Ann. Geophys., 25, 971–987, 2007
www.ann-geophys.net/25/971/2007/
© European Geosciences Union 2007

Annales Geophysicae

Least-squares gradient calculation from multi-point observations of scalar and vector fields: methodology and applications with Cluster in the plasmasphere

J. De Keyser[1], **F. Darrouzet**[1], **M. W. Dunlop**[2], **and P. M. E. Décréau**[3]

[1]Belgian Institute for Space Aeronomy (BIRA-IASB), Ringlaan 3, 1180 Brussels, Belgium
[2]Rutherford Appleton Laboratory (RAL), Chilton, Didcot, Oxfordshire OX11 0QX, UK
[3]Laboratoire de Physique et Chimie de l'Environnement (LPCE/CNRS), 3A Avenue de la Recherche Scientifique, 45071 Orléans Cedex 02, France

Received: 26 September 2006 – Revised: 30 March 2007 – Accepted: 4 April 2007 – Published: 8 May 2007

Abstract. This paper describes a general-purpose algorithm for computing the gradients in space and time of a scalar field, a vector field, or a divergence-free vector field, from in situ measurements by one or more spacecraft. The algorithm provides total error estimates on the computed gradient, including the effects of measurement errors, the errors due to a lack of spatio-temporal homogeneity, and errors due to small-scale fluctuations. It also has the ability to diagnose the conditioning of the problem. Optimal use is made of the data, in terms of exploiting the maximum amount of information relative to the uncertainty on the data, by solving the problem in a weighted least-squares sense. The method is illustrated using Cluster magnetic field and electron density data to compute various gradients during a traversal of the inner magnetosphere. In particular, Cluster is shown to cross azimuthal density structure, and the existence of field-aligned currents in the plasmasphere is demonstrated.

Keywords. Magnetospheric physics (Magnetospheric configuration and dynamics; Plasmasphere; Instruments and techniques)

1 Introduction

This paper deals with the computation of gradients of physical quantities (scalar or vector fields) that are measured in situ at different times and positions. This topic has gained importance in the context of recent magnetospheric multi-spacecraft missions, in particular the Cluster mission consisting of four identical spacecraft, flying in formation. The rationale behind this mission is the idea that exactly four simultaneous measurements are needed to determine the three spatial gradient components, at least if the four spacecraft are not coplanar. Methods to do so have been developed (Harvey, 1998; Chanteur, 1998; Chanteur and Harvey, 1998; Robert et al., 1998a). Such gradient computations are difficult for a number of reasons: (1) Four simultaneous measurement points are required. A smaller number of spacecraft is insufficient, while the method also cannot properly take advantage of a larger number of spacecraft if available. For those Cluster instruments that are not operating on the four spacecraft, this instantaneous spatial gradient computation is therefore precluded. Also, data are never obtained exactly simultaneously, a problem that is usually dealt with by time averaging or interpolation. (2) One never computes a gradient, but rather a spatial difference representing the average gradient over a length scale fixed by the spacecraft separation distances, usually different in each space dimension. This scale often does not correspond to the physical scale size of interest. (3) The method applies only if the so-called spatial homogeneity condition is satisfied, that is, it computes an average gradient over the spacecraft separation length scale, but that is only useful if the true gradient does not differ too much from the average one. (4) Computing differences is notoriously difficult. Subtracting two similar data values leads to a relative error on the difference that may be much larger than the error on the original data. This is especially true for spacecraft that are very close together, a situation often required to satisfy the spatial homogeneity condition. A difference can therefore be reliably computed only when both random and systematic errors on the data are very small. This necessitates accurate data intercalibration between the spacecraft, something that can be difficult to achieve as the operating conditions on each spacecraft tend to be different. Systematic gradient computations with Cluster have therefore been applied only to magnetic field data (FGM instrument; Balogh et al., 1997, 2001) with its low measurement error and good intercalibration (especially the curlometer; Dunlop et al., 2001, 2006; Dunlop and Balogh, 2005; Vallat et al., 2005) and to electron density data obtained from the plasma frequency (WHISPER instrument; Décréau et al.,

Correspondence to: J. De Keyser
(johan.dekeyser@bira-iasb.oma.be)

Published by Copernicus GmbH on behalf of the European Geosciences Union.

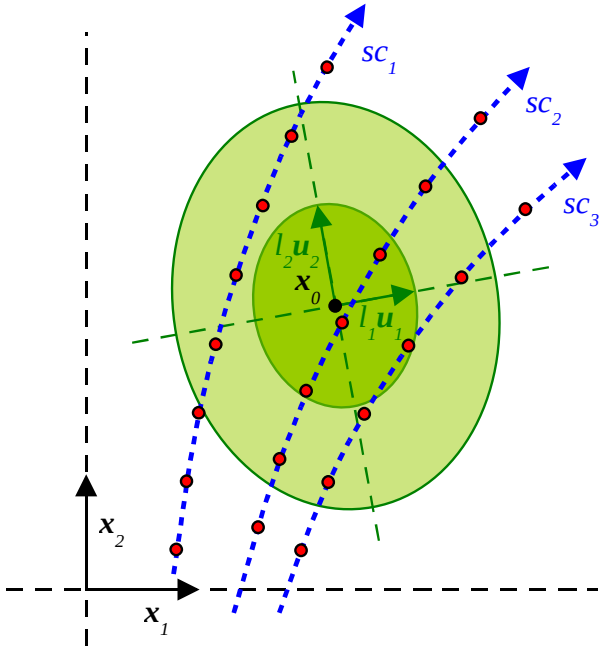

Fig. 1. The least-squares gradient algorithm uses data acquired in a set of points in space-time, represented here as a 2-dimensional space ($\boldsymbol{x}_1, \boldsymbol{x}_2$). In this example, the data are obtained along the trajectories of three spacecraft (red dots on the dotted lines), although that does not matter for the method. The homogeneity condition is expressed by associating with each data point an error that grows with the distance from $\boldsymbol{x}_0$, the point where the gradient is computed. This distance, however, is measured in a frame ($l_1\boldsymbol{u}_1$, $l_2\boldsymbol{u}_2$) that may be rotated and scaled relative to the original frame. Points on the ellipse with semi-axes l_1 and l_2 are assigned a unit distance. Points inside the ellipse (dark shaded area) correspond to smaller distances and therefore a smaller error. Points outside that ellipse (lightly shaded region) will have a larger error so that they are less relevant, thus reflecting the homogeneity condition. Points outside the shaded regions are considered irrelevant.

1997, 2001; Trotignon et al., 2003) because of their absolute calibration (Darrouzet et al., 2006b). (5) Additional errors arise due to the lack of synchronization of the spacecraft clocks, the nonzero duration over which data are gathered, and the uncertainties in the spacecraft positions. The impact of these errors may be hard to quantify, especially because they may not be statistically independent. Such errors tend to be relevant for short time scales and small separations only.

This paper describes an alternative way to determine the gradient obtained by relaxing the requirement of simultaneity of the observations. This is achieved by formulating the concepts of temporal and spatial homogeneity in a more general way. By using all observations in a region of space-time over which the spatial and temporal gradients are essentially constant over prescribed length and time scales, an overdetermined system of equations is obtained from which the gradient can be computed in a least-squares sense. In principle, data from an arbitrary number of spacecraft can be exploited. It is possible to attempt to compute gradients on length and time scales that match the physical scales of interest. In any given practical situation the method will find out whether such gradients can actually be computed with the available data. For in situ measurements by the Cluster spacecraft in a medium that is at rest, for instance, the scales along the spacecraft orbit will usually be limited by the time resolution of the data, while the scales perpendicular to the velocity will be dictated by the orbital separations. Proper error estimates are derived that account for the measurement errors, for the errors due to the fact that the gradient is not constant over the region in which measurements are available, and for the effect of small-scale structures and perturbing wave fields. It is also possible to take into account geometrical and spatio-temporal constraints.

In Sects. 2–5, the method is presented in a formal mathematical way. We use linear algebra techniques (such as eigen-decomposition and singular value decomposition) and therefore adopt the standard linear algebra notation: Bold lower-case symbols represent vectors, bold upper-case symbols are matrices, other symbols denote scalars. Sections 6–8 illustrate the application of the method for scalar fields, for vector fields, and for divergence-free vector fields with Cluster examples for a pass through the inner magnetosphere.

2 The problem for scalar fields

One or more spacecraft sample a scalar field $f(\boldsymbol{x}, t)$ at positions and times $\boldsymbol{x}_i{=}[x_i; y_i; z_i; t_i]$, $i{=}1, \ldots, N$, in 4-dimensional space-time. The case of vector fields is discussed later. The measurements f_i have known error variances $\delta f^2_{\mathrm{meas},i}$. Data intercalibration removes all systematic errors (we disregard clock synchronization and spacecraft position errors).

The gradient $\nabla_{xt} f{=}[\partial f/\partial x; \partial f/\partial y; \partial f/\partial z; \partial f/\partial t]$ at a point $\boldsymbol{x}_0$ can be computed by combining all measurements made inside a region in which the gradient does not change appreciably. This region is the "homogeneity domain". Its size is determined by physical considerations. It can be described by a 4-dimensional ellipsoid, that is, by four mutually orthonormal directions $\boldsymbol{u}_j$ and the corresponding scales l_j, specifying the axes of the hyperellipsoid. This is illustrated in Fig. 1 by means of an analogy in 2-dimensional space. In this example, the points (red dots) are obtained along the trajectories of three spacecraft (dotted lines), although that does not matter for the method. The points inside the homogeneity domain (represented by the dark shaded ellipse) can safely be used to compute the gradient in $\boldsymbol{x}_0$. Rather than accepting points inside the homogeneity domain for the computation and rejecting points outside, we will use a more gradual approach. With each data point, we will associate an error that grows with the distance from $\boldsymbol{x}_0$. This distance, however, is measured in a frame ($l_1\boldsymbol{u}_1$, $l_2\boldsymbol{u}_2$) that may be rotated and

scaled relative to the original (x_1, x_2) frame. This frame will be called the β-frame, and the unit coordinate vectors of this frame are exactly the semi-axes of the elliptical homogeneity domain. Points on the border of the homogeneity domain are assigned a unit distance. Points inside the ellipse correspond to smaller distances and therefore a smaller error. Points farther outside that ellipse have a larger error so that they are of progressively diminishing relevance, thus reflecting the homogeneity condition.

The transformation $\boldsymbol{x}^\beta = \mathbf{P}\boldsymbol{x}$, where $\mathbf{P} = (\mathbf{UL})^{-1}$ with orthonormal matrix $\mathbf{U} = [\ldots \boldsymbol{u}_j \ldots]$ and diagonal matrix $\mathbf{L} = \mathrm{diag}(l_j)$ (this notation means: the diagonal matrix with the l_j on the diagonal) represents the transition from the original frame to the new frame, denoted by superscript β, in which the homogeneity domain is the unit hypersphere. The corresponding gradient operator is $\boldsymbol{\nabla}_{xt}^\beta = \mathbf{P}^{-1\top}\boldsymbol{\nabla}_{xt} = \mathbf{L}\mathbf{U}^\top\boldsymbol{\nabla}_{xt}$. The distance measure that is needed to express the homogeneity condition assigns to each vector $\boldsymbol{x}$ a length or norm $\|\boldsymbol{x}\|_\beta = \sqrt{\boldsymbol{x}^{\beta\top}\boldsymbol{x}^\beta} = \sqrt{\boldsymbol{x}^\top\mathbf{U}\mathbf{L}^{-2}\mathbf{U}^\top\boldsymbol{x}}$.

Homogeneity considerations in time are often separable from spatial homogeneity. One of the $\boldsymbol{u}_j$ must then be $[0; 0; 0; 1]$. If the time scale is $l_t = \tau$ and if the three spatial homogeneity scales are $l_x = l_y = l_z = \xi$, the transformation is simply a rescaling $\mathbf{P} = \mathrm{diag}([1/\xi, 1/\xi, 1/\xi, 1/\tau])$. The β-norm of a given vector $\boldsymbol{x}$ then is $\|\boldsymbol{x}\|_\beta = \sqrt{(x/l_x)^2 + (y/l_y)^2 + (z/l_z)^2 + (t/l_t)^2}$. We will refer to this particular case as the standard isotropic homogeneity case.

If f satisfies the appropriate analyticity conditions near $\boldsymbol{x}_0$, it can be locally approximated by a Taylor expansion. With $\boldsymbol{\Delta x}_i = \boldsymbol{x}_i - \boldsymbol{x}_0$, and denoting the function value and the gradient at $\boldsymbol{x}_0$ by f_0 and $\boldsymbol{\nabla}_{xt} f_0$, this becomes

$$f_i = f_0 + \boldsymbol{\Delta x}_i{}^\top \boldsymbol{\nabla}_{xt} f_0 + r_i, \quad (1)$$

where the residual is $r_i = \mathcal{O}(\|\boldsymbol{\Delta x}_i\|^2)$. This leads to a system of N equations for f_0 and $\boldsymbol{\nabla}_{xt} f_0$, i.e., $M = 5$ unknowns:

$$\boldsymbol{r} = f_0 + \boldsymbol{\Delta X}^\top \boldsymbol{\nabla}_{xt} f_0 - \boldsymbol{f} = f_0 + \boldsymbol{\Delta X}^{\beta\top} \boldsymbol{\nabla}_{xt}{}^\beta f_0 - \boldsymbol{f} = \boldsymbol{0} \quad (2)$$

where $\boldsymbol{\Delta X} = [\ldots \boldsymbol{\Delta x}_i \ldots]$ groups all relative positions and $\boldsymbol{f}$ denotes all measurements. The homogeneity domain can usually be chosen big enough so that $N \gg M$: The system is overdetermined and can never be satisfied exactly. However, the gradient can be computed in a least-squares sense (minimization of $\boldsymbol{r}^\top\boldsymbol{r}$). Expressing this system in the β-frame (the dimensionless form) is preferable from the numerical point of view (since all system matrix elements then are of order unity). Although system (2) is usually overdetermined, it may still be ill-conditioned. Such ill-conditioning can often be avoided by adding a priori knowledge in the form of constraints. In the case of Cluster it may become possible to compute gradients with only three, two, or even a single spacecraft, depending on the number of constraints. We will not use such constraints in the present paper, but Appendix A explains how they can be incorporated.

3 Approximation errors

In order to obtain the gradient, we approximate the scalar field f, with all its space-time variations, by a local linear approximation. There are two aspects to this approximation: The function might have variations at scales smaller than the specified homogeneity scale (this variability usually cannot be evaluated completely from the measurements, so a statistical approach is needed to estimate its effect), and there is also the curvature of f at the homogeneity scale itself, which is why the linear approximation is only valid inside the homogeneity domain. Both contribute to the total "approximation error"; the small-scale errors are the "fluctuation errors" and the errors due to the linear approximation are the "curvature errors". We will therefore write the scalar field as $f = f_{\mathrm{hs}} + \delta f_{\mathrm{ls}} + \delta f_{\mathrm{ss}}$, the sum of the linear field f_{hs} at the homogeneity scale (i.e. the linear approximation), a deviation δf_{ls} due to variations at larger scales, and a small-scale field δf_{ss}.

3.1 Structure at large scales

Structure at large scales ($\Delta x^\beta = \|\boldsymbol{\Delta x}\|_\beta \geq 1$) is properly represented by the Taylor expansion of Eq. (1). With the constant f_c indicating how much f changes over the homogeneity domain due to the higher-order terms in the expansion (function curvature), the curvature error is estimated as

$$\delta f_{\mathrm{ls}}^2(\boldsymbol{\Delta x}_i) \leq f_c^2 (\Delta x_i^\beta)^4,$$

so that the linear approximation is valid when $\Delta x_i^\beta < 1$ but not much beyond that. The curvature error is completely determined by the homogeneity conditions through the β-norm. The user must specify the value of f_c based on physical considerations. This may not be trivial, as will be discussed in Sect. 6 and in the Conclusions. Note that the value of f_c is linked to the homogeneity lengths: halving the homogeneity lengths is equivalent to multiplying f_c by a factor of four. In principle, the curvature errors at the various sampling positions are not statistically independent, but since nothing is known about them a priori, their cross-correlations are ignored here. This is justified even more so because, as explained in Sect. 5, only little weight will be given to data points far from $\boldsymbol{x}_0$ for which the cross-correlations would be large.

If the system is intrinsically changing on a short time scale, shorter than the sampling time scale (time-separable case with $l_t < t_{\mathrm{sample}}$), successive measurements cannot be related to each other since the system has changed in between. The homogeneity condition therefore indicates that only simultaneous measurements can be used for computing the gradient. Indeed, data taken at a different time have a large curvature error because t_{sample}/l_t and therefore Δx^β is large. When using only simultaneous data, the last row in $\boldsymbol{\Delta X}$ vanishes so that $\partial f/\partial t$ remains undetermined. For four spacecraft

this corresponds to the classical spatial gradient computation (Harvey, 1998; Chanteur, 1998; Chanteur and Harvey, 1998).

3.2 Structure at small scales

Small-scale structure is often present, for instance in the form of small-amplitude waves or turbulence. Small-scale perturbations are usually under-sampled, so their influence on gradient precision must be characterized with a stochastic model.

The $\delta f_{\rm ss}(\boldsymbol{x})$ can be thought of as the superposition of individual perturbations, each with length scales λ_j along mutually orthogonal directions, which we take to be the homogeneity directions $\boldsymbol{u}_j$ for the sake of simplicity, and with amplitude $\delta f_{{\rm ss}\boldsymbol{\lambda}}(\boldsymbol{\lambda}, \boldsymbol{x})$:

$$\delta f_{\rm ss}(\boldsymbol{x}) = \int_{\boldsymbol{\lambda}} \delta f_{{\rm ss}\boldsymbol{\lambda}}(\boldsymbol{\lambda}, \boldsymbol{x})\, d\boldsymbol{\lambda}.$$

Let the population of perturbations $\delta f_{\rm ss}$ be characterized by typical length scales $\hat{\boldsymbol{\lambda}}$. We can then introduce a new reference frame γ, similar to frame β but with different axis scaling, which leads to a norm $\|\boldsymbol{x}\|_\gamma = \sqrt{\boldsymbol{x}^\top \mathbf{U}\boldsymbol{\Lambda}^{-2}\mathbf{U}^\top \boldsymbol{x}}$ where $\boldsymbol{\Lambda}=\mathrm{diag}(\hat{\boldsymbol{\lambda}})$. Restricting ourselves to distributions that are isotropic in γ-space and assuming that the perturbation amplitudes do not vary appreciably over the homogeneity domain, one has $\langle(\delta f_{{\rm ss}\boldsymbol{\lambda}}(\boldsymbol{\lambda}, \boldsymbol{x}))^2\rangle = \delta f_\lambda^2(\lambda^\gamma)$ everywhere, with $\lambda^\gamma = \|\boldsymbol{\lambda}\|_\gamma$ and where the acute brackets identify the expected value for the population of perturbations. Because of the locality of the perturbations, $\langle \delta f_{{\rm ss}\boldsymbol{\lambda}}(\boldsymbol{\lambda}, \boldsymbol{x}_i)\delta f_{{\rm ss}\boldsymbol{\lambda}}(\boldsymbol{\lambda}, \boldsymbol{x}_j)\rangle \approx \delta f_\lambda^2(\lambda^\gamma)$ when $\Delta x_{ij}^\gamma = \|\boldsymbol{x}_j - \boldsymbol{x}_i\|_\gamma \ll \lambda^\gamma$ and zero when $\Delta x_{ij}^\gamma \gg \lambda^\gamma$. Appendix B computes that, for a distribution with perturbation strength $\delta f_\lambda^2(\lambda^\gamma)$ decreasing exponentially, the covariances are

$$\langle \delta f_{\rm ss}(\boldsymbol{x}_i)\delta f_{\rm ss}(\boldsymbol{x}_j)\rangle = f^{*2} \mathrm{e}^{-\Delta x_{ij}^\gamma},$$

with f^{*2} the total perturbation variance. The cross-correlation is large between nearby points and vanishes as their distance exceeds the perturbation length. Better models are possible if the type of perturbation (e.g. a particular wave mode) is known a priori. In such cases it would be best to compute the gradients of several wave field components simultaneously, coupled through the wave relations.

4 The problem for vector fields

The gradients of the individual components of a vector field can be obtained by treating each component individually as a separate scalar field under the simplifying assumption that the curvature errors are not correlated (or that one does not know a priori how) and that the small-scale perturbations for the different components are uncorrelated as well (which is not really true if they are due to a particular wave mode). Different homogeneity parameters for each of the components (different $\boldsymbol{u}_j$, l_j, f_c, $\hat{\lambda}_j$, f^*) could be used. Here, the discussion is limited to the case of identical values. The number of unknowns at each point is $M=3\times5=15$. For divergence-free vector fields (such as the magnetic field) the gradients of the vector components must be computed simultaneously, subject to the constraint that the divergence vanishes, so that $M=14$.

Computing the curl and the divergence of the vector field poses an additional difficulty. As the divergence and each of the components of the curl are sums of terms of the same order of magnitude, but possibly with opposite sign, the relative error on the result can be larger than the relative errors on the individual gradient components, which themselves are differences of similar values and have a significant uncertainty.

5 Solving the overdetermined system

The overdetermined system (2) expressing N measurements of a scalar field (the number of equations) can be written as

$$\boldsymbol{r} = \mathbf{A}\boldsymbol{q} - \boldsymbol{f} = \mathbf{0}, \tag{3}$$

with $\mathbf{A} = \left[1,\, \boldsymbol{\Delta}\mathbf{X}^{\beta\top}\right]$ and where $\boldsymbol{q}=[f_0;\, \nabla^\beta_{xI} f_0]$ is the vector of $M=5$ unknowns. The total error on f_i as used in the gradient computation at $\boldsymbol{x}_0$ is $\delta f_i^2 = (\delta f_{{\rm meas},i})^2 + f_c^2 (\Delta x_i^\beta)^4 + f^{*2}$, in which the terms represent the independent contributions of measurement error, curvature error and fluctuation error. In addition, there are the cross-correlations $\delta f_{ij}^2 = f^{*2}\mathrm{e}^{-\Delta x_{ij}^\gamma}$. These estimates give the $N\times N$ correlation matrix $\mathbf{C}_f = \langle \boldsymbol{\delta f}\, \boldsymbol{\delta f}^\top\rangle$, which is real, symmetric, and positive definite. It is strongly diagonal dominant if the $\delta f_{{\rm meas},i}$ and f_c are large, if f^* is small, or if the $\hat{\lambda}_j$ are small. As N may be large, significant computing time and storage savings can be achieved by setting the off-diagonals δf_{ij}^2 to zero if Δx_{ij}^γ is above an appropriately chosen limit, thus ignoring small cross-correlations (typically $\mathrm{e}^{-\Delta x_{ij}^\gamma} < 10^{-4}$).

The eigen-decomposition $\mathbf{C}_f = \mathbf{W}^\top \mathbf{D}^2 \mathbf{W}$ is computed, where $\mathbf{W}$ contains the orthonormal eigen-vectors and where the non-negative eigen-values, which are denoted by d_i^2, constitute the diagonal matrix $\mathbf{D}^2 = \mathrm{diag}(d_i^2)$. This is trivial if $\mathbf{C}_f$ is diagonal, i.e., when we do not consider small-scale fluctuations. Applying operator $\mathbf{W}$ on the left to the vectors $\boldsymbol{r}$, $\boldsymbol{f}$, and $\mathbf{A}\boldsymbol{q}$ in the overdetermined system (3), the errors on the transformed residuals are now statistically independent, so that the i-th equation represents an individual piece of information corresponding to a covariance d_i^2. A further operation by $\mathbf{D}^{-1}$ on the left normalizes the residuals so that they all have unit variance, by dividing each by its error estimate. The resulting weighted system is

$$\tilde{\boldsymbol{r}} = \tilde{\mathbf{A}}\boldsymbol{q} - \tilde{\boldsymbol{f}} = \mathbf{0}, \tag{4}$$

where $\tilde{\mathbf{A}} = \mathbf{D}^{-1}\mathbf{W}\mathbf{A}$, $\tilde{\boldsymbol{f}} = \mathbf{D}^{-1}\mathbf{W}\boldsymbol{f}$, and $\tilde{\boldsymbol{r}} = \mathbf{D}^{-1}\mathbf{W}\boldsymbol{r}$. Performing the equivalent operations on the correlation matrix $\mathbf{C}_f$ gives $\tilde{\mathbf{C}}_f = \mathbf{D}^{-1}\mathbf{W}\mathbf{C}_f\mathbf{W}\mathbf{D}^{-1} = \mathbf{I}$: There are no cross-correlations, and the variances are unity. The least-squares

method minimizes $\tilde{\boldsymbol{r}}^\top\tilde{\boldsymbol{r}}$ so that equations with less relevant information (d_i^2 large) will hardly play a role. Large measurement errors and short-scale fluctuations reduce the weights, and data points outside the homogeneity domain have such important curvature errors that their weight is negligible, thus ensuring that the information used for computing the gradient comes from within the homogeneity domain. Consider the situation depicted in Fig. 1. Spacecraft 2 passes near $\boldsymbol{x}_0$ and has three points inside the homogeneity domain (the dark shaded ellipse), the middle of which carries only measurement and fluctuation errors while the curvature error is small there. Spacecraft 1 does not have any point inside the homogeneity domain. Nevertheless, its data points just outside the homogeneity domain (in the light shaded region) will also appear in the overdetermined system. As their curvature errors are larger, their weights will be smaller. They therefore contribute only a limited amount of information in the computation of the gradient.

The least-squares solution is obtained by formally solving the overdetermined system as

$$\boldsymbol{q} = (\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^\top\tilde{\boldsymbol{f}}.$$

In practice, computing $\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}}$ (an $M\times M$ symmetric positive semidefinite matrix) is avoided because it implies sums with many terms and therefore a potential loss of precision. The standard method is to compute a decomposition $\tilde{\mathbf{A}}=\mathbf{QR}$, where $\mathbf{Q}$ is a unitary $N\times M$ matrix and $\mathbf{R}$ is an $M\times M$ upper triangular matrix with the so-called economy-size QR-algorithm. One can then easily compute $\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}}=\mathbf{R}^\top\mathbf{R}$. This symmetric $M\times M$ matrix can be inverted by computing its singular value decomposition $\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}}=\mathbf{V}^\top\mathbf{S}^2\mathbf{V}$, where $\mathbf{S}=\mathrm{diag}(s_j)$ are the singular values and $\mathbf{V}$ is unitary, so that

$$\boldsymbol{q} = \mathbf{V}^\top\mathbf{S}^{-2}\mathbf{V}\tilde{\mathbf{A}}^\top\tilde{\boldsymbol{f}}. \tag{5}$$

The correlation of errors on the result is $\mathbf{C}_q=(\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}})^{-1}$, from which the errors on the result can be estimated as $\boldsymbol{\delta q}=\sqrt{\mathrm{diag}(\mathbf{C}_q)}$ by ignoring the cross-correlations between the solution components. From $\boldsymbol{q}=[f_0;\nabla^\beta_{xt}f_0]$ and $\boldsymbol{\delta q}=[\delta f_0;\delta(\nabla^\beta_{xt}f_0)]$, the gradient and the error estimates are found as $\nabla_{xt}f_0=\mathbf{UL}^{-1}\nabla^\beta_{xt}f_0$ and $\delta(\nabla_{xt}f_0)=\mathbf{UL}^{-1}\delta(\nabla^\beta_{xt}f_0)$.

As can be seen from Eq. (5), the solution is obtained by applying the transformation $\mathbf{V}^\top\mathbf{S}^{-2}\mathbf{V}$ to the weighted data $\tilde{\mathbf{A}}^\top\tilde{\boldsymbol{f}}$. Since $\mathbf{V}$ is unitary, the conditioning of the problem is completely determined by the singular values s_j. If a singular value is small, the propagation of errors in the associated direction is important. For spatial gradient computations using simultaneous data from the four Cluster spacecraft, the concepts of planarity and elongation of the spacecraft tetrahedron have been used as a diagnostic for the well-posedness of the problem (Robert et al., 1998a,b), which have the advantage of being easily visualized. The singular values, however, provide an abstract but general diagnostic that works for an arbitrary number of spacecraft and in the presence of constraints (Appendix A). We therefore define the condition number

$$\mathrm{cond} = \min_j s_j^2 / \max_j s_j^2. \tag{6}$$

Even when the problem is close to singular (condition number small) the method produces a valid result, but the error estimates on the result (or, at least, on some of its components) will be large. As results with too large errorbars are useless, computations with cond$<10^{-5}$ are ignored.

The technique described here is based on an unconstrained approximation of f. If the scalar field is strictly positive, it is best to apply the technique to $\bar{f}=\log f$, with $\delta\bar{f}=\delta f/f$. The results are transformed back using $\nabla_{xt}f=\mathrm{e}^{\bar{f}}\ \nabla_{xt}\bar{f}$ and $\delta(\nabla_{xt}f)=\mathrm{e}^{\bar{f}}[\delta(\nabla_{xt}\bar{f})+\delta\bar{f}\ \nabla_{xt}\bar{f}]$.

When computing the gradients of the three components of a vector field $\boldsymbol{B}$, supplemented by the condition $\nabla\cdot\boldsymbol{B}=0$, the overdetermined system is

$$\begin{bmatrix} \mathbf{1}\ \Delta\mathbf{X}^{\beta\top} & \mathbf{0}_{N\times5} & \mathbf{0}_{N\times5} \\ \mathbf{0}_{N\times5} & \mathbf{1}\ \Delta\mathbf{X}^{\beta\top} & \mathbf{0}_{N\times5} \\ \mathbf{0}_{N\times5} & \mathbf{0}_{N\times5} & \mathbf{1}\ \Delta\mathbf{X}^{\beta\top} \\ 0\ \boldsymbol{u}_1^\top/l_1 & 0\ \boldsymbol{u}_2^\top/l_2 & 0\ \boldsymbol{u}_3^\top/l_3 \end{bmatrix} \begin{bmatrix} B_{x0} \\ \nabla^\beta_{xt}B_{x0} \\ B_{y0} \\ \nabla^\beta_{xt}B_{y0} \\ B_{z0} \\ \nabla^\beta_{xt}B_{z0} \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}_{x,N\times1} \\ \boldsymbol{B}_{y,N\times1} \\ \boldsymbol{B}_{z,N\times1} \\ 0 \end{bmatrix}$$

with a 3×3 block-diagonal coefficient matrix supplemented by the zero-divergence condition.

In the particular situation in which all data have been acquired simultaneously, the coefficients of the time derivatives are all zero and no information about the time variations can be extracted. The corresponding unknowns can be removed from the system, so that $M=4$ for the gradient of a scalar field, $M=12$ for a vector field, and $M=11$ for a divergence-free vector field.

6 Gradients of a scalar field

In this section, we consider the passage of the four Cluster spacecraft through the inner magnetosphere on 7 August 2003, from 06:00–11:00 UT, with perigee around 08:05 UT. A subinterval of this passage has been studied earlier by Darrouzet et al. (2006b). We focus on the gradients of the magnetic field strength $|\boldsymbol{B}|$ obtained by the FGM magnetometer and of the electron density n_e derived from the plasma frequency f_p measured by the WHISPER instrument. The spacecraft separation distances were on the order of $200\times400\times1000$ km in the GSE X, Y, Z directions near perigee. The spacecraft cross the inner magnetosphere from the south to the north. Figure 2 shows that $|\boldsymbol{B}|$ reaches a local minimum and n_e a maximum around perigee, at a geocentric distance of about 4.53 R_E. Near perigee, the spacecraft enter the outer regions of the plasmasphere ($K_p=2^+$, down from 6^- one day before, indicating post-storm recovery, a

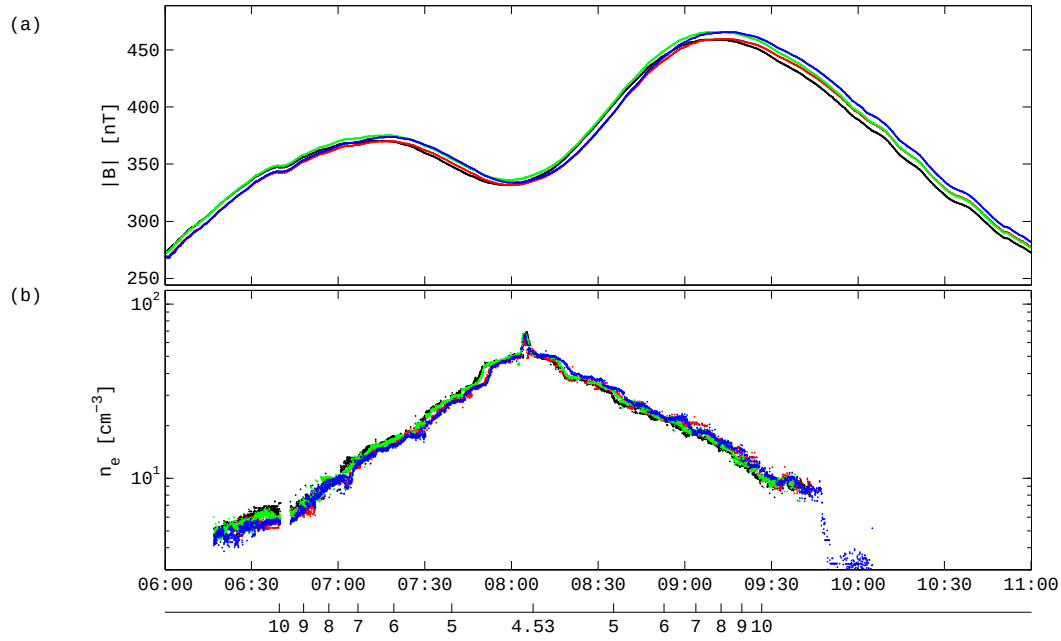

Fig. 2. Cluster observations during a pass through the inner magnetosphere on 7 August 2003, from 06:00–11:00 UT, with perigee around 08:02 UT: **(a)** magnetic field strength $|\boldsymbol{B}|$ obtained from the FGM magnetometer, and **(b)** electron density n_e computed from the plasma frequency as identified by the WHISPER instrument. The spacecraft separation distances were $200\times400\times1000$ km in the GSE X, Y, Z directions near perigee. $|\boldsymbol{B}|$ reaches a local minimum and n_e a maximum around perigee (C1 – black, C2 – red, C3 – green, C4 – blue). The bottom scale gives the L-shell position of the center of the Cluster tetrahedron (for $L<10$, elsewhere L cannot be determined accurately).

situation in which the plasmapause typically is located rather close to Earth). The plasma encountered near perigee (between 07:40 and 08:40 UT) is of plasmaspheric origin, as indicated by the Cluster plasma spectrometers. CIS/CODIF on Cluster 4, for instance, detects He^+ and O^+. The spectrometers also indicate corotating flow of a few km/s, although the measurements are not very precise since the instruments miss a major fraction of the cold plasma distributions due to spacecraft potential effects. The $|\boldsymbol{B}|$ profiles are smooth with minor variations at the begin and the end of the pass, when the spacecraft are outside the plasmasphere and sample higher L-values at higher magnetic latitudes, and where n_e is low. FGM is very precise and well calibrated. Spin-averaged magnetic field data (4 s time resolution) are used here. The measurement error on the components is 0.1 nT while the uncertainty on $|\boldsymbol{B}|$ is 0.15 nT (< 0.05 %). These data appear to allow an accurate gradient determination since the magnetic field values registered by the four spacecraft at any given time differ by up to 10 nT, larger than the measurement errors. The measurement error on f_p is the 163 Hz discretization error (half of the frequency resolution of WHISPER). Additional errors due to the possible misidentification of the plasma frequency line in the WHISPER spectrograms have been kept to a minimum as the algorithms for plasma frequency detection have matured (Trotignon et al., 2001, 2003, 2006; Rauch et al., 2006). The relation between density n_e and plasma frequency f_p is

$$n_e[\mathrm{cm}^{-3}] = (f_p[\mathrm{kHz}]/9.0)^2,$$

from which $\delta n_e/n_e = 2\delta f_p/f_p$. The peak density is almost 70 cm^{-3} with an error of 0.3 cm^{-3}. The relative error is smallest for these high densities, typically 0.4 %, and increases up to 20 % for lower densities near the detection limit.

First consider the standard isotropic homogeneity case with a characteristic spatial scale $l_x=l_y=l_z=500$ km. The homogeneity time scale depends on the time it takes structures to convect over these spatial scales as well as on intrinsic temporal changes. For typical plasmaspheric convection velocities of a few km/s, a homogeneity time scale $l_t=60$ s seems to be a natural choice compatible with the spatial scale. This is also sufficient to resolve the intrinsic temporal behaviour of plasmaspheric refilling (hours) or electric field reconfigurations involved in the creation of medium and large structures (tens of minutes). The terrestrial dipole field strength at the equator is $B=B_{\mathrm{eq}}/L^3$ (where B_{eq} is the equatorial field strength at the surface), so that $d^2B/dL^2=12B_{\mathrm{eq}}/L^5$ from which we obtain a typical curvature error estimate $f_c=12B_{\mathrm{eq}}(l_x/R_E)^2/L^5\sim1$ nT for $l_x=500$ km and $L=5$. A limitation of the present approach is that only a constant f_c can be specified, while in reality it can vary from point to point. No small-scale fluctuations are considered, $f^*=0$. In principle, all data points

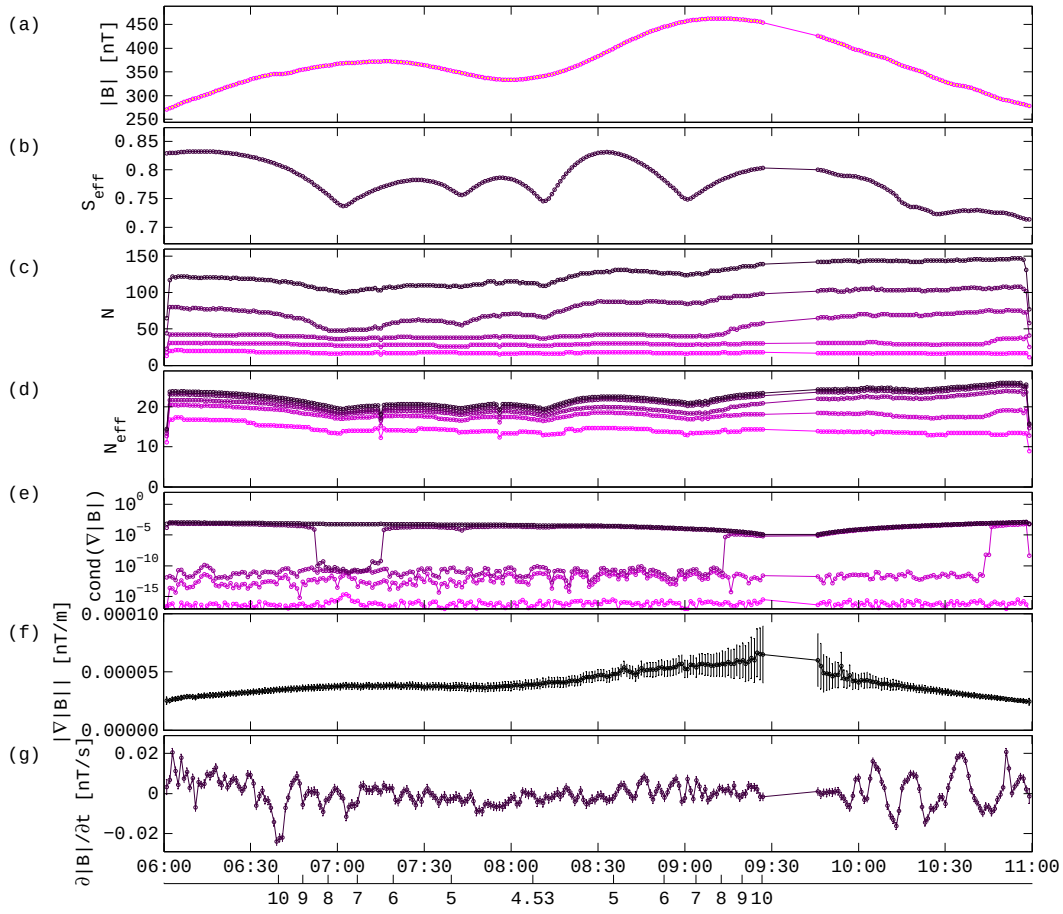

Fig. 3. Computation of $\nabla_{xt}|\boldsymbol{B}|$ every 60 s with isotropic homogeneity domain (see text). **(a)** magnetic field strength $|\boldsymbol{B}|$ at the moving center of the homogeneity domain; **(b)** effective scale factor S_{eff}; **(c)** number of equations N in the overdetermined system, where each curve refers to a threshold σ_k, with darker color shade as σ_k increases; **(d)** effective number of equations N_{eff}; **(e)** problem condition number for each σ_k; **(f)** magnitude of the spatial gradient $|\nabla|\boldsymbol{B}||$, with growing error bars and gap in the computed gradient due to spacecraft coplanarity; **(g)** temporal gradient $\partial|\boldsymbol{B}|/\partial t$. The bottom scale gives the L-shell position of the center of the Cluster tetrahedron.

can be included in the weighted system for determining the gradient at a given point. As the weights of most points are negligible, however, it is computationally more efficient to include only those points with a relative variance satisfying

$$\rho_i = \delta f_i^2 / \min_j \delta f_j^2 < \sigma_k,$$

where the σ_k is a series of increasing threshold values. In the example sketched in Fig. 1 this threshold corresponds to the outer ellipse: All points included in the computation form a set that is larger than the homogeneity domain (assuming that the measurement and fluctuation errors for all points are the same, otherwise there exists no simple geometrical representation). In practice, we use 5 logarithmically spaced values σ_k from 2 up to 100, for each of which the gradient is computed. The gradient computation is thus repeated with the outer ellipse being progressively expanded. The result is most precise for the largest threshold (largest N). Solving the problem for all σ_k allows us to assess whether the largest threshold is too low (a better result could be obtained) or unnecessarily large (computationally inefficient). Figure 3 illustrates the computation of $\nabla_{xt}|\boldsymbol{B}|$. It has been determined every 60 s. The magnetic field strength at the moving center of the homogeneity domain is a spatio-temporal average of the observed values (panel a). In order to find the physical scales of this averaging process, note that the region from which the weighted least-squares method will take most of its information (where homogeneity-scale error $\leq$ measurement error + fluctuation error) has a diameter or effective scale size D_j in the j-th homogeneity direction that can be computed from

$$(D_j/l_j)^4 = 16(\delta f_{\text{meas}}^2 + f^{*2})/f_c^2 = S_{\text{eff}}^4,$$

where S_{eff} is called the effective scale factor. It is about 0.8 in this example (panel b). If $\delta f_{\text{meas},i}$, f_c, and f^* are of the same order of magnitude, $S_{\text{eff}} \sim 1$ so that the effective scales correspond to the homogeneity scales. A general property

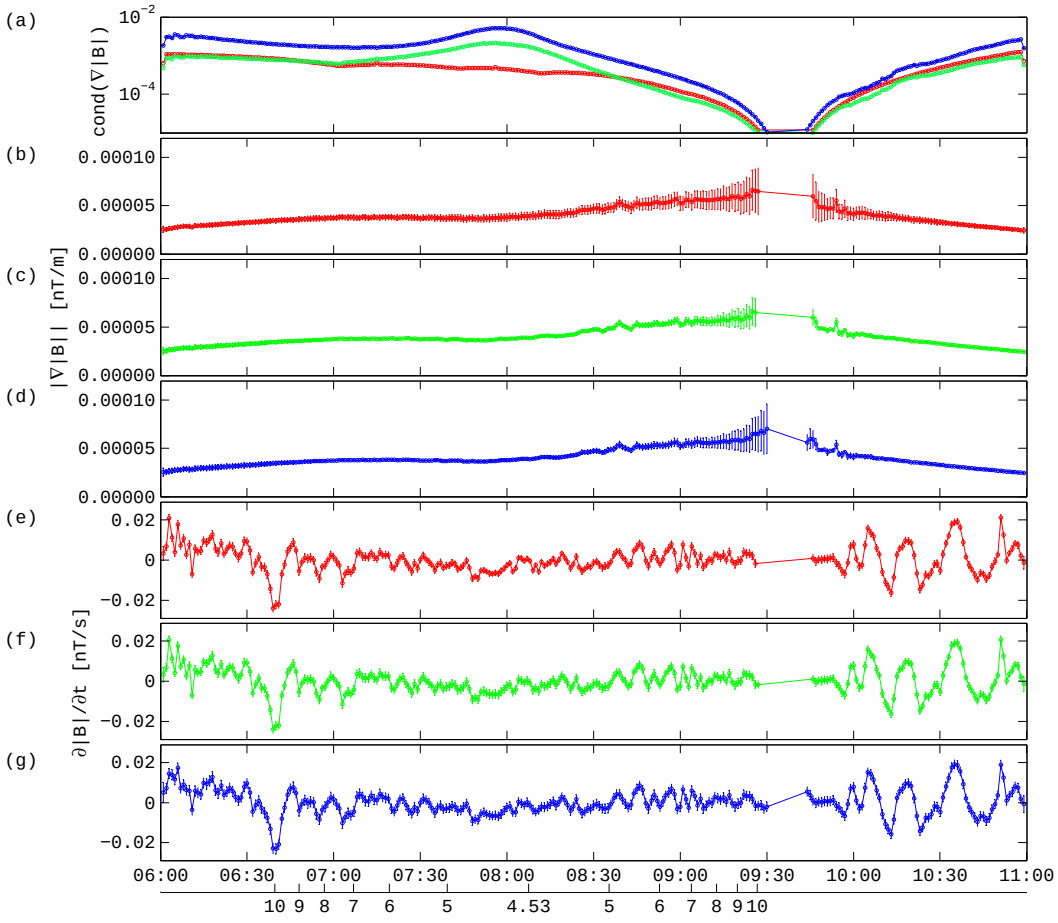

Fig. 4. Computation of $\nabla_{xt}|\boldsymbol{B}|$ with isotropic homogeneity domain (red), anisotropic domain (green), and anisotropic domain with small-scale fluctuations (blue) (see text). **(a)** condition number; **(b–d)** magnitude of spatial gradient $|\nabla|\boldsymbol{B}||$; **(e–g)** temporal gradient $\partial|\boldsymbol{B}|/\partial t$. The bottom scale gives the L-shell position of the center of the Cluster tetrahedron.

is that S_{eff} becomes smaller as f_c becomes larger, reflecting the equivalence between small homogeneity scales with a small f_c on the one hand, and larger homogeneity scales with a larger f_c on the other hand. The number of equations N actually used to compute the gradient (panel c, each curve refers to a σ_k) increases with σ_k. A measure for the amount of information used is the effective number of equations $N_{\text{eff}}=\sum_i 1/\rho_i$, which is the sum of the inverse relative variances of the corresponding equations, so that $N_{\text{eff}}\leq N$ (panel d). For ever larger σ_k, N_{eff} increases, but the relative gain decreases as the added equations have progressively lower weights. The convergence of the N_{eff}-curves indicates that $\max_k \sigma_k=100$ is well-chosen. Figure 3e presents an overview of the condition number as defined by Eq. (6). For low σ_k (N small, not necessarily using data from the four spacecraft) the problem is almost singular (cond$\sim 10^{-16}$), but for larger σ_k the condition number becomes very reasonable, typically 10^{-3}. It remains below 10^{-5} in the time interval 09:30–09:45 UT, so that the gradient cannot be sensibly computed there. The growing errorbars on the magnitude of the spatial gradient, $|\nabla|\boldsymbol{B}||$, correspond to this ill-conditioning (panel f), while no such large errorbars are present for the temporal gradient, $\partial|\boldsymbol{B}|/\partial t$ (panel g). This ill-conditioning is due to a purely spatial effect, namely the near-coplanarity of the spacecraft during this period. Elsewhere, the spatial gradient is well computed, with a value around 0.04 nT/km. The errorbars are on the order of 10%, but growing where the condition number gets worse. The temporal gradient varies around zero with an amplitude of $<$0.01 nT/s.

Medium-scale structures in the plasmasphere are often aligned with the magnetic field, especially at the rather low latitudes that Cluster is sampling here. It therefore makes sense to use an anisotropic homogeneity domain that is aligned with the magnetic field, with $l_x=l_y=500$ km perpendicular to $\boldsymbol{B}$ and $l_z=2000$ km parallel to $\boldsymbol{B}$. In addition, we may consider small-scale fluctuations. Their characteristic scales are taken to be 1/5th of the homogeneity scales ($\hat{\lambda}_x=\hat{\lambda}_y=100$ km, $\hat{\lambda}_z=400$ km, $\hat{\lambda}_t=12$ s) and $f^*=0.2$ nT. Figure 4 summarizes the results for the isotropic case, the anisotropic case, and the anisotropic case with fluctuations

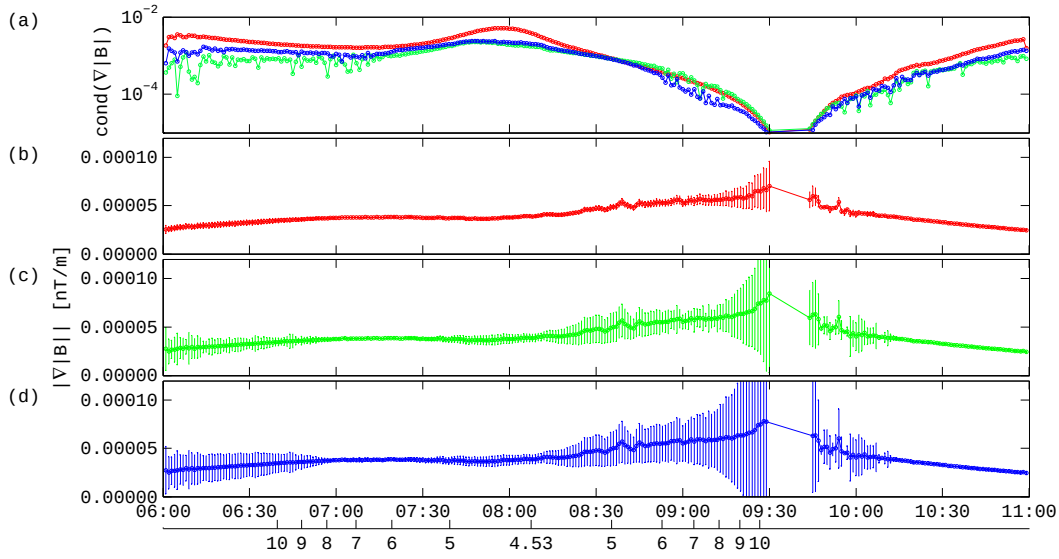

Fig. 5. Computation of $\nabla|\boldsymbol{B}|$ with anisotropic homogeneity domain and fluctuations (red), with the instantaneous gradient version (green), and with the instantaneous version with fluctuations (blue) (see text). **(a)** condition number; **(b–d)** magnitude of spatial gradient $|\nabla|\boldsymbol{B}||$. The bottom scale gives the L-shell position of the center of the Cluster tetrahedron.

(panels b, c, d and e, f, g, respectively). There is not much difference between the gradients themselves. Yet, more data points are available wherever the spacecraft trajectories are aligned with the direction of largest extent of the homogeneity domain (the magnetic field direction), i.e. near perigee. Consequently, N is larger there and the condition number improves (panel a). The error margins are smaller (down to 8% on the spatial gradient magnitude) because larger N implies more accurate averaging and because the total errors on the data are smaller. Adding fluctuations increases N and leads to a systematic improvement of the condition number. The error margins increase a little to about 10% on the spatial gradient magnitude.

As discussed in Sect. 5, the least-squares method can also be used to obtain the traditional instantaneous spatial gradient. To this end, we first time-average and resample the data onto a common time scale (60 s resolution in the present case), so as to obtain simultaneous data points. Note that time-averaging requires an appropriate evaluation of the error margins. The error on a time-average $f = \sum_{i=1}^{n} f_i/n$ can be estimated by

$$\delta f^2 = \frac{1}{n}\left[\frac{1}{n}\sum_{i=1}^{n}\delta f_{\mathrm{meas},i}^2 + \frac{1}{n-1}\sum_{i=1}^{n}(f_i - f)^2\right],$$

which takes into account both the measurement errors (of diminishing importance as the number of data points grows) and the time-variability of the observed quantity (assuming this variability to be gaussian, requiring an estimate of the standard deviation). Figure 5 compares the spatio-temporal gradient (anisotropic case with fluctuations) to the instantaneous spatial gradients for f^*=0 nT and f^*=0.2 nT (panels b, c, and d, respectively). For the 4-spacecraft Cluster case N=M=4. The weights then do not matter and the classical instantaneous spatial gradient method is recovered. The gradients obtained with or without fluctuation error are identical, but their error margins are not. There is a significant difference from the space-time gradient only near the coplanarity region. The condition number (panel a) is slightly different for the three computations. The error margins are smallest for the spatio-temporal gradient.

The gradient of the plasma density can be computed in similar ways. Figure 6 compares ∇n_e and $\partial n_e/\partial t$ obtained with f_c=0.5 cm^{-3} for the isotropic case, for the anisotropic case, and for the anisotropic case with f^*=0.2 cm^{-3} (panels c, d, e and f, g, h, respectively). As plasma density is a strictly positive quantity, a logarithmic scaling has been used. Computing gradients in data gaps is, in principle, always possible. However, the error associated with far-away data points will be large, depending on the homogeneity scales, so that the gradient will carry a large error. Often also the condition number becomes low. For instance, there is a data gap for the four spacecraft around 06:42. As it lasts only a few times the homogeneity time scale, the gradient can still be computed, but the condition number temporarily decreases (panel a). Something similar happens around 07:18 UT, due to a 5-min data gap for Cluster 2 only. Also note that the condition number has improved again for the anisotropic case near perigee, where the orbit is along the direction of largest extent of the homogeneity domain. The spatial gradient is well-computed except close to the coplanarity region around 09:40 UT. The error margins are larger there but still so small that they can hardly be seen in the figure. There

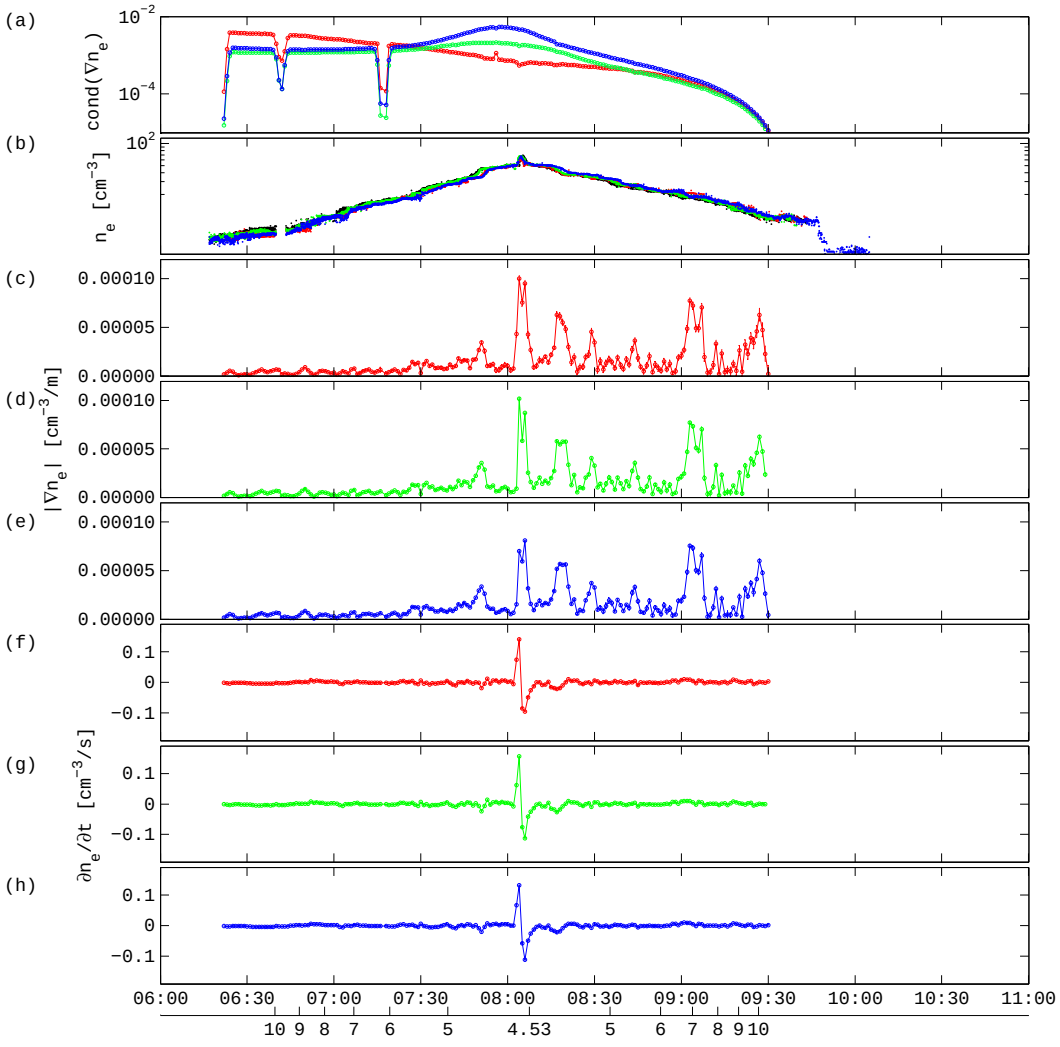

Fig. 6. Computation of $\nabla_{xt} n_e$ with isotropic homogeneity domain (red), anisotropic domain (green), and anisotropic domain with small-scale fluctuations (blue) (see text). **(a)** condition number; **(b)** n_e observed by the four spacecraft; **(c–e)** magnitude of spatial gradient $|\nabla n_e|$; **(f–h)** temporal gradient $\partial n_e/\partial t$. The bottom scale gives the L-shell position of the center of the Cluster tetrahedron.

is a double-peaked spatial gradient around 08:05 UT corresponding to the rising and falling slopes around the density peak observed near perigee (compare with panel b). There are also important gradients ($\sim$0.05 cm^{-3}/km, relative precision 10%) near 07:50 UT and near 08:15 UT. These strong gradients are identical to those reported by Darrouzet et al. (2006b) and interpreted there as proof of azimuthal structure. As the densities drop farther away from Earth, the gradients tend to become negligible well before and after perigee. The temporal gradient is pretty small ($<$0.01 cm^{-3}/s with a relative precision of 20% at best), except for the bipolar structure near perigee, which corresponds to the convection of the density structure. The three results are almost identical, but the error margins are again smallest ($<$8%) for the case of an anisotropic homogeneity domain.

Comparing $\nabla_{xt} n_e$ for the anisotropic case to the instantaneous spatial gradients without or with fluctuations (Fig. 7), the two instantaneous spatial gradients are necessarily found to be equal and there are only minor differences with the space-time gradient (panel c, d, and e). The condition number is not much different between the three computations (panel a), but it clearly is best for the space-time gradient. The space-time gradient also has the smallest error margins.

7 Gradients of a vector field

The gradients of the magnetic field vector components have been computed with the same anisotropic homogeneity domain, with f_c=1 nT and f^*=0.2 nT, treating the three

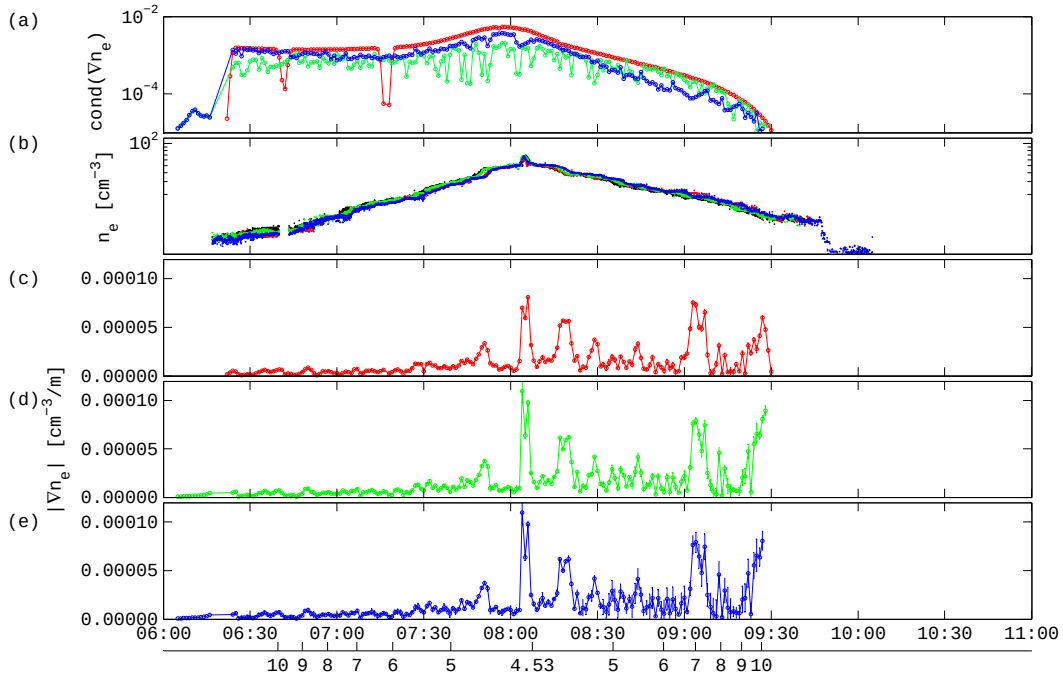

Fig. 7. Computation of ∇n_e with anisotropic homogeneity domain and fluctuations (red), with the instantaneous gradient version (green), and with the instantaneous version with fluctuations (blue) (see text). **(a)** condition number; **(b)** n_e observed by the four spacecraft; **(c–e)** magnitude of spatial gradient $|\nabla n_e|$. The bottom scale gives the L-shell position of the center of the Cluster tetrahedron.

components independently or coupling them through the zero-divergence constraint. From them, $\nabla\times\boldsymbol{B}$ and $\nabla\cdot\boldsymbol{B}$ are obtained, as shown in Fig. 8. The effective scale factor is close to unity (as for the $\nabla_{xt}|\boldsymbol{B}|$ computation in Sect. 6). $S_{\rm eff}$, N, and $N_{\rm eff}$ are identical for both computations. The number of equations N (panel a) is three times larger than for a scalar field. Both N and $N_{\rm eff}$ (panel b) peak near perigee as a consequence of the alignment between the spacecraft orbits and the direction with the longest homogeneity scale length (the magnetic field direction). The condition number is identical for both computations (panel c), although the overdetermined system sizes are different (same data used, but only 14 unknowns in the divergence-free case, rather than 15). The progressively deteriorating condition number toward 09:40 UT reflects the spacecraft coplanarity issue. The condition number is quite good elsewhere. The curl (panels d and e) and the divergence (panels f and g) from the uncoupled and the divergence-free computations have absolute error margins of typically 0.005 nT/km (a relative error of 10% on $|\nabla\times\boldsymbol{B}|$). Away from the coplanarity time interval, $\nabla\cdot\boldsymbol{B}$ does not significantly deviate from zero. The correction of the solution implied by requiring $\nabla\cdot\boldsymbol{B}=0$ is rather small. While the least-squares method minimizes the differences between the observations and the linear approximation, adding a constraint limits the solution search space so that the error margins become slightly larger. Nevertheless, adding physically relevant constraints obviously improves the realism of the solution.

Figure 9 compares the curl (panels b, c, and d) and divergence (panels e, f, and g) for the divergence-free spatio-temporal gradient computation as well as for the divergence-free instantaneous gradient computation, applied to 60 s averaged data, without and with small-scale fluctuations, respectively. While the results of the computations are essentially the same, the condition number (panel a) is best for the spatio-temporal gradient. It occasionally drops below 10^{-5} for the instantaneous gradient without fluctuations. The error margins are smallest for the spatio-temporal gradient despite the fact that only the spatio-temporal gradient error estimates account for the three error sources (measurement error, curvature error, and fluctuation error).

8 Physical relevance

In order to illustrate the usefulness of these gradients for scientific analysis, Fig. 10a shows the angles $\alpha_{\boldsymbol{B},\nabla|\boldsymbol{B}|}$ and $\alpha_{\boldsymbol{B},\nabla n_e}$ between the ambient magnetic field $\boldsymbol{B}$ on the one hand and $\nabla|\boldsymbol{B}|$ and ∇n_e on the other hand, in blue and red, respectively (anisotropic case with fluctuations). The Cluster spacecraft are sampling the outer regions of the plasmasphere, which happen to be the most dynamic ones where erosion can be important (e.g., Carpenter and Lemaire, 1997; Lemaire and Gringauz, 1998; Carpenter and Lemaire, 2004; Décréau et al., 2005). Cluster has therefore been used intensively to study these regions (e.g., Darrouzet et al., 2004,

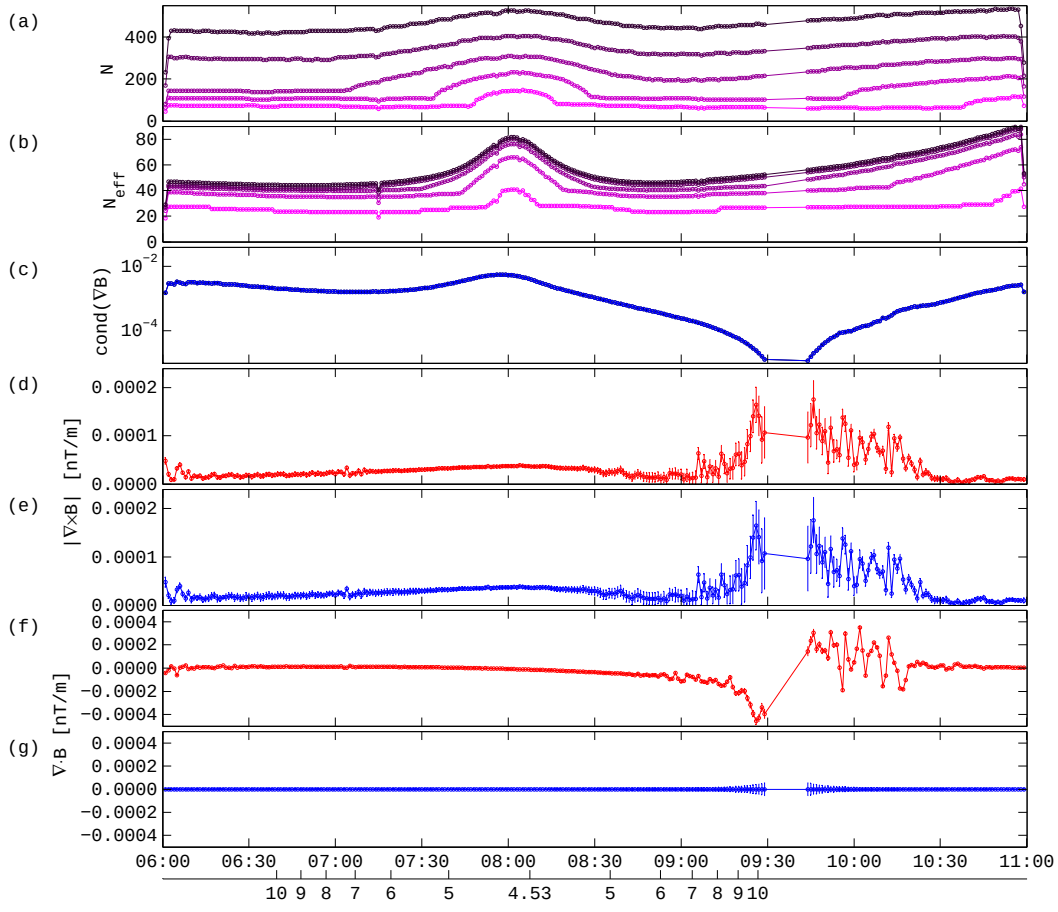

Fig. 8. Computation of $\nabla\times\boldsymbol{B}$ and $\nabla\cdot\boldsymbol{B}$ every 60 s with anisotropic homogeneity domain and fluctuations (see text) without (red) and with (blue) the $\nabla\cdot\boldsymbol{B}=0$ constraint. **(a)** number of equations N in the overdetermined system, with each curve referring to a threshold σ_k; **(b)** effective number of equations N_{eff}; **(c)** problem condition number; **(d–e)** magnitude of the curl $|\nabla\times\boldsymbol{B}|$, with growing error bars and gap in the computed gradient due to spacecraft coplanarity; **(f–g)** divergence $\nabla\cdot\boldsymbol{B}$. The bottom scale gives the L-shell position of the center of the Cluster tetrahedron.

2006a,b). As we are only interested in the direction of these gradients, not their sense, the angles are reduced to the interval $[0°, 90°]$. By definition, the magnetic field strength gradient is perpendicular to $\boldsymbol{B}$ at the magnetic equator, corresponding to $\alpha_{\boldsymbol{B},\nabla|\boldsymbol{B}|}=90°$, near 08:02 UT (close to, but not exactly at Cluster perigee). Before and after that time, as the Cluster spacecraft are at higher magnetic latitudes, that angle decreases rapidly because of the progressively more important field-aligned gradient. The error margins are large away from perigee as both $\boldsymbol{B}$ and $\nabla|\boldsymbol{B}|$ are small there. Angle $\alpha_{\boldsymbol{B},\nabla n_e}$ remains quite large throughout the time interval, indicating that $\nabla_{||}n_e \ll \nabla_{\perp}n_e$ at the relatively low latitudes Cluster is sampling, something that has also been found with radio sounding techniques (Reinisch et al., 2001). This is due to a small longitudinal gradient within each flux tube, but also because of the existence of strong radial and azimuthal density structure on the transverse homogeneity scale of 500 km adopted here. The gradient orientations obtained here compare very well to the instantaneous gradient directions reported by Darrouzet et al. (2006b).

Assuming that there are no time changes in the electric field, the current density $\boldsymbol{j}=\nabla\times\boldsymbol{B}/\mu_0$ can be readily computed from the curl of the magnetic field. The angle $\alpha_{\boldsymbol{B},\boldsymbol{j}}$ between $\boldsymbol{B}$ and $\boldsymbol{j}$ as well as the current density magnitude $|\boldsymbol{j}|$ are given in Figs. 10b and c (anisotropic divergence-free vector case with fluctuations). The current density vector $\boldsymbol{j}$ is perpendicular to $\boldsymbol{B}$ somewhat northward of the magnetic equator, around 08:10 UT. Close to perigee, $\alpha_{\boldsymbol{B},\boldsymbol{j}}$ is $<90°$ south of the equator, $>90°$ north of it, with a current density of 30 nA/m^2. The existence of field-aligned currents, away from the equator itself, is clearly established in the plasmasphere but also on auroral field lines (e.g., just after 06:00 UT). For a dipolar magnetic field, $\boldsymbol{j}\equiv\boldsymbol{0}$, but this is definitely not true in the present situation.

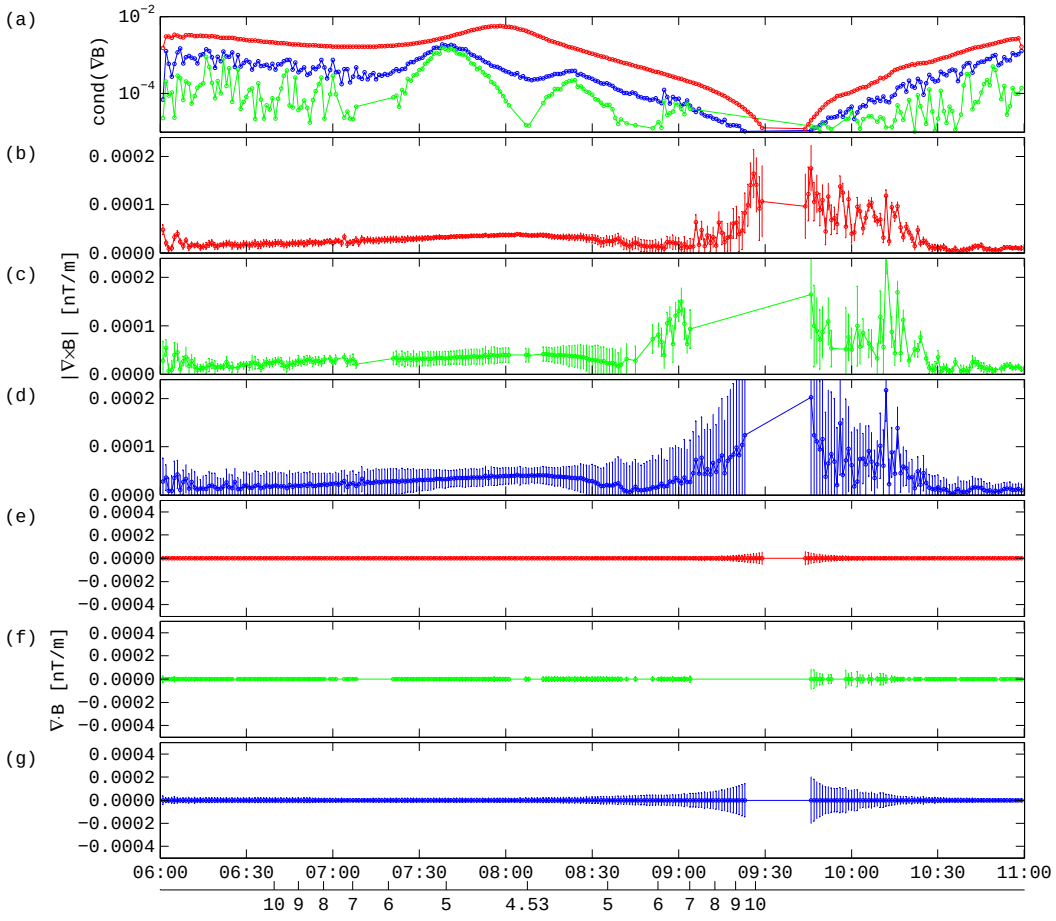

Fig. 9. Computation of $\nabla \times \boldsymbol{B}$ and $\nabla \cdot \boldsymbol{B}$, subject to the $\nabla \cdot \boldsymbol{B} = 0$ constraint, with anisotropic homogeneity domain and fluctuations (red), with the instantaneous gradient version (green), and with the instantaneous version with fluctuations (blue) (see text). **(a)** condition number; **(b–d)** magnitude of the curl $|\nabla \times \boldsymbol{B}|$; **(f–g)** divergence $\nabla \cdot \boldsymbol{B}$. The bottom scale gives the L-shell position of the center of the Cluster tetrahedron.

9 Conclusions

This paper describes a general-purpose method for computing gradients of scalar and vector fields in space and time. It has been shown that (1) The weighted least-squares method for computing gradients is a very robust one. (2) The method provides reliable error estimates that include the effects of measurement errors and approximation errors due to structure at scales that are larger and/or smaller than the physical scale of interest. (3) The method provides diagnostics to assess the quality of the computation, in particular by monitoring the singular values of the problem as a generalization of the concepts of planarity or elongation of a 4-spacecraft configuration. The role of the different parameters of the gradient computation algorithm has been illustrated. The relative importance of the different types of errors and their effect on the quality of the results has been discussed.

The method has been found to be superior to the traditional instantaneous gradient computation. Its primary advantage is its generality and its robustness. It correctly applies the principle of locality of information since only local data are used to compute the gradient at any given point, in accordance with the homogeneity condition. It also yields more stringent error margins on the obtained gradients. A disadvantage is its mathematical complexity. Implementing the method is not trivial. Computing the gradients is time-consuming when one considers small-scale fluctuations ($f^* \neq 0$), because then the (possibly large) error covariance matrices must be diagonalized. While the gradients obtained with this new method typically do not differ very much from those obtained with the traditional instantaneous gradient method, one now obtains a quantitative estimate of the total error on the results.

A prerequisite for a correct application of this method (and of any other method) is the ability to specify realistic values for the different error contributions. The measurement error is usually well-known, the fluctuation error is often only a minor correction, but providing an estimate for the curvature error may be more difficult. A posteriori verification,

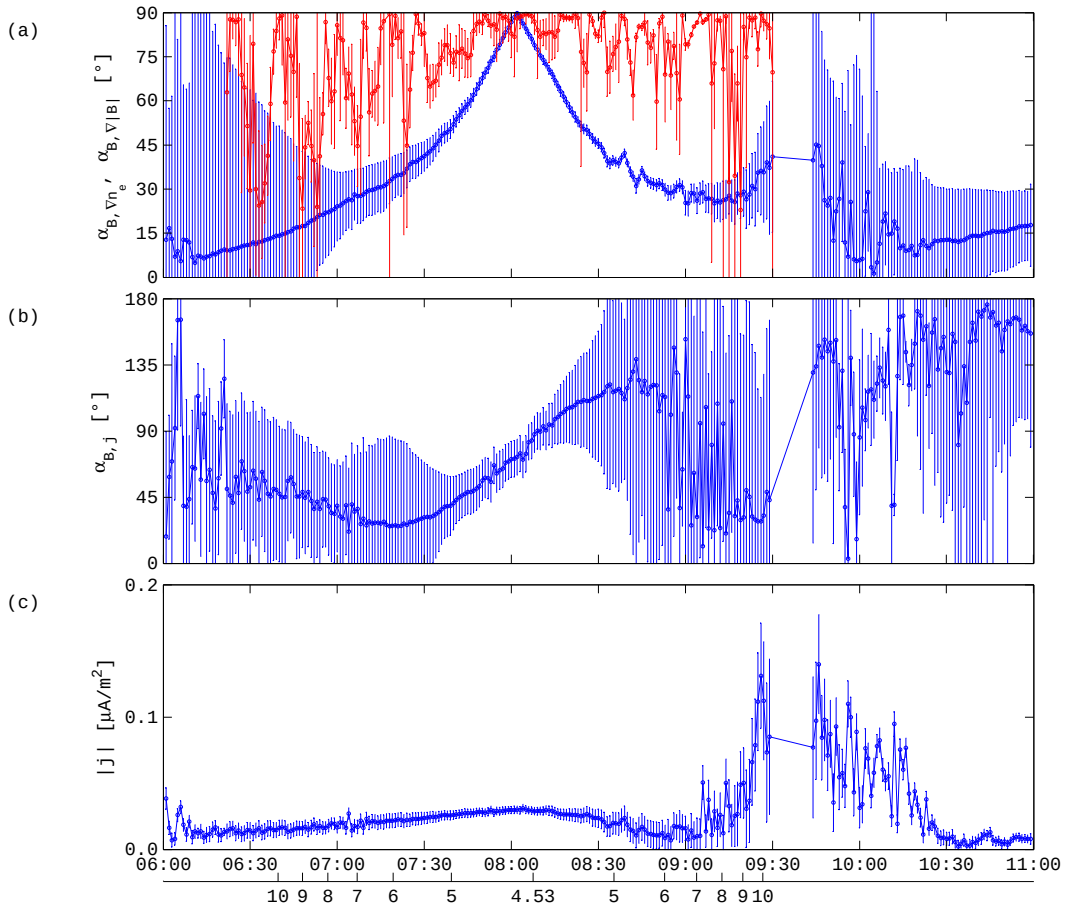

Fig. 10. Orientation of gradients during the Cluster inner magnetosphere pass on 7 August 2003, as a function of time and of L-shell. **(a)** Angle $\alpha_{\boldsymbol{B},\nabla|\boldsymbol{B}|}$ (blue) and $\alpha_{\boldsymbol{B},\nabla n_e}$ (red) between $\boldsymbol{B}$ and $\nabla|\boldsymbol{B}|$ and ∇n_e, respectively (anisotropic homogeneity with fluctuations, see text), reduced to [0°, 90°]. The magnetic equator corresponds to $\alpha_{\boldsymbol{B},\nabla|\boldsymbol{B}|}$=90°, near 08:02 UT. Both angles reflect the relative proportion of parallel and perpendicular gradients; $\nabla_{||}|\boldsymbol{B}|$ is rather strong away from the equator, while $\nabla_{||}n_e \ll \nabla_{\perp}n_e$ due to small longitudinal gradients within each flux tube and because of radial and azimuthal density structure. **(b)** Angle $\alpha_{\boldsymbol{B},\boldsymbol{j}}$ between $\boldsymbol{B}$ and current density $\boldsymbol{j}$ (where $\boldsymbol{j}=\nabla\times\boldsymbol{B}/\mu_0$ in a steady situation). **(c)** Current density magnitude $|\boldsymbol{j}|$ is significantly different from zero in the plasmasphere, indicating deviation from a dipolar field. The current density vector contains a significant field-aligned component inside the plasmasphere (around perigee) and also on auroral field lines (just after 06:00 UT).

however, is always possible. Once the gradient is computed along the spacecraft trajectory, one can check how it changes with position and/or with time, at least to a certain extent, so that an evaluation can be made of the curvature error. As long as there are enough points within the homogeneity domain, the value and the precision of the gradient are determined mainly by the measurement and fluctuation errors, and the exact value of the curvature error is not too important. A limitation of the present method is that a single, fixed value for the curvature error parameter f_c is used throughout the domain. Another limitation is that we have not accounted for timing errors or spacecraft position errors.

As an illustration, this method was used to analyze magnetic field and plasma density data obtained by Cluster during a pass through the inner magnetosphere. The relative importance of the perpendicular and field-aligned gradients of the magnetic field strength and of the plasma density have been discussed. In addition, nonzero current densities have been found, indicating that the field is not dipolar. Field-aligned currents appear to exist in the outer regions of the plasmasphere and on auroral field lines. The correct evaluation of the error margins on the gradients offered by the proposed method is absolutely necessary to ascertain the reliability of these findings.

The homogeneity scales must be adapted to the physical structures that one intends to study. In the present example, with a density structure that does not seem to exhibit too fine scales, homogeneity lengths of a few hundreds of kilometers and a time scale of 1 min were fine. In situations of stronger geomagnetic activity, finer-scale plasmaspheric structures may be formed more rapidly, necessitating smaller homogeneity scales in space and time. The least-squares

method will always produce a result, but whether the computed gradients are accurate depends on the nature of the data. With Cluster, a good gradient can usually be obtained when the homogeneity scales are on the order of, or larger than, the spacecraft separations in space and time.

Appendix A

Constraints

The overdetermined system (2) may still be ill-conditioned as the redundancy that stems from repeatedly measuring the same quantity over time on different spacecraft may be rather limited. For example, if the spacecraft are all in the same orbital plane, it is impossible to extract information about variations in the direction perpendicular to that plane. This ill-conditioning can be avoided if one adds new information about the problem in the form of constraints.

We discuss here two types of (linear) constraints that may be very useful in practice: geometrical ones, which state that one or more spatial gradient components are zero, and the stationarity constraint, which specifies that the total time derivative is zero.

A1 Geometrical constraints

Geometrical constraints are introduced by specifying an orthonormal set of vectors $\boldsymbol{c}_j$, $j=1,\ldots,m$ ($m\leq 3$) to which $\boldsymbol{\nabla}_{xt} f_0$ must be perpendicular. For example, the gradient might be required to be perpendicular to the local magnetic field vector $\boldsymbol{B}$. In that case, $m=1$ and $\boldsymbol{c}_1=\boldsymbol{B}/\|\boldsymbol{B}\|$. A set of orthonormal vectors $\boldsymbol{d}_i$, $i=1,\ldots,4-m$ can then be constructed, so that $\boldsymbol{d}_i{}^\top \boldsymbol{c}_j=0$. Transforming any vector as $\boldsymbol{x}^\beta=[\ldots \boldsymbol{d}_i \ldots \boldsymbol{c}_j \ldots]^\top \boldsymbol{x}$, the gradient itself becomes $[\ldots \boldsymbol{d}_i \ldots \boldsymbol{c}_j \ldots]^\top \boldsymbol{\nabla}_{xt} f_0$. Since $\boldsymbol{c}_j{}^\top \boldsymbol{\nabla}_{xt} f_0=0$, the m last components of the gradient vanish. The directions $\boldsymbol{c}_j$ can thus be regarded as homogeneity directions corresponding to infinite homogeneity scales, since the gradient must be invariant (identically zero) in each of those directions, so that $\mathbf{U}=[\ldots \boldsymbol{d}_j \ldots \boldsymbol{c}_j \ldots]$ and $\mathbf{L}=\mathrm{diag}([\ldots l_j \ldots +\infty \ldots])$. Transformation $\mathbf{P}$ is now a projection rather than a rotation and scaling (the space spanned by the $\boldsymbol{c}_j$ is its null space), its target space being $(4-m)$-dimensional. The constraint generally improves problem conditioning, but leads to larger residuals as there is less freedom to mimimize $\boldsymbol{r}^\top \boldsymbol{r}$.

As an example, consider the situation in which the gradient direction in space is known, say, that it lies along x. Then $\boldsymbol{c}_1=[0;\,1;\,0;\,0]$ and $\boldsymbol{c}_2=[0;\,0;\,1;\,0]$. For time-separable homogeneity with length scale $l_x=l_y=l_z=\xi$ and time scale $l_t=\tau$, one finds

$$\mathbf{U}=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \text{ and } \mathbf{L}=\begin{bmatrix} \xi & 0 & 0 & 0 \\ 0 & \tau & 0 & 0 \\ 0 & 0 & +\infty & 0 \\ 0 & 0 & 0 & +\infty \end{bmatrix},$$

which leads to a projection $x^\beta=x/\xi$, $t^\beta=t/\tau$.

A2 Stationarity constraint

For a structure moving with a given constant velocity $\boldsymbol{v}_0$, time-stationarity is expressed by

$$\frac{df_0}{dt}=[\ \boldsymbol{v}_0\ \ 1\ ]^\top \boldsymbol{\nabla}_{xt} f_0 = 0.$$

This constraint corresponds to an infinite homogeneity scale along direction $\boldsymbol{u}_4 \propto [\ \boldsymbol{v}_0\ \ 1\ ]$. Homogeneity in time is not separable from homogeneity in space unless $\boldsymbol{v}_0=\boldsymbol{0}$.

As an example, consider $\boldsymbol{v}_0=[v_0;\,0;\,0]$. Homogeneity directions $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, and $\boldsymbol{u}_3$ must form an orthonormal set together with $\boldsymbol{u}_4$. One particular choice produces

$$\mathbf{U}=\begin{bmatrix} 1/\sqrt{v_0^2+1} & 0 & 0 & v_0/\sqrt{v_0^2+1} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -v_0/\sqrt{v_0^2+1} & 0 & 0 & 1/\sqrt{v_0^2+1} \end{bmatrix} \text{ and}$$

$$\mathbf{L}=\begin{bmatrix} \xi/\sqrt{v_0^2+1} & 0 & 0 & 0 \\ 0 & \xi & 0 & 0 \\ 0 & 0 & \xi & 0 \\ 0 & 0 & 0 & +\infty \end{bmatrix},$$

where ξ is a length scale. The projection turns out to be $x^\beta=(x-v_0 t)/\xi$, $y^\beta=y/\xi$, $z^\beta=z/\xi$, so that one actually computes the spatial gradient with isotropic homogeneity length scale $l_x=l_y=l_z=\xi$ in a reference frame that moves with $\boldsymbol{v}_0$.

Appendix B

Cross-correlations of small-scale perturbations

We restrict ourselves to distributions of small-scale perturbations that are isotropic in γ-space with perturbation strength dropping off exponentially. The mean perturbation amplitude is assumed constant over the homogeneity domain, such that

$$\langle (\delta f_{\mathrm{ss}\boldsymbol{\lambda}}(\boldsymbol{\lambda},\boldsymbol{x}))^2 \rangle = \delta f_\lambda^2(\lambda^\gamma) = 16 f^{*2} S_4 \Big(\prod_{k=1}^{4} \hat{\lambda}_k\Big) \frac{\mathrm{e}^{-\lambda^\gamma}}{(\lambda^\gamma)^3},$$

with $\lambda^\gamma=\|\boldsymbol{\lambda}\|_\gamma$, f^* a constant, and S_4 the surface of a 4-dimensional sphere. Because of the locality of the perturbations, $\langle \delta f_{\mathrm{ss}\boldsymbol{\lambda}}(\boldsymbol{\lambda},\boldsymbol{x}_i)\delta f_{\mathrm{ss}\boldsymbol{\lambda}}(\boldsymbol{\lambda},\boldsymbol{x}_j)\rangle \approx \delta f_\lambda^2(\lambda^\gamma)$ when $\Delta x_{ij}^\gamma=\|\boldsymbol{x}_j-\boldsymbol{x}_i\|_\gamma \ll \lambda^\gamma$ and zero when $\Delta x_{ij}^\gamma \gg \lambda^\gamma$. For simplicity, the switch between both situations is taken to be an abrupt one. The covariances of δf_{ss} at points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ can then be computed as

$$\langle \delta f_{\mathrm{ss}}(\boldsymbol{x}_i)\delta f_{\mathrm{ss}}(\boldsymbol{x}_j)\rangle = \int_{\boldsymbol{\lambda}}\int_{\boldsymbol{\lambda}'} \langle \delta f_{\mathrm{ss}\boldsymbol{\lambda}}(\boldsymbol{\lambda},\boldsymbol{x}_i)\delta f_{\mathrm{ss}\boldsymbol{\lambda}}(\boldsymbol{\lambda}',\boldsymbol{x}_j)\rangle\, d\boldsymbol{\lambda} d\boldsymbol{\lambda}' \quad \text{(B1)}$$

$$= \int_{\boldsymbol{\lambda}} \langle \delta f_{\mathrm{ss}\boldsymbol{\lambda}}(\boldsymbol{\lambda}, \boldsymbol{x}_i) \delta f_{\mathrm{ss}\boldsymbol{\lambda}}(\boldsymbol{\lambda}, \boldsymbol{x}_j) \rangle \, d\boldsymbol{\lambda}$$

$$\approx (\prod_{k=1}^{4} \hat{\lambda}_k) \frac{S_4}{16} \int_{\Delta x_{ij}^{\gamma}}^{+\infty} (\lambda^{\gamma})^3 \delta f_{\lambda}^2(\lambda^{\gamma}) d\lambda^{\gamma} \quad \text{(B2)}$$

$$= f^{*2} \mathrm{e}^{-\Delta x_{ij}^{\gamma}}, \quad \text{(B3)}$$

where f^{*2} is the total perturbation variance. Whatever the specific choice of perturbation amplitude distribution, it must decrease faster than $1/(\lambda^{\gamma})^3$ in order to obtain a finite total perturbation, and the end result will always be that the cross-correlation is large between nearby points, and becomes zero as the distance between both points exceeds the perturbation length scale.

Acknowledgements. The authors thank M. Hamrin for fruitful discussions. This work was supported by the Belgian Federal Office for Scientific, Technical and Cultural Affairs through ESA (PRODEX/Cluster and PRODEX/Solar Drivers of Space Weather).

Topical Editor I. A. Daglis thanks M. L. Adrian and C. Harvey for their help in evaluating this paper.

References

Balogh, A., Dunlop, M. W., Cowley, S. W. H., Southwood, D. J., Thomlinson, J. G., Glassmeier, K.-H., Musmann, G., Lühr, H., Buchert, S., Acuña, M. H., Fairfield, D. H., Slavin, J. A., Riedler, W., Schwingenschuh, K., Kivelson, M. G., and the Cluster magnetometer team: The Cluster Magnetic Field Investigation, Space Sci. Rev., 79, 65–91, 1997.

Balogh, A., Carr, C. M., Acuña, M. H., Dunlop, M. W., Beek, T. J., Brown, P., Fornaçon, K.-H., Georgescu, E., Glassmeier, K.-H., Harris, J., Musmann, G., Oddy, T., and Schwingenschuh, K.: The Cluster Magnetic Field Investigation: overview of in-flight performance and initial results, Ann. Geophys., 19, 1207–1217, 2001, http://www.ann-geophys.net/19/1207/2001/.

Carpenter, D. L. and Lemaire, J.: Erosion and recovery of the plasmasphere in the plasmapause region, Space Sci. Rev., 80, 153–179, 1997.

Carpenter, D. L. and Lemaire, J.: The Plasmasphere Boundary Layer, Ann. Geophys., 22, 4291–4298, 2004, http://www.ann-geophys.net/22/4291/2004/.

Chanteur, G.: Spatial Interpolation for Four Spacecraft: Theory, in: Analysis Methods for Multi-Spacecraft Data, edited by Paschmann, G. and Daly, P. W., pp. 349–369, ISSI Scientific Report SR-001, 1998.

Chanteur, G. and Harvey, C. C.: Spatial Interpolation for Four Spacecraft: Application to Magnetic Gradients, in: Analysis Methods for Multi-Spacecraft Data, edited by: Paschmann, G. and Daly, P. W., pp. 371–393, ISSI Scientific Report SR-001, 1998.

Darrouzet, F., Décréau, P. M. E., De Keyser, J., Masson, A., Gallagher, D. L., Santolik, O., Sandel, B. R., Trotignon, J. G., Rauch, J. L., Le Guirriec, E., Canu, P., Sedgemore, F., André, M., and Lemaire, J. F.: Density structures inside the plasmasphere: Cluster observations, Ann. Geophys., 22, 2577–2585, 2004, http://www.ann-geophys.net/22/2577/2004/.

Darrouzet, F., De Keyser, J., Décréau, P. M. E., Gallagher, D. L., Pierrard, V., Lemaire, J. F., Sandel, B. R., Dandouras, I., Matsui, H., Dunlop, M. W., Cabrera, J., Masson, A., Canu, P., Trotignon, J.-G., Rauch, J.-L., and André, M.: Analysis of plasmaspheric plumes: CLUSTER and IMAGE observations, Ann. Geophys., 24, 1737–1758, 2006a.

Darrouzet, F., De Keyser, J., Décréau, P. M. E., Lemaire, J. F., and Dunlop, M. W.: Spatial gradients in the plasmasphere from Cluster, Geophys. Res. Lett., 33, L08 105, doi:10.1029/2006GL025727, 2006b.

Décréau, P. M. E., Fergeau, P., Krasnoselskikh, V., Lévêque, M., Martin, P., Randriamboarison, O., Sené, F. X., Trotignon, J. G., Canu, P., Mögensen, P. B., and Whisper investigators: WHISPER, A Resonance Sounder and Wave Analyser: Performances and Perspectives for the Cluster Mission, Space Sci. Rev., 79, 157–193, 1997.

Décréau, P. M. E., Fergeau, P., Krasnoselskikh, V., Le Guirriec, E., Lévêque, M., Martin, P., Randriamboarison, O., Rauch, J. L., Sené, F. X., Séran, H. C., Trotignon, J. G., Canu, P., Cornilleau, N., de Féraudy, H., Alleyne, H., Yearby, K., Mögensen, P. B., Gustafsson, G., André, M., Gurnett, D. A., Darrouzet, F., Lemaire, J., Harvey, C. C., Travnicek, P., and Whisper experimenters: Early results from the Whisper instrument on Cluster: an overview, Ann. Geophys., 19, 1241–1258, 2001, http://www.ann-geophys.net/19/1241/2001/.

Décréau, P. M. E., Le Guirriec, E., Rauch, J. L., Trotignon, J. G., Canu, P., Darrouzet, F., Lemaire, J., Masson, A., Sedgemore, F., and André, M.: Density irregularities in the plasmasphere boundary player: Cluster observations in the dusk sector, Adv. Space Res., 36, 1964–1969, 2005.

Dunlop, M. W. and Balogh, A.: Magnetopause current as seen by Cluster, Ann. Geophys., 23, 901–907, 2005, http://www.ann-geophys.net/23/901/2005/.

Dunlop, M. W., Balogh, A., Glassmeier, K.-H., and Robert, P.: Four-point Cluster application of magnetic field analysis tools: The Curlometer, J. Geophys. Res., 107, 1384, doi:10.1029/2001JA0050088, 2001.

Dunlop, M. W., Balogh, A., Shi, Q.-Q., Pu, Z., Vallat, C., Robert, P., Haaland, S., Shen, C., Davies, J. A., Glassmeier, K.-H., Cargill, P., Darrouzet, F., and Roux, A.: The Curlometer and other gradient measurements with Cluster, Proceedings of the Cluster and Double Star Symposium, 5th Anniversary of Cluster in Space, ESA SP-598, 2006.

Harvey, C. C.: Spatial Gradients and the Volumetric Tensor, in: Analysis Methods for Multi-Spacecraft Data, edited by: Paschmann, G. and Daly, P. W., pp. 307–322, ISSI Scientific Report SR-001, 1998.

Lemaire, J. F. and Gringauz, K. I.: The Earth's Plasmasphere, Cambridge University Press, Cambridge, 1998.

Rauch, J. L., Suraud, X., Décréau, P. M. E., Trotignon, J. G., Ledée, R., Lemercier, G., El-Lemdani Mazouz, F., Grimald, S., Bozan, G., Vallières, X., Canu, P., and Darrouzet, F.: Automatic determination of the plasma frequency using image processing on WHISPER data, Proceedings of the Cluster and Double Star Symposium, 5th Anniversary of Cluster in Space, ESA SP-598, 2006.

Reinisch, B. W., Huang, X., Song, P., Sales, G. S., Fung, S. F., Green, J. L., Gallagher, D. L., and Vasyliunas, V. M.: Plasma Density Distribution Along the Magnetospheric Field: RPI Ob-

servations From IMAGE, Geophys. Res. Lett., 28, 4521–4524, doi:10.1029/2001GL013684, 2001.

Robert, P., Dunlop, M. W., Roux, A., and Chanteur, G.: Accuracy of Current Density Determination, in: Analysis Methods for Multi-Spacecraft Data, edited by: Paschmann, G. and Daly, P. W., pp. 395–418, ISSI Scientific Report SR-001, 1998a.

Robert, P., Roux, A., Harvey, C. C., Dunlop, M. W., Daly, P. W., and Glassmeier, K.-H.: Tetrahedron Geometric Factors, in: Analysis Methods for Multi-Spacecraft Data, edited by: Paschmann, G. and Daly, P. W., pp. 323–348, ISSI Scientific Report SR-001, 1998b.

Trotignon, J. G., Décréau, P. M. E., Rauch, J. L., Randriamboarison, O., Krasnoselskikh, V., Canu, P., Alleyne, H., Yearby, K., Le Guirriec, E., Séran, H. C., Sené, F. X., Martin, P., Lévêque, M., and Fergeau, P.: How to determine the thermal electron density and the magnetic field strength from the CLUSTER/WHISPER observations around the Earth, Ann. Geophys., 19, 1711–1720, 2001, http://www.ann-geophys.net/19/1711/2001/.

Trotignon, J. G., Décréau, P. M. E., Rauch, J. L., Le Guirriec, E., Canu, P., and Darrouzet, F.: The Whisper Relaxation Sounder Onboard Cluster: A Powerful Tool for Space Plasma Diagnosis around the Earth, Cosmic Research, 41, 369–372, 2003.

Trotignon, J. G., Décréau, P. M. E., Rauch, J. L., Suraud, X., Grimald, S., El-Lemdani Mazouz, F., Vallières, X., Canu, P., Darrouzet, F., and Masson, A.: The electron density around the Earth, a high level product of the CLUSTER/WHISPER relaxation sounder, Proceedings of the Cluster and Double Star Symposium, 5th Anniversary of Cluster in Space, ESA SP-598, 2006.

Vallat, C., Dandouras, I., Dunlop, M., Balogh, A., Lucek, E., Parks, G. K., Wilber, M., Roelof, E. C., Chanteur, G., and Rème, H.: First current density measurements in the ring current region using simultaneous multi-spacecraft CLUSTER-FGM data, Ann. Geophys., 23, 1849–1865, 2005, http://www.ann-geophys.net/23/1849/2005/.