eq. (20) we have:

$$\begin{aligned}
\frac{\partial \text{trace}\boldsymbol{\Sigma}_\theta}{\partial H_1(i,k)} &= \text{trace}\left\{\frac{\partial\boldsymbol{\Sigma}_\theta}{\partial H_1(i,k)}\right\} = \text{trace}\left\{\frac{\partial(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}}{\partial H_1(i,k)}\right\}\\
&= \text{trace}\left\{\frac{\partial(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1}{\partial H_1(i,k)}\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1} + (\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\frac{\partial\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}}{\partial H_1(i,k)}\right\}\\
&= \text{trace}\left\{\begin{array}{l}\left(\frac{\partial(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}}{\partial H_1(i,k)}\mathbf{H}_1 + (\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\frac{\partial\mathbf{H}_1}{\partial H_1(i,k)}\right)\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\\ +(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\left(\frac{\partial\mathbf{H}_1^{\text{T}}}{\partial H_1(i,k)}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1} + \mathbf{H}_1^{\text{T}}\frac{\partial(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}}{\partial H_1(i,k)}\right)\end{array}\right\}\\
&= \text{trace}\left\{\begin{array}{l}\left(-(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\frac{\partial\mathbf{H}_1\mathbf{E}_1^{\text{T}}}{\partial H_1(i,k)}(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1 + (\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\frac{\partial\mathbf{H}_1}{\partial H_1(i,k)}\right)\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\\ +(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\left(\frac{\partial\mathbf{H}_1^{\text{T}}}{\partial H_1(i,k)}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1} - \mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\frac{\partial\mathbf{E}_1\mathbf{H}_1^{\text{T}}}{\partial H_1(i,k)}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\right)\end{array}\right\}\\
&= \text{trace}\left\{\begin{array}{l}\left(-(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{e}_i\mathbf{e}_k^{\text{T}}\mathbf{E}_1^{\text{T}}(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1 + (\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{e}_i\mathbf{e}_k^{\text{T}}\right)\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\\ +(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\left(\mathbf{e}_k\mathbf{e}_i^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1} - \mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\mathbf{E}_1\mathbf{e}_k\mathbf{e}_i^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\right)\end{array}\right\}\\
&= \text{trace}\left\{\begin{array}{l}-(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{e}_i\mathbf{e}_k^{\text{T}}\mathbf{E}_1^{\text{T}}(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\\ -(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\mathbf{E}_1\mathbf{e}_k\mathbf{e}_i^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\end{array}\right\}\\
&\quad+\text{trace}\left\{\begin{array}{l}(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{e}_i\mathbf{e}_k^{\text{T}}\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\\ +(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{e}_k\mathbf{e}_i^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\end{array}\right\}\\
&= \text{trace}\left\{\begin{array}{l}\left(-(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{e}_i\mathbf{e}_k^{\text{T}}\mathbf{E}_1^{\text{T}}(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\right)^{\text{T}}\\ -(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\mathbf{E}_1\mathbf{e}_k\mathbf{e}_i^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\end{array}\right\}\\
&\quad+\text{trace}\left\{\begin{array}{l}\left((\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{e}_i\mathbf{e}_k^{\text{T}}\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\right)^{\text{T}}\\ +(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{e}_k\mathbf{e}_i^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\end{array}\right\}\\
&= -2\text{trace}\left\{(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\mathbf{E}_1\mathbf{e}_k\mathbf{e}_i^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\right\}\\
&\quad+2\text{trace}\left\{(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{e}_k\mathbf{e}_i^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\right\}
\end{aligned} \tag{81}$$

which is equivalent to

$$\begin{aligned}
&\frac{\partial \text{trace}\boldsymbol{\Sigma}_\theta}{\partial H_1(i,k)}\\
&= -2\mathbf{e}_i^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\mathbf{E}_1\mathbf{e}_k\\
&\quad+2\mathbf{e}_i^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{e}_k\\
&= -2[(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\mathbf{E}_1]_{ik}\\
&\quad+2[(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}]_{ik}
\end{aligned} \tag{82}$$

Generalizing the above result for all elements of the MC submatrix $\mathbf{H}_1$, we get

$$\begin{aligned}\frac{\partial \text{trace}\boldsymbol{\Sigma}_\theta}{\partial \mathbf{H}_1} &= 2(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\\ &\quad\times\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}(\mathbf{I}-\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\mathbf{E}_1)\end{aligned} \tag{83}$$

The solution of the optimization problem $\min\limits_{\mathbf{H}_1}\text{trace}\boldsymbol{\Sigma}_\theta$ is obtained by setting the above expression equal to zero, i.e.

$$(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}(\mathbf{H}_1\mathbf{E}_1^{\text{T}})^{-1}\mathbf{H}_1\boldsymbol{\Sigma}_x^{\text{prior}}(\mathbf{I}-\mathbf{H}_1^{\text{T}}(\mathbf{E}_1\mathbf{H}_1^{\text{T}})^{-1}\mathbf{E}_1) = \mathbf{0} \tag{84}$$

which is algebraically equivalent to

$$\mathbf{H}_1 \boldsymbol{\Sigma}_x^{\text{prior}} (\mathbf{I} - \mathbf{H}_1^{\mathrm{T}} (\mathbf{E}_1 \mathbf{H}_1^{\mathrm{T}})^{-1} \mathbf{E}_1) = \mathbf{0} \tag{85}$$

and it is obviously satisfied when the MC submatrix has the form

$$\mathbf{H}_1 = \mathbf{E}_1 (\boldsymbol{\Sigma}_x^{\text{prior}})^{-1} \tag{86}$$

The above result is identical to the one obtained in Sect. 4 under the alternate optimization principle $\min\limits_{\mathbf{H}_1} \text{trace} \boldsymbol{\Sigma}_{\hat{x}}^{\text{mc}}$.

References

Altamimi Z, Boucher C, Sillard P (2002) New trends for the realization of the International Terrestrial Reference System. Adv Spac Res 30(2):175–184

Angermann D, Drewes H, Krugel M, Meisel B, Gerstl M, Kelm R, Muller H, Seemuller W, Tesmer V (2004) ITRS Combination Center at DGFI: a terrestrial reference frame realization 2003. Deutsche Geodätische Kommission, Reihe B, Heft Nr 313

Blaha G (1971) Inner adjustment constraints with emphasis on range observations, OSU report no. 148. Department of Geodetic Science, The Ohio State University, Columbus

Blaha G (1982) Free networks: minimum norm solution as obtained by the inner adjustment constraint method. Bull Geod 56:209–219

Coulot D, Pollet A, Collilieux X, Berio P (2010) Global optimization of core station networks for space geodesy: application to the referencing of the SLR EOP with respect to ITRF. J Geod 84:31–50

Dermanis A (1985) Optimization problems in geodetic networks with signals. In: Grafarend EW, Sanso F (eds) Optimization and design of geodetic networks. Springer, Berlin, pp 221–256

Dermanis A (1998) Generalized inverses of nonlinear mappings and the nonlinear geodetic datum problem. J Geod 72:71–100

Dermanis A (2012) Personal communication

Grafarend EW (1974) Optimization of geodetic networks. Boll Geod Sci Affi XXXIII:351–406

Grafarend EW, Schaffrin B (1974) Unbiased free net adjustment. Surv Rev XXII(171):200–218

Grafarend EW, Schaffrin B (1976) Equivalence of estimable quantities and invariants in geodetic networks. ZfV 101:485–491

Harville DA (1997) Matrix algebra from a statistician perspective. Springer, New York

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

Kotsakis C (2012) Reference frame stability and nonlinear distortion in minimum-constrained network adjustment. J Geod 86(9):755–774

Meissl P (1969) Zusammengfassung und Ausbau der inneren Fehlertheoric eines Punkthaufens. Deutsche Geodätische Kommission, Reihe A 61:8–21

Papo HB, Perelmuter A (1981) Datum definition by free net adjustment. Bull Geod 55:218–226

Pope AJ (1971) Transformation of covariance matrices due to changes in minimal control. Presented at the AGU Fall Meeting, San Francisco, 9 Dec 1971

Schaffrin B (1985) Aspects of network design. In: Grafarend EW, Sanso F (eds) Optimization and design of geodetic networks. Springer, Berlin, pp 549–597

Schmitt G (1982) Optimization of geodetic networks. Rev Geophys Space Phys 20(4):877–884

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

Teunissen P (1985) Zero order design: generalized inverse, adjustment, the datum problem and S-transformations. In: Grafarend EW, Sanso F (eds) Optimization and design of geodetic networks. Springer, Berlin, pp 11–55