First International Symposium
on the Mathematical and Physical Foundations of Theoretical Geodesy
September 7th-9th, 1993, Stuttgart, Germany

Quadratic Collocation and Robust Quadratic Collocation

Athanasios Dermanis

Department of Geodesy and Surveying
The Aristotle University of Thessaloniki

1. Introduction

Estimation and prediction in geodesy is typically limited to seeking optimal (minimum mean square error) estimators and/or predictors in the class of linear functions of the parameters. This can be justified in view of the well known fundamental result (see e.g. Rao, 1973) that when the random parameters of the model follow a Gaussian distribution, the linear optimal estimators and predictors cannot be improved, i.e. they are indeed the optimal non-linear ones. Since Gaussianity is rather a mathematical convenience that an absolute observational fact, it seems worthwhile to seek non-linear estimators and predictors which minimize the mean square error, without any restrictive assumptions on the distribution of the relevant random variables.

The idea of using nonlinear predictors goes back to Grafarend (1972). Dermanis and Sanso (1993) have studied some basic characteristics of non linear estimators and came to the conclusion that estimation of deterministic model parameters is feasible only within a Bayesian point of view. There is no need to seek rescue in Bayesian ideas in the case of random effects models which involve no deterministic but only random parameters. Prediction of random parameters, either directly present in the model or stochastically related ones, is of much geodetic interest especially for the determination of the gravity field of the earth conceived as a stochastic process in the framework of the statistical approach of collocation (Moritz, 1980, Sanso, 1986).

Among the various types of non-linear predictors the simplest and perhaps the only practically tractable ones are the quadratic ones. Quadratic estimation is widely used for the determination of variance components (e.g. Schaffrin, 1983) and rarely for combined variance component and parameter estimation (Kubáček, 1985).

Dermanis and Sanso (1993) have derived Bayesian quadratic estimators which can be directly applied to random effect models such as the models used in geodetic collocation. Here optimal quadratic predictors will be derived directly for the simplest collocation model and then extended to non-linear ones. As expected the results extend the usual (linear) collocation equations in a way which involves the effect of third and fourth order central moments which describe the deviation of the distribution of the random parameters from the Gaussian assumption.

2. The simple collocation model

We start with the simplest possible collocation model of direct observation of random signals $\mathbf{s}$ under additive zero-mean observational errors $\mathbf{v}$

$$\mathbf{y}=\mathbf{s}+\mathbf{v}\,, \qquad E\{\mathbf{s}\}=\mathbf{m}\,, \qquad E\{\mathbf{v}\}=\mathbf{0}\,. \tag{1}$$

We assume that $\mathbf{s}$ and $\mathbf{v}$ are stochastically independent, but no further assumptions are made about their probability distributions.

The objective is to seek optimal predictors for any single random variable s', which is stochastically related to $\mathbf{s}$, within the class of quadratic functions of the observations $\mathbf{y}$, i.e. of the form

$$\hat{s}'=\gamma+\mathbf{d}^T\mathbf{y}+\mathbf{y}^T\mathbf{Q}\mathbf{y} \qquad \text{(inhomogeneous)} \tag{2a}$$

or

$$\hat{s}'=\mathbf{d}^T\mathbf{y}+\mathbf{y}^T\mathbf{Q}\mathbf{y} \qquad \text{(homogeneous)} \tag{2b}$$

We shall call the predictor (2a) inhomogeneous and the predictor (2b) homogeneous, not in respect to the property of homogeneity itself (of course both predictors are inhomogeneous functions of $\mathbf{y}$) but in order to establish a correspondence with the terminology used in the linear case obtained by letting $\mathbf{Q}=\mathbf{0}$. These two predictors can be combined with the possibility of introducing the property of unbiasedness or not, in order to produce four basic classes of best quadratic predictors:

Best inhomogeneous Quadratic Prediction *(inhomBQP)*
Best homogeneous Quadratic Prediction *(homBQP)*
Best inhomogeneous Quadratic Unbiased Prediction *(inhomBQUP)*
Best homogeneous Quadratic Unbiased Prediction *(homBQUP)*

Best (optimal) prediction here refers to the minimization of the ***mean square prediction error*** :

$$\phi=E\{(\hat{s}'-s')^2\}=E\{\varepsilon^2\}=\min \tag{3}$$

where $\varepsilon=\hat{s}'-s'$ is the prediction error. The optional property of unbiased prediction is

$$\beta = E\{\hat{s}' - s'\} = E\{\varepsilon\} = 0 \,. \tag{4}$$

Since $\hat{s}'$ is quadratic in $\mathbf{y}$, ε^2 will be of the 4th order in $\mathbf{y}$, and consequently ϕ will depend on up to 4th order (central) moments of $\mathbf{s}$, $\mathbf{v}$, and cross-moments of s' and $\mathbf{s}$. To describe these moments in matrix notation we introduce the following basic definitions :

$$\mathbf{m} = E\{\mathbf{s}\}\,, \qquad \delta\mathbf{s} = \mathbf{s} - \mathbf{m}\,, \qquad \delta\mathbf{y} = \mathbf{y} - \mathbf{m} = \delta\mathbf{s} + \mathbf{v} \tag{5}$$

$$m' = E\{s'\}\,, \quad \delta s' = s' - m' \tag{6}$$

$$\mathbf{C}_{\mathbf{s}} = E\{\delta\mathbf{s}\delta\mathbf{s}^T\}\,, \quad \mathbf{C}_{\mathbf{v}} = E\{\mathbf{v}\mathbf{v}^T\}\,, \tag{7}$$

$$\mathbf{C} = E\{\delta\mathbf{y}\delta\mathbf{y}^T\} = \mathbf{C}_{\mathbf{s}} + \mathbf{C}_{\mathbf{v}}\,, \tag{8}$$

$$\mathbf{c}_{\mathbf{s}s'} = E\{\delta\mathbf{s}\delta s'\} = E\{\delta\mathbf{y}\delta s'\} = \mathbf{c}^T_{s'\mathbf{s}}\,, \tag{9}$$

$$\sigma^2_{s'} = E\{(\delta s')^2\}\,, \tag{10}$$

$$\mathbf{\Phi} = E\{vec(\delta\mathbf{y}\delta\mathbf{y}^T)\delta\mathbf{y}^T\} = \begin{bmatrix} \mathbf{\Phi}_1 \\ \vdots \\ \mathbf{\Phi}_n \end{bmatrix} = \mathbf{\Phi}_{\mathbf{s}} + \mathbf{\Phi}_{\mathbf{v}}\,, \tag{11}$$

$$\mathbf{\Phi}_{s'\mathbf{s}} = E\{\delta s'\delta\mathbf{y}\delta\mathbf{y}^T\} = E\{\delta s'\delta\mathbf{s}\delta\mathbf{s}^T\}\,, \tag{12}$$

$$\mathbf{\Psi} = E\{vec(\delta\mathbf{y}\delta\mathbf{y}^T)vec^T(\delta\mathbf{y}\delta\mathbf{y}^T)\} =$$

$$= \begin{bmatrix} \mathbf{\Psi}_{11} & \cdots & \mathbf{\Psi}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{\Psi}_{n1} & \cdots & \mathbf{\Psi}_{nn} \end{bmatrix} = \mathbf{\Psi}_{\mathbf{s}} + \mathbf{\Psi}_{\mathbf{v}} \tag{13}$$

3. Solution

For the determination of the optimal quadratic predictors it is more convenient to express (2a) in the equivalent form ($\mathbf{y} = \mathbf{m} + \delta\mathbf{y}$)

$$\hat{s}' = \gamma + (\mathbf{d} + 2\mathbf{Q}\mathbf{m})^T\mathbf{y} + vec^T\mathbf{Q}\,vec(\delta\mathbf{y}\delta\mathbf{y}^T - \mathbf{m}\mathbf{m}^T) =$$

$$= \gamma + \begin{bmatrix} \mathbf{d}^T + 2\mathbf{m}^T\mathbf{Q} & vec^T\mathbf{Q} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ vec(\delta\mathbf{y}\delta\mathbf{y}^T - \mathbf{m}\mathbf{m}^T) \end{bmatrix} =$$

$$= \gamma + \mathbf{a}^T\mathbf{z} \tag{14}$$

where $\mathbf{a}$ is a new unknown vector,

$$\mathbf{z} = \begin{bmatrix} \mathbf{y} \\ vec(\delta\mathbf{y}\delta\mathbf{y}^T - \mathbf{m}\mathbf{m}^T) \end{bmatrix}, \tag{15}$$

$$\mathbf{m}_{\mathbf{z}} = E\{\mathbf{z}\} = \begin{bmatrix} \mathbf{m} \\ vec(\mathbf{C} - \mathbf{m}\mathbf{m}^T) \end{bmatrix}, \tag{16}$$

$$\mathbf{C}_{\mathbf{z}} = E\{(\mathbf{z} - \mathbf{m}_{\mathbf{z}})(\mathbf{z} - \mathbf{m}_{\mathbf{z}})^T\} = \begin{bmatrix} \mathbf{C} & \mathbf{\Phi}^T \\ \mathbf{\Phi} & \mathbf{\Psi} - vec\mathbf{C}\,vec^T\mathbf{C} \end{bmatrix} \tag{17}$$

$$\mathbf{c}_{\mathbf{z}s'} = E\{(\mathbf{z} - \mathbf{m}_{\mathbf{z}})(s' - m')\} = \begin{bmatrix} \mathbf{c}_{\mathbf{s}s'} \\ vec\mathbf{\Phi}_{s'\mathbf{s}} \end{bmatrix} \tag{18}$$

while for the homogeneous case (2b) becomes $\hat{s}' = \mathbf{a}^T\mathbf{z}$.

With the above "vectorized" formulation it is easy to calculate the bias and the mean square error

$$\beta = \gamma + \mathbf{a}^T\mathbf{m}_{\mathbf{z}} - m' \tag{19}$$

$$\phi = \beta^2 + \mathbf{a}^T\mathbf{C}_{\mathbf{z}}\mathbf{a} + \sigma^2_{s'} - 2\mathbf{a}^T\mathbf{c}_{\mathbf{z}s'}\,. \tag{20}$$

The problem of prediction has now the same structure as the linear problem with well known solution (Schaffrin, 1985a). A simple derivation for the various types of predictors is given in appendix A, following Dermanis (1991). It has the general form

$$\hat{s}' = \alpha\,m' + \mathbf{c}^T_{\mathbf{z}s'}\mathbf{C}^{-1}_{\mathbf{z}}(\mathbf{z} - \alpha\,\mathbf{m}_{\mathbf{z}}) \tag{21}$$

with (minimum) mean square error

$$\phi = \sigma^2_{s'} - \mathbf{c}^T_{\mathbf{z}s'}\mathbf{C}^{-1}_{\mathbf{z}}\mathbf{c}_{\mathbf{z}s'} + \delta\phi \tag{22}$$

where α and $\delta\phi$ are different for different predictors:

inhomBQP = *inhomBQUP*:

$$\alpha = 1\,, \qquad \delta\phi = 0\,, \qquad \beta = 0\,. \tag{23}$$

homBQP:

$$\alpha = \frac{\mathbf{m}^T_{\mathbf{z}}\mathbf{C}^{-1}_{\mathbf{z}}\mathbf{z}}{1 + \mathbf{m}^T_{\mathbf{z}}\mathbf{C}^{-1}_{\mathbf{z}}\mathbf{m}_{\mathbf{z}}}\,, \tag{24a}$$

$$\delta\phi = \frac{(m' - \mathbf{c}^T_{\mathbf{z}s'}\mathbf{C}^{-1}_{\mathbf{z}}\mathbf{m}_{\mathbf{z}})^2}{1 + \mathbf{m}^T_{\mathbf{z}}\mathbf{C}^{-1}_{\mathbf{z}}\mathbf{m}_{\mathbf{z}}}\,, \tag{24b}$$

$$\beta = \frac{\mathbf{c}^T_{\mathbf{z}s'}\mathbf{C}^{-1}_{\mathbf{z}}\mathbf{m}_{\mathbf{z}} - m'}{1 + \mathbf{m}^T_{\mathbf{z}}\mathbf{C}^{-1}_{\mathbf{z}}\mathbf{m}_{\mathbf{z}}}\,. \tag{24c}$$

homBQUP:

$$\alpha = \frac{\mathbf{m}^T_{\mathbf{z}}\mathbf{C}^{-1}_{\mathbf{z}}\mathbf{z}}{\mathbf{m}^T_{\mathbf{z}}\mathbf{C}^{-1}_{\mathbf{z}}\mathbf{m}_{\mathbf{z}}}\,, \qquad \delta\phi = \frac{(m' - \mathbf{c}^T_{\mathbf{z}s'}\mathbf{C}^{-1}_{\mathbf{z}}\mathbf{m}_{\mathbf{z}})^2}{\mathbf{m}^T_{\mathbf{z}}\mathbf{C}^{-1}_{\mathbf{z}}\mathbf{m}_{\mathbf{z}}}\,, \quad \beta = 0\,. \tag{25}$$

In order to obtain expicit equations we only need to replace $\mathbf{z}$ from (15), $\mathbf{m}$ from (16), $\mathbf{C}_{\mathbf{z}}$ from (17) and $\mathbf{c}_{\mathbf{z}s'}$ from (18). Also the following identity needs to be used

$$\mathbf{C}^{-1}_{\mathbf{z}} = \begin{bmatrix} \mathbf{C} & \mathbf{\Phi}^T \\ \mathbf{\Phi} & \mathbf{\Psi} - vec\mathbf{C}vec^T\mathbf{C} \end{bmatrix}^{-1} =$$

$$=\begin{bmatrix} \mathbf{C}^{-1}+\mathbf{C}^{-1}\mathbf{\Phi}^T\mathbf{N}^{-1}\mathbf{\Phi}\mathbf{C}^{-1} & -\mathbf{C}^{-1}\mathbf{\Phi}^T\mathbf{N}^{-1} \\ -\mathbf{N}^{-1}\mathbf{\Phi}\mathbf{C}^{-1} & \mathbf{N}^{-1} \end{bmatrix} \quad (26)$$

where

$$\mathbf{N}=\mathbf{\Psi}-vec\mathbf{C}vec^T\mathbf{C}-\mathbf{\Phi}\mathbf{C}^{-1}\mathbf{\Phi}^T \,. \quad (27)$$

The general solution has the form

$$\hat{s}'=\alpha\, m'+\mathbf{c}_{\mathbf{s}s'}^T\mathbf{C}^{-1}(\mathbf{s}-\alpha\,\mathbf{m})+\mathbf{n}_{\mathbf{s}s'}^T\mathbf{N}^{-1}[\mathbf{q}(\mathbf{y})-\alpha\,\mathbf{m}_\mathbf{q}] \quad (28)$$

with

$$\mathbf{n}_{\mathbf{s}s'}=\mathbf{\Phi}\mathbf{C}^{-1}\mathbf{c}_{\mathbf{s}s'}-vec\mathbf{\Phi}_{s'\mathbf{s}}\,, \quad (29)$$

$$\mathbf{q}=\mathbf{q}(\mathbf{y})=\mathbf{\Phi}\mathbf{C}^{-1}\mathbf{y}-vec[(\mathbf{y}-\mathbf{m})(\mathbf{y}-\mathbf{m})^T-\mathbf{m}\mathbf{m}^T]\,, \quad (30)$$

$$\mathbf{m}_\mathbf{q}=\mathbf{\Phi}\mathbf{C}^{-1}\mathbf{m}-vec(\mathbf{C}-\mathbf{m}\mathbf{m}^T)=E\{\mathbf{q}(\mathbf{y})\}\,. \quad (31)$$

The mean square error of the prediction is

$$\phi=\sigma_{s'}^2-\mathbf{c}_{\mathbf{s}s'}^T\mathbf{C}^{-1}\mathbf{c}_{\mathbf{s}s'}-\mathbf{n}_{\mathbf{s}s'}^T\mathbf{N}^{-1}\mathbf{n}_{\mathbf{s}s'}+\delta\phi\,. \quad (32)$$

For the particular prediction types we have

inhomBQP = *inhomBQUP*:

$$\alpha=1\,, \qquad \delta\phi=0\,, \qquad \beta=0\,. \quad (33)$$

homBQP:

$$\alpha=\frac{\mathbf{m}^T\mathbf{C}^{-1}\mathbf{y}+\mathbf{m}_\mathbf{q}^T\mathbf{N}^{-1}\mathbf{q}}{1+\mathbf{m}^T\mathbf{C}^{-1}\mathbf{m}+\mathbf{m}_\mathbf{q}^T\mathbf{N}^{-1}\mathbf{m}_\mathbf{q}}\,, \quad (34a)$$

$$\delta\phi=\frac{(m'-\mathbf{c}_{\mathbf{s}s'}^T\mathbf{C}^{-1}\mathbf{m}-\mathbf{n}_{\mathbf{s}s'}^T\mathbf{N}^{-1}\mathbf{q})^2}{1+\mathbf{m}^T\mathbf{C}^{-1}\mathbf{m}+\mathbf{m}_\mathbf{q}^T\mathbf{N}^{-1}\mathbf{m}_\mathbf{q}} \quad (34b)$$

$$\beta=\frac{\mathbf{c}_{\mathbf{s}s'}^T\mathbf{C}^{-1}\mathbf{m}-\mathbf{n}_{\mathbf{s}s'}^T\mathbf{N}^{-1}\mathbf{q}-m'}{1+\mathbf{m}^T\mathbf{C}^{-1}\mathbf{m}+\mathbf{m}_\mathbf{q}^T\mathbf{N}^{-1}\mathbf{m}_\mathbf{q}} \quad (34c)$$

homBQUP:

$$\alpha=\frac{\mathbf{m}^T\mathbf{C}^{-1}\mathbf{y}+\mathbf{m}_\mathbf{q}^T\mathbf{N}^{-1}\mathbf{q}}{\mathbf{m}^T\mathbf{C}^{-1}\mathbf{m}+\mathbf{m}_\mathbf{q}^T\mathbf{N}^{-1}\mathbf{m}_\mathbf{q}}\,, \quad (35a)$$

$$\delta\phi=\frac{(m'-\mathbf{c}_{\mathbf{s}s'}^T\mathbf{C}^{-1}\mathbf{m}-\mathbf{n}_{\mathbf{s}s'}^T\mathbf{N}^{-1}\mathbf{q})^2}{\mathbf{m}^T\mathbf{C}^{-1}\mathbf{m}+\mathbf{m}_\mathbf{q}^T\mathbf{N}^{-1}\mathbf{m}_\mathbf{q}} \qquad \beta=0\,. \quad (35b)$$

The values from equations (33) are a quadratic extension of what is usually described as "best linear prediction" and is commonly called "collocation" in the geodetic literature. The other solutions offer some alternatives which are "robust" with respect to incorrect (prior) information about the signal mean $\mathbf{m}$. The homBQUP solution in particular given by equations (35) is a quadratic extension of what has been called "robust collocation" by Schaffrin (1985b). Note that the quadratic predictors differ from the linear ones in (a) the additional third tern in equation (28) and (b) in the additional terms appearing in the definitions of α, β and $\delta\phi$ (terms where the matrix $\mathbf{N}^{-1}$ appears).

4. Vector generalization

For the prediction of a vector of new signals $\mathbf{s}'$ we can predict each component s_i' independently and combine the solutions $\hat{s}_i'$ in a single matrix equation for $\hat{\mathbf{s}}'$. We also need to introduce

$$\mathbf{m}_{\mathbf{s}'}=E\{\mathbf{s}'\}\,, \quad (36)$$

$$\mathbf{C}_{\mathbf{s}'\mathbf{s}}=E\{(\mathbf{s}'-\mathbf{m}_{\mathbf{s}'})(\mathbf{s}-\mathbf{m}_\mathbf{s})^T\}\,, \quad (37)$$

$$\mathbf{C}_{\mathbf{s}'}=E\{(\mathbf{s}'-\mathbf{m}_{\mathbf{s}'})(\mathbf{s}'-\mathbf{m}_{\mathbf{s}'})^T\} \quad (38)$$

$$\mathbf{\Phi}_{\mathbf{s}'\mathbf{s}}=E\{vec[(\mathbf{s}-\mathbf{m}_\mathbf{s})(\mathbf{s}-\mathbf{m}_\mathbf{s})^T](\mathbf{s}'-\mathbf{m}_{\mathbf{s}'})^T\}\,, \quad (39)$$

$$\mathbf{N}_{\mathbf{s}'\mathbf{s}}=\mathbf{\Phi}\mathbf{C}^{-1}\mathbf{C}_{\mathbf{s}'\mathbf{s}}-\mathbf{\Phi}_{\mathbf{s}'\mathbf{s}} \quad (40)$$

(using the notation $\mathbf{m}_\mathbf{s}=E\{\mathbf{s}\}$ instead of $\mathbf{m}$ for the sake of distinction), as well as the bias vector $\boldsymbol{\beta}$ and the mean square error matrix $\mathbf{M}$

$$\boldsymbol{\beta}=E\{\hat{\mathbf{s}}'-\mathbf{s}'\}=E\{\boldsymbol{\varepsilon}\}\,, \qquad (\boldsymbol{\varepsilon}=\hat{\mathbf{s}}'-\mathbf{s}') \quad (41)$$

$$\mathbf{M}=E\{(\hat{\mathbf{s}}'-\mathbf{s}')(\hat{\mathbf{s}}'-\mathbf{s}')^T\}=E\{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\}\,. \quad (42)$$

With the above notation the general solution becomes

$$\hat{\mathbf{s}}'=\alpha\,\mathbf{m}_{\mathbf{s}'}+\mathbf{C}_{\mathbf{s}'\mathbf{s}}\mathbf{C}^{-1}(\mathbf{s}-\alpha\,\mathbf{m}_\mathbf{s})+\mathbf{N}_{\mathbf{s}'\mathbf{s}}\mathbf{N}^{-1}[\mathbf{q}(\mathbf{y})-\alpha\,\mathbf{m}_\mathbf{q}] \quad (43)$$

with α and $\mathbf{q}(\mathbf{y})$ as before and mean square error matrix

$$\mathbf{M}=\mathbf{C}_{\mathbf{s}'}-\mathbf{C}_{\mathbf{s}'\mathbf{s}}^T\mathbf{C}^{-1}\mathbf{C}_{\mathbf{s}'\mathbf{s}}-\mathbf{N}_{\mathbf{s}'\mathbf{s}}^T\mathbf{N}^{-1}\mathbf{N}_{\mathbf{s}'\mathbf{s}}+\delta\mathbf{M}\,. \quad (44)$$

For the particular prediction types we have

inhomBQP = *inhomBQUP*:

$$\alpha=1\,, \qquad \delta\mathbf{M}=\mathbf{0}\,, \qquad \boldsymbol{\beta}=\mathbf{0}\,. \quad (45)$$

homBQP:

$$\alpha=\frac{\mathbf{m}_\mathbf{s}^T\mathbf{C}^{-1}\mathbf{y}+\mathbf{m}_\mathbf{q}^T\mathbf{N}^{-1}\mathbf{q}}{1+\mathbf{m}_\mathbf{s}^T\mathbf{C}^{-1}\mathbf{m}_\mathbf{s}+\mathbf{m}_\mathbf{q}^T\mathbf{N}^{-1}\mathbf{m}_\mathbf{q}}\equiv\frac{\alpha_N}{\alpha_D} \quad (46a)$$

$$\boldsymbol{\beta}=\frac{1}{\alpha_D}(\mathbf{C}_{\mathbf{s}'\mathbf{s}}\mathbf{C}^{-1}\mathbf{m}_\mathbf{s}+\mathbf{N}_{\mathbf{s}'\mathbf{s}}\mathbf{N}^{-1}\mathbf{q}-\mathbf{m}_{\mathbf{s}'}) \quad (46b)$$

$$\delta\mathbf{M}=\alpha_D\boldsymbol{\beta}\boldsymbol{\beta}^T \quad (46c)$$

homBQUP:

$$\alpha=\frac{\mathbf{m}_\mathbf{s}^T\mathbf{C}^{-1}\mathbf{y}+\mathbf{m}_\mathbf{q}^T\mathbf{N}^{-1}\mathbf{q}}{\mathbf{m}_\mathbf{s}^T\mathbf{C}^{-1}\mathbf{m}_\mathbf{s}+\mathbf{m}_\mathbf{q}^T\mathbf{N}^{-1}\mathbf{m}_\mathbf{q}}\equiv\frac{\alpha_N}{\alpha_D}\,, \quad \boldsymbol{\beta}=\mathbf{0}\,. \quad (47a)$$

$$\delta\mathbf{M} = \frac{1}{\alpha_D}\mathbf{m}_0\mathbf{m}_0^T ,$$
$$\mathbf{m}_0 = \mathbf{C}_{\mathbf{s}'\mathbf{s}}\mathbf{C}^{-1}\mathbf{m}_\mathbf{s} + \mathbf{N}_{\mathbf{s}'\mathbf{s}}\mathbf{N}^{-1}\mathbf{q} - \mathbf{m}_{\mathbf{s}'} \tag{47b}$$

Recall that $\mathbf{C}=\mathbf{C}_\mathbf{s}+\mathbf{C}_\mathbf{v}$ and $\mathbf{N}=\boldsymbol{\Psi}-vec\mathbf{C}vec^T\mathbf{C}-\boldsymbol{\Phi}\mathbf{C}^{-1}\boldsymbol{\Phi}^T$ where also $\boldsymbol{\Phi}=\boldsymbol{\Phi}_\mathbf{s}+\boldsymbol{\Phi}_\mathbf{v}$ and $\boldsymbol{\Psi}=\boldsymbol{\Psi}_\mathbf{s}+\boldsymbol{\Psi}_\mathbf{v}$.

5. Extension to the general model

A more general model involves observables $\mathbf{a}$, which are non-linear functions $\mathbf{a}(\mathbf{s})$ of the signals $\mathbf{s}$, for which up to fourth order central moments are known. In this case the general non-linear collocation model with additive noise is

$$\mathbf{y}=\mathbf{a}(\mathbf{s})+\mathbf{v} \tag{48}$$

Furthermore, we want to predict a stochastic signal g which is also a nonlinear function $g(\mathbf{s})$ of the same fundamental signals $\mathbf{s}$.

The solution is straightforward if the marginal and joint central moments (up to 4th order) are known for the random parameters $\mathbf{a}$ and g: we merely need to replace $\mathbf{s}$ and s' and their moments, by $\mathbf{a}$ and g and their moments, respectively. In this way the problem reduces to one of "moment propagation", i.e. to the determination of $\mathbf{m}_\mathbf{a}$, $\mathbf{C}_\mathbf{a}$, $\boldsymbol{\Phi}_\mathbf{a}$, $\boldsymbol{\Psi}_\mathbf{a}$, m_g, $\mathbf{c}_{g\mathbf{a}}$ and $\boldsymbol{\Phi}_{g\mathbf{a}}$ from the known $\mathbf{m}=\mathbf{m}_\mathbf{s}$, $\mathbf{C}_\mathbf{s}$, $\boldsymbol{\Phi}_\mathbf{s}$, $\boldsymbol{\Psi}_\mathbf{s}$. If the joint distribution of $\mathbf{s}$ and s' were known, the joint distribution of $\mathbf{a}$ and g, and therefore the required moments, could be derived in principle though tedious numerical techniques need to be used. However this requirement departs from the spirit of a "fourth order theory" and an approximate propagation law should be used instead. Linearization of $\mathbf{a}(\mathbf{s})$ and $g(\mathbf{s})$ provides such an approximation, as in the case of linearized models and "second order theory". A better and hopefully sufficient approximation can be provided by propagation laws in quadratic approximations, which are derived from quadratic approximations to the nonlinear functions $\mathbf{a}(\mathbf{s})$ and $g(\mathbf{s})$. Using Taylor expansions up to the second order about the known (Taylor point) $\mathbf{m}=E\{\mathbf{s}\}$ we obtain

$$\mathbf{a}=\mathbf{a}(\mathbf{m})+\mathbf{A}\delta\mathbf{s}+\mathbf{H}_\mathbf{a}vec(\delta\mathbf{s}\delta\mathbf{s}^T) , \tag{49}$$

$$g=g(\mathbf{m})+\mathbf{g}^T\delta\mathbf{s}+\mathbf{h}_g^T vec(\delta\mathbf{s}\delta\mathbf{s}^T) , \tag{50}$$

where $\delta\mathbf{s}=\mathbf{s}-\mathbf{m}$ and

$$\mathbf{A}=\left[\frac{\partial\mathbf{a}}{\partial\mathbf{s}}\right]_\mathbf{m} , \qquad \mathbf{g}^T=\left[\frac{\partial g}{\partial\mathbf{s}}\right]_\mathbf{m} , \tag{51}$$

$$\mathbf{H}_{a_i}=\left[\frac{1}{2}\frac{\partial}{\partial\mathbf{s}}\left(\frac{\partial a_i}{\partial\mathbf{s}}\right)^T\right]_\mathbf{m} , \qquad \mathbf{H}_\mathbf{a}=\begin{bmatrix} vec^T\mathbf{H}_{a_1} \\ \vdots \\ vec^T\mathbf{H}_{a_n}\end{bmatrix}, \tag{52}$$

$$\mathbf{h}_g = vec\mathbf{H}_g , \qquad \mathbf{H}_g=\left[\frac{1}{2}\frac{\partial}{\partial\mathbf{s}}\left(\frac{\partial g}{\partial\mathbf{s}}\right)^T\right]_\mathbf{m} , \tag{53}$$

$$\mathbf{d}=-\mathbf{H}_\mathbf{a}vec\mathbf{C}_\mathbf{s} , \qquad c=-\mathbf{h}_g^T vec\mathbf{C}_\mathbf{s} . \tag{54}$$

Using these relations in the definitions of $\mathbf{m}_\mathbf{a}$, $\mathbf{C}_\mathbf{a}$, $\boldsymbol{\Phi}_\mathbf{a}$, $\boldsymbol{\Psi}_\mathbf{a}$, m_g, $\mathbf{c}_{\mathbf{a}g}$ and $\boldsymbol{\Phi}_{g\mathbf{a}}$, while making use of the definitions of $\mathbf{m}$, $\mathbf{C}_\mathbf{s}$, $\boldsymbol{\Phi}_\mathbf{s}$, $\boldsymbol{\Psi}_\mathbf{s}$, the following propagation laws are derived

$$\mathbf{m}_\mathbf{a}=\mathbf{a}(\mathbf{m})+\mathbf{H}_\mathbf{a}vec\mathbf{C}_\mathbf{s} \tag{55}$$

$$\mathbf{C}_\mathbf{a}=\mathbf{A}\mathbf{C}_\mathbf{s}\mathbf{A}^T+\mathbf{A}\boldsymbol{\Phi}_\mathbf{s}^T\mathbf{H}_\mathbf{a}^T+\mathbf{H}_\mathbf{a}\boldsymbol{\Phi}_\mathbf{s}\mathbf{A}^T+ +\mathbf{H}_\mathbf{a}(\boldsymbol{\Psi}_\mathbf{s}-vec\mathbf{C}_\mathbf{s}vec^T\mathbf{C}_\mathbf{s})\mathbf{H}_\mathbf{a}^T \tag{56}$$

$$\boldsymbol{\Phi}_\mathbf{a}=\mathbf{d}\overset{s}{\otimes}\mathbf{C}_\mathbf{a}+(\mathbf{A}\otimes\mathbf{A})[\boldsymbol{\Phi}_\mathbf{s}\mathbf{A}^T+\boldsymbol{\Psi}_\mathbf{s}\mathbf{H}_\mathbf{a}^T]+(\mathbf{A}\overset{s}{\otimes}\mathbf{H}_\mathbf{a})\boldsymbol{\Psi}_{vec}^\mathbf{s}+ +vec(\mathbf{C}_\mathbf{a}+\mathbf{d}\mathbf{d}^T)\mathbf{d}^T \tag{57}$$

$$\boldsymbol{\Psi}_\mathbf{a}=(\mathbf{A}\otimes\mathbf{A})\boldsymbol{\Psi}_\mathbf{s}(\mathbf{A}\otimes\mathbf{A})^T-\mathbf{d}^T\overset{s}{\otimes}(\mathbf{C}_\mathbf{a}\overset{s}{\otimes}\mathbf{d})- -(\mathbf{d}\otimes\mathbf{d})(\mathbf{d}\otimes\mathbf{d})^T+\boldsymbol{\Phi}_\mathbf{a}\overset{s}{\otimes}\mathbf{d}^T+\boldsymbol{\Phi}_\mathbf{a}^T\overset{s}{\otimes}\mathbf{d}- -vec\mathbf{C}_\mathbf{a}(\mathbf{d}\otimes\mathbf{d})^T-(\mathbf{d}\otimes\mathbf{d})vec^T\mathbf{C}_\mathbf{a} \tag{58}$$

$$m_g=g(\mathbf{m})+\mathbf{h}_g^T vec\mathbf{C}_\mathbf{s} \tag{59}$$

$$\mathbf{c}_{\mathbf{a}g}=\mathbf{A}\mathbf{C}_\mathbf{s}\mathbf{g}+\mathbf{H}_\mathbf{a}\boldsymbol{\Phi}_\mathbf{s}\mathbf{g}+\mathbf{A}\boldsymbol{\Phi}_\mathbf{s}^T\mathbf{h}_g+ +\mathbf{H}_\mathbf{a}(\boldsymbol{\Psi}_\mathbf{s}-vec\mathbf{C}_\mathbf{s}vec^T\mathbf{C}_\mathbf{s})\mathbf{h}_g \tag{60}$$

$$vec\boldsymbol{\Phi}_{g\mathbf{a}}=\mathbf{d}\overset{s}{\otimes}\mathbf{c}_{\mathbf{a}g}+(\mathbf{A}\otimes\mathbf{A})[\boldsymbol{\Phi}_\mathbf{s}\mathbf{g}+\boldsymbol{\Psi}_\mathbf{s}\mathbf{h}_g]+ +(\mathbf{A}\overset{s}{\otimes}\mathbf{H}_\mathbf{a})\boldsymbol{\Psi}_{vec}^\mathbf{s}\mathbf{g}+c\,vec(\mathbf{C}_\mathbf{a}+\mathbf{d}\mathbf{d}^T) \tag{61}$$

where

$$\boldsymbol{\Psi}_{vec}^\mathbf{s}=\begin{bmatrix} vec\boldsymbol{\Psi}_{11}^\mathbf{s} & \cdots & vec\boldsymbol{\Psi}_{1n}^\mathbf{s} \\ \vdots & \ddots & \vdots \\ vec\boldsymbol{\Psi}_{n1}^\mathbf{s} & \cdots & vec\boldsymbol{\Psi}_{nn}^\mathbf{s}\end{bmatrix} \tag{62}$$

while use of the "symmetric" Kronecker matrix product has been made, defined for any two matrices $\mathbf{A}$ and $\mathbf{B}$ by

$$\mathbf{A}\overset{s}{\otimes}\mathbf{B}=\mathbf{A}\otimes\mathbf{B}+\mathbf{B}\otimes\mathbf{A} \tag{63}$$

The above propagation laws can be interpreted either as quadratic approximations of the true propagation laws for the original non-linear model $\mathbf{y}=\mathbf{a}(\mathbf{s})+\mathbf{v}$, $g(\mathbf{s})$, or as exact propagation laws for the "quadraticized" model

$$\mathbf{y}=\mathbf{a}(\mathbf{m})+\mathbf{A}\delta\mathbf{s}+\mathbf{H}_\mathbf{a}vec(\delta\mathbf{s}\delta\mathbf{s}^T)+\mathbf{v} , \tag{64}$$

$$g=g(\mathbf{m})+\mathbf{g}^T\delta\mathbf{s}+\mathbf{h}_g^T vec(\delta\mathbf{s}\delta\mathbf{s}^T) . \tag{65}$$

6. Discussion

The question that remains open is whether the newly derived predictors are of any practical significance. The answer depends of course on the distribution of the signals involved i.e. in whether this distribution departs from the Gaussian one. If the distribution is Gaussian linear predictors cannot be improved and the additional terms in the quadratic predictors will have to vanish. Within the framework of the "fourth order" approach followed here, the distributions themselves play no role and the question of departure from the Gaussian distribution can be confined in two specific points: (a) the existence of non-vanishing third order central moments and (b) the departure of fourth order central moments from the corresponding values in the Gaussian case which are known functions of the lower order moments.

A practical way to evaluate the actual effect of such departures, when they exist, is to evaluate in each case the mean square errors ϕ_L and ϕ_Q from the linear predictors, respectively, and to look whether the relative improvement $\frac{\phi_L - \phi_Q}{\phi_L}$ is significant or not.

Turning to problems related to the gravity field, the question is directly related to the stochastic characteristics of the underlying stochastic process, namely the disturbing potential of the earth T. Sampling techniques can be used to derive the "third order central moment function" $\Phi_T(P,Q,R)$, which is a three-point function, and the "fourth order central moment function" $\Psi_T(P,Q,R,S)$, which is a four-point function, in a way similar to the one used for the estimation of the covariance (second order central moment) function. Under the usual assumptions of stationarity and isotropy the above functions become functions of spherical distances ψ between the points in question.

The function $\Phi_T(P,Q,R) = \Phi_T(\psi_{PQ}, \psi_{PR}, \psi_{QR})$, e.g., is approximated by fitting a function $\Phi_T(\psi_1, \psi_2, \psi_3)$ to the 3-dimensional histogram (3-parameter step function) derived by taking averages of all the products

$$[T(P) - m_T][T(Q) - m_T][T(R) - m_T]$$

where $m_T = E\{T\}$, corresponding to triads of points P, Q, R, such that the values ψ_{PQ}, ψ_{PR}, ψ_{QR} fall within a prescribed cube.

The results of such a sampling investigation have a local character, i.e. they cannot be taken to apply to different areas of the earth surface.

In conclusion laborious numerical investigations in various parts of the earth need to be undertaken before quadratic collocation can be evaluated as a worthwhile alternative to the usual (linear) one.

References

Dermanis, A. (1991): *A unified approach to linear estimation and prediction*. IUGG General Assembly, Vienna Aug. 1991.

Dermanis, A. and F. Sanso (1993): A Study of Nonlinear Estimation. General Meeting of the International Association of Geodesy, Beijing, China, August 8-14, 1993.

Grafarend, E. (1972): Nichtlineare Prädiktion. *ZfV*, 97, 6, 245-255.

Kubáček, L. (1985): Locally best quadratic estimators. *Math. Slovaca*, vol. 35, no. 4, 393-408.

Moritz, H. (1980): *Advanced Physical Geodesy*. H. Wichmann Verlag, Karlsruhe.

Rao, C.R. (1973): *Linear statistical inference and its applications*. 2nd edition. Wiley, New York.

Sanso, F. (1986): Statistical Methods in Physical Geodesy. In: H. Sünkel (ed.) "*Mathematical and Numerical Techniques in Physical Geodesy*", Lecture Notes in Earth Sciences, Vol. 7, 49-155. Springer-Verlag, Berlin.

Schaffrin, B. (1983): *Varianz-kovarianz-komponenten-Schätzung bei der Ausgleichung heterogener Wiederholungsmessungen*. DGK Reihe C, Heft Nr. 282.

Schaffrin, B. (1985a): *Das geodätische Datum mit stochastischer Vorinformation*. DGK Reihe C, Heft Nr. 313.

Schaffrin, B. (1985b): On robust collocation. In: F. Sanso (ed). Proceedings First Hotine-Marussi Symposium on Mathematical Geodesy (Roma, 1985), Milano, 1986, pp. 343-361.

Appendix A

The various types of quadratic predictors can more easily be derived in the "vectorized" form starting with equations (14), (19) and (20)

$$\hat{s}' = \gamma + \mathbf{a}^T \mathbf{z} \qquad \text{(inhom)},$$

$$\hat{s}' = \mathbf{a}^T \mathbf{z} \qquad \text{(hom)} \tag{A1}$$

$$\beta = \gamma + \mathbf{a}^T \mathbf{m}_\mathbf{z} - m' \qquad \text{(inhom)},$$

$$\beta = \mathbf{a}^T \mathbf{m}_\mathbf{z} - m' \qquad \text{(hom)} \tag{A2}$$

$$\phi = \beta^2 + \mathbf{a}^T \mathbf{C}_\mathbf{z} \mathbf{a} + \sigma_{s'}^2 - 2\mathbf{a}^T \mathbf{c}_{\mathbf{z}s'} . \tag{A3}$$

We note that

$$\frac{\partial \beta}{\partial \mathbf{a}} = \mathbf{m}_\mathbf{z}^T , \qquad \frac{\partial \beta}{\partial \gamma} = 1 . \tag{A4}$$

Best quadratic inhomogeneous prediction (inhomBQP):

Unconditional minimization of ϕ follows from the set of equations

$$\frac{1}{2}\left(\frac{\partial \phi}{\partial \mathbf{a}}\right)^T = \beta \left(\frac{\partial \beta}{\partial \mathbf{a}}\right)^T + \mathbf{C}_\mathbf{z}\mathbf{a} - \mathbf{c}_{\mathbf{z}s'} = \beta\, \mathbf{m}_\mathbf{z} + \mathbf{C}_\mathbf{z}\mathbf{a} - \mathbf{c}_{\mathbf{z}s'} = \mathbf{0} , \tag{A5}$$

$$\frac{1}{2}\frac{\partial \phi}{\partial \gamma} = \beta \frac{\partial \beta}{\partial \gamma} = \beta = \gamma + \mathbf{a}^T \mathbf{m}_\mathbf{z} - m' = 0 , \tag{A6}$$

with solution

$$\mathbf{a}=\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{c}_{\mathbf{z}s'}\,, \qquad \gamma=m'-\mathbf{c}_{\mathbf{z}s'}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{z}\,, \tag{A7}$$

$$\hat{s}'=m'+\mathbf{c}_{\mathbf{z}s'}^{T}\mathbf{C}_{\mathbf{z}}^{-1}(\mathbf{z}-\mathbf{m}_{\mathbf{z}}) \tag{A8}$$

$$\phi=\mathbf{a}^{T}\mathbf{C}_{\mathbf{z}}\mathbf{a}+\sigma_{s'}^{2}-2\mathbf{a}^{T}\mathbf{c}_{\mathbf{z}s'}=\sigma_{s'}^{2}-\mathbf{c}_{\mathbf{z}s'}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{c}_{\mathbf{z}s'}\,. \tag{A9}$$

Since the above prediction is already unbiased from (A6) it follows that it is identical with the Best Quadratic Unbiased inhomogeneous prediction (inhomBQUP = inhomBQP).

Best quadratic homogeneous prediction (homBQP):

Unconditional minimization of ϕ follows from

$$\frac{1}{2}\left(\frac{\partial\phi}{\partial\mathbf{a}}\right)^{T}=\beta\left(\frac{\partial\beta}{\partial\mathbf{a}}\right)^{T}+\mathbf{C}_{\mathbf{z}}\mathbf{a}-\mathbf{c}_{\mathbf{z}s'}=$$

$$=(\mathbf{m}_{\mathbf{z}}^{T}\mathbf{a}-m')\,\mathbf{m}_{\mathbf{z}}+\mathbf{C}_{\mathbf{z}}\mathbf{a}-\mathbf{c}_{\mathbf{z}s'}=\mathbf{0}\,, \tag{A10}$$

with solution

$$\mathbf{a}=(\mathbf{C}_{\mathbf{z}}+\mathbf{m}_{\mathbf{z}}\mathbf{m}_{\mathbf{z}}^{T})^{-1}(\mathbf{c}_{\mathbf{z}s'}+m'\mathbf{m}_{\mathbf{z}})=$$

$$=\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{c}_{\mathbf{z}s'}+\frac{m'-\mathbf{c}_{\mathbf{z}s'}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}}}{1+\mathbf{m}_{\mathbf{z}}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}}}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}}\,, \tag{A11}$$

$$\hat{s}'=\frac{\mathbf{m}_{\mathbf{z}}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{z}}{1+\mathbf{m}_{\mathbf{z}}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}}}m'+\mathbf{c}_{\mathbf{z}s'}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\left(\mathbf{z}-\frac{\mathbf{m}_{\mathbf{z}}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{z}}{1+\mathbf{m}_{\mathbf{z}}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}}}\mathbf{m}_{\mathbf{z}}\right), \tag{A12}$$

$$\beta=\frac{\mathbf{c}_{\mathbf{z}s'}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}}-m'}{1+\mathbf{m}_{\mathbf{z}}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}}}\,, \tag{A13}$$

$$\phi=\sigma_{s'}^{2}-\mathbf{c}_{\mathbf{z}s'}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{c}_{\mathbf{z}s'}+\frac{(m'-\mathbf{c}_{\mathbf{z}s'}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}})^{2}}{1+\mathbf{m}_{\mathbf{z}}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}}}\,. \tag{A14}$$

Best quadratic unbiased homogeneous prediction (homBQUP):

Minimization of ϕ under the condition $\beta=0$, requires the formulation of the Lagrangean function $L=\phi-2\lambda\beta$ and the solution is provided from the equations

$$\frac{1}{2}\left(\frac{\partial L}{\partial\mathbf{a}}\right)^{T}=\beta\left(\frac{\partial\beta}{\partial\mathbf{a}}\right)^{T}+\mathbf{C}_{\mathbf{z}}\mathbf{a}-\mathbf{c}_{\mathbf{z}s'}-\lambda\left(\frac{\partial\beta}{\partial\mathbf{a}}\right)^{T}=$$

$$=(\beta-\lambda)\mathbf{m}_{\mathbf{z}}+\mathbf{C}_{\mathbf{z}}\mathbf{a}-\mathbf{c}_{\mathbf{z}s'}=\mathbf{0}\,, \tag{A15}$$

$$\frac{1}{2}\frac{\partial L}{\partial\lambda}=-\beta=m'-\mathbf{a}^{T}\mathbf{m}_{\mathbf{z}}=0\,, \tag{A16}$$

Since $\beta=0$ (A15) gives

$$\mathbf{a}=\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{c}_{\mathbf{z}s'}+\lambda\,\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{z} \tag{A17}$$

which can be inserted into (A16) to provide

$$\lambda=\frac{m'-\mathbf{c}_{\mathbf{z}s'}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}}}{\mathbf{m}_{\mathbf{z}}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}}} \tag{A18}$$

with the above values of $\mathbf{a}$ and λ the prediction and its mean square error become

$$\hat{s}'=\frac{\mathbf{m}_{\mathbf{z}}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{z}}{\mathbf{m}_{\mathbf{z}}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}}}m'+\mathbf{c}_{\mathbf{z}s'}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\left(\mathbf{z}-\frac{\mathbf{m}_{\mathbf{z}}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{z}}{\mathbf{m}_{\mathbf{z}}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}}}\mathbf{m}_{\mathbf{z}}\right) \tag{A19}$$

$$\phi=\sigma_{s'}^{2}-\mathbf{c}_{\mathbf{z}s'}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{c}_{\mathbf{z}s'}+\frac{(m'-\mathbf{c}_{\mathbf{z}s'}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}})^{2}}{\mathbf{m}_{\mathbf{z}}^{T}\mathbf{C}_{\mathbf{z}}^{-1}\mathbf{m}_{\mathbf{z}}}\,. \tag{A20}$$