# Error Propagation of Capon's Minimum Variance Estimator

S. Toepfer[1]*, Y. Narita[2,3], D. Heyner[3] and U. Motschmann[1,4]

[1]Institut für Theoretische Physik, Technische Universität Braunschweig, Braunschweig, Germany, [2]Space Research Institute, Austrian Academy of Sciences, Graz, Austria, [3]Institut für Geophysik und Extraterrestrische Physik, Technische Universität Braunschweig, Braunschweig, Germany, [4]DLR Institute of Planetary Research, Berlin, Germany

The error propagation of Capon's minimum variance estimator resulting from measurement errors and position errors is derived within a linear approximation. It turns out, that Capon's estimator provides the same error propagation as the conventionally used least square fit method. The shape matrix which describes the location depence of the measurement positions is the key parameter for the error propagation, since the condition number of the shape matrix determines how the errors are amplified. Furthermore, the error resulting from a finite number of data samples is derived by regarding Capon's estimator as a special case of the maximum likelihood estimator.

Keywords: Capon's method, error propagation, least-squares method, maximum likelihood, condition number

## 1 INTRODUCTION

The reconstruction of model parameters from a given set of measurements is one of the most important tasks in geophysical and space science studies. The measurements are always affected by measurement errors which result in estimation errors for the wanted model parameters. The Capon method [1–3], also known as minimum variance distortionless response estimator (MVDR), is currently being considered as a robust inversion method for the analysis of planetary magnetic fields. In the past, the method has successfully been applied to the analysis of waves [1, 4] and therefore, specific attention has been paid to errors of the spectrum resulting from random perturbations in the amplitude and phase of sensor arrays [5]. Since Capon's method is based on the evaluation of statistically averaged data, the spectrum is also affected by errors resulting from a finite number of samples [6]. Within the estimation of the frequency-wavenumber spectrum, the difference between static structures and waves or their combination can be discerned through dispersion relation analysis [7–12]. Concerning the application of the Capon method for the analysis of planetary magnetic fields, the error propagation of Capon's estimator itself is of major importance for assessing the quality of the reconstructed model parameters. In general, two essential types of errors are expectable. On the one hand, the measured magnetic field data are affected by e.g., offsets, gains resulting from thermal variations and spacecraft magnetic disturbances [13]. On the other hand, the determination of the spacecraft's positions can be defective (measurement position errors), which results in a defective shape matrix. As a follow-up of the generalized derivation of Capon's method [3] and the error estimation of the power spectrum [6], in this work the effects of measurement errors, measurement position errors as well as finite sample sizes on Capon's estimator are considered.

## 2 CAPON'S METHOD

Before deducing the error propagation of Capon's method, the main ideas of the method are shortly revisited [2, 3]. Due to the complexity of several physical problems, the entire parametrization of experimental data is unrealizable. Thus, it is useful to decompose the measurements $\underline{B}$ into parametrized parts $\underline{H}\,\underline{g}$, where $\underline{g}$ contains the corresponding wanted model coefficients and the shape matrix $\underline{H}$ describes the distribution of the measurement positions with respect to the underlying model, non-parametrized parts $\underline{v}$ as well as measurement noise $\underline{n}$, so that

$$\underline{B} = \underline{H}\,\underline{g} + \underline{v} + \underline{n} \tag{1}$$

is valid. The measurement noise is assumed to be Gaussian with variance $\sigma_n$ and zero mean ($\langle \underline{n} \rangle = 0$). Since the shape matrix $\underline{H}$ is not invertible and the non-parametrized parts are unknown, the exact solution for the wanted model coefficients $\underline{g}$ is not available in general. Capon's method delivers an estimator $\underline{g}_C$ for the ideal solution $\underline{g}$. The method is based on the construction of a filter matrix $\underline{w}$, that minimizes the output power

$$\mathrm{tr}\left[ \underline{w}^{\dagger} \underline{M}\,\underline{w} \right] \tag{2}$$

with respect to the distortionless constraint

$$\underline{w}^{\dagger} \underline{H} = \underline{I} \tag{3}$$

where $\mathrm{tr}\left[ \underline{w}^{\dagger} \underline{M}\,\underline{w} \right]$ is the trace of the matrix $\underline{w}^{\dagger} \underline{M}\,\underline{w}$ and $\underline{I}$ is the identity matrix. The matrix $\underline{M} = \langle \underline{B} \circ \underline{B} \rangle$ denotes the data covariance matrix. Capon's estimator realizing the minimal output power results in

$$\underline{g}_C = \underline{w}^{\dagger} \langle \underline{B} \rangle = \left[ \underline{H}^{\dagger} \underline{M}^{-1} \underline{H} \right]^{-1} \underline{H}^{\dagger} \underline{M}^{-1} \langle \underline{B} \rangle. \tag{4}$$

The robustness of the method can be improved by the diagonal loading technique $\underline{M} \to \underline{M} + \sigma_d^2 \underline{I}$, where $\sigma_d$ is the so called diagonal loading parameter [3]. Thus, the estimator depends on the measurements and on the measurement positions so that measurement errors as well as measurement position errors transfer onto the estimator.

## 3 ERROR PROPAGATION

In the following, the error propagation of Capon's estimator is deduced. Since it is expectable that the errors are much smaller than the measurements themselves, the error propagation is derived by making use of a linearization. For the approximation of the matrix inversions that are necessary for calculating Capon's estimator, the Neumann series bears of essential meaning for which reason we discuss it in a seperate section.

### 3.1 Neumann Series

The Neumann series is a special case of the functional calculus for linear operators or matrices, respectively, [14] and enables the approximation of matrix inversions.

Let $\underline{T}$ denote a bounded matrix with norm $\|\underline{T}\| < 1$. Then,

$$\left[ \underline{I} - \underline{T} \right]^{-1} = \sum_{k=0}^{\infty} \underline{T}^k = \underline{I} + \underline{T} + \mathcal{O}\left( \|\underline{T}\|^2 \right) \tag{5}$$

where $\underline{I}$ is the identity matrix and $\underline{T}^k$ denotes the $k$'th power of the matrix $\underline{T}$. The demand $\|\underline{T}\| < 1$ guarantees the convergence of the series. Thus, the Neumann series can be understood as a generalization of the geometric series for linear operators.

As a direct consequence it follows that

$$\left[ \underline{I} + \underline{T} \right]^{-1} = \left[ \underline{I} - (-\underline{T}) \right]^{-1} = \sum_{k=0}^{\infty} (-1)^k \underline{T}^k = \underline{I} - \underline{T} + \mathcal{O}\left( \|\underline{T}\|^2 \right). \tag{6}$$

For the estimation of Capon's error propagation a generalized formulation of the Neumann series is required. To expand the inverse of the sum of a matrix $\underline{S}$ and a matrix $\underline{T}$, where it is assumed that $\underline{S}$ is invertible, the sum can be rewritten as

$$\underline{S} + \underline{T} = \underline{S}\left( \underline{I} + \underline{S}^{-1} \underline{T} \right). \tag{7}$$

If $\|\underline{S}^{-1}\underline{T}\| < 1$, the Neumann series can be applied to the sum $\underline{I} + \underline{S}^{-1}\underline{T}$ resulting in

$$\begin{aligned}
\left[ \underline{S} + \underline{T} \right]^{-1} &= \left[ \underline{S}\left( \underline{I} + \underline{S}^{-1}\underline{T} \right) \right]^{-1} = \left( \underline{I} + \underline{S}^{-1}\underline{T} \right)^{-1} \underline{S}^{-1} \\
&= \left[ \sum_{k=0}^{\infty} (-1)^k \left( \underline{S}^{-1}\underline{T} \right)^k \right] \underline{S}^{-1} \\
&= \underline{S}^{-1} - \underline{S}^{-1}\underline{T}\,\underline{S}^{-1} + \mathcal{O}\left( \|\underline{S}^{-1}\underline{T}\|^2 \right).
\end{aligned} \tag{8}$$

### 3.2 Measurement Errors

In the following, the error of Capon's estimator resulting from measurement errors is derived. The model for the (temporal) statistically averaged accurate data $\langle \underline{B} \rangle$ without measurement errors is given by

$$\langle \underline{B} \rangle = \underline{H}\,\underline{g} + \langle \underline{v} \rangle, \tag{9}$$

where $\underline{g}$ is the wanted coefficient vector, $\underline{H}$ denotes the shape matrix which describes the spatial dependence of the measurement positions with respect to the underlying model and $\langle \underline{v} \rangle$ denotes the (temporal) statistically averaged parts of the measurements that are not parametrized by the model $\underline{H}\,\underline{g}$ [3]. The corresponding accurate estimator resulting from Capon's method is given by

$$\underline{g}_C = \underline{w}^{\dagger} \langle \underline{B} \rangle \tag{10}$$

where

$$\underline{w}^{\dagger} = \left[ \underline{H}^{\dagger} \underline{M}^{-1} \underline{H} \right]^{-1} \underline{H}^{\dagger} \underline{M}^{-1} \tag{11}$$

denotes the accurate filter matrix composed of the shape matrix $\underline{H}$ and the data covariance matrix $\underline{M} = \langle \underline{B} \circ \underline{B} \rangle$ [3].

The perturbed measurements $\tilde{\underline{B}}$ can be rewritten as

$$\tilde{\underline{B}} = \underline{B} + \delta\underline{B}, \tag{12}$$

where $\delta\underline{B}$ denotes the measurement error. Because of the linearity of the averaging process, the perturbed averaged measurements $\langle \tilde{\underline{B}} \rangle$ are given by

$$\langle \underline{\tilde{B}} \rangle = \langle \underline{B} \rangle + \langle \delta \underline{B} \rangle \tag{13}$$

where $\langle \delta \underline{B} \rangle$ denotes the statistically averaged measurement error. The corresponding perturbed estimator results in

$$\underline{\tilde{g}}_C = \underline{\tilde{w}}^\dagger \langle \underline{\tilde{B}} \rangle \tag{14}$$

where

$$\underline{\tilde{w}}^\dagger = \left[ \underline{H}^\dagger \underline{\tilde{M}}^{-1} \underline{H} \right]^{-1} \underline{H}^\dagger \underline{\tilde{M}}^{-1} \tag{15}$$

denotes the perturbed filter matrix. Thus, the difference between the accurate and the perturbed estimator results in

$$\underline{g}_C - \underline{\tilde{g}}_C = \underline{w}^\dagger \langle \underline{B} \rangle - \underline{\tilde{w}}^\dagger \langle \underline{\tilde{B}} \rangle = \underline{w}^\dagger \langle \underline{B} \rangle - \underline{w}^\dagger \langle \delta \underline{B} \rangle - \underline{\tilde{w}}^\dagger \langle \underline{\tilde{B}} \rangle$$
$$= \left( \underline{w}^\dagger - \underline{\tilde{w}}^\dagger \right) \langle \underline{B} \rangle - \underline{w}^\dagger \langle \delta \underline{B} \rangle \tag{16}$$

Within the practical application only the perturbed measurements $\langle \underline{\tilde{B}} \rangle$ are known and thus, only the filter matrix $\underline{\tilde{w}}^\dagger$ is known. For the calculation of the difference between the two filter matrices, in the following a linearized approximation is applied. By means of **Eq. 12**, the unknown accurate data covariance matrix can be rewritten as

$$\underline{M} = \langle \underline{B} \circ \underline{B} \rangle = \left\langle \left( \underline{\tilde{B}} - \delta \underline{B} \right) \circ \left( \underline{\tilde{B}} - \delta \underline{B} \right) \right\rangle$$
$$= \underline{\tilde{M}} - 2 \langle \underline{\tilde{B}} \rangle \circ \langle \delta \underline{B} \rangle + \langle \delta \underline{B} \circ \delta \underline{B} \rangle \tag{17}$$

where $\underline{\tilde{M}} = \langle \underline{\tilde{B}} \circ \underline{\tilde{B}} \rangle$. We assume, that the measurement errors are much smaller than the measurements themselves, i.e., $|\delta \underline{B}| \ll |\underline{\tilde{B}}|$. This assumption is surely justified in the majority of applications. Considering for example the analysis of planetary magnetic fields, the measurement errors are smaller than 1 nT, so that $|\delta \underline{B}|/|\underline{\tilde{B}}| < 1\%$. Thus, in the following all terms being quadratic within the errors (e.g., $\langle \delta \underline{B} \circ \delta \underline{B} \rangle$) will be neglected, so that the data covariance matrix results in

$$\underline{M} = \underline{\tilde{M}} - 2 \langle \underline{\tilde{B}} \rangle \circ \langle \delta \underline{B} \rangle =: \underline{\tilde{M}} + \Delta \underline{M}. \tag{18}$$

In the case of $\left\| \underline{\tilde{M}}^{-1} \Delta \underline{M} \right\|_2 < 1$, where $\|.\|_2$ denotes the spectral norm, the application of the Neumann series (**Section 3.1**) delivers

$$\underline{M}^{-1} = \left( \underline{\tilde{M}} + \Delta \underline{M} \right)^{-1} \approx \underline{\tilde{M}}^{-1} - \underline{\tilde{M}}^{-1} \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \tag{19}$$

as well as

$$\underline{P} = \left[ \underline{H}^\dagger \underline{M}^{-1} \underline{H} \right]^{-1} \overset{(19)}{=} \left[ \underline{H}^\dagger \underline{\tilde{M}}^{-1} \underline{H} - \underline{H}^\dagger \underline{\tilde{M}}^{-1} \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \underline{H} \right]^{-1}$$
$$= \left[ \underline{H}^\dagger \underline{\tilde{M}}^{-1} \underline{H} \right]^{-1} + \left[ \underline{H}^\dagger \underline{\tilde{M}}^{-1} \underline{H} \right]^{-1} \underline{H}^\dagger \underline{\tilde{M}}^{-1} \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \underline{H} \left[ \underline{H}^\dagger \underline{\tilde{M}}^{-1} \underline{H} \right]^{-1}$$
$$= \underline{\tilde{P}} + \underline{\tilde{P}} \, \underline{H}^\dagger \underline{\tilde{M}}^{-1} \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \underline{H} \, \underline{\tilde{P}} \tag{20}$$

for the unknown coefficient matrix $\underline{P}$ [3]. Using

$$\underline{H}^\dagger \underline{M}^{-1} = \underline{H}^\dagger \underline{\tilde{M}}^{-1} - \underline{H}^\dagger \underline{\tilde{M}}^{-1} \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \tag{21}$$

it follows that

$$\underline{w}^\dagger = \underline{P} \, \underline{H}^\dagger \underline{M}^{-1}$$
$$= \left[ \underline{\tilde{P}} + \underline{\tilde{P}} \, \underline{H}^\dagger \underline{\tilde{M}}^{-1} \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \underline{H} \, \underline{\tilde{P}} \right] \cdot \left[ \underline{H}^\dagger \underline{\tilde{M}}^{-1} - \underline{H}^\dagger \underline{\tilde{M}}^{-1} \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \right]$$
$$= \underline{\tilde{P}} \, \underline{H}^\dagger \underline{\tilde{M}}^{-1} - \underline{\tilde{P}} \, \underline{H}^\dagger \underline{\tilde{M}}^{-1} \Delta \underline{M} \, \underline{\tilde{M}}^{-1} +$$
$$\underline{\tilde{P}} \, \underline{H}^\dagger \underline{\tilde{M}}^{-1} \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \underline{H} \, \underline{\tilde{P}} \, \underline{H}^\dagger \underline{\tilde{M}}^{-1} + \mathcal{O}\left( \left\| \Delta \underline{\tilde{M}} \right\|^2 \right)$$
$$= \underline{\tilde{w}}^\dagger - \underline{\tilde{w}}^\dagger \Delta \underline{M} \, \underline{\tilde{M}}^{-1} + \underline{\tilde{w}}^\dagger \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \underline{H} \, \underline{\tilde{w}}^\dagger$$
$$= \underline{\tilde{w}}^\dagger + \underline{\tilde{w}}^\dagger \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \left( \underline{H} \, \underline{\tilde{w}}^\dagger - \underline{I} \right), \tag{22}$$

where $\underline{I}$ again denotes the identity matrix. Thus, the difference between the filter matrices can be approximated by

$$\underline{w}^\dagger - \underline{\tilde{w}}^\dagger = \underline{\tilde{w}}^\dagger \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \left( \underline{H} \, \underline{\tilde{w}}^\dagger - \underline{I} \right). \tag{23}$$

Inserting this approximation into **Eq. 16** delivers

$$\underline{g}_C - \underline{\tilde{g}}_C = \underline{\tilde{w}}^\dagger \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \left( \underline{H} \, \underline{\tilde{w}}^\dagger - \underline{I} \right) \langle \underline{B} \rangle - \underline{w}^\dagger \langle \delta \underline{B} \rangle \tag{24}$$

where

$$\underline{w}^\dagger \langle \delta \underline{B} \rangle = \left[ \underline{\tilde{w}}^\dagger + \underline{\tilde{w}}^\dagger \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \left( \underline{H} \, \underline{\tilde{w}}^\dagger - \underline{I} \right) \right] \langle \delta \underline{B} \rangle = \underline{\tilde{w}}^\dagger \langle \delta \underline{B} \rangle \tag{25}$$

within the linear approximation. Further transformation for estimating the relative error of the estimator delivers

$$\underline{g}_C - \underline{\tilde{g}}_C = \underline{\tilde{w}}^\dagger \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \left( \underline{H} \, \underline{\tilde{w}}^\dagger - \underline{I} \right) \langle \underline{B} \rangle - \underline{\tilde{w}}^\dagger \langle \delta \underline{B} \rangle$$
$$= \underline{\tilde{w}}^\dagger \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \left( \underline{H} \, \underline{\tilde{w}}^\dagger \langle \underline{B} \rangle - \langle \underline{B} \rangle \right) - \underline{\tilde{w}}^\dagger \langle \delta \underline{B} \rangle$$
$$= \underline{\tilde{w}}^\dagger \Delta \underline{M} \, \underline{\tilde{M}}^{-1} \left( \underline{H} \, \underline{\tilde{g}}_C - \langle \underline{B} \rangle \right) - \underline{\tilde{w}}^\dagger \langle \delta \underline{B} \rangle$$
$$= \underline{\tilde{w}}^\dagger \Delta \underline{M} \, \underline{\tilde{M}}^{-1/2} \, \underline{\tilde{M}}^{-1/2} \left( \underline{H} \, \underline{\tilde{g}}_C - \langle \underline{B} \rangle \right) - \underline{\tilde{w}}^\dagger \langle \delta \underline{B} \rangle. \tag{26}$$

Making use of the Cauchy-Schwarz inequality yields

$$\left| \underline{g}_C - \underline{\tilde{g}}_C \right|^2 \leq \left\| \underline{\tilde{w}}^\dagger \Delta \underline{M} \, \underline{\tilde{M}}^{-1/2} \right\|_2^2 \cdot \left\| \underline{H} \, \underline{\tilde{g}}_C - \langle \underline{B} \rangle \right\|_{\underline{\tilde{M}}^{-1}}^2 + \left| \underline{\tilde{w}}^\dagger \langle \delta \underline{B} \rangle \right|^2, \tag{27}$$

where $\left\| \underline{H} \, \underline{\tilde{g}}_C - \langle \underline{B} \rangle \right\|_{\underline{\tilde{M}}^{-1}}^2 = \left| \underline{\tilde{M}}^{-1/2} \left( \underline{\tilde{H}} \, \underline{\tilde{g}}_C - \langle \underline{\tilde{B}} \rangle \right) \right|^2$. For the calculation of Capon's estimator, the weighted difference $\left| \underline{\tilde{M}}^{-1/2} \left( \underline{\tilde{H}} \, \underline{\tilde{g}}_C - \langle \underline{\tilde{B}} \rangle \right) \right|^2$ is minimized with respect to the unknown set of model parameters $\underline{\tilde{g}}$, resulting in $\left| \underline{\tilde{M}}^{-1/2} \left( \underline{\tilde{H}} \, \underline{\tilde{g}}_C - \langle \underline{\tilde{B}} \rangle \right) \right|^2 \sim 10^{-7}$ [3]. A discussion about the calculation of the matrix $\underline{M}^{-1/2}$ is given within the Appendix. Assuming that the weighted model mismatches are negligibly small compared to the measurement errors, the deviation can be estimated upwards via

$$\left| \underline{g}_C - \underline{\tilde{g}}_C \right|^2 \leq \left| \underline{\tilde{w}}^\dagger \langle \delta \underline{B} \rangle \right|^2 \leq \left\| \underline{\tilde{w}}^\dagger \right\|_2^2 \cdot \left| \langle \delta \underline{B} \rangle \right|^2, \tag{28}$$

or equivalently
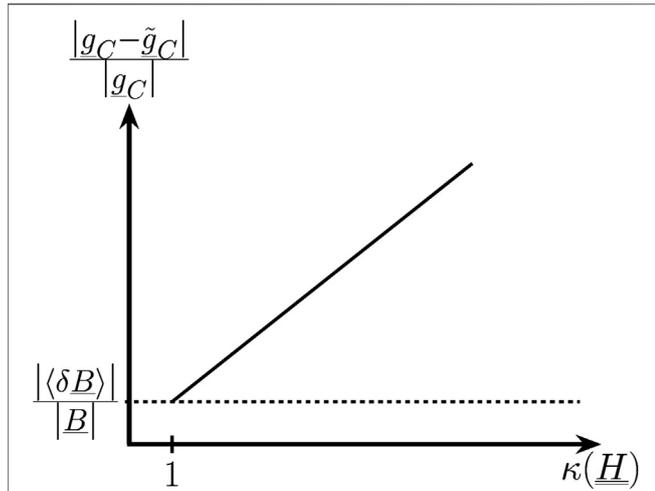
**FIGURE 1** | Sketch of the upper bound for the relative estimation error resulting from measurement errors with respect to the condition number $\kappa(\underline{H})$ of the shape matrix.

$$\frac{|\underline{g}_C - \underline{\tilde{g}}_C|}{|\underline{g}_C|} \le \frac{\left\|\underline{\underline{\tilde{w}}}^\dagger\right\|_2 \cdot |\langle \delta \underline{B} \rangle|}{|\underline{g}_C|} = \frac{\left\|\underline{\underline{\tilde{w}}}^\dagger\right\|_2 \cdot |\langle \delta \underline{B} \rangle|}{|\underline{g}_C|} \frac{|\langle \underline{B} \rangle|}{|\langle \underline{B} \rangle|} \quad (29)$$

for the relative error. When the measurements are adequately described by the underlying model, i.e., $\langle \underline{B} \rangle \approx \underline{H} \, \underline{g}_C$, it follows that

$$\frac{|\underline{g}_C - \underline{\tilde{g}}_C|}{|\underline{g}_C|} \le \left\|\underline{\underline{\tilde{w}}}^\dagger\right\|_2 \cdot \|\underline{H}\|_2 \frac{|\langle \delta \underline{B} \rangle|}{|\langle \underline{B} \rangle|}. \quad (30)$$

Comparing this expression with the relative error of the least square fit estimator [15].

$$\frac{|\underline{g}_L - \underline{\tilde{g}}_L|}{|\underline{g}_L|} \le \|\underline{\underline{H}}^+\|_2 \cdot \|\underline{H}\|_2 \frac{|\langle \delta \underline{B} \rangle|}{|\langle \underline{B} \rangle|}, \quad (31)$$

where $\underline{H}^+$ denotes the pseudoinverse of the shape matrix $\underline{H}$, shows that the error propagation of Capon's estimator follows the structure of the error propagation of the least square fit estimator. Since Capon's method is based on the evaluation of averaged data, it can be seen that Gaussian errors with a vanishing mean value (i.e., $\langle \delta \underline{B} \rangle = 0$) do not influence the estimator. Making use of the distortionless constraint [3].

$$\underline{\tilde{w}}^\dagger \underline{H} = \underline{I} \quad (32)$$

delivers

$$\begin{aligned}
\left\|\underline{\underline{\tilde{w}}}^\dagger\right\|_2 &= \max_{\underline{B}} \frac{|\underline{\tilde{w}}^\dagger \underline{B}|}{|\underline{B}|} = \max_{\underline{H}\underline{\tilde{g}}} \frac{|\underline{\tilde{w}}^\dagger \underline{H} \underline{\tilde{g}}|}{|\underline{H} \underline{\tilde{g}}|} \\
&= \max_{\underline{\tilde{g}}} \frac{|\underline{\tilde{g}}|}{|\underline{H} \underline{\tilde{g}}|} = \left( \min_{\underline{\tilde{g}}} \frac{|\underline{H} \underline{\tilde{g}}|}{|\underline{\tilde{g}}|} \right)^{-1} \\
&= \frac{1}{\Sigma_{\min}} = \|\underline{\underline{H}}^+\|_2,
\end{aligned} \quad (33)$$

where $\Sigma_{\min}$ denotes the smallest singular value of the shape matrix $\underline{H}$ [15], i.e. $1/\Sigma_{\min}$ describes the largest singular value of $\underline{H}^+$. Using the definition of the condition number

$$\kappa(\underline{H}) = \frac{\Sigma_{\max}}{\Sigma_{\min}} = \left\|\underline{\underline{H}}^+\right\|_2 \cdot \left\|\underline{\underline{H}}\right\|_2 = \left\|\underline{\underline{w}}^\dagger\right\|_2 \cdot \left\|\underline{\underline{H}}\right\|_2 \ge 1, \quad (34)$$

where $\Sigma_{\max}$ denotes the largest singular value of the shape matrix $\underline{H}$, yields

$$\frac{|\underline{g}_C - \underline{\tilde{g}}_C|}{|\underline{g}_C|} \le \kappa(\underline{H}) \frac{|\langle \delta \underline{B} \rangle|}{|\langle \underline{B} \rangle|}. \quad (35)$$

Thus, Capon's estimator propagates the measurement errors in the same way as the least square fit estimator. This mathematical property can be interpreted as follows: These two methods differ in the filter matrices $\underline{H}^+$ and $\underline{w}^\dagger$ which weight or eliminate parts of the data in different ways. Since these matrices are applied to the same set of measurements, which is characterized by a given measurement error, the different weighting of the data or the elimination of subsets does not reduce the errors, since the weighted data originate from the same ensemble. The upper bound of the relative estimation error is qualitatively sketched in **Figure 1**.

Furthermore, it should be noted that the condition number of the shape matrix determines how the measurement errors are amplified. Thereby, the condition number depends on the underlying model and the measurement positions, whereas the measurement error is produced by the sensor. For a given underlying model, the condition number solely depends on the measurement positions. Thus, the estimation errors can be reduced by analyzing measurements from suitable data points. Considering the analysis of Mercury's magnetic field, the underlying model describes the geometry of the magnetic field, for example the internal dipole and quadrupole field [2, 3]. For the estimation of the corresponding Gauss coefficients [16, 17], the data points have to cover the geometry of the field properly. For example, the superposition of Mercury's internal dipole field with the quadrupole field can equivalently be described as a northward shifted dipole field. When only measurement positions in the northern hemisphere are available, the condition number of the shape matrix increases, resulting in large estimation errors so that the internal field can be misinterpreted as a strong dipole field. The analysis of measurement points covering the southern and the northern hemisphere symmetrically, like that of the BepiColombo mission, decreases the condition number and enables a more detailed characterization of the geometry of Mercury's internal magnetic field [18]. The smaller the condition number of the shape matrix becomes, the better the geometry of the field is covered by the data points [18].

Before discussing the errors resulting from the perturbed determination of the measurement positions, let us make a general comment about the error of Capon's estimator resulting from measurement errors. The perturbed data can be rewritten in the form

$$\langle \underline{\tilde{B}} \rangle = \langle \underline{B} \rangle + \langle \delta \underline{B} \rangle = \underline{H} \, \underline{g} + \langle \underline{v} \rangle + \langle \delta \underline{B} \rangle = \underline{H} \, \underline{g} + \langle \underline{\tilde{v}} \rangle, \quad (36)$$

where $\langle \tilde{\underline{v}} \rangle = \langle \underline{v} \rangle + \langle \delta \underline{B} \rangle$. Since the filter matrix $\underline{\tilde{w}}^{\dagger}$ eliminates the non-parametrized parts from the total measured field [3], one might suppose that the measurement errors do not influence the estimation since these errors can as well be interpreted as non-parametrized parts. From the elimination of the non-parametrized parts

$$0 = \underline{\tilde{w}}^{\dagger} \langle \tilde{\underline{v}} \rangle = \underline{\tilde{w}}^{\dagger} \langle \underline{v} \rangle + \underline{\tilde{w}}^{\dagger} \langle \delta \underline{B} \rangle \tag{37}$$

it *does not* follow that $\underline{\tilde{w}}^{\dagger} \langle \delta \underline{B} \rangle = 0$. Since this term dominates the error of the estimator (cf. **Eq. 28**), the measurement errors are not eliminated by the filter matrix in general.

## 3.3 Measurement Position Errors

The model for the correctly determined measurement positions is given by

$$\langle \underline{B} \rangle = \underline{H} \, \underline{g} + \langle \underline{v} \rangle. \tag{38}$$

The corresponding accurate estimator results in

$$\underline{g}_C = \underline{w}^{\dagger} \langle \underline{B} \rangle, \tag{39}$$

where

$$\underline{w}^{\dagger} = \left[ \underline{H}^{\dagger} \underline{M}^{-1} \underline{H} \right]^{-1} \underline{H}^{\dagger} \underline{M}^{-1}. \tag{40}$$

The perturbed determination of the sensor's positions transfers onto a perturbed shape matrix $\underline{\tilde{H}}$. Thus, the underlying model $\underline{H} \, \underline{g}$ and the non-parametrized parts $\langle \underline{v} \rangle$ are perturbed so that the noisy model is given by

$$\langle \underline{B} \rangle = \underline{\tilde{H}} \, \tilde{\underline{g}} + \langle \tilde{\underline{v}} \rangle \tag{41}$$

and the corresponding perturbed estimator results in

$$\tilde{\underline{g}}_C = \underline{\tilde{w}}^{\dagger} \langle \underline{B} \rangle, \tag{42}$$

where

$$\underline{\tilde{w}}^{\dagger} = \left[ \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} \underline{\tilde{H}} \right]^{-1} \underline{\tilde{H}}^{\dagger} \underline{M}^{-1}. \tag{43}$$

The deviation between the accurate and the perturbed estimator results in

$$\underline{g}_C - \tilde{\underline{g}}_C = \left( \underline{w}^{\dagger} - \underline{\tilde{w}}^{\dagger} \right) \langle \underline{B} \rangle. \tag{44}$$

As discussed above, only the perturbed filter matrix $\underline{\tilde{H}}$ is known. Within a linear approximation, the unknown matrix $\underline{H}$ can be rewritten as $\underline{H} = \underline{\tilde{H}} + \Delta \underline{H}$, where $\left\| \Delta \underline{H} \right\|_2 \ll \left\| \underline{\tilde{H}} \right\|_2$. For example, consider the shape matrix

$$\underline{H} = \begin{bmatrix} 1 & \left( \dfrac{R_M}{r_1} \right)^3 \\[2ex] 1 & \left( \dfrac{R_M}{r_2} \right)^3 \end{bmatrix} \tag{45}$$

which describes the magnetic field of a current sheet superposed with Mercury's internal dipolar field [6], where $R_M$ indicates the planetary radius of Mercury and $r_1$ and $r_2$ are the measurement positions. The perturbed filter matrix is given by

$$\underline{\tilde{H}} = \begin{bmatrix} 1 & \left( \dfrac{R_M}{r_1 + \Delta r_1} \right)^3 \\[2ex] 1 & \left( \dfrac{R_M}{r_2 + \Delta r_2} \right)^3 \end{bmatrix} = \begin{bmatrix} 1 & \dfrac{R_M^3}{r_1^3 \left( 1 + \Delta r_1 / r_1 \right)^3} \\[2ex] 1 & \dfrac{R_M^3}{r_2^3 \left( 1 + \Delta r_2 / r_2 \right)^3} \end{bmatrix}. \tag{46}$$

Assuming that $\Delta r_i \ll r_i$, for $i = 1, 2$, the perturbed matrix can be rewritten as

$$\underline{\tilde{H}} = \begin{bmatrix} 1 & \left( \dfrac{R_M}{r_1} \right)^3 \\[2ex] 1 & \left( \dfrac{R_M}{r_2} \right)^3 \end{bmatrix} + \begin{bmatrix} 0 & -3 \dfrac{R_M^3}{r_1^3} \dfrac{\Delta r_1}{r_1} \\[2ex] 0 & -3 \dfrac{R_M^3}{r_2^3} \dfrac{\Delta r_2}{r_2} \end{bmatrix} = \underline{H} - \Delta \underline{H} \tag{47}$$

by making use of a Taylor series expansion. It should be noted that the shape matrix can be linearized for any underlying model by performing a Taylor series expansion for the functions within the model. Linearization of the error terms for the estimation of the unknown accurate filter matrix delivers

$$\begin{aligned}
\underline{w}^{\dagger} &\overset{(40)}{=} \left[ \underline{H}^{\dagger} \underline{M}^{-1} \underline{H} \right]^{-1} \underline{H}^{\dagger} \underline{M}^{-1} \\
&= \left[ \left( \underline{\tilde{H}}^{\dagger} + \Delta \underline{H}^{\dagger} \right) \underline{M}^{-1} \left( \underline{\tilde{H}} + \Delta \underline{H} \right) \right]^{-1} \left( \underline{\tilde{H}}^{\dagger} + \Delta \underline{H}^{\dagger} \right) \underline{M}^{-1} \\
&= \left[ \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} \left( \underline{\tilde{H}} + \Delta \underline{H} \right) + \Delta \underline{H}^{\dagger} \underline{M}^{-1} \left( \underline{\tilde{H}} + \Delta \underline{H} \right) \right]^{-1} \left( \underline{\tilde{H}}^{\dagger} + \Delta \underline{H}^{\dagger} \right) \underline{M}^{-1} \\
&= \left[ \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} \underline{\tilde{H}} + \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} \Delta \underline{H} + \Delta \underline{H}^{\dagger} \underline{M}^{-1} \underline{\tilde{H}} \right]^{-1} \left( \underline{\tilde{H}}^{\dagger} + \Delta \underline{H}^{\dagger} \right) \underline{M}^{-1}.
\end{aligned} \tag{48}$$

Making use of the Neumann series results in

$$\begin{aligned}
&\left[ \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} \underline{\tilde{H}} + \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} \Delta \underline{H} + \Delta \underline{H}^{\dagger} \underline{M}^{-1} \underline{\tilde{H}} \right]^{-1} \\
&= \left[ \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} \underline{\tilde{H}} \right]^{-1} - \left[ \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} \underline{\tilde{H}} \right]^{-1} \cdot \\
&\quad \left[ \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} \Delta \underline{H} + \Delta \underline{H}^{\dagger} \underline{M}^{-1} \underline{\tilde{H}} \right] \left[ \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} \underline{\tilde{H}} \right]^{-1} \\
&= \underline{\tilde{P}} - \underline{\tilde{P}} \, \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} \Delta \underline{H} \, \underline{\tilde{P}} - \underline{\tilde{P}} \, \Delta \underline{H}^{\dagger} \underline{M}^{-1} \underline{\tilde{H}} \, \underline{\tilde{P}},
\end{aligned} \tag{49}$$

so that

$$\begin{aligned}
\underline{w}^{\dagger} &= \underline{\tilde{P}} \, \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} - \underline{\tilde{P}} \, \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} \Delta \underline{H} \, \underline{\tilde{P}} \, \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} \\
&\quad - \underline{\tilde{P}} \, \Delta \underline{H}^{\dagger} \underline{M}^{-1} \underline{\tilde{H}} \, \underline{\tilde{P}} \, \underline{\tilde{H}}^{\dagger} \underline{M}^{-1} + \underline{\tilde{P}} \, \Delta \underline{H}^{\dagger} \underline{M}^{-1} \\
&= \underline{\tilde{w}}^{\dagger} - \underline{\tilde{w}}^{\dagger} \Delta \underline{H} \, \underline{\tilde{w}}^{\dagger} - \underline{\tilde{P}} \, \Delta \underline{H}^{\dagger} \underline{M}^{-1} \underline{\tilde{H}} \, \underline{\tilde{w}}^{\dagger} + \underline{\tilde{P}} \, \Delta \underline{H}^{\dagger} \underline{M}^{-1} \\
&= \underline{\tilde{w}}^{\dagger} - \underline{\tilde{w}}^{\dagger} \Delta \underline{H} \, \underline{\tilde{w}}^{\dagger} - \underline{\tilde{P}} \, \Delta \underline{H}^{\dagger} \underline{M}^{-1} \left( \underline{\tilde{H}} \, \underline{\tilde{w}}^{\dagger} - \underline{I} \right),
\end{aligned} \tag{50}$$

or equivalently

$$\underline{w}^{\dagger} - \underline{\tilde{w}}^{\dagger} = -\underline{\tilde{w}}^{\dagger} \Delta \underline{H} \, \underline{\tilde{w}}^{\dagger} - \underline{\tilde{P}} \, \Delta \underline{H}^{\dagger} \underline{M}^{-1} \left( \underline{\tilde{H}} \, \underline{\tilde{w}}^{\dagger} - \underline{I} \right). \tag{51}$$

Thus, the deviation of the estimator can be approximated via

$$\underline{g}_C - \tilde{\underline{g}}_C = -\underline{\tilde{w}}^{\dagger} \Delta \underline{H} \, \underline{\tilde{w}}^{\dagger} \langle \underline{B} \rangle - \underline{\tilde{P}} \, \Delta \underline{H}^{\dagger} \underline{M}^{-1} \left( \underline{\tilde{H}} \, \underline{\tilde{w}}^{\dagger} \langle \underline{B} \rangle - \langle \underline{B} \rangle \right). \tag{52}$$

By $\underline{\tilde{w}}^{\dagger} \langle \underline{B} \rangle = \tilde{\underline{g}}_C$ it follows that

$$\underline{g}_C - \tilde{\underline{g}}_C = -\underline{\tilde{w}}^{\dagger} \Delta \underline{H} \, \tilde{\underline{g}}_C - \underline{\tilde{P}} \, \Delta \underline{H}^{\dagger} \underline{M}^{-1} \left( \underline{\tilde{H}} \, \tilde{\underline{g}}_C - \langle \underline{B} \rangle \right), \tag{53}$$

**FIGURE 2 |** Sketch of the upper bound for the relative estimation error resulting from measurement position errors with respect to the condition number $\kappa(\underline{\tilde{H}})$ of the shape matrix.
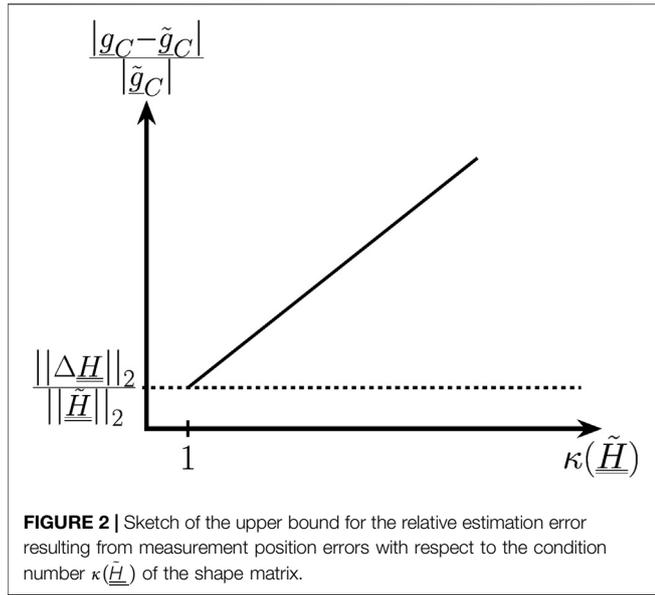
and therefore

$$
\left|\underline{g}_C - \underline{\tilde{g}}_C\right|^2 \leq \left\|\underline{\tilde{w}}^\dagger\right\|_2^2 \cdot \left\|\Delta\underline{\underline{H}}\right\|_2^2 \cdot \left|\underline{\tilde{g}}_C\right|^2 + \left\|\underline{\tilde{P}}\,\Delta\underline{\underline{H}}^\dagger\underline{\underline{M}}^{-1/2}\right\|_2^2 \\
\cdot \left|\underline{\underline{M}}^{-1/2}\left(\underline{\tilde{\underline{H}}}\,\underline{\tilde{g}}_C - \langle\underline{B}\rangle\right)\right|^2.
$$

(54)

Within the application of Capon's method to Mercury's magnetic field analysis, the first summand on the right hand side of **Eq. 54** is of the order of about $10\,\mathrm{nT}^2$, whereas $\left\|\underline{\tilde{P}}\,\Delta\underline{\underline{H}}^\dagger\underline{\underline{M}}^{-1/2}\right\|_2^2 \cdot \left|\underline{\underline{M}}^{-1/2}\left(\underline{\tilde{\underline{H}}}\,\underline{\tilde{g}}_C - \langle\underline{B}\rangle\right)\right|^2 \sim 500 \cdot 10^{-7}\,\mathrm{nT}^2$ and thus, the weighted model mismatches are negligibly small compared to the measurement errors, so that the deviation can be estimated upwards via

$$
\left|\underline{g}_C - \underline{\tilde{g}}_C\right|^2 \leq \left\|\underline{\tilde{w}}^\dagger\right\|_2^2 \cdot \left\|\Delta\underline{\underline{H}}\right\|_2^2 \cdot \left|\underline{\tilde{g}}_C\right|^2.
$$

(55)

The relative error results in

$$
\frac{\left|\underline{g}_C - \underline{\tilde{g}}_C\right|}{\left|\underline{\tilde{g}}_C\right|} \leq \left\|\underline{\tilde{w}}^\dagger\right\|_2 \cdot \left\|\underline{\tilde{\underline{H}}}\right\|_2 \cdot \frac{\left\|\Delta\underline{\underline{H}}\right\|_2}{\left\|\underline{\tilde{\underline{H}}}\right\|_2} = \kappa\left(\underline{\tilde{\underline{H}}}\right) \cdot \frac{\left\|\Delta\underline{\underline{H}}\right\|_2}{\left\|\underline{\tilde{\underline{H}}}\right\|_2},
$$

(56)

where $\kappa\left(\underline{\tilde{\underline{H}}}\right)$ denotes the condition number of the perturbed shape matrix $\underline{\tilde{\underline{H}}}$. The upper bound for the relative error with respect to the condition number is qualitatively sketched in **Figure 2**. Since $\left\|\underline{\tilde{w}}^\dagger\right\|_2 = \left\|\underline{\tilde{\underline{H}}}^+\right\|_2$ (cf. **Eq. 33**), the error propagation of Capon's estimator again equals the error propagation of the least square fit estimator [15].

$$
\frac{\left|\underline{g}_L - \underline{\tilde{g}}_L\right|}{\left|\underline{\tilde{g}}_L\right|} \leq \left\|\underline{\tilde{\underline{H}}}^+\right\|_2 \cdot \left\|\underline{\tilde{\underline{H}}}\right\|_2 \cdot \frac{\left\|\Delta\underline{\underline{H}}\right\|_2}{\left\|\underline{\tilde{\underline{H}}}\right\|_2} = \kappa\left(\underline{\tilde{\underline{H}}}\right) \cdot \frac{\left\|\Delta\underline{\underline{H}}\right\|_2}{\left\|\underline{\tilde{\underline{H}}}\right\|_2}.
$$

(57)

## 3.4 Measurement Errors and Measurement Position Errors

Within the former sections the influences of measurement errors and measurement position errors have been discussed separately. Within the practical application it is expectable that both errors occur simultanously. The relative error resulting from the noisy measurements and the perturbed measurement positions is given by the quadratic sum of the two cases discussed above

$$
\frac{\left|\underline{g}_C - \underline{\tilde{g}}_C\right|^2}{\left|\underline{\tilde{g}}_C\right|^2} \leq \kappa^2\left(\underline{\tilde{\underline{H}}}\right)\left(\frac{\left\|\Delta\underline{\underline{H}}\right\|_2^2}{\left\|\underline{\tilde{\underline{H}}}\right\|_2^2} + \frac{\left|\langle\delta\underline{B}\rangle\right|^2}{\left|\langle\underline{\tilde{B}}\rangle\right|^2}\right).
$$

(58)

This can be derived as follows:
The accurate model is given by

$$
\langle\underline{B}\rangle = \underline{\underline{H}}\,\underline{g} + \langle\underline{v}\rangle,
$$

(59)

so that the corresponding accurate estimator results in

$$
\underline{g}_C = \underline{w}^\dagger\langle\underline{B}\rangle
$$

(60)

where

$$
\underline{w}^\dagger = [\underline{\underline{H}}^\dagger\underline{\underline{M}}^{-1}\underline{\underline{H}}]^{-1}\underline{\underline{H}}^\dagger\underline{\underline{M}}^{-1}.
$$

(61)

The measured data are not affected by the perturbed determination of the sensor's positions. The data are solely affected by measurement errors $\langle\delta\underline{B}\rangle$ so that the perturbed model can be written as

$$
\langle\underline{\tilde{B}}\rangle = \langle\underline{B}\rangle + \langle\delta\underline{B}\rangle = \underline{\tilde{\underline{H}}}\,\underline{\tilde{g}} + \langle\underline{\tilde{v}}\rangle + \langle\delta\underline{B}\rangle
$$

(62)

The corresponding perturbed estimator results in

$$
\underline{\tilde{g}}_C = \underline{\tilde{w}}^\dagger\langle\underline{\tilde{B}}\rangle,
$$

(63)

where

$$
\underline{\tilde{w}}^\dagger = \left[\underline{\tilde{\underline{H}}}^\dagger\underline{\tilde{\underline{M}}}^{-1}\underline{\tilde{\underline{H}}}\right]^{-1}\underline{\tilde{\underline{H}}}^\dagger\underline{\tilde{\underline{M}}}^{-1}
$$

(64)

and $\underline{\underline{M}} = \underline{\tilde{\underline{M}}} + \Delta\underline{\underline{M}}$. The noisy shape matrix can again be written as $\underline{\underline{H}} = \underline{\tilde{\underline{H}}} + \Delta\underline{\underline{H}}$ within a linear approximation. Thus, the deviation between the accurate and the perturbed estimator is given by

$$
\begin{aligned}
\underline{g}_C - \underline{\tilde{g}}_C &= \underline{w}^\dagger\langle\underline{B}\rangle - \underline{\tilde{w}}^\dagger\langle\underline{\tilde{B}}\rangle = \underline{w}^\dagger\langle\underline{\tilde{B}}\rangle - \underline{w}^\dagger\langle\delta\underline{B}\rangle - \underline{\tilde{w}}^\dagger\langle\underline{\tilde{B}}\rangle \\
&= \left(\underline{w}^\dagger - \underline{\tilde{w}}^\dagger\right)\langle\underline{\tilde{B}}\rangle - \underline{w}^\dagger\langle\delta\underline{B}\rangle.
\end{aligned}
$$

(65)

By making use of the Neumann series, the unknown filter matrix $\underline{w}^\dagger$ can be rewritten within a linear approximation resulting in

$$
\begin{aligned}
\underline{w}^\dagger &= [\underline{\underline{H}}^\dagger\underline{\underline{M}}^{-1}\underline{\underline{H}}]^{-1}\underline{\underline{H}}^\dagger\underline{\underline{M}}^{-1} \\
&= \left[\left(\underline{\tilde{\underline{H}}}^\dagger + \Delta\underline{\underline{H}}^\dagger\right)\left(\underline{\tilde{\underline{M}}} + \Delta\underline{\underline{M}}\right)^{-1}\left(\underline{\tilde{\underline{H}}} + \Delta\underline{\underline{H}}\right)\right]^{-1}\left(\underline{\tilde{\underline{H}}}^\dagger + \Delta\underline{\underline{H}}^\dagger\right)\left(\underline{\tilde{\underline{M}}} + \Delta\underline{\underline{M}}\right)^{-1} \\
&= \left[\left(\underline{\tilde{\underline{H}}}^\dagger + \Delta\underline{\underline{H}}^\dagger\right)\left(\underline{\tilde{\underline{M}}}^{-1} - \underline{\tilde{\underline{M}}}^{-1}\Delta\underline{\underline{M}}\,\underline{\tilde{\underline{M}}}^{-1}\right)\left(\underline{\tilde{\underline{H}}} + \Delta\underline{\underline{H}}\right)\right]^{-1}\left(\underline{\tilde{\underline{H}}}^\dagger + \Delta\underline{\underline{H}}^\dagger\right)\left(\underline{\tilde{\underline{M}}}^{-1} - \underline{\tilde{\underline{M}}}^{-1}\Delta\underline{\underline{M}}\,\underline{\tilde{\underline{M}}}^{-1}\right).
\end{aligned}
$$

Using

$$
\begin{aligned}
&\left(\underline{\tilde{H}}^{\dagger} + \Delta \underline{H}^{\dagger}\right)\left(\underline{\tilde{M}}^{-1} - \underline{\tilde{M}}^{-1}\Delta\underline{M}\,\underline{\tilde{M}}^{-1}\right)\left(\underline{\tilde{H}} + \Delta\underline{H}\right) \\
&= \left(\underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1} - \underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1}\Delta\underline{M}\,\underline{\tilde{M}}^{-1} + \Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1}\right)\left(\underline{\tilde{H}} + \Delta\underline{H}\right) \\
&= \underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1}\underline{\tilde{H}} - \underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1}\Delta\underline{M}\,\underline{\tilde{M}}^{-1}\underline{\tilde{H}} + \Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1}\underline{\tilde{H}} + \underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1}\Delta\underline{H} \\
&= \underline{\tilde{P}}^{-1} - \underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1}\Delta\underline{M}\,\underline{\tilde{M}}^{-1}\underline{\tilde{H}} + \Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1}\underline{\tilde{H}} + \underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1}\Delta\underline{H}
\end{aligned}
\tag{66}
$$

and again making use of the Neumann series delivers

$$
\begin{aligned}
\underline{w}^{\dagger} &= \left[\underline{\tilde{P}} + \underline{\tilde{P}}\,\underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1}\Delta\underline{M}\,\underline{\tilde{M}}^{-1}\underline{\tilde{H}}\,\underline{\tilde{P}} - \underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1}\underline{\tilde{H}}\,\underline{\tilde{P}} - \underline{\tilde{P}}\,\underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1}\Delta\underline{H}\,\underline{\tilde{P}}\right] \\
&\quad \cdot \left(\underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1} - \underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1}\Delta\underline{M}\,\underline{\tilde{M}}^{-1} + \Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1}\right) \\
&= \underline{\tilde{P}}\,\underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1} - \underline{\tilde{P}}\,\underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1}\Delta\underline{M}\,\underline{\tilde{M}}^{-1} + \underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1} \\
&\quad + \underline{\tilde{P}}\,\underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1}\Delta\underline{M}\,\underline{\tilde{M}}^{-1}\underline{\tilde{H}}\,\underline{\tilde{P}}\,\underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1} - \underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1}\underline{\tilde{H}}\,\underline{\tilde{P}}\,\underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1} \\
&\quad - \underline{\tilde{P}}\,\underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1}\Delta\underline{H}\,\underline{\tilde{P}}\,\underline{\tilde{H}}^{\dagger}\underline{\tilde{M}}^{-1} \\
&= \underline{\tilde{w}}^{\dagger} - \underline{\tilde{w}}^{\dagger}\Delta\underline{M}\,\underline{\tilde{M}}^{-1} + \underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1} \\
&\quad + \underline{\tilde{w}}^{\dagger}\Delta\underline{M}\,\underline{\tilde{M}}^{-1}\underline{\tilde{H}}\,\underline{\tilde{w}}^{\dagger} - \underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1}\underline{\tilde{H}}\,\underline{\tilde{w}}^{\dagger} - \underline{\tilde{w}}^{\dagger}\Delta\underline{H}\,\underline{\tilde{w}}^{\dagger}.
\end{aligned}
\tag{67}
$$

Taking into account that

$$
\underline{w}^{\dagger}\langle\delta\underline{B}\rangle = \underline{\tilde{w}}^{\dagger}\langle\delta\underline{B}\rangle
\tag{68}
$$

within the linear approximation as well as

$$
\begin{aligned}
\underline{w}^{\dagger} - \underline{\tilde{w}}^{\dagger} &= -\underline{\tilde{w}}^{\dagger}\Delta\underline{M}\,\underline{\tilde{M}}^{-1} + \underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1} + \underline{\tilde{w}}^{\dagger}\Delta\underline{M}\,\underline{\tilde{M}}^{-1}\underline{\tilde{H}}\,\underline{\tilde{w}}^{\dagger} \\
&\quad - \underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1}\underline{\tilde{H}}\,\underline{\tilde{w}}^{\dagger} - \underline{\tilde{w}}^{\dagger}\Delta\underline{H}\,\underline{\tilde{w}}^{\dagger},
\end{aligned}
\tag{69}
$$

the deviation between the estimators can be approximated by

$$
\begin{aligned}
\underline{g}_C - \underline{\tilde{g}}_C &= \left(\underline{w}^{\dagger} - \underline{\tilde{w}}^{\dagger}\right)\langle\underline{\tilde{B}}\rangle - \underline{w}^{\dagger}\langle\delta\underline{B}\rangle \\
&= -\underline{\tilde{w}}^{\dagger}\Delta\underline{M}\,\underline{\tilde{M}}^{-1}\langle\underline{\tilde{B}}\rangle + \underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1}\langle\underline{\tilde{B}}\rangle + \underline{\tilde{w}}^{\dagger}\Delta\underline{M}\,\underline{\tilde{M}}^{-1}\underline{\tilde{H}}\,\underline{\tilde{w}}^{\dagger}\langle\underline{\tilde{B}}\rangle \\
&\quad - \underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1}\underline{\tilde{H}}\,\underline{\tilde{w}}^{\dagger}\langle\underline{\tilde{B}}\rangle - \underline{\tilde{w}}^{\dagger}\Delta\underline{H}\,\underline{\tilde{w}}^{\dagger}\langle\underline{\tilde{B}}\rangle - \underline{w}^{\dagger}\langle\delta\underline{B}\rangle
\end{aligned}
\tag{70}
$$

Using $\underline{\tilde{w}}^{\dagger}\langle\underline{\tilde{B}}\rangle = \underline{\tilde{g}}_C$, it follows that

$$
\begin{aligned}
\underline{g}_C - \underline{\tilde{g}}_C &= -\underline{\tilde{w}}^{\dagger}\Delta\underline{M}\,\underline{\tilde{M}}^{-1}\langle\underline{\tilde{B}}\rangle + \underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1}\langle\underline{\tilde{B}}\rangle + \underline{\tilde{w}}^{\dagger}\Delta\underline{M}\,\underline{\tilde{M}}^{-1}\underline{\tilde{H}}\,\underline{\tilde{g}}_C \\
&\quad - \underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1}\underline{\tilde{H}}\,\underline{\tilde{g}}_C - \underline{\tilde{w}}^{\dagger}\Delta\underline{H}\,\underline{\tilde{g}}_C - \underline{w}^{\dagger}\langle\delta\underline{B}\rangle \\
&= \underline{\tilde{w}}^{\dagger}\Delta\underline{M}\,\underline{\tilde{M}}^{-1}\left(\underline{\tilde{H}}\,\underline{\tilde{g}}_C - \langle\underline{\tilde{B}}\rangle\right) + \underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1}\left(\langle\underline{\tilde{B}}\rangle - \underline{\tilde{H}}\,\underline{\tilde{g}}_C\right) \\
&\quad - \underline{\tilde{w}}^{\dagger}\Delta\underline{H}\,\underline{\tilde{g}}_C - \underline{w}^{\dagger}\langle\delta\underline{B}\rangle
\end{aligned}
\tag{71}
$$

and thus,

$$
\begin{aligned}
\left|g_C - \tilde{g}_C\right|^2 &= \left\|\underline{\tilde{w}}^{\dagger}\Delta\underline{M}\,\underline{\tilde{M}}^{-1/2}\underline{\tilde{M}}^{-1/2}\left(\underline{\tilde{H}}\,\underline{\tilde{g}}_C - \langle\underline{\tilde{B}}\rangle\right) + \underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1/2}\underline{\tilde{M}}^{-1/2}\left(\langle\underline{\tilde{B}}\rangle - \underline{\tilde{H}}\,\underline{\tilde{g}}_C\right)\right. \\
&\quad \left. -\underline{\tilde{w}}^{\dagger}\Delta\underline{H}\,\underline{\tilde{g}}_C - \underline{\tilde{w}}^{\dagger}\langle\delta\underline{B}\rangle\right\|_2^2 \le \left\|\underline{\tilde{w}}^{\dagger}\Delta\underline{M}\,\underline{\tilde{M}}^{-1/2}\right\|_2^2 \cdot \left|\underline{\tilde{M}}^{-1/2}\left(\underline{\tilde{H}}\,\underline{\tilde{g}}_C - \langle\underline{\tilde{B}}\rangle\right)\right|^2 + \left\|\underline{\tilde{P}}\,\Delta\underline{H}^{\dagger}\underline{\tilde{M}}^{-1/2}\right\|_2^2 \\
&\quad \cdot\left|\underline{\tilde{M}}^{-1/2}\left(\langle\underline{\tilde{B}}\rangle - \underline{\tilde{H}}\,\underline{\tilde{g}}_C\right)\right|^2 + \left\|\underline{\tilde{w}}^{\dagger}\right\|_2^2 \cdot \|\Delta\underline{H}\|_2^2 \cdot |\underline{\tilde{g}}_C|^2 + \left\|\underline{\tilde{w}}^{\dagger}\right\|_2^2 \cdot |\langle\delta\underline{B}\rangle|^2 \\
&\to \left\|\underline{\tilde{w}}^{\dagger}\right\|_2^2 \cdot \|\Delta\underline{H}\|_2^2 \cdot |\underline{\tilde{g}}_C|^2 + \left\|\underline{\tilde{w}}^{\dagger}\right\|_2^2 \cdot |\langle\delta\underline{B}\rangle|^2.
\end{aligned}
\tag{72}
$$

Assuming that $\langle\underline{\tilde{B}}\rangle \approx \underline{\tilde{H}}\,\underline{\tilde{g}}_C$ the relative error can be estimated by

$$
\begin{aligned}
\frac{\left|g_C - \underline{\tilde{g}}_C\right|^2}{\left|\underline{\tilde{g}}_C\right|^2} &= \left\|\underline{\tilde{w}}^{\dagger}\right\|_2^2 \cdot \|\Delta\underline{H}\|_2^2 + \left\|\underline{\tilde{w}}^{\dagger}\right\|_2^2 \cdot \frac{|\langle\delta\underline{B}\rangle|^2}{|\langle\underline{\tilde{B}}\rangle|^2}\frac{|\langle\underline{\tilde{B}}\rangle|^2}{|\underline{\tilde{g}}_C|^2} \\
&= \left\|\underline{\tilde{w}}^{\dagger}\right\|_2^2 \cdot \left\|\underline{\tilde{H}}\right\|_2^2 \cdot \frac{\|\Delta\underline{H}\|_2^2}{\left\|\underline{\tilde{H}}\right\|_2^2} + \left\|\underline{\tilde{w}}^{\dagger}\right\|_2^2 \cdot \frac{|\langle\delta\underline{B}\rangle|^2}{|\langle\underline{\tilde{B}}\rangle|^2}\frac{\left|\underline{\tilde{H}}\,\underline{\tilde{g}}_C\right|^2}{|\underline{\tilde{g}}_C|^2} \\
&\le \left\|\underline{\tilde{w}}^{\dagger}\right\|_2^2 \cdot \left\|\underline{\tilde{H}}\right\|_2^2 \cdot \frac{\|\Delta\underline{H}\|_2^2}{\left\|\underline{\tilde{H}}\right\|_2^2} + \left\|\underline{\tilde{w}}^{\dagger}\right\|_2^2 \cdot \left\|\underline{\tilde{H}}\right\|_2^2 \cdot \frac{|\langle\delta\underline{B}\rangle|^2}{|\langle\underline{\tilde{B}}\rangle|^2} \\
&= \left\|\underline{\tilde{w}}^{\dagger}\right\|_2^2 \cdot \left\|\underline{\tilde{H}}\right\|_2^2 \cdot \left(\frac{\|\Delta\underline{H}\|_2^2}{\left\|\underline{\tilde{H}}\right\|_2^2} + \frac{|\langle\delta\underline{B}\rangle|^2}{|\langle\underline{\tilde{B}}\rangle|^2}\right) = \kappa^2\left(\underline{\tilde{H}}\right)\left(\frac{\|\Delta\underline{H}\|_2^2}{\left\|\underline{\tilde{H}}\right\|_2^2} + \frac{|\langle\delta\underline{B}\rangle|^2}{|\langle\underline{\tilde{B}}\rangle|^2}\right)
\end{aligned}
\tag{73}
$$

It is important to note that the right sides of **Eqs. 35**, **56** and **73** represent upper bounds for the error of the estimator. These bounds can be much larger than the true errors. The great advantage of the bounds lies in the fact that they solely depend on known quantities, wheras the accurate estimator for calculating the true estimation error is unknown within the practical application of the method. Thus, the above derived upper bounds are conservative and guarantee that the true estimation errors cannot exceed these bounds as long as the measurement errors as well as the measurement position errors are much smaller than the measurements themselves so that the linear approximation is valid.

# 4 FINITE SAMPLE AVERAGING

For the calculation of Capon's estimator, averaged magnetic field data

$$
\langle\underline{B}\rangle = \frac{1}{Q}\sum_{\alpha=1}^{Q}\underline{B}^{\alpha}
\tag{74}
$$

are required [3]. Here, $Q$ denotes the number of measurements at a fixed set of data points. Within the practical application of the method only a finite number $Q < \infty$ of samples is available, which yields a standard deviation of Capon's estimator $\underline{g}_C$. In the following, an approximation for the variance of Capon's estimator is derived by regarding Capon's method as a special case of the maximum likelihood method [6].

In the vicinity of its maximum value the likelihood function can be approximated as

$$
\mathcal{L} \sim \exp\left[-\frac{1}{2}\sum_{\alpha=1}^{Q}\left(\underline{B}^{\alpha} - \underline{\underline{H}}\,\underline{g}\right)^{\dagger}\underline{N}^{-1}\left(\underline{B}^{\alpha} - \underline{\underline{H}}\,\underline{g}\right)\right],
\tag{75}
$$

where $\underline{N}$ denotes the noise covariance matrix [6, 19]. Before deducing the error resulting from the finite number of samples let us make a general comment about the construction of the likelihood function in terms of the averaging process. The noise matrix is already statistically evaluated so that the likelihood function is constructed by making use of two different averaging processes. Within the first averaging process the noise matrix $\underline{N}$ is calculated. This averaging process does not incorporate the underlying model since only the distribution of the data around the mean value $\langle\underline{B}\rangle$ is determined which allows to assess the general quality of the measurements. When the distribution of the measurements is large, significant errors within the subsequent data fitting are expectable. Especially, these errors are independent of the chosen model. Within the second averaging process a specific underlying model $\underline{H}\,\underline{g}$ is fitted to the data $\underline{B}^{\alpha}$.

For most practical applications, the noise matrix $\underline{N}$ is unknown and has to be approximated. In the special case of Gaussian errors with zero mean and variance $\sigma_n$, the noise covariance matrix can be written as $\underline{N} = \sigma_n^2\underline{I}$, where $\underline{I}$ denotes the identity matrix. Substituting the noise covariance matrix $\underline{N}$ by the data covariance matrix $\underline{M} = \langle\underline{B}\circ\underline{B}\rangle$, the maximum

likelihood estimator may be converted into Capon's estimator [6]. The corresponding likelihood function is modified to

$$\mathcal{L} \sim \exp\left[ -\frac{1}{2}\sum_{\alpha=1}^{Q}\left( \underline{B}^{\alpha} - \underline{H}\,\underline{g} \right)^{\dagger} \underline{M}^{-1}\left( \underline{B}^{\alpha} - \underline{H}\,\underline{g} \right) \right]. \qquad (76)$$

The weighted difference between the model and the data can be rewritten as

$$\left( \underline{B}^{\alpha} - \underline{H}\,\underline{g} \right)^{\dagger}\underline{M}^{-1}\left( \underline{B}^{\alpha} - \underline{H}\,\underline{g} \right) = \left( \underline{B}^{\alpha\dagger} - \underline{g}^{\dagger}\underline{H}^{\dagger} \right)\underline{M}^{-1}\left( \underline{B}^{\alpha} - \underline{H}\,\underline{g} \right)$$
$$= \underline{B}^{\alpha\dagger}\underline{M}^{-1}\underline{B}^{\alpha} - 2\underline{g}^{\dagger}\underline{H}^{\dagger}\underline{M}^{-1}\underline{B}^{\alpha} + \underline{g}^{\dagger}\underline{H}^{\dagger}\underline{M}^{-1}\underline{H}\,\underline{g}. \qquad (77)$$

As discussed above, the data covariance matrix $\underline{M}$ is already statistically evaluated so that the averaging process results in

$$\sum_{\alpha=1}^{Q}\left( \underline{B}^{\alpha} - \underline{H}\,\underline{g} \right)^{\dagger}\underline{M}^{-1}\left( \underline{B}^{\alpha} - \underline{H}\,\underline{g} \right) = Q\langle \underline{B}^{\dagger}\underline{M}^{-1}\underline{B}\rangle - 2Q\underline{g}^{\dagger}\underline{H}^{\dagger}\underline{M}^{-1}\langle \underline{B}\rangle$$
$$+ Q\underline{g}^{\dagger}\underline{H}^{\dagger}\underline{M}^{-1}\underline{H}\,g. \qquad (78)$$

Insertion into **Eq. 76** yields

$$\mathcal{L} \sim \exp\left\{ -\frac{Q}{2}\left[ \langle \underline{B}^{\dagger}\underline{M}^{-1}\underline{B}\rangle - 2\underline{g}^{\dagger}\underline{H}^{\dagger}\underline{M}^{-1}\langle \underline{B}\rangle + \underline{g}^{\dagger}\underline{H}^{\dagger}\underline{M}^{-1}\underline{H}\,\underline{g} \right] \right\} \qquad (79)$$

or equivalently

$$\mathcal{L} \sim \exp\left\{ -\frac{Q}{2}\left[ \langle B_i M_{ij}^{-1} B_j\rangle - 2g_k H_{ki}^{\dagger}M_{ij}^{-1}\langle B_j\rangle + g_l H_{lk}^{\dagger}M_{ki}^{-1}H_{ij}g_j \right] \right\}. \qquad (80)$$

The $m$'th component of Capon's estimator $\underline{g}_C$ corresponds with the maximum value of the likelihood function [6, 19]. The corresponding standard deviation ($1\sigma$-error) is described by the width of the likelihood function at its maximum value which is given by [6, 19].

$$\sigma_{g_m} = \left[ -\partial_{g_m}^2 \ln\mathcal{L} \right]^{-1/2}, \qquad (81)$$

where

$$\partial_{g_m}^2 \ln\mathcal{L} = \partial_{g_m}\left( \frac{1}{\mathcal{L}}\partial_{g_m}\mathcal{L} \right). \qquad (82)$$

Since

$$\partial_{g_m}\mathcal{L} = -\frac{Q}{2}\mathcal{L}\,\partial_{g_m}\left[ \langle B_i M_{ij}^{-1} B_j\rangle - 2g_k H_{ki}^{\dagger}M_{ij}^{-1}\langle B_j\rangle + g_l H_{lk}^{\dagger}M_{ki}^{-1}H_{ij}g_j \right]$$
$$= -\frac{Q}{2}\mathcal{L}\left[ -2\delta_{km}H_{ki}^{\dagger}M_{ij}^{-1}\langle B_j\rangle + \delta_{lm}H_{lk}^{\dagger}M_{ki}^{-1}H_{ij}g_j + g_l H_{lk}^{\dagger}M_{ki}^{-1}H_{ij}\delta_{jm} \right]$$
$$= -\frac{Q}{2}\mathcal{L}\left[ -2H_{mi}^{\dagger}M_{ij}^{-1}\langle B_j\rangle + H_{mk}^{\dagger}M_{ki}^{-1}H_{ij}g_j + g_l H_{lk}^{\dagger}M_{ki}^{-1}H_{im} \right] \qquad (83)$$

where $\delta_{ij}$ denotes the Kronecker delta, the $m$'th component of Capon's estimator $\underline{g}_C$ can be calculated via [6].

$$\partial_{g_m}\mathcal{L} = 0 \qquad (84)$$
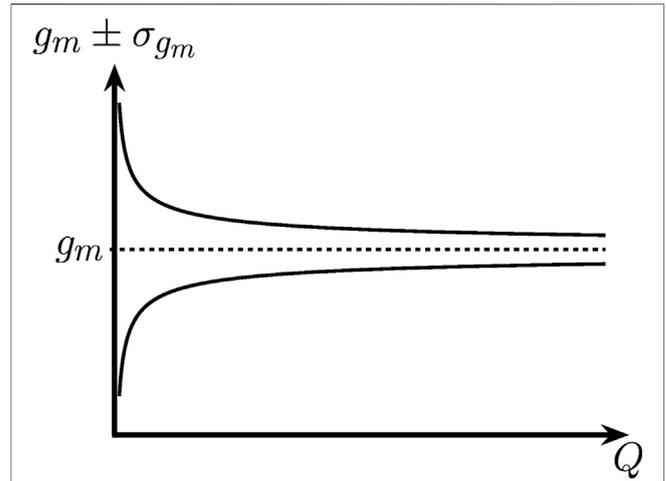


**FIGURE 3** | Sketch of the $m$'the component of Capon's estimator and the corresponding $1\sigma$-error subject to the sample size $Q$. The error declines as $1/\sqrt{Q}$.

yielding

$$-2H_{mi}^{\dagger}M_{ij}^{-1}\langle B_j\rangle + H_{mk}^{\dagger}M_{ki}^{-1}H_{ij}g_j + g_l H_{lk}^{\dagger}M_{ki}^{-1}H_{im} = 0. \qquad (85)$$

Because of $M_{ki}^{-1} = M_{ik}^{-1}$, as well as $H_{lk}^{\dagger} = H_{kl}$ and $H_{im} = H_{mi}^{\dagger}$ in the case of real shape matrices [3], it follows that

$$0 = -2H_{mi}^{\dagger}M_{ij}^{-1}\langle B_j\rangle + H_{mk}^{\dagger}M_{ki}^{-1}H_{ij}g_j + g_l H_{kl}M_{ik}^{-1}H_{mi}^{\dagger}$$
$$= -2H_{mi}^{\dagger}M_{ij}^{-1}\langle B_j\rangle + H_{mk}^{\dagger}M_{ki}^{-1}H_{ij}g_j + H_{mi}^{\dagger}M_{ik}^{-1}H_{kl}g_l \qquad (86)$$
$$= -2H_{mi}^{\dagger}M_{ij}^{-1}\langle B_j\rangle + 2H_{mk}^{\dagger}M_{ki}^{-1}H_{ij}g_j$$

and therefore,

$$\underline{g}_C = \left[ \underline{H}^{\dagger}\underline{M}^{-1}\underline{H} \right]^{-1}\underline{H}^{\dagger}\underline{M}^{-1}\langle \underline{B}\rangle \qquad (87)$$

which is in agreement with the estimator resulting from the linear algebraic formulation of Capon's method [3]. The second derivative results in

$$\partial_{g_m}^2 \ln\mathcal{L} = -\frac{Q}{2}\partial_{g_m}\left[ -2H_{mi}^{\dagger}M_{ij}^{-1}\langle B_j\rangle + H_{mk}^{\dagger}M_{ki}^{-1}H_{ij}g_j + g_l H_{lk}^{\dagger}M_{ki}^{-1}H_{im} \right]$$
$$= -\frac{Q}{2}\left[ H_{mk}^{\dagger}M_{ki}^{-1}H_{ij}\delta_{jm} + \delta_{lm}H_{lk}^{\dagger}M_{ki}^{-1}H_{im} \right]$$
$$= -\frac{Q}{2}\left[ H_{mk}^{\dagger}M_{ki}^{-1}H_{im} + H_{mk}^{\dagger}M_{ki}^{-1}H_{im} \right]$$
$$= -QH_{mk}^{\dagger}M_{ki}^{-1}H_{im}. \qquad (88)$$

Using the definition of the coefficient matrix [3].

$$\underline{P} = \left[ \underline{H}^{\dagger}\underline{M}^{-1}\underline{H} \right]^{-1} \qquad (89)$$

or equivalently

$$\left( \underline{P}^{-1} \right)_{mm} = H_{mk}^{\dagger}M_{ki}^{-1}H_{im} \qquad (90)$$

delivers

$$\partial_{g_m}^2 \ln\mathcal{L} = -Q\left( \underline{P}^{-1} \right)_{mm}. \qquad (91)$$

Thus, the error of the $m$'th component of Capon's estimator results in

$$\sigma_{g_m} = \frac{1}{\sqrt{Q \cdot \left(\underline{\underline{P}}^{-1}\right)_{mm}}}, \tag{92}$$

which declines as $1/\sqrt{Q}$. The functional dependence of the error is qualitatively illustrated in **Figure 3**.

## 5 SUMMARY AND OUTLOOK

The analysis of the error propagation is of major importance for the application of linear inversion methods. Within a linear approximation, upper bounds for the errors of Capon's estimator resulting from measurement errors and measurement position errors are derived. These upper bounds solely depend on known quantities, i.e., measurements and measurement positions, whereas the true estimation error cannot be calculated within the practical application of the method, since the accurate estimator is unavailable. It turns out that Capon's method provides the same error propagation as the least square fit method. These two methods differ in the filter matrices which weight or eliminate parts of the data in different ways. Since these matrices are applied to the same set of measurements, the different weighting of the data or the elimination of subsets does not reduce the errors. The condition number of the shape matrix is the key parameter for the error propagation, since it determines how the errors are amplified. The measurement errors as well as the measurement position errors have to be estimated from the measurements. For a given underlying model, the condition number of the shape matrix solely depends on the measurement positions. Thus, the amplification of the errors can be reduced by choosing preferred data points.

Furthermore, Capon's method is based on the evaluation of statistically averaged data. For the practical application of the method only a finite number $Q$ of samples is available. This limited number of samples results in an error of Capon's estimator which can be derived by regarding Capon's method as a special case of the maximum likelihood estimator. The more samples are available, the smaller the error becomes, since the error declines as $1/\sqrt{Q}$.

As a follow-up of the generalized derivation of Capon's method, the present work establishes the mathematical basis of Capon's error propagation for the practical application of the method.

## Appendix: Matrix Power of the Data Covariance Matrix

For the calculation of Capon's estimator the inverse data covariance matrix $\underline{\underline{M}}^{-1}$ is necessary. First of all it should be noted that the inverse exists due to the averaging process, whereas the matrix $\underline{\underline{B}} \circ \underline{\underline{B}}$ is not invertible [3]. In the case of a

vanishing standard deviation ($\sigma_n = 0$), the application of the diagonal loading technique [3] guarantees the existence of the inverse data covariance matrix, since the condition number of $\underline{\underline{M}}$ is close to unity at the optimal diagonal loading parameter.

Furthermore, there exists a unitary transformation $\underline{\underline{V}}$ so that

$$\underline{\underline{M}} = \underline{\underline{V}}\,\underline{\underline{D}}_{\underline{M}}\,\underline{\underline{V}}^{-1} \tag{93}$$

where

$$\underline{\underline{D}}_{\underline{M}} = \mathrm{diag}\,[\lambda_1, \dots, \lambda_n] \tag{94}$$

is a diagonal matrix that contains the eigenvalues of $\underline{\underline{M}}$. Therefore, the inverse data covariance matrix results in

$$\underline{\underline{M}}^{-1} = \left[\underline{\underline{V}}\,\underline{\underline{D}}_{\underline{M}}\,\underline{\underline{V}}^{-1}\right]^{-1} = \underline{\underline{V}}\,\underline{\underline{D}}_{\underline{M}}^{-1}\,\underline{\underline{V}}^{-1}, \tag{95}$$

where

$$\underline{\underline{D}}_{\underline{M}}^{-1} = \mathrm{diag}\,\left[\lambda_1^{-1}, \dots, \lambda_n^{-1}\right]. \tag{96}$$

Further transformation delivers

$$\begin{aligned}
\underline{\underline{M}}^{-1} &= \underline{\underline{V}}\,\underline{\underline{D}}_{\underline{M}}^{-1}\,\underline{\underline{V}}^{-1} \\
&= \underline{\underline{V}}\,\underline{\underline{D}}_{\underline{M}}^{-1/2}\,\underline{\underline{D}}_{\underline{M}}^{-1/2}\,\underline{\underline{V}}^{-1} \\
&= \underline{\underline{V}}\,\underline{\underline{D}}_{\underline{M}}^{-1/2}\,\underline{\underline{V}}^{-1}\,\underline{\underline{V}}\,\underline{\underline{D}}_{\underline{M}}^{-1/2}\,\underline{\underline{V}}^{-1} \\
&= \underline{\underline{M}}^{-1/2}\,\underline{\underline{M}}^{-1/2}
\end{aligned} \tag{97}$$

so that the square root of the inverse data covariance matrix is defined as

$$\underline{\underline{M}}^{-1/2} = \underline{\underline{V}}\,\underline{\underline{D}}_{\underline{M}}^{-1/2}\,\underline{\underline{V}}^{-1} \tag{98}$$

where

$$\underline{\underline{D}}_{\underline{M}}^{-1/2} = \mathrm{diag}\,\left[\lambda_1^{-1/2}, \dots, \lambda_n^{-1/2}\right]. \tag{99}$$

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

## AUTHOR CONTRIBUTIONS

All authors contributed conception and design of the study; ST and UM wrote the first draft of the manuscript; All authors contributed to manuscript revision, read and approved the submitted version.

## FUNDING

## ACKNOWLEDGMENTS

# REFERENCES

1. Capon J. High-resolution Frequency-Wavenumber Spectrum Analysis. *Proc IEEE* (1969) 57:1408–18. doi:10.1109/PROC.1969.7278

2. Toepfer S, Narita Y, Heyner D, and Motschmann U. The Capon Method for Mercury's Magnetic Field Analysis. *Front Phys* (2020) 8:249. doi:10.3389/fphy.2020.00249

3. Toepfer S, Narita Y, Heyner D, Kolhey P, and Motschmann U. Mathematical Foundation of Capon's Method for Planetary Magnetic Field Analysis. *Geosci Instrum Method Data Syst* (2020) 9:471–81. doi:10.5194/gi-9-471-2020

4. Motschmann U, Woodward TI, Glassmeier KH, Southwood DJ, and Pinçon JL. Wavelength and Direction Filtering by Magnetic Measurements at Satellite Arrays: Generalized Minimum Variance Analysis. *J Geophys Res* (1996) 101: 4961–5. doi:10.1029/95JA03471

5. Su W, Gu H, Ni J, and Liu G. Performance Analysis of MVDR Algorithm in the Presence of Amplitude and Phase Errors. *IEEE Trans Antennas Propagat* (2001) 49(12):1875–7. doi:10.1109/8.982472

6. Narita Y. A Note on Capon's Minimum Variance Projection for Multi-Spacecraft Data Analysis. *Front Phys* (2019) 7:121. doi:10.3398/fphy.2019.00008

7. Sahraoui F, Belmont G, Rezeau L, Cornilleau-Wehrlin N, Pinçon JL, and Balogh A. Anisotropic Turbulent Spectra in the Terrestrial Magnetosheath as Seen by the Cluster Spacecraft. *Phys Rev Lett* (2006) 96:7. doi:10.1103/PhysRevLett.96.075002

8. Huang SY, Zhou M, Sahraoui F, Deng XH, Pang Y, Yuan ZG, et al. Wave Properties in the Magnetic Reconnection Diffusion Region with Highβ: Application of Thek-Filtering Method to Cluster Multispacecraft Data. *J Geophys Res* (2010) 115:17–9. doi:10.1029/2010JA015335

9. Narita Y, Comişel H, and Motschmann U. Spatial Structure of Ion-Scale Plasma Turbulence. *Front Phys* (2014) 2:13. doi:10.3389/fphy.2014.00013

10. Roberts OW, Li X, and Jeska L. A Statistical Study of the Solar Wind Turbulence at Ion Kinetic Scales Using the K-Filtering Technique and Cluster Data. *ApJ* (2015) 802:1. doi:10.1088/0004-637X/802/1/2

11. Wang T, Cao J, Fu H, Meng X, and Dunlop M. Compressible Turbulence with Slow-Mode Waves Observed in the Bursty Bulk Flow of Plasma Sheet. *Geophys Res Lett* (2016) 43(5):1854–61. doi:10.1002/2016GL068147

12. Zhang L, He J, Narita Y, and Feng X. Reconstruction Test of Turbulence Power Spectra in 3D Wavenumber Space with at Most 9 Virtual Spacecraft Measurements. *J Geophys Res Space Phys* (2021) 126:e2019JA027413. doi:10.1029/2019JA027413

13. Narita Y, Plaschke F, Magnes W, Fischer D, and Schmid D. Error Estimate for Fluxgate Magnetometer In-Flight Calibration on a Spinning Spacecraft. *Geosci Instrum Method Data Syst* (2021) 10:13–24. doi:10.5194/gi-10-13-202

14. Werner D. *Funktionalanalysis*. Berlin: Springer (2007)

15. Belsley DA, Kuh E, and Welsch RE. *The Condition Number, Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: Wiley (1980). p. 100–4.

16. Gauß CF. *(1839) Allgemeine Theorie des Erdmagnetismus: Resultate aus den Beobachtungen des magnetischen Vereins im Jahre 1838*, editorsCF Gauss and W Weber. Leipzig: Weidmannsche Buchhandlung (1839). p. 1–57.

17. Glassmeier K-H, and Tsurutani BT. Carl Friedrich Gauss—General Theory of Terrestrial Magnetism—A Revised Translation of the German Text. *Hist Geo Space Sci* (2014) 5:11–62. doi:10.5194/hgss-5-11-2014

18. Toepfer S, Narita Y, Glassmeier K-H, Heyner D, Kolhey P, Motschmann U, et al. The Mie Representation for Mercury's Magnetic Field. *Earth Planets Space* (2021) 73:65. doi:10.1186/s40623-021-01386-4

19. Dodelson S. *Modern Cosmology*. London: Academic Press (2003).