



# Mathematical foundation of Capon's method for planetary magnetic field analysis

Simon Toepfer<sup>1</sup>, Yasuhito Narita<sup>2,3</sup>, Daniel Heyner<sup>3</sup>, Patrick Kolhey<sup>3</sup>, and Uwe Motschmann<sup>1,4</sup>

<sup>1</sup>Institut für Theoretische Physik, Technische Universität Braunschweig, Braunschweig, Germany

<sup>2</sup>Space Research Institute, Austrian Academy of Sciences, Graz, Austria

<sup>3</sup>Institut für Geophysik und extraterrestrische Physik, Technische Universität Braunschweig, Braunschweig, Germany

<sup>4</sup>DLR Institute of Planetary Research, Berlin, Germany

**Correspondence:** Simon Toepfer (s.toepfer@tu-braunschweig.de)

## Abstract.

Minimum variance distortionless projection, the so-called Capon method, serves as a powerful and robust data analysis tool when working on various kinds of ill-posed inverse problems. The method has not only successfully been applied to multi-point wave and turbulence studies in the context of space plasma physics, but also is currently being considered as a technique to perform the multipole expansion of planetary magnetic fields from a limited data set, such as Mercury's magnetic field analysis. The practical application and limits of the Capon method are discussed in a rigorous fashion by formulating its linear-algebraic derivation in view of planetary magnetic field studies. Furthermore, the optimization of Capon's method by making use of diagonal loading is considered.

## 1 Introduction

Nonlinear and adaptive filter technique has a wider range of applications in geophysical and space science studies to find the most likely parameter set describing the measurement data or to decompose the data into a set of signals and noise. Above all, the minimum variance distortionless projection introduced by Capon (1969) (hereafter, Capon's method) has successfully been applied to multi-point data analyses for the waves, turbulence fields, and current sheets (Motschmann et al., 1996; Glassmeier et al., 2001; Narita et al., 2003, 2013; Contantinescu et al., 2006; Plaschke et al., 2008). The strength of Capon's method lies in the fact that the method performs a robust data fitting even when the spatial sampling or data amount is limited in the measurement (e.g., successfully applied to four-spacecraft data (Motschmann et al., 1996)). Capon's method is currently being considered for planetary magnetic field studies in which the data (i.e., magnetic field samples) are more limited (e.g., non-ideal orbits), and has recently been applied to Mercury's magnetic field data in view of the BepiColombo mission (Toepfer et al., 2020).

From a theoretical point of view there exist several origins for the derivation of the method. The first derivation of Capon's method (Capon, 1969), constructed for the analysis of seismic waves, is based on the estimation of frequency-wavenumber spectra. Later on, this approach has been reformulated in terms of matrix algebra (Motschmann et al., 1996). In the light



of mathematical statistics Capon's estimator can be regarded as a special case of the maximum likelihood estimator (Narita,  
 25 2019). In this work the linear-algebraic formulation of the method (Motschmann et al., 1996) with specific attention to the  
 magnetic field analysis is extended and the application of diagonal loading is discussed to improve the quality of data analysis  
 with more justified applications and limits.

## 2 Motivation of Capon's method

The analysis of planetary magnetic fields is of great interest and one of the main tasks in space science. Here we pay special  
 30 attention to the analysis of Mercury's internal magnetic field which is one of the primary goals of the BepiColombo mission  
 (Benkhoff et al., 2010). The magnetometer on board the Mercury Planetary Orbiter (MPO) (Glassmeier et al., 2010) measures  
 the magnetic field vectors  $\mathbf{b}^i = (b_x^i, b_y^i, b_z^i)^T \in \mathbb{R}^3$  at  $N$  data points  $\mathbf{x}^i$ ,  $i = 1, \dots, N$  along the orbit in the vicinity of Mercury.  
 The magnetic fields around Mercury are considered as a composition or superposition of internal fields generated by the dy-  
 namo process, crustal and induced fields, which are mainly dominated by dipole and quadrupole fields and external fields  
 35 generated by the currents flowing in the magnetosphere. For Mercury the external fields contribute a significant amount to the  
 total magnetic field within the magnetosphere (Anderson et al., 2011) and therefore, a robust method is required for separating  
 the internal fields from the total measured field. Yet, each component has to be properly modeled and parameterized when  
 decomposing the field.

40 For example, when only data in current-free regions are analyzed, the planetary magnetic field is irrotational and can be  
 parameterized via the Gauss representation (Gauss, 1839). Within these current-free regions, the internal magnetic field can  
 be expressed as the gradient of a scalar potential  $\Phi$ , which can be expanded into a set of basis functions. Considering the  
 expansion into spherical harmonics, the potential for the planetary dipole and quadrupole fields results in

$$\Phi(\mathbf{x}^i) = R_M \sum_{l=1}^2 \left( \frac{R_M}{r^i} \right)^{l+1} \sum_{m=0}^l [g_l^m \cos(m\lambda^i) + h_l^m \sin(m\lambda^i)] P_l^m(\cos(\theta^i)), \quad (1)$$

45 where planetary centered coordinates with radius  $r^i \in [R_M, \infty)$ , azimuth angle  $\lambda^i \in [0, 2\pi]$  and polar angle  $\theta^i \in [0, \pi]$  are  
 chosen.  $R_M$  indicates the radius of Mercury and  $P_l^m$  are the Schmidt-normalized associated Legendre polynomials of de-  
 gree  $l$  and order  $m$ . Within this series expansion there occurs a set of expansion coefficients  $g_l^m$  and  $h_l^m$ , named internal  
 Gauss coefficients. By constructing the Gauss coefficients in a vectorial fashion and defining the true coefficient vector as  
 $\mathbf{g} = (g_1^0, g_1^1, h_1^1, g_2^0, g_2^1, h_2^1, g_2^2, h_2^2)^T$ , the magnetic field vectors at every data point  $\mathbf{x}^i$  and the expansion coefficients are related  
 50 via

$$\mathbf{b}^i = -\partial_{\mathbf{x}^i} \Phi(\mathbf{x}^i) + \mathbf{v}^i + \mathbf{n}^i \quad (2)$$

$$= \mathbf{H}^i \mathbf{g} + \mathbf{v}^i + \mathbf{n}^i, \quad (3)$$

where the terms of the series expansion are arranged in the matrix  $\mathbf{H}^i(r^i, \theta^i, \lambda^i)$ , called shape matrix. The vector  $\mathbf{v}^i$  describes  
 the parts that are not parameterized by the underlying model, e.g. the external parts, closing the void between the parameterized



55 field and the measurement noise of the sensors, which is symbolized by the vector  $\mathbf{n}^i$ . Especially, the measurement noise is neither correlated with the parameterized part  $\mathbf{H}^i \mathbf{g}$  nor with the non-parameterized part  $\mathbf{v}^i$ .

Summarizing the magnetic field measurements for all  $N$  data points into a vector  $\mathbf{B} = \left( (\mathbf{b}^1)^T \dots (\mathbf{b}^N)^T \right)^T \in \mathbb{R}^{3N}$ , the field can be written as

$$60 \quad \mathbf{B} = \mathbf{H} \mathbf{g} + \mathbf{v} + \mathbf{n}, \quad (4)$$

where  $\mathbf{n} = (\mathbf{n}^1 \dots \mathbf{n}^N)^T \in \mathbb{R}^{3N}$ ,  $\mathbf{v} = (\mathbf{v}^1 \dots \mathbf{v}^N)^T \in \mathbb{R}^{3N}$  and  $\mathbf{H} = [\mathbf{H}^1 \dots \mathbf{H}^N]^T \in \mathbb{R}^{3N \times G}$ .  $G$  indicates the number of expansion coefficients.

Within Eq. (4) the magnetic field vector  $\mathbf{B}$  and the shape matrix  $\mathbf{H}$ , given by the underlying model, are known. The coefficient vector  $\mathbf{g}$  is to be determined by data fitting. Since in most applications the number of known magnetic data points is much larger than the number of wanted expansion coefficients ( $G \ll 3N$ ),  $\mathbf{H}$  is a rectangular matrix in general. Furthermore, the non-parameterized parts of the field and the noise are unknown. Therefore, the direct inversion of Eq. (4) is impossible and  $\mathbf{g}$  has to be estimated. In this case, Capon's method establishes a robust and useful tool to find the estimated solution for the expansion coefficients in Eq. (4). Since the method does not require the orthogonality of the basis functions, it has a wider range of applications when decomposing the measured data into a set of superposed signals, especially when the number of data points is limited.

### 3 Derivation of Capon's method

The following derivation of Capon's method is based on the linear-algebraic formulation (Motschmann et al., 1996) which was formerly applied to the analysis of plasma waves in the terrestrial magnetosphere. Now we are focussing on the analysis of planetary magnetic fields.

As illustrated in the previous section, the magnetic field  $\mathbf{b}^i$  measured at data point  $\mathbf{x}^i$  in the vicinity of Mercury and the wanted expansion coefficients  $\mathbf{g}$  are related via

$$\mathbf{b}^i = \mathbf{H}^i \mathbf{g} + \mathbf{v}^i + \mathbf{n}^i \quad (5)$$

80 or in a compact form for all  $N$  data points

$$\mathbf{B} = \mathbf{H} \mathbf{g} + \mathbf{v} + \mathbf{n}, \quad (6)$$

where the shape matrix  $\mathbf{H}$  describes the underlying model.



For every data point  $x^i$  the noise vector  $n^i$  is assumed to be Gaussian with variance  $\sigma_n$  and zero mean, so that  $\langle n^i \rangle = 0$  and  $\langle n^i \circ n^i \rangle = \sigma_n^2 \mathbf{I}$ , where  $\mathbf{I}$  is the identity matrix. The angular brackets indicate averaging over an ensemble, e.g., different samples, realizations or measurements. Therefore,  $\langle b^i \rangle = \mathbf{H}^i \mathbf{g} + \langle v^i \rangle$  holds and equivalently

$$\langle \mathbf{B} \rangle = \mathbf{H} \mathbf{g} + \langle \mathbf{v} \rangle, \quad (7)$$

since the model  $\mathbf{H}$  and the true coefficient vector  $\mathbf{g}$  are not affected by the averaging.

Because  $\mathbf{H}$  is not always a square matrix but is in general a rectangular matrix with different sizes between rows and columns, the direct inversion of Eq. (7) is not guaranteed. Let us ignore for the moment the non-existence of  $\mathbf{H}^{-1}$  and write down the equation

$$\mathbf{g} = \mathbf{H}^{-1} (\langle \mathbf{B} \rangle - \langle \mathbf{v} \rangle). \quad (8)$$

Despite its simplicity it is obviously incorrect. As  $\mathbf{H}^{-1}$  does not exist, let us look for another matrix  $\mathbf{w}$  called filter matrix, which follows the structure of this equation and fulfills in principle the resolution of Eq. (7) with respect to  $\mathbf{g}$ . Capon's method is just the procedure to construct the filter matrix  $\mathbf{w}$  and to calculate or better to say estimate  $\mathbf{g}$ . To do so, some helpful quantities are introduced. In order to distinguish between the true coefficient vector  $\mathbf{g}$  and the estimated solution, in the following Capon's estimator will be symbolized by  $\mathbf{g}_C$ .

For the inversion of Eq. (4) the data covariance matrix  $\mathbf{M}$  and the coefficient matrix  $\mathbf{P}$  are introduced as follows:

$$\mathbf{M} = \langle \mathbf{B} \circ \mathbf{B} \rangle = \frac{1}{Q} \sum_{\alpha=1}^Q \mathbf{B}^\alpha \circ \mathbf{B}^\alpha \in \mathbb{R}^{3N \times 3N} \quad (9)$$

$$\mathbf{P} = \langle \mathbf{g}_C \circ \mathbf{g}_C \rangle \in \mathbb{R}^{G \times G} \quad (10)$$

Here  $Q$  indicates the number of measurements and the circle 'o' symbolizes the outer product, which is defined by

$$\mathbf{x} \circ \mathbf{y} = \mathbf{x} \cdot \mathbf{y}^\dagger \in \mathbb{R}^{n \times m} \quad (11)$$

for any pair of vectors  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{y} \in \mathbb{R}^m$ . The dagger  $\dagger$  indicates the Hermitian conjugate and the dot stands for the multiplication of the matrices  $\mathbf{x} \in \mathbb{R}^{n \times 1}$  and  $\mathbf{y}^\dagger \in \mathbb{R}^{1 \times m}$ . Therefore, the diagonal of the matrix  $\mathbf{P}$  contains the quadratic averaged components of the wanted estimator. It is important to note here that the covariance matrix  $\mathbf{M}$  must be statistically averaged, otherwise the matrix is singular with a vanishing determinant and the further analysis cannot be achieved.



If the non-invertibility of the matrix  $B^\alpha \circ B^\alpha$  is neglected, for every measurement  $\alpha = 1, \dots, Q$  an estimator  $g_C^\alpha$  for the true  
 110 coefficient vector  $g$  can be determined, so that

$$B^\alpha = \mathbf{H} g_C^\alpha + v^\alpha + n^\alpha \quad (12)$$

is valid. Thereby, each estimator deviates from the true coefficient vector by an error vector  $\varepsilon^\alpha = g - g_C^\alpha$ . Note that because of  
 the non-invertibility of the matrix  $B^\alpha \circ B^\alpha$  the single estimator  $g_C^\alpha$  cannot be calculated. Since the invertibility is solely given  
 by the averaging over  $Q$  measurements, only the averaged estimator

$$115 \quad g_C = \frac{1}{Q} \sum_{\alpha=1}^Q g_C^\alpha, \quad (13)$$

with its related error

$$\langle \varepsilon \rangle = \frac{1}{Q} \sum_{\alpha=1}^Q \varepsilon^\alpha \quad (14)$$

is available.

120 In contrast to the estimator, the true coefficient vector is a theoretical given vector, that is not affected by the averaging  
 ( $g \equiv \langle g \rangle$ , Eq. (7)). This property directly links the estimator to the true coefficient vector, which can be rewritten as

$$g = g_C^\alpha + \varepsilon^\alpha. \quad (15)$$

Averaging over  $Q$  measurements and using  $g \equiv \langle g \rangle$  results in

$$g = g_C + \langle \varepsilon \rangle \quad (16)$$

125 and

$$g \circ g = \langle g_C \circ g_C \rangle + \langle g_C \circ \varepsilon \rangle + \langle \varepsilon \circ g_C \rangle + \langle \varepsilon \circ \varepsilon \rangle \quad (17)$$

analogously for the second order moments.

In the limit of vanishing errors  $\langle \varepsilon \rangle \rightarrow 0$ ,  $\langle \varepsilon \circ \varepsilon \rangle \rightarrow 0$  respectively, Capon's estimator converges to the true coefficient vector

$$130 \quad g_C \rightarrow g \quad (18)$$

and therefore

$$\langle g_C \circ g_C \rangle \rightarrow g \circ g. \quad (19)$$



For the further evaluation of Eq. (4) the definition of the outer product (Eq. 11) is utilized. Matrix multiplication of Eq. (4) with its Hermitian adjoint and averaging yields

$$135 \quad \langle \mathbf{B} \cdot \mathbf{B}^\dagger \rangle = \langle (\mathbf{H}\mathbf{g} + \mathbf{v} + \mathbf{n}) \cdot (\mathbf{H}\mathbf{g} + \mathbf{v} + \mathbf{n})^\dagger \rangle \quad (20)$$

and therefore

$$\langle \mathbf{B} \circ \mathbf{B} \rangle = \mathbf{H} \cdot \mathbf{g} \circ \mathbf{g} \cdot \mathbf{H}^\dagger + 2(\mathbf{H}\mathbf{g}) \circ \langle \mathbf{v} \rangle + \langle \mathbf{v} \circ \mathbf{v} \rangle + \langle \mathbf{n} \circ \mathbf{n} \rangle, \quad (21)$$

assuming, that  $\mathbf{n}$  is Gaussian with variance  $\sigma_n$  and zero mean ( $\langle \mathbf{n} \rangle = 0$ ). By means of the limit  $\langle \varepsilon \rangle \rightarrow 0$  in Eq. (19) the unknown matrix  $\mathbf{g} \circ \mathbf{g}$  and the true coefficient vector  $\mathbf{g}$  can be approximated by Capon's estimator, resulting in

$$140 \quad \langle \mathbf{B} \circ \mathbf{B} \rangle = \mathbf{H} \cdot \langle \mathbf{g}_C \circ \mathbf{g}_C \rangle \cdot \mathbf{H}^\dagger + 2(\mathbf{H}\mathbf{g}_C) \circ \langle \mathbf{v} \rangle + \langle \mathbf{v} \circ \mathbf{v} \rangle + \langle \mathbf{n} \circ \mathbf{n} \rangle. \quad (22)$$

Taking into account that  $\langle \mathbf{n} \circ \mathbf{n} \rangle = \sigma_n^2 \mathbf{I}$  and using the above defined abbreviations, Eq. (22) can be rewritten as

$$\mathbf{M} = \mathbf{H}\mathbf{P}\mathbf{H}^\dagger + 2(\mathbf{H}\mathbf{g}_C) \circ \langle \mathbf{v} \rangle + \langle \mathbf{v} \circ \mathbf{v} \rangle + \sigma_n^2 \mathbf{I}. \quad (23)$$

Since Eq. (23) cannot be directly solved for  $\mathbf{P}$ , the goal is to find the best estimator  $\mathbf{g}_C$  for  $\mathbf{g}$ , so that  $\mathbf{P} = \langle \mathbf{g}_C \circ \mathbf{g}_C \rangle$  is obtained as an approximate solution of Eq. (23). Therefore, a filter matrix  $\mathbf{w}$  is constructed, that separates the parameterized  
 145 field from the noise by projecting the measured data onto the parameter space

$$\mathbf{w}^\dagger \langle \mathbf{B} \rangle = \mathbf{g}_C \quad (24)$$

and simultaneously truncates the non-parameterized parts, i.e.

$$\mathbf{w}^\dagger \langle \mathbf{v} \rangle = 0. \quad (25)$$

Applying the filter matrix to the average of the non-parameterized parts of the field in Eq. (4)

$$150 \quad 0 = \mathbf{w}^\dagger \langle \mathbf{v} \rangle = \mathbf{w}^\dagger (\langle \mathbf{B} \rangle - \mathbf{H}\mathbf{g}_C) = \mathbf{w}^\dagger \langle \mathbf{B} \rangle - \mathbf{w}^\dagger \mathbf{H}\mathbf{g}_C = \mathbf{g}_C - \mathbf{w}^\dagger \mathbf{H}\mathbf{g}_C, \quad (26)$$

where the true coefficient vector has been replaced by Capon's estimator (Eq. 16) and taking into account that  $\langle \mathbf{n} \rangle = 0$  leads to the distortionless constraint

$$\mathbf{w}^\dagger \mathbf{H} = \mathbf{I}. \quad (27)$$

This equation is one of the important constraints for the construction of the wanted filter matrix  $\mathbf{w}$  but it is not enough.  
 155 Let us look for another criterion. The basic idea is that in Eq. (4) there may be contributions  $\langle \mathbf{v} \rangle$  in the data  $\mathbf{B}$  which are not caused by the internal magnetic field and thus, these contributions are not modeled by  $\mathbf{H}\mathbf{g}$ . Although the filter matrix  $\mathbf{w}$  is already designed to truncate these parts, i.e.  $\mathbf{w}^\dagger \langle \mathbf{v} \rangle = 0$ , their contributions to the data are unknown. This yields the following procedure.



Conferring to Eq. (24) the average output power, which is defined as the sum of the quadratic averaged components of the  
 160 estimator, can be rewritten as (Pillai , 1989)

$$\text{tr} \mathbf{P} = \text{tr} \langle \mathbf{g}_C \circ \mathbf{g}_C \rangle = \text{tr} (\mathbf{w}^\dagger \langle \mathbf{B} \circ \mathbf{B} \rangle \mathbf{w}), \quad (28)$$

where  $\text{tr} \langle \mathbf{g}_C \circ \mathbf{g}_C \rangle$  is the trace of the matrix  $\langle \mathbf{g}_C \circ \mathbf{g}_C \rangle$ .

Using Eq. (28), the coefficient matrix  $\mathbf{P}$  can be expressed by the weight  $\mathbf{w}$  and the data covariance  $\mathbf{M}$  as

$$165 \quad \mathbf{P} = \langle \mathbf{g}_C \circ \mathbf{g}_C \rangle = \mathbf{w}^\dagger \langle \mathbf{B} \circ \mathbf{B} \rangle \mathbf{w} = \mathbf{w}^\dagger \mathbf{M} \mathbf{w}. \quad (29)$$

Since the amount of the data's noise and the amount of the non-parameterized parts are unknown, we assume here, that a large  
 part of the data are influenced by the noise and the non-parameterized parts, that have to be truncated by the matrix  $\mathbf{w}$  and the  
 underlying model keeps the minimal contribution to the data. This minimal contribution has to be extracted.

170 Therefore, the output power  $P = \text{tr} \mathbf{P}$  has to be minimized with respect to  $\mathbf{w}^\dagger$ , subject to the distortionless constraint  $\mathbf{w}^\dagger \mathbf{H} = \mathbf{I}$   
 or equivalently  $\mathbf{H}^\dagger \mathbf{w} = \mathbf{I}$ . Using the Lagrange multiplier method this minimization problem can be formulated as

$$\text{minimize} \quad \text{tr} [\mathbf{w}^\dagger \mathbf{M} \mathbf{w} + \Lambda (\mathbf{I} - \mathbf{H}^\dagger \mathbf{w})], \quad (30)$$

where  $\Lambda$  are the related Lagrange multipliers and the minimum is taken with respect to  $\mathbf{w}$ . Since the components  $w_{ij}$  and  $w_{ij}^\dagger$   
 of the matrix  $\mathbf{w}$  and  $\mathbf{w}^\dagger$  respectively, are independent of each other, Eq. (30) can be expanded as

$$175 \quad \text{minimize} \quad P = \text{tr} [\mathbf{w}^\dagger \mathbf{M} \mathbf{w} + \Lambda (\mathbf{I} - \mathbf{H}^\dagger \mathbf{w}) + (\mathbf{I} - \mathbf{w}^\dagger \mathbf{H}) \Gamma] \quad (31)$$

or equivalently

$$\text{minimize} \quad P = w_{ij}^\dagger M_{jk} w_{ki} + \Lambda_{ii} - \Lambda_{ij} H_{jk}^\dagger w_{ki} + \Gamma_{ii} - w_{ij}^\dagger H_{jk} \Gamma_{ki} \quad (32)$$

with related additional Lagrange multipliers  $\Gamma$ . Taking the derivatives with respect to  $w_{ki}$  and  $w_{ij}^\dagger$  results in

$$0 = \partial_{w_{ki}} P = w_{ij}^\dagger M_{jk} - \Lambda_{ij} H_{jk}^\dagger \quad (33)$$

180 yielding

$$\mathbf{w}^\dagger \mathbf{M} = \Lambda \mathbf{H}^\dagger \quad (34)$$

and

$$0 = \partial_{w_{ij}^\dagger} P = M_{jk} w_{ki} - H_{jk} \Gamma_{ki} \quad (35)$$



resulting in

$$185 \quad \mathbf{M}\mathbf{w} = \mathbf{H}\mathbf{\Gamma} \quad . \quad (36)$$

Multiplication of Eq. (34) with  $\mathbf{w}$  from the right and multiplication of Eq. (36) with  $\mathbf{w}^\dagger$  from the left considering the distortionless constraint delivers

$$\mathbf{P} = \mathbf{w}^\dagger \mathbf{M}\mathbf{w} = \mathbf{\Gamma} = \mathbf{\Lambda} \quad (37)$$

and therefore

$$190 \quad P = \text{tr}\mathbf{P} = \text{tr}\mathbf{\Gamma} = \text{tr}\mathbf{\Lambda}. \quad (38)$$

Because  $P = \text{tr}\mathbf{P}$  is a convex function,  $\mathbf{\Lambda}$  and  $\mathbf{\Gamma}$  are realizing the minimal output power.

Due to the ensemble averaging the matrix  $\mathbf{M}$  is invertible and  $\mathbf{M}^{-1}$  exists. Multiplying Eq. (36) with  $\mathbf{H}^\dagger \mathbf{M}^{-1}$  and again considering the distortionless constraint yields

$$195 \quad \mathbf{P} = \mathbf{\Lambda} = \mathbf{\Gamma} = [\mathbf{H}^\dagger \mathbf{M}^{-1} \mathbf{H}]^{-1}. \quad (39)$$

By means of Eq. (34) the filter matrix results in

$$\mathbf{w}^\dagger = \mathbf{P} \mathbf{H}^\dagger \mathbf{M}^{-1} \quad (40)$$

and therefore Capon's estimator is given by

$$g_C = [\mathbf{H}^\dagger \mathbf{M}^{-1} \mathbf{H}]^{-1} \mathbf{H}^\dagger \mathbf{M}^{-1} \langle \mathbf{B} \rangle. \quad (41)$$

200 Regarding the expensive derivation, this compact formular for Capon's estimator is surprising. The same expression can be derived by treating Capon's method as a special case of the maximum likelihood estimator (Narita, 2019).

205 Since Capon's method has formerly been applied to the analysis of waves, the existing derivations treat the non-parameterized parts of the field as Gaussian noise. Considering the analysis of planetary magnetic fields this assumption is indefensible. For example, the external parts of the field are systematic noise and cannot be modeled by a Gaussian distribution. Therefore, the above presented derivation generalizes the previous derivations of Capon's method.

#### 4 Diagonal Loading

The filter matrix  $\mathbf{w}$  is the key parameter to distinguish between the parameterized and the non-parameterized parts of the field. Conferring to Eq. (23) the ratio of these parts defines the input signal-noise-ratio

$$210 \quad SNR_i = \frac{\text{tr}(\mathbf{H}\mathbf{P}\mathbf{H}^\dagger)}{\text{tr}(2(\mathbf{H}g_C) \circ \langle \mathbf{v} \rangle + \langle \mathbf{v} \circ \mathbf{v} \rangle + \sigma_n^2 \mathbf{I})}. \quad (42)$$



The filter matrix is applied to the disturbed data for estimating the output power, that is related to Capon's estimator. Thus, the output signal-noise-ratio can be expressed as (Haykin, 2014; Van Trees, 2002)

$$SNR_o = \frac{\text{tr}(\mathbf{w}^\dagger \mathbf{H} \mathbf{P} \mathbf{H}^\dagger \mathbf{w})}{\text{tr}(2 \mathbf{w}^\dagger (\mathbf{H} \mathbf{g}_C) \circ \langle \mathbf{v} \rangle \mathbf{w} + \mathbf{w}^\dagger \langle \mathbf{v} \circ \mathbf{v} \rangle \mathbf{w}) + \sigma_n^2 \text{tr}(\mathbf{w}^\dagger \mathbf{w})} = \frac{\text{tr}(\mathbf{P})}{\sigma_n^2 \text{tr}(\mathbf{w}^\dagger \mathbf{w})}, \quad (43)$$

since the filter fulfills the distortionless constraint  $\mathbf{w}^\dagger \mathbf{H} = \mathbf{I}$  and truncates the non-parameterized parts of the field, i.e.  $\mathbf{w}^\dagger \langle \mathbf{v} \rangle = 0$ . The ratio of the output and the input signal-noise-ratio is the so-called array gain (Van Trees, 2002)

$$\frac{SNR_o}{SNR_i} = \frac{1}{\text{tr}(\mathbf{w}^\dagger \mathbf{w})} \frac{\text{tr}(\mathbf{P})}{\sigma_n^2 \text{tr}(\mathbf{H} \mathbf{P} \mathbf{H}^\dagger)} \text{tr}(2(\mathbf{H} \mathbf{g}_C) \circ \langle \mathbf{v} \rangle + \langle \mathbf{v} \circ \mathbf{v} \rangle + \sigma_n^2 \mathbf{I}) \sim \frac{1}{\text{tr}(\mathbf{w}^\dagger \mathbf{w})}, \quad (44)$$

which is controlled by  $\text{tr}(\mathbf{w}^\dagger \mathbf{w})$ .

The input signal-noise-ratio is given by the data and the underlying model and therefore,  $SNR_i = \text{const.}$  in Eq. (44). In contrast to the input signal-noise-ratio, the output signal-noise-ratio depends on the filtering. When  $\text{tr}(\mathbf{w}^\dagger \mathbf{w})$  is large, the output signal-noise-ratio can decrease ( $SNR_o \rightarrow 0$ ), resulting in signal elimination and thus, the performance of Capon's estimator degrades. To prevent the signal elimination and to improve the robustness of Capon's estimator it is desirable to restrict  $\text{tr}(\mathbf{w}^\dagger \mathbf{w})$  with an upper boundary  $T_0 = \text{const.}$  (Van Trees, 2002), which can be expressed by the additional quadratic constraint

$$\text{tr}(\mathbf{w}^\dagger \mathbf{w}) = T_0. \quad (45)$$

For reasons of mathematical aesthetics, the constant  $T_0$  is expressed as the trace of a matrix  $\mathbf{T}$ . For example, one can choose

$$\mathbf{T} = \frac{T_0}{G} \mathbf{I} \in \mathbb{R}^{G \times G}, \quad (46)$$

where  $G$  again indicates the number of wanted expansion coefficients and  $\mathbf{I} \in \mathbb{R}^{G \times G}$  is the identity matrix, so that

$$\text{tr}(\mathbf{T}) = \frac{T_0}{G} \text{tr}(\mathbf{I}) = T_0 \quad (47)$$

holds. Thus, Eq. (45) can be rewritten as

$$\text{tr}(\mathbf{w}^\dagger \mathbf{w} - \mathbf{T}) = 0. \quad (48)$$

It should be noted that the matrix  $\mathbf{T}$  can be chosen arbitrarily as long as it is independent of  $\mathbf{w}$  and  $\mathbf{w}^\dagger$ .

Conferring to the previous section and considering the additional quadratic constraint, the filter matrix can be calculated by solving

$$\text{minimize} \quad \text{tr}[\mathbf{w}^\dagger \mathbf{M} \mathbf{w} + \sigma_d^2 (\mathbf{w}^\dagger \mathbf{w} - \mathbf{T}) + \boldsymbol{\Lambda} (\mathbf{I} - \mathbf{H}^\dagger \mathbf{w}) + (\mathbf{I} - \mathbf{w}^\dagger \mathbf{H}) \boldsymbol{\Gamma}] \quad (49)$$

with respect to  $\mathbf{w}$ , where  $\sigma_d^2$  is the related additional Lagrange multiplier. Carrying out the same procedure as described in section 3, the constrained minimizer is given by

$$\mathbf{w} = (\mathbf{M} + \sigma_d^2 \mathbf{I})^{-1} \mathbf{H} [\mathbf{H}^\dagger (\mathbf{M} + \sigma_d^2 \mathbf{I})^{-1} \mathbf{H}]^{-1}. \quad (50)$$



240 The comparison of Eq. (40) with Eq. (50) shows that the quadratic constraint results in the addition of the constant value  $\sigma_d^2$  to the diagonal of the data covariance matrix  $\mathbf{M}$ , which is known as diagonal loading (Van Trees, 2002). Consequently, the filter matrix is designed for a higher Gaussian background noise than is actually present (Van Trees, 2002).

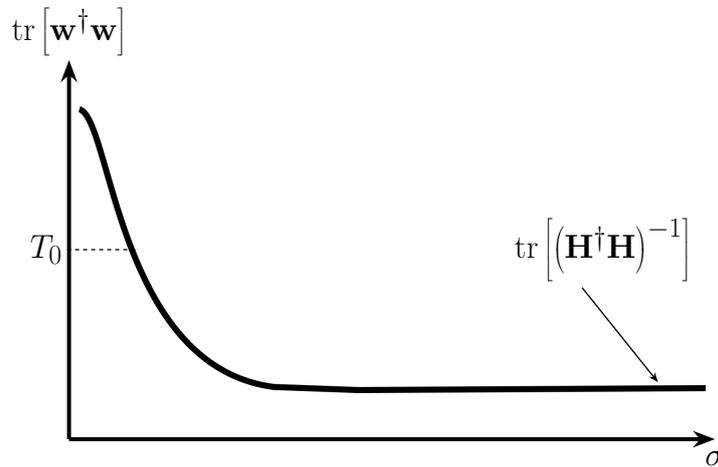
In Figure 1  $\text{tr}(\mathbf{w}^\dagger \mathbf{w})$  is displayed with respect to  $\sigma = \sqrt{\sigma_d^2 + \sigma_n^2}$ . For  $\sigma_d \rightarrow 0$ ,  $\text{tr}(\mathbf{w}^\dagger \mathbf{w})$  is large resulting in a small output  
 245 signal-noise ratio, which can cause signal elimination. For increasing values of  $\sigma_d$ ,  $\text{tr}(\mathbf{w}^\dagger \mathbf{w})$  decreases and for  $\sigma_d \rightarrow \infty$  Capon's filter converges to the least square fit filter

$$\mathbf{w}^\dagger \xrightarrow{\sigma_d \rightarrow \infty} [\mathbf{H}^\dagger \mathbf{H}]^{-1} \mathbf{H}^\dagger \quad (51)$$

or equivalently

$$\mathbf{w}^\dagger \mathbf{w} \xrightarrow{\sigma_d \rightarrow \infty} [\mathbf{H}^\dagger \mathbf{H}]^{-1}, \quad (52)$$

250 that treats all data equally.



**Figure 1.** Sketch of the trace of the filter matrix with respect to  $\sigma = \sqrt{\sigma_d^2 + \sigma_n^2}$ . For increasing values of the diagonal loading parameter  $\sigma_d$  the output signal-noise ratio increases and Capon's filter converges to the least square fit filter, if the diagonal loading parameter is large.

Since  $\text{tr}(\mathbf{w}^\dagger \mathbf{w})$  is a monotonically decreasing function,  $\sigma_d \rightarrow \infty$  might be the best choice for the diagonal loading parameter. To check this expectation we have to take a look at the output power  $P_d$  for the synthetic increased noise, which can be estimated by replacing  $\mathbf{M} \rightarrow \mathbf{M} + \sigma_d^2 \mathbf{I}$  in Eq. (39), resulting in (Pajovic, 2019; Richmond et al., 2005)

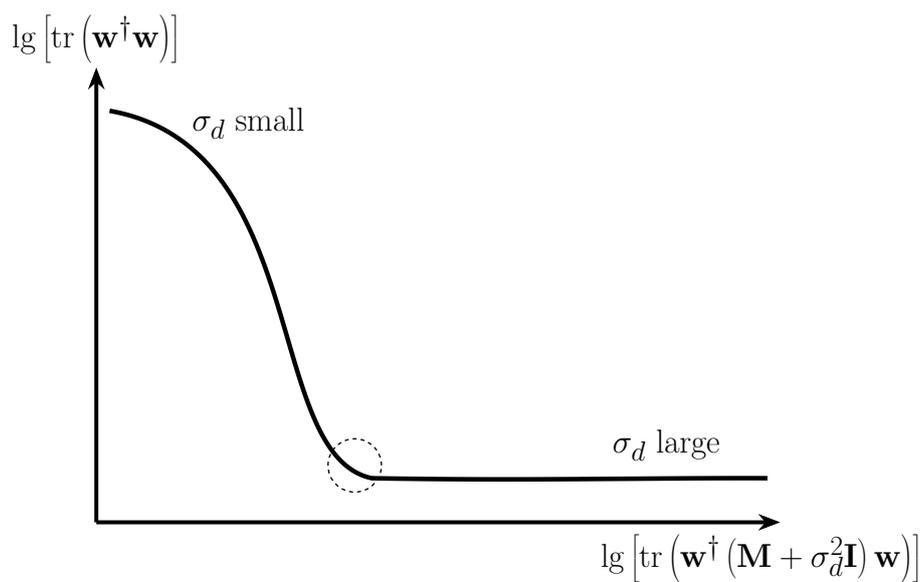
$$255 \quad P_d = \text{tr}[\mathbf{w}^\dagger (\mathbf{M} + \sigma_d^2 \mathbf{I}) \mathbf{w}] = \text{tr}[\mathbf{H}^\dagger (\mathbf{M} + \sigma_d^2 \mathbf{I})^{-1} \mathbf{H}]^{-1} \xrightarrow{\sigma_d \rightarrow \infty} \sigma_d^2 \text{tr}[\mathbf{H}^\dagger \mathbf{H}]^{-1}. \quad (53)$$



Thus,  $P_d$  is an increasing function of  $\sigma_d$ . Since the output power has to be minimized one would expect that  $\sigma_d = 0$  is the best choice, which is in contradiction to the argumentation above. Therefore, the maximization of the array gain is not equivalent to the minimization of the output power (Van Trees, 2002). Since  $\sigma_d$  cannot be directly calculated within the minimization procedure (Eq. 49), it is not clear how to choose the optimal diagonal loading parameter  $\sigma_{opt.}$ , that lies somewhere between those extrema.

In the literature there exist several methods for determining the optimal diagonal loading parameter (Pajovic, 2019). In contrast to the analysis of waves, we favor measurements that are stationary up to the Gaussian noise for the analysis of planetary magnetic fields. Comparing the measurement times with planetary geology time scales this assumption is surely valid for the internal magnetic field. For the external parts of the field this can be realized by choosing data sets of preferred situations, e.g. calm solar wind conditions. By virtue of the stationarity, the data covariance matrix  $\mathbf{M} = \langle \mathbf{B} \rangle \langle \mathbf{B} \rangle + \sigma_n^2 \mathbf{I}$  contains only one non-trivial eigenvalue  $\lambda_1 = |\langle \mathbf{B} \rangle|^2 + \sigma_n^2$  and  $\lambda_i = \sigma_n^2$ , for  $i = 2, \dots, 3N$  elsewhere. Therefore, estimators for the diagonal loading parameter, that are related to eigenvalues corresponding to interference and noise (Carlson, 1988) or estimators taking into account the standard deviation of the diagonal elements of the data covariance matrix (Ma and Goh, 2003) cannot be applied. When simulated data are analyzed the deviation between Capon's estimator and the true coefficient vector, implemented in the simulation, is a useful metric for estimating the optimal diagonal loading parameter (Pajovic, 2019; Toepfer et al., 2020). But when the method is applied to real spacecraft data no true coefficient vector is available and therefore, another estimation method for the diagonal loading parameter, that solely depends on the data and the underlying model is required.

The additional quadratic constraint (Eq. 45), resulting in diagonal loading, bounds the trace of the filter matrix, that is the solution of the minimization procedure (Eq. 49). To prevent signal elimination,  $\text{tr}(\mathbf{w}^\dagger \mathbf{w})$  and  $P_d$  have to be minimized simultaneously. Since  $\text{tr}(\mathbf{w}^\dagger \mathbf{w})$  is decreasing for higher values of  $\sigma_d$  (cf. Figure 1),  $P_d$  is an increasing function and thus, they act as competitors. At the optimal value  $\sigma_{opt.}$  the two competitors compromise. Therefore, the diagonal loading of the data covariance matrix is equivalent to the Tikhonov regularization for ill-posed problems and the optimal diagonal loading parameter can be estimated analogously by the method of the L-curve (Hiemstra et al. (2002)). The L-curve arises by plotting  $\lg[\text{tr}(\mathbf{w}^\dagger \mathbf{w})]$  versus  $\lg[\text{tr}(\mathbf{w}^\dagger (\mathbf{M} + \sigma_d^2 \mathbf{I}) \mathbf{w})]$  for different values of  $\sigma_d$  and is displayed in Figure 2. The optimal value  $\sigma_{opt.}$  is located in the vicinity of the L-curve's knee, which is defined by the maximum curvature of the L-curve (Hiemstra et al., 2002).



**Figure 2.** Sketch of the L-curve for estimating the optimal diagonal loading parameter  $\sigma_d$ . The optimal value is located in the vicinity of the L-curves knee (dashed circle).



## 5 Practical application of Capon's method

285 When Capon's method is applied to the reconstruction of Mercury's internal magnetic field (Toepfer et al., 2020), several computationally burdensome matrix inversions are necessary to calculate the estimator

$$g_C = [\mathbf{H}^\dagger \mathbf{M}^{-1} \mathbf{H}]^{-1} \mathbf{H}^\dagger \mathbf{M}^{-1} \langle \mathbf{B} \rangle. \quad (54)$$

Thus, for the practical application of Capon's method it is useful to rewrite the method in terms of the least square fit method.

290 Capon's estimator inherits the matrix operation structure of the least square fit estimator, which is given by (Haykin, 2014)

$$g_L = [\mathbf{H}^\dagger \mathbf{H}]^{-1} \mathbf{H}^\dagger \langle \mathbf{B} \rangle. \quad (55)$$

Substituting  $\mathbf{H} \rightarrow \mathbf{M}^{-1/2} \mathbf{H}$  and  $\langle \mathbf{B} \rangle \rightarrow \mathbf{M}^{-1/2} \langle \mathbf{B} \rangle$  in Eq. (55), the least square fit estimator may be converted into Capon's estimator.

295 The least square fit method minimizes the deviation between the disturbed measurements  $\mathbf{B}$  and the underlying model  $\mathbf{H}\mathbf{g}$ , measured in the Euclidian norm  $\|\cdot\|_2$ , so that

$$g_L = \arg \min_g \|\mathbf{H}\mathbf{g} - \mathbf{B}\|_2^2. \quad (56)$$

Conferring to the above mentioned substitutions, Capon's method can be interpreted as measuring the deviation in Eq. (56) in a different norm

$$300 \|\mathbf{H}\mathbf{g} - \mathbf{B}\|_{\mathbf{M}^{-1}}^2 = \|\mathbf{M}^{-1/2}(\mathbf{H}\mathbf{g} - \mathbf{B})\|_2^2, \quad (57)$$

so that Capon's estimator is given by

$$g_C = \arg \min_g \|\mathbf{H}\mathbf{g} - \mathbf{B}\|_{\mathbf{M}^{-1}}^2. \quad (58)$$

Thus, Capon's method can be regarded as a special case of the least square fit method or more precisely of the weighted least square fit method, where the data and the model are measured and weighted with the inverse data covariance matrix.

305 This property is useful for the practical application of Capon's method. In contrast to the computationally burdensome matrix inversions in Eq. (41), the usage of minimization algorithms, such as gradient descent or conjugate gradient method, for solving Eq. (58) is more stable and computationally inexpensive.



To illustrate the above presented mathematical foundations, Capon's method is applied to simulated magnetic field data in order to reconstruct Mercury's internal magnetic field. As a proof of concept, the underlying model is restricted to the internal dipole and quadrupole contributions to the magnetic field as discussed in section 2.

For the reconstruction of Mercury's internal dipole and quadrupole field the internal Gauss coefficients  $g_1^0 = -190$  nT and  $g_2^0 = -78$  nT (Anderson et al., 2012; Wardinski et al., 2019), defining the non-vanishing components of the true coefficient vector  $\mathbf{g}$  are implemented in the hybrid code AIKEF (Müller et al., 2011) and the magnetic field data resulting from the plasma interaction of Mercury with the solar wind are simulated. The data are evaluated along selected parts of the prospective trajectories of the BepiColombo mission on the night side of Mercury within a distance of  $0.2 R_M$  up to  $0.4 R_M$  from Mercury's surface. Since simulated data are analyzed, the deviation between the true coefficient vector  $\mathbf{g}$  and Capon's estimator  $\mathbf{g}_C$  can be used as a metric to verify the estimation of the optimal diagonal loading parameter by making use of the L-curve technique. The optimal diagonal loading parameter results in  $\sigma_{opt.} \approx 800$  nT which corresponds with the vicinity of the L-curve's knee. The reconstructed Gauss coefficients for the internal dipole and quadrupole field are presented in Table 1.

**Table 1.** Implemented and reconstructed Gauss coefficients for the internal dipole and quadrupole field.

Gauss coefficient	input in nT	output Capon in nT
$g_1^0$	-190.0	-189.2
$g_1^1$	0.0	1.9
$h_1^1$	0.0	0.2
$g_2^0$	-78.0	-68.4
$g_2^1$	0.0	26.1
$h_2^1$	0.0	11.4
$g_2^2$	0.0	-2.4
$h_2^2$	0.0	0.0

The deviation  $|\mathbf{g}_C - \mathbf{g}|$  between Capon's estimator and the true coefficient vector results in 30.2 nT or 14.7%, respectively, for the optimal diagonal loading parameter. When the magnetic field data are evaluated at an ensemble of data points with a distance of  $0.2 R_M$  from Mercury's surface this deviation is of the same order (Toepfer et al., 2020). Extending the underlying model by a parameterization of the external parts of the magnetic field improves Capon's estimator especially when the magnetic field data are evaluated in some distance above Mercury's surface.



## 6 Conclusions

Capon's method is a robust and useful tool for various kinds of ill-posed inverse problems, such as Mercury's planetary magnetic field analysis. The derivation of the method can be regarded from different mathematical perspectives. Here we revisited the linear-algebraic matrix formulation of the method and extended the derivation for Mercury's magnetic field analysis.

330 Capon's method becomes even more robust by incorporating the diagonal loading technique. Thereby, the construction of a filter matrix is vital to the derivation of Capon's estimator.

Especially the trace of the filter matrix determines the array gain, which is defined as the ratio of the output and the input signal-noise-ratio. If the trace is large, the output signal-noise-ratio can decrease resulting in signal elimination and thus, the performance of Capon's estimator degrades. Bounding the trace of the filter matrix results in diagonal loading of the data covariance matrix, which improves the robustness of the method. The main problem of the diagonal loading technique is that in general it is not clear how to choose the optimal diagonal loading parameter. Since for the analysis of planetary magnetic fields we prefer measurements that are stationary up to the Gaussian noise, estimators for the diagonal loading parameter, that are related to eigenvalues of the data covariance matrix corresponding to interference and noise cannot be applied. Making use

335 of the L-curve's technique enables a robust procedure for estimating the optimal diagonal loading parameter.

340

For the calculation of Capon's estimator several computationally burdensome matrix inversions are necessary. Interpretation of Capon's method as a special case of the least square fit method enables the usage of numerically more stable and less burden minimization algorithms, e.g. gradient descent or conjugate gradient method, for calculating Capon's estimator.

345

It should be noted that the parameterization of Mercury's internal magnetic field via the Gauss representation, as mentioned in section 2, is only one of several possibilities for modeling the magnetic field in the vicinity of Mercury. The underlying model can be extended by other parameterizations, for example the Mie representation (toroidal-poloidal decomposition) or magnetospheric models and Capon's method can be applied to estimate the related model coefficients. Besides the analysis

350 of Mercury's internal magnetic field, the extension of the model also enables the reconstruction of current systems flowing in the magnetosphere. Concerning the BepiColombo mission this work establishes a mathematical basis for the application of Capon's method to analyze Mercury's internal magnetic field in a robust and manageable way.

*Data availability.* The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

*Author contributions.* All authors contributed conception and design of the study; ST, YN and UM wrote the first draft of the manuscript;

355 All authors contributed to manuscript revision, read and approved the submitted version.



*Funding:* We acknowledge support by the German Research Foundation and the Open Access Publication Funds of the Technische Universität Braunschweig.

The work by Y. Narita is supported by the Austrian Space Applications Programme at the Austrian Research Promotion Agency under contract 865967.

360 D. Heyner was supported by the German Ministerium für Wirtschaft und Energie and the German Zentrum für Luft- und Raumfahrt under contract 50 QW1501.

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* The authors are grateful for stimulating discussions and helpful suggestions by Karl-Heinz Glassmeier and Alexander Schwenke.



## 365 References

- Anderson, B. J., Johnson, C. L., Korth, H., Purucker, M. E., Winslow, R. M., Slavin, J. A., Solomon, S. C., McNutt Jr., R. L., Raines, J. M., and Zurbuchen, T. H.: The global magnetic field of Mercury from MESSENGER orbital observations, *Science*, 333, 1859–1862, <https://doi.org/10.1126/science.1211001>, 2011.
- Anderson, B. J., Johnson, C. L., Korth, H., Winslow, R. M., Borovsky, J. E., Purucker, M. E., et al.: Low-degree structure in Mercury's planetary magnetic field, *J. Geophys. Res.*, 117, E00L12, <https://doi.org/10.1029/2012JE004159>, 2012.
- 370 Benkhoff, J., van Casteren, J., Hayakawa, H., Fujimoto, M., Laakso, H., Novara, M., Ferri, P., Middleton, H. R., and Ziethe, R.: BepiColombo–Comprehensive exploration of Mercury: Mission overview and science goals, *Planet. Space Sci.*, 85, 1–2, 2–20, <https://doi.org/10.1016/j.pss.2009.09.020>, 2010.
- Capon, J.: High resolution frequency-wavenumber spectrum analysis, *Proc. IEEE*, 57, 1408–1418, <https://doi.org/10.1109/PROC.1969.7278>,  
375 1969.
- Carlson, B. D.: Covariance matrix estimation errors and diagonal loading in adaptive arrays, *IEEE Trans. Aerosp. Electron. Syst.*, 24, 397–401, <https://doi.org/10.1109/7.7181>, 1988.
- Constantinescu, O. D., Glassmeier, K.-H., Motschmann, U., Treumann, R. A., Fornaçon, K. -H., and Fränz, M.: Plasma wave source location using CLUSTER as a spherical wave telescope, *J. Geophys. Res.*, 111, A09221, <https://doi.org/10.1029/2005JA011550>, 2006.
- 380 Gauß, C. F.: Allgemeine Theorie des Erdmagnetismus: Resultate aus den Beobachtungen des magnetischen Vereins im Jahre 1838, edited by: Gauss, C. F. and Weber, W., 1–57, Weidmannsche Buchhandlung, Leipzig, 1839.
- Glassmeier, K.-H., Motschmann, U., Dunlop, M., Balogh, A., Acuña, M. H., Carr, C., Musmann, G., Fornaçon, K.-H., Schweda, K., Vogt, J., Georgescu, E., and Buchert, S.: Cluster as a wave telescope – first results from the fluxgate magnetometer, *Ann. Geophys.*, 19, 1439–1447, <https://doi.org/10.5194/angeo-19-1439-2001>, 2001.
- 385 Glassmeier, K.-H., Auster, H.-U., Heyner, D., Okrafka, K., Carr, C., Berghofer, G., Anderson, B. J., et al.: The fluxgate magnetometer of the BepiColombo Mercury Planetary Orbiter, *Planet. Space Sci.*, 58, 287–299, <https://doi.org/10.1016/j.pss.2008.06.018>, 2010.
- Haykin, S.: Adaptive Filter Theory (5th ed.), International edition, Pearson, 2014.
- Hiemstra, J. H., Wippert, M. W., Goldstein, J. S., and Pratt, T.: Application of the L-curve technique to loading level determination in adaptive beamforming, Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA,  
390 2002, 2, 1261–1266, 2002.
- Ma, N., and Goh, J. T.: Efficient method to determine diagonal loading value, *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, 5, 341–344, <https://doi.org/10.1109/ICASSP.2003.1199948>, 2003.
- Motschmann, U., Woodward, T. I., Glassmeier, K.-H., Southwood, D. J., and Pinçon, J.-L.: Wavelength and direction filtering by magnetic measurements at satellite arrays: Generalized minimum variance analysis, *J. Geophys. Res.*, 101, 4961–4966,  
395 <https://doi.org/10.1029/95JA03471>, 1996.
- Müller, J., Simon, S., Motschmann, U., Schüle, J., Glassmeier, K.-H., and Pringle, G. J.: A.I.K.E.F.: Adaptive hybrid model for space plasma simulations, *Comput. Phys. Commun.*, 182, 946–966, <https://doi.org/10.1016/j.cpc.2010.12.033>, 2011.
- Narita, Y.: A note on Capon's minimum variance projection for multi-spacecraft data analysis, *Front. Phys.*, 7, 8, <https://doi.org/10.3398/fphy.2019.00008>, 2019.



- 400 Narita, Y., Glassmeier, K.-H., Schäfer, S., Motschmann, U., Sauer, K., Dandouras, I., Fornaçon, K. H., Georgescu, E., and Rème, H.: Dispersion analysis of ULF waves in the foreshock using cluster data and the wave telescope technique, *Geophys. Res. Lett.*, 30, <https://doi.org/10.1029/2003GL017432>, 2003.
- Narita, Y., Nakamura, R., and Baumjohann, W.: Cluster as current sheet surveyor in the magnetotail, *Ann. Geophys.*, 31, 1605–1610, <https://doi.org/10.5194/angeo-31-1605-2013>, 2013.
- 405 Pajovic, M., Preisig, J. C., and Baggeroer, A. B.: Analysis of optimal diagonal loading for MPDR-based spatial power estimators in the snapshot deficient regime, *IEEE Journal of Oceanic Engineering*, 44, 451–465, <https://doi.org/10.1109/JOE.2018.2815480>, 2018.
- Pillai, S. U.: *Array Signal Processing*, Springer Verlag, New York, 1989.
- Plaschke, F., Glassmeier, K.-H., Constantinescu, O. D., Mann, I. R., Milling, D. K., Motschmann, U., and Rae, I. J.: Statistical analysis of ground based magnetic field measurements with the field line resonance detector, *Ann. Geophys.*, 26, 3477–3489, <https://doi.org/10.5194/angeo-26-3477-2008>, 2008.
- 410 Richmond, C. D., Nadakuditi, R. R., and Edelman, A.: Asymptotic mean squared error performance of diagonally loaded Capon-MVDR processor, *Conference Record of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computers 2005*, Pacific Grove, CA, 1711–1716, 2005.
- Toepfer, S., Narita, Y., Heyner, D., and Motschmann, U.: The Capon method for Mercury’s magnetic field analysis, *Front. Phys.*, 8, 249, <https://doi.org/10.3389/fphy.2020.00249>, 2020.
- 415 Van Trees, H. L.: *Detection, Estimation, and Modulation Theory, Optimum Array Processing*, Wiley, New York, 2002.
- Wardinski, I., Langlais, B., and Thébault, E.: Correlated time-varying magnetic field and the core size of Mercury, *J. Geophys. Res.*, 124, 2178–2197, <https://doi.org/10.1029/2018JE005835>, 2019.