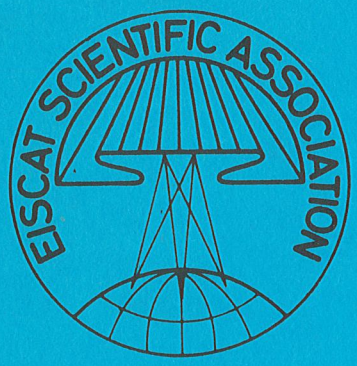


89/49

EISCAT
TECHNICAL
NOTE

ERROR ANALYSIS OF INCOHERENT SCATTER RADAR MEASUREMENTS

by

Matti Vallinkoski
Finnish Meteorological Institute
Box 503, SF-00101 Helsinki, Finland

KIRUNA
Sweden

ERROR ANALYSIS OF INCOHERENT SCATTER RADAR MEASUREMENTS

by
Matti Vallinkoski
Finnish Meteorological Institute
Box 503, SF-00101 Helsinki, Finland

Ph.D. Thesis,
University of Helsinki, 1989

EISCAT Scientific Association
Box 812, S-98128 Kiruna, Sweden
April 1989
EISCAT Technical Note 89/49
Printed in Sweden
ISSN 0349-2710
ISBN 952-90081-8-X (diss. Helsinki)

Preface

The *possibility* of a systematic and rigorous framework to study the various sources of error occurring in incoherent scatter measurements, or in the measurements of any kind of physical theory, arises naturally in connection with the applications of statistical inversion theory. For three years, the Academy of Finland supported an inversion project called 'Optimising accuracy in incoherent scatter measurements' at the Finnish Meteorological Institute with Markku Lehtinen as the responsible leader. This thesis is a product of that project. Although it may be considered as a normal type of hybrid consisting of articles and texts written during the past three years on the analysis of accuracies obtainable in incoherent scatter measurements with the EISCAT UHF radar as the prime example, the thesis nevertheless forms a consistent unit whose contents transcend the usual rules of thumb and which – I hope – will be useful as a reference work later on in the design and evaluation of incoherent scatter experiments.

The parturition of this work may be thanked to Markku Lehtinen's clever midwifery. My sincere thanks also go to the Finnish Meteorological Institute for providing financial support to finish this thesis. I also thank several people for help in the preparation of this work. Markku Mäkelä has been the chief decorator of the figures produced by MATLAB (except Chapter 2 which is due to Tuomo Rantanen). The figures in Chapter 5 were produced by Vesa Meskanen, Figs. 21–23 are due to Ingemar Häggström, and Figs. 24–26 to Markku Lehtinen. Tuomo Nygrén and Noralv Bjørnå have pointed out a not insignificant amount of oversights and inaccuracies in the manuscript. My last humble thanks go to those people whose serious scientific attitude has provided inspiration and confidence in the practical work presented here. Among those not listed above, I want to mention Asko Huuskonen, Lasse Jalonen and Jorma Kangas from Oulu and Tauno Turunen himself from Sodankylä.

Helsinki, April 1989 Matti Vallinkoski

Contents

Appendix. IS spectrum, ACF, derivatives and error routines

Vallinkoski, Matti K.: *Error analysis of incoherent scatter radar measurements.*
University of Helsinki, Department of Theoretical Physics, Helsinki 1989. 111 pp.
(EISCAT Scientific Association, EISCAT Technical Note 89/49,
ISSN 0349-2710.)
ISBN 952-90081-8-X (diss. Helsinki)

Abstract

Statistical inversion (SI) theory is applied to incoherent scatter (IS) error estimation. Numerical examples are given for the EISCAT UHF radar operating at about 933 MHz. The effects of drifts, the geomagnetic field or non–Maxwellian velocity distributions are neglected. A general approach in linearised SI is presented by which it is possible to determine the shortest lag extent and the longest lag resolution in a specific experiment in order to obtain optimal results. The results of these and the following calculations are applied to the cases when the collision frequency is zero, $T_e/T_i = 1$ or 2 and $T_i = 300$ K. The validity interval in ion temperature, when all other variables are kept fixed, is determined by the scaling properties of the IS spectra. The effect of a priori knowledge on one parameter on the accuracies of the other parameters is solved, and it is shown that under the conditions considered in this paper, composition can be determined to an accuracy of 10% when the ion temperature is known to within 7%, if we can assume that the collision frequency is zero. The effects of parameter mixing, *i.e.*, the parameters varying during integration, are shown to lead to the theory being corrected by the second central covariances of the parameter variations coupled to the second derivatives of the theory function. If these covariances are known exactly, then the resulting bias can be estimated. An uncertainty in the covariances leads to a new source of error and it is seen that the most important second moment is that of the ion temperature which will have to be known to better than 5% if one wants to determine composition as accurately as possible. Posteriori distributions of difficult parameters like the ion composition are calculated in semilinear theories, and a formula including the effects of postintegration of intermediary results to improve statistical accuracy in the difficult parameter is presented. Finally, the ion composition of the ionospheric plasma is determined by calculating the composition a posteriori marginal distributions in the first alternating code experiments ever run, using the EISCAT UHF radar.

Key words: (statistical) inversion, incoherent scatter, scattering, ionosphere, (ionospheric) plasma, ion composition, ion temperature, collision frequency, EISCAT

0. Introduction

0.1. The nature of incoherent scatter measurements

Incoherent scatter is radar power scattering from the electron density fluctuations of the ionospheric plasma. This scattering is analogous to Rayleigh scattering of sunlight by density fluctuations in the neutral atmoshere. In incoherent scatter, the scattering mechanism is Thomson scattering, *i.e.*, the radar waves are scattered independently of each other. Due to the thermal velocity distribution of the electrons, each scattered wave suffers a Doppler shift which depends on the individual scattering. However, as is well known, the scattering cross section is not the same as a Gaussian with a width corresponding to electron thermal velocities, as predicted by Gordon (1958). Instead, it was found (Bowles, 1958, 1959) that the spectrum was more like two peaks with widths corresponding to *ion* thermal velocities so close to the origin that there was a small dip at the central frequency corresponding to no Doppler shift.

The theory explaining the full spectra has been developed by several authors independently. (Dougherty and Farley, 1960; Fejer, 1960; Hagfors, 1961; Farley *et al.*, 1961; Dougherty and Farley, 1963; Farley, 1964; Farley, 1966; Swartz and Farley, 1979). The theory showed that under the conditions of the Bowles experiment (radar wavelength much longer than the characteristic Debye length of the plasma), the spectrum should in fact reflect the ion velocity distribution. The incident wave interacts with the electrons which in turn shield the ions, so that there is a reduction in the scattering cross section. The true incoherent scatter is only possible when the radar wavelength is much smaller than the Debye length. Normally one sees with the EISCAT UHF radar the typical double-humped ion line spectrum. The ion lines are situated at the frequencies corresponding to ion-acoustic waves. They are broadened because of Landau damping; when the electron temperature increases, these peaks separate and become narrower, since the damping gets weaker. The electron component is due to electron-acoustic or plasma waves. As they suffer very little damping they appear in the spectrum as two sharp plasma resonance lines displaced from the transmitted frequency by about the plasma frequency ω_p. A sample spectrum is shown in Fig. 1; general review articles are Evans (1969), Beynon and Williams (1978) and Rishbeth and Williams (1985). The derivation of the IS spectrum is sketched in Chapter 1.

Typical examples containing the ion lines starting from the collisionless F–region double-humped spectra down to the collision-dominated lower E–region, where most of the spectral mass is concentrated in a single central peak, are shown in Figures 2, which show the effects of the variation of plasma parameters.

In incoherent scatter measurements, one usually fits a set of parameters to

the measured autocorrelation function (ACF). The measured ACF is normally in principle not the Fourier transform of the whole spectrum but of the filtered spectrum which usually only comprises the ion lines; a simultaneous detection of ion and electron components is difficult, but not impossible (see *e.g.* Hagen and Behnke, 1976).

Since the theory is known, it is in priciple possible to determine these parameters, like the electron density N_e, ion temperature T_i, electron and ion temperature ratio T_e/T_i, ion–neutral collision frequencies ν_{in} and composition parameters c_s or drift velocities $\mathbf{v}_s$.

The bulk velocity of the plasma may first be fitted on the basis of the imaginary part of the ACF; the remaining parameters are then fitted simultaneously. Fitting the electron density is easy, since the spectrum is approximately linear in it. Also the ion temperature is routinely fitted, and the temperature ratio or the collision frequency with assumptions about the value of the temperature ratio. These parameters N_e, T_i and T_e/T_i may be called ordinary parameters, whereas the collision frequency or the ion composition are the difficult ones. For UHF radars, the collision frequency is so small that it may be usually set equal to zero above 120 km, and the temperature ratio to about one in the region where the effects of the collision frequency may be seen. The most difficult task is fitting the ion composition, or the fraction of the atomic oxygen in a plasma consisting of atomic oxygen and a mixture of the molecular ions (oxygen and nitric oxide). Usually this requires such long integration times or height integrations that one cannot be sure how well the plasma remains stable within the limits of the integration.

So far, quantitative measures of estimating the inherent uncertainties have been missing. However, with the advent of the concepts and methods of *statistical inversion theory* (Tarantola, 1987; Lehtinen, 1988; Menke, 1984), a whole new mathematically sound framework to estimate the different kinds of uncertainties occurring in practice has been created. The present work applies these methods to clarify and answer the following important questions:

a) What are the uncertainties in the fitted parameters? In other words, how good the results can be. This and the following two questions are discussed at length in Chapter 2 which differs from a separate publication (Vallinkoski, 1988) by the method used to calculate the numerical results. These may differ slightly from those in the publication.

b) What is the longest lag separation that still gives essentially the same results as an infinitely short lag separation? In practical terms, how densely we ought to sample an ACF to get optimal results.

c) What is the shortest ACF that still gives essentially the same results as an infinitely long ACF? In practical terms, how long we must sample an ACF.

d) How accurately must one know the value of one important parameter in order

to be able to resolve another one? In practical terms, how accurately one must know *e.g.* the ion temperature in order to be able to fit the composition sufficiently accurately. Chapter 3 shows the solution to this problem. This is a separate publication (Vallinkoski and Lehtinen, 1989a).

e) How large variation of the plasma parameters within a measuring volume may be allowed? In practical terms, how much the ion temperature may vary but still give reasonable results for the ion composition. Integration of the data improves the statistical accuracy of the results by the square root of the integration time, so that this is an important practical question. Chapter 4 is a separate publication (Vallinkoski and Lehtinen, 1989b)

f) How can we postintegrate *fitted results* in order to get better resolution for the difficult parameters? To do this, a whole new approach is developed in Chapter 5.

After this, in the last Chapter, all of these methods are applied to the first experiment in which the ion composition is determined accurately (Häggström *et al.*, 1989) and in which one can determine the accuracy of the results. The Chapters 5 and 6 are essentially this forthcoming publication.

All calculations in this thesis have been performed in a programming environment known as MATLAB[1] using microcomputers. Finally, all of the standard MATLAB routines used in the calculations are documented in the Appendix. They form an integral part of a general and complete MATLAB analysis package of IS measurements currently under development.

[1] A reference work is *e.g.* PC–MATLAB User's Guide by *The MathWorks, Inc.*

1. The calculation of the incoherent scatter spectrum

1.1. General

Since the error analysis will be theoretically based on calculations done on the spectra, it is for reference purposes useful to sketch the derivation and to write down explicitly the expressions of the plasma spectrum as a function of its parameters in the case of a Maxwellian plasma. Non-Maxwellian features need not be discussed, since the error analysis will be applied to situations where the effects of departure from Maxwellian plasmas cannot be seen.

The spectrum may be derived from the general fluctuation-dissipation theorem of linear systems by using a generalised version the Nyquist theorem on electrical circuits (Nyquist, 1928; Callen and Welton, 1951; Sitenko, 1967). This is the line followed most notably by Dougherty and Farley (1960, 1963), Farley *et al.* (1961, 1966) and Farley and Swartz (1979). We start by a general linear system with an external time harmonic force $V(t) = V_0 e^{-i\omega t}$ producing a linear response $I(t) = I_0 e^{-i\omega t}$ such that the instantaneous power dissipated by the system is $V(t)I(t)$. Let us define a generalised admittance function $Y(\omega)$ such that

$$I_0 = Y(\omega)V_0. \tag{1.1}$$

If the system is in a thermodynamical equilibrium at a temperature T, then there will be even in the absence external forces spontaneous fluctuations of force and response with spectra

$$\langle |V(\omega)|^2 \rangle d\omega = k_B T \,\mathrm{Re}[Z(\omega)] \frac{d\omega}{\pi} \tag{1.2}$$

$$\langle |I(\omega)|^2 \rangle d\omega = k_B T \,\mathrm{Re}[Y(\omega)] \frac{d\omega}{\pi}, \tag{1.3}$$

where $Z = Y^{-1}$ is a generalised impedance function and k_B the Boltzmann constant; ω may attain also negative values. In other words: if we know the response of a linear system to an external disturbance, the fluctuations of the undisturbed system in equilibrium may be calculated. In this original form, the results were applied to the transmission line noise resulting from voltage and current fluctuations with the impedances and admittances in their usual meanings.

If the components of the system have the same temperatures, the above equations may still be used provided Z and Y are interpreted as the total impedances and admittances of the system. In a multicomponent plasma, the different species may have different temperatures T_j. The linearity of the system then means that $V(\omega) = \sum_j H_j V_j(\omega)$, where the functions H_j are the transfer

functions of the plasma components, and the thermodynamical independence of the subsystems then means that equation (1.2) may be written by coupling the impedances in series:

$$\begin{aligned}\langle|V(\omega)|^2\rangle d\omega &= \sum_j \langle|H_j V_j|^2\rangle d\omega \\ &= \sum_j |H_j|^2 \langle|V_j|^2\rangle d\omega \\ &= \frac{k_B}{\pi} \sum_j |H_j|^2 T_j \,\mathrm{Re}[Z_j(\omega)] d\omega.\end{aligned} \tag{1.4}$$

In a plasma, the components j are subjected to total Lorentz forces $\mathbf{F}_j$ with responses I_j being the current densities $N_j \mathbf{u}_j$:

$$N_j \mathbf{u}_j = Y_j \mathbf{F}_j. \tag{1.5}$$

The admittance function Y_j may be derived from Boltzmann equations by linearising and integrating in the equations thus obtained over velocities perpendicular to the scattering wave vector $\mathbf{k}$. This is possible, since the velocity distribution functions are assumed to be Maxwellian. Then we are left with an integral which is customarily called *the* plasma dispersion function. When the distribution functions are not Maxwellian, the integration is not straightforward, but in some typical cases one may end up with one single integral (Raman *et al.*, 1981; Suvanto, 1987). When we denote $u_z = u_e$, the electron density fluctuations in the plasma will be described by

$$\langle|\Delta N_e|^2\rangle = \frac{k^2}{\omega^2} \langle|N_e u_e(\mathbf{k},\omega)|^2\rangle. \tag{1.6}$$

The connection of the above formalism to the plasma spectrum is in the equation above. According to the general theory, the spectrum $S(\mathbf{k},\omega)$ in incoherent scatter is given by

$$S(\mathbf{k},\omega) d\omega = \sigma_0 \, sin^2\chi \, \langle|\Delta N_e|^2\rangle d\omega, \tag{1.7}$$

where $\sigma_0 = 4\pi r_e^2 \approx 10^{-28}$ m^2 is the single electron radar cross section and $r_e = 1/4\pi\epsilon_0 m_e c^2 \approx 2.8 \cdot 10^{-15}$ m the classical electron radius. The angle χ is the polarisation angle, and we shall restrict ourselves to the case of backscatter where $\chi = \pi/2$, so that $sin^2\chi$ may be replaced by 1. Then also the wave number $k = 4\pi/\lambda$ is twice the radar wave number, when λ is the radar wavelength.

The transfer functions H_j may after this be calculated from the Lorentz equations, continuity equations and Maxwell's first equation $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$. When the admittance functions Y_j are expressed in terms of the dimensionless normalised functions y_j:

$$Y_j = \frac{N_j \omega}{k^2 k_B T_j} y_j, \tag{1.8}$$

we have the general expression for the incoherent scatter spectrum for the simplified case without magnetic fields or plasma drifts:

$$\frac{S(\mathbf{k},\omega)}{N_e\sigma_0}dw = \frac{|y_e|^2\sum_i \eta_i \mathrm{Re}[y_i] + |\sum_i \mu_i y_i + ik^2\lambda_D^2|^2 \mathrm{Re}[y_e]}{|y_e + \sum_i \mu_i y_i + ik^2\lambda_D^2|^2}\frac{d\omega}{\pi\omega}, \tag{1.9}$$

where λ_D is the electron Debye length $\sqrt{\epsilon_0 k_B T_e/N_e e^2}$ and

$$\begin{aligned} \eta_i &= \frac{N_i}{N_e} \\ \mu_i &= \eta_i \frac{T_e}{T_i} \end{aligned} \tag{1.10}$$

for the ion species i. The expressions of the normalised admittances will be presented in the next section, where we start from a collisionless plasma. The more general case of a collisional plasma with differential or global drifts will be presented later. For the purpose of this thesis, we shall need the real and imaginary parts of the admittances, which will be substituted in equation (1.9). The effect of the magnetic field may be neglected as long as we are far away from perpendicularity with the magnetic field (Farley, 1964); the most important cases in practice with the EISCAT UHF radar are backscatter measurements along the magnetic field lines. In the same cases, the non-Maxwellian features of the plasma velocity distribution are not present. In short, there is not much loss of generality, when we restrict our study to backscatter from a Maxwellian non-magnetic plasma.

1.2. Collisionless plasma

In a two-component collisionless plasma, the normalised admittance functions of the electrons and ions are y_e and y_i, where

$$y_s(\theta_s) = i(1 - \theta_s G(\theta_s)). \qquad s = e, i \tag{1.11}$$

The arguments θ_j are dimensionless variables

$$\theta_j = \frac{\omega}{\omega_j}, \tag{1.12}$$

where $\omega_j = kv_j$ and v_j is the thermal speed $v_j = \sqrt{(2k_B T_j/m_j)}$, and $G(z)$ the plasma dispersion function

$$G(z) = i\sqrt{\pi}e^{-z^2} + 2e^{-z^2}\int_0^z e^{t^2}dt = 2e^{-z^2}\int_{-i\infty}^z e^{t^2}dt. \tag{1.13}$$

The value of G at origin is $G(0) = i\sqrt{\pi}$, and for small values of the argument z $G(z) \simeq i\sqrt{\pi} + 2z$. Asymptotically, $G(z) \approx 1/z + 1/2z^3$ for large values of $|z|$ with

$\mathrm{Im}z < 0$. The name plasma dispersion function is also applied to the function $Z(z)$ tabulated by Fried and Conte (1961), which is related to G by $G(z) = Z(-z)$.

By substituting (1.13) into (1.11), we have the alternative form

$$y_s(\theta_s) = \theta_s\sqrt{\pi}e^{-\theta_s^2} + i(1 - 2\theta_s I(\theta_s)), \tag{1.14}$$

where I is the Dawson integral $I(x) = e^{-x^2}\int_0^x e^{t^2}dt$. The admittance function $y(\omega)$ may be obtained by calculating the total plasma impedance $z = 1/y$ by connecting the partial impedances z_s in series. For the electrons, $z_e = 1/y_e$; the ion impedance z_i includes an interaction term: $z_i = 1/(y_i + ik^2\lambda_D^2)$, and hence

$$\frac{1}{y} = \frac{1}{y_e} + \frac{1}{y_i + ik^2\lambda_D^2}, \tag{1.15}$$

from which it is possible to calculate $\mathrm{Re}[y(\omega)]$, when $T_e = T_i$:

$$\mathrm{Re}[y(\omega)] = \frac{|y_e|^2\mathrm{Re}[y_i] + |y_i + ik^2\lambda_D^2|^2\mathrm{Re}[y_e]}{|y_e + y_i + ik^2\lambda_D^2|^2}. \tag{1.16}$$

For a collisionless plasma (*i.e.*, for real ω), we always have

$$\begin{aligned}\mathrm{Re}[y_s] &= \theta_s\sqrt{\pi}e^{-\theta_s^2} = \theta_s\mathrm{Im}G(\theta_s)\\ \mathrm{Im}[y_s] &= 1 - \theta_s\mathrm{Re}G(\theta_s)\\ &= 1 - 2\theta_s I_s.\end{aligned} \tag{1.17}$$

In numerical calculations, both the real and imaginary parts of the plasma dispersion function will have to be calculated. When ω is real, one may directly substitute (1.17) into (1.16).

Near the ion line, $0 \le \theta_i \le 3$, $\theta_e \ll \theta_i$, $y_e \approx i$, and equation (1.16) will go over to

$$\mathrm{Re}[y(\omega)] = \frac{\mathrm{Re}[y_i]}{|i + y_i + ik^2\lambda_D^2|^2}. \tag{1.18}$$

When $\omega \to 0$, $y_s \to i$; in the limit $\omega \to 0$, $\mathrm{Re}[y(\theta_s)] \sim \theta_s = \omega/\omega_s$, so that $S(\mathbf{k}, 0)$ is finite. It is, however, essential to note that the value of the spectrum for $\omega = 0$ *has to be calculated separately*, since $y(0)$ is purely imaginary:

$$y(0) = i\frac{1 + k^2\lambda_D^2}{2 + k^2\lambda_D^2}. \tag{1.19}$$

When the ion and electron temperatures are equal, $T_e = T_i$, we have

$$S(\mathbf{k}, 0)d\omega = \frac{N_e\sigma_0}{(2 + k^2\lambda_D^2)^2}\left(\frac{1}{\omega_i}\frac{1}{1/\mathrm{Im}G(0)} + \frac{1}{\omega_e}(1 + k^2\lambda_D^2)^2\frac{1}{1/\mathrm{Im}G(0)}\right)\frac{d\omega}{\pi}. \tag{1.20}$$

The total scattering cross section σ_{tot} is obtained by integrating ω over the whole real axis. The integral may then be extended over the whole lower half plane, since

the plasma is stable so that there will be no singularities within the integration path. The remaining principal value integral will give the residue at the origin $\omega = 0$, from which an 'optical theorem' follows:

$$\begin{aligned} \sigma_{tot} &= N_e \sigma_0 \mathrm{Im}[y(0)] \\ &= N_e \sigma_0 \frac{1 + k^2\lambda_D^2}{2 + k^2\lambda_D^2} \cdot \end{aligned} \tag{1.21}$$

It may be noted that in the long wavelength limit, $k \to 0$ (the usual radar situation), and we have $\sigma_{tot} = \frac{1}{2} N_e \sigma_0$. In the short wavelength limit, $k \to \infty$, and we have $\sigma_{tot} = N_e \sigma_0$ which is the cross section of free non-interacting electrons. (This was the original idea of truly incoherent scatter, and the early planning of the Arecibo Observatory was based on this assumption.)

Since the plasma scattering cross section is a measure of the energy scattered, it may be argued that the *total* cross section does not depend on the magnetic field (which does no work) or the collisions (which only redistribute the energy into different frequencies). It may indeed be shown (Farley *et al.*, 1961; Dougherty and Farley, 1963) that for $T_e = T_i$, the total cross section will be independent of the magnetic field and of the collisions.

If there are several species of ions in the plasma at the same temperatures $T_i = T_e$, the ion admittance y_i will by linearity be replaced by the sum

$$y_i \to \sum_i \eta_i y_i, \tag{1.22}$$

where $\eta_i = N_i/N_e$ is the relative abundance of ion species i. We have the first result: for $T_e = T_i$, the admittance function of the collisionless plasma is given by

$$\mathrm{Re}[y(\omega)] = \frac{|y_e|^2 \sum_i \eta_i \mathrm{Re}[y_i] + |\sum \eta_i y_i + ik^2\lambda_D^2|^2 \mathrm{Re}[y_e]}{|y_e + \sum \eta_i y_i + ik^2\lambda_D^2|^2}. \tag{1.23}$$

The total cross section is still given by (1.21).

When the electron and ion temperatures are different, the Nyquist theorem can still be applied without difficulty. When we denote $\mu = T_e/T_i$, one obtains in the case of one ion species

$$\mathrm{Re}[y(\omega)] = \frac{|y_e|^2 \mathrm{Re}[y_i] + |\mu y_i + ik^2\lambda_D^2|^2 \mathrm{Re}[y_e]}{|y_e + \mu y_i + ik^2\lambda_D^2|^2}. \tag{1.24}$$

If there are more ions at different temperatures T_i, one has to make the replacement

$$\frac{y_i}{T_i} \to \sum_i \frac{N_i}{N_e} \frac{y_i}{T_i} = \sum_i \eta_i \frac{y_i}{T_i}, \tag{1.25}$$

so that the general expression for the admittance function of a collisionless plasma is given by

$$\mathrm{Re}[y(\omega)] = \frac{|y_e|^2 \sum_i \eta_i \mathrm{Re}[y_i] + |\sum_i \mu_i y_i + ik^2\lambda_D^2|^2 \mathrm{Re}[y_e]}{|y_e + \sum_i \mu_i y_i + ik^2\lambda_D^2|^2}. \tag{1.26}$$

The total cross section may only be obtained by a numerical integration, since (1.26) differs from the real part of an analytic function by terms which are proportional to $(1 - T_e/T_i)$. For values $1 \leq T_e/T_i < 4$, an often quoted good approximation for one ion species is

$$\frac{\sigma_{tot}}{N_e \sigma_0} = \frac{k^2\lambda_D^2}{1+k^2\lambda_D^2} + \frac{1}{(1+k^2\lambda_D^2)(1+T_e/T_i+k^2\lambda_D^2)}. \tag{1.27}$$

given first by Buneman (1962).

For $\omega = 0$, one obtains an exact result. When all ion temperatures are equal, the result may be written in the form

$$\begin{aligned} S(\mathbf{k},0)d\omega = & \frac{N_e\sigma_0}{(1+T_e/T_i+k^2\lambda_D^2)^2} \\ & \cdot \Big(\sum_i \frac{\eta_i}{\omega_i}\frac{1}{1/\mathrm{Im}G(0)} + \frac{1}{\omega_e}\Big(\frac{T_e}{T_i}+k^2\lambda_D^2\Big)^2\frac{1}{1/\mathrm{Im}G(0)}\Big)\frac{d\omega}{\pi}, \end{aligned} \tag{1.28}$$

which will be of the same type as in the collisional case.

1.3. Collisional plasma

In the previous sections, the collisions between neutral molecules and electrons or ions were neglected. The small–angle collisions between charged particles need not be taken into account, since the mean free path of the ions and electrons will be larger than the radar wavelength; see Farley (1964a,b). They would affect the gyroresonances, but this effect is only significant when the wave vector is nearly perpendicular to the magnetic field. On the other hand, the effect of the collisions cannot be neglected in the lower ionosphere. This effect is described by a suitably defined phenomenological parameter ν. The normal relaxation substitution $\omega \to \omega - i\nu$ will be true, but also the admittance functions will have to be modified, not only their arguments (otherwise the admittances would become singular for zero frequency). Physically, the pure relaxation approximation in the Boltzmann equation is not enough, since the particle number will no longer be locally conserved. Since the charge-neutral collisions are elastic, they can only affect the distribution of particles in velocity space. The relaxation approximation should therefore not be used in connection with density fluctuations. Instead, one needs a kinetic equation which also conserves particles locally. This is called the BGK approximation (Bhatnagar *et al.*, 1954). It does not, however, conserve energy or momentum; a modification of the collision term to do this leads to corrections which are negligible for small collision frequencies and UHF radars above about 100 km (Grassmann, 1988). The main results of this thesis will be applied to *precisely* this region for this radar, so that we can safely use the BGK collision term.

The effect of the collisions is then to identify the collision frequencies ν_s with the imaginary parts of Doppler frequencies ω_s, so that the variables θ_s will be

replaced by

$$\theta_s \rightarrow \frac{\omega - i\nu_s}{\omega_s} = \theta_s - i\psi_s \equiv \theta_s'. \tag{1.29}$$

The admittances y_s will be functions of the complex variables θ_s' so that

$$y_s(\theta_s) \rightarrow y'(\theta_s') \equiv i + \theta_s \frac{J(\theta_s')}{1 - \psi_s J(\theta_s')}, \tag{1.30}$$

where the function J is defined by

$$J(\theta_s') = J(\theta_s - i\psi_s) = -iG(\theta_s - i\psi_s) = -iG(\theta_s'). \tag{1.31}$$

Again, G is the plasma dispersion function (1.13) with a complex argument. If the magnetic field was included, J would be a so-called normalised Gordeyev integral. It is again true that $y_s'(\omega = 0) = i$, because J is purely real and positive:

$$J(-i\psi) = 2e^{\psi^2} \int_{\psi}^{\infty} e^{-t^2} dt. \tag{1.32}$$

In principle, the value of the spectrum at origin has to be calculated separately again. For that purpose, the admittances have to be developed in a Taylor series with θ_s around the origin. The derivative is

$$\left.\frac{\partial y_s}{\partial \theta_s}\right|_{\theta_s=0} = \frac{1}{1/A_s - \psi_s}, \tag{1.33}$$

where $A_s = J(-i\psi_s) = \mathrm{Im}G(-i\psi_s)$ is the integral (1.32). Then

$$y_s \simeq i + \theta_s \frac{1}{1/A_s - \psi_s}, \tag{1.34}$$

and the result for the value of the spectrum at origin is for equal ion temperatures T_i:

$$\begin{aligned} S(\mathbf{k},0)d\omega = & \frac{N_e\sigma_0}{(1 + T_e/T_i + k^2\lambda_D^2)^2} \\ & \cdot \Big(\sum_i \frac{\eta_i}{\omega_i} \frac{1}{1/A_i - \psi_i} + \frac{1}{\omega_e}\Big(\frac{T_e}{T_i} + k^2\lambda_D^2\Big)^2 \frac{1}{1/A_e - \psi_e}\Big) \frac{d\omega}{\pi}. \end{aligned} \tag{1.35}$$

The admittances will be expressed in terms of ψ, θ and the real and imaginary parts of the plasma dispersion function; the values of the plasma dispersion function will have to be calculated numerically. A standard method in the analysis packages used in Finland (*e.g.* in the University of Oulu) is to write $G_r = \mathrm{Re}G(\theta_s - i\psi_s)$ and $G_i = \mathrm{Im}G(\theta_s - i\psi_s)$ and use as a starting point in the programs the following results:

$$\mathrm{Re}[y_s(\theta_s')] = \frac{\theta_s(G_i - \psi_s|G|^2)}{(1 - \psi_s G_i)^2 + \psi_s^2 G_r^2} \tag{1.36}$$

$$\mathrm{Im}[y_s(\theta_s')] = 1 - \frac{\theta_s G_r}{(1 - \psi_s G_i)^2 + \psi_s^2 G_r^2}, \tag{1.37}$$

where the arguments of the plasma dispersion function have been suppressed.

If the plasma as a whole moves with respect to the observer or if its components move relative to each other, the velocity distributions will be shifted by the amount of the drift velocities v_s. The sole effect of this is that the frequencies ω in the denominators of equation (1.9) will have to be replaced by

$$\omega \rightarrow \omega' = \omega - \mathbf{k} \cdot \mathbf{v}_s \tag{1.38}$$

and the the ratios $\mathrm{Re}[y_s]/\omega$ replaced by

$$\frac{\mathrm{Re}[y_s]}{\omega} \rightarrow \frac{\mathrm{Re}[y_s(\theta'_s)]}{\omega - \mathbf{k} \cdot \mathbf{v}_s} \tag{1.39}$$

where $\theta'_s = \omega'/\omega_s$. If all velocities $\mathbf{v}_s$ are equal to $\mathbf{v}_d$, the whole spectrum will be shifted, and will no longer be symmetrical with respect to the origin. Then the plasma ACF will develop a complex phase which is just $exp(i\mathbf{k} \cdot \mathbf{v}_d\, t)$.

To summarise: the general expression for a Maxwellian plasma with arbitrary drifts and temperatures is

$$\begin{aligned} S(\mathbf{k},\omega)d\omega = & N_e\sigma_0\Big(|y_e|^2 \sum_i \eta_i \frac{\mathrm{Re}[y_i]}{\omega - \mathbf{k}\cdot\mathbf{v}_i} + \Big|\sum_i \mu_i y_i + ik^2\lambda_D^2\Big|^2 \frac{\mathrm{Re}[y_e]}{\omega - \mathbf{k}\cdot\mathbf{v}_e}\Big) \\ & \frac{1}{\left|y_e + \sum_i \mu_i y_i + ik^2\lambda_D^2\right|^2} \frac{d\omega}{\pi}, \end{aligned} \tag{1.40}$$

where the summation is over the ion species i

$$\begin{aligned} \eta_i &= \frac{N_i}{N_e}\frac{q_i^2}{e^2} \\ \mu_i &= \eta_i \frac{T_e}{T_i} \end{aligned} \tag{1.41}$$

and q_i is the charge of the ion species i. The admittances y_s for an unmagnetised plasma are defined by Equations (1.29)–(1.31). The above form is also valid for multiply ionised or negative ions. The actual computation of the incoherent scatter spectrum is documented in the Appendix. It is also seen that by virtue of Eq. (1.30), it is actually not necessary to calculate the value of the spectrum at origin separately.

2. Parameter estimation uncertainties in linear statistical inversion

2.1. Introduction

The behaviour of estimation errors in fitting theoretical ACFs obtained by Fourier transforming the spectra to a set of measured ACF lag values can be studied theoretically by supposing that the errors are so small that the variation of the plasma spectrum or autocorrelation function with respect to the plasma parameters can be approximated by using the first partial derivatives of the theoretical spectra. Then, by using Bayesian parameter estimation in the inversion theoretical approach (Box and Tiao, 1973; Tarantola and Valette, 1982a, 1982b; Menke, 1984; Tarantola, 1987 and references therein; Lehtinen, 1988), the estimation errors can be calculated as linearised *a posteriori* errors. These errors coincide with the errors of the correctly weighted least squares estimator (Tarantola, 1987, Section 4; Lehtinen, 1988). We shall call this approach simply the statistical inversion method.

The variation of different plasma parameters may give rise to quite similar changes in the theoretical spectrum; *e.g.* the width of the spectrum for fixed T_e/T_i is proportional to $\sqrt{T_i/m_i}$, where m_i is the ion mass, so it is very difficult to tell the difference between an increase in the ion temperature and a decrease in the ion mass. Also, when there is more than one ion species present, it is difficult to see the difference between changes in ion composition and the temperature ratio T_e/T_i for some ion mixtures.

In Section 2, we present the general inversion theoretical formalism. In Section 3, we treat the special case of a linear theory; in Section 4, we present the linear theory of IS measurements. In Section 5, we present some results of the calculations done for the EISCAT UHF radar, applied to those regions of the E- and F-layers where the effect of collisions is unimportant.

Since the inversion theoretical approach makes it possible to estimate the probable errors in fitting any number of parameters in a unified manner, the errors associated with *all* possible parameter fitting combinations will be determined *i*) without prior information on the parameters *ii*) with a 10% a priori accuracy for the ion temperature. The possibilities to determine the ion composition sufficiently accurately will be briefly discussed. Then we study the effects of varying the lag resolution and the lag extent of the ACF, *i.e.*, the sampling interval and the longest lag used in sampling the known theoretical ACF, for two important multiparameter fits and study how much the lag resolution requirement is affected, when we do not use the zero lag data in the fits. These results will be used to determine the stable ion temperature intervals for the 3– and 5–parameter fits in a specific experiment.

2.2. The inversion theoretical approach

Let us in all generality suppose that the theory of some phenomenon is describable in terms of a single function f depending on a set of parameters $x = (x_1, \ldots, x_N)^{\mathrm{T}}$, where $^{\mathrm{T}}$ denotes the transpose. In incoherent scatter theory, this function is the spectrum or its Fourier transform, the ACF. The vector x may for a fixed wavelength in the absence of drifts and magnetic field be given as

$$x = (N_e, T_i, T_e/T_i, \nu_{in}, p)^{\mathrm{T}}, \tag{2.1}$$

where p is the fractional density of the second positive ion, when there are only two ion species. For a given set of ionospheric parameters, the ACF uniquely decribes the incoherent scattering.

The result of a measurement is described by a random variable, the vector $m = (m_1, \ldots m_M)^{\mathrm{T}}$ of measured data points. In IS mesurements, m consists of the values of the ACF at different lags or those of the spectrum at different frequencies. Usually, we suppose that the measurements are corrupted by noise, *i.e.*, the measurement m depends on the parameter x and the theory f through the vector equation

$$m = f(x) + \epsilon, \tag{2.2}$$

where ϵ is a Gaussian zero-mean measurement noise vector with covariance matrix Σ_m given by the expectation value $\langle \epsilon \epsilon^{\mathrm{T}} \rangle$. In the most general case, ϵ and thus Σ_m may depend on x. (Non-zero mean would indicate a bias; in most practical cases, the errors may well be approximated by Gaussians.) The diagonal elements of Σ_m are the measurement variances σ_i^2 and the non-diagonal elements σ_{ij} are given by $\sigma_i \sigma_j \rho_{ij}$ where ρ_{ij} is the the correlation coefficient of the measurement errors.

In the direct theory, we suppose that we know the parameter values x. The application of the function f on x then predicts the values of a measurement m according to (2.2) where the errors ϵ are supposed to have a known probability distribution. In inversion theory, the measured values $\overline{m}$ and the distribution of the errors ϵ are known from the experiment, and we want to *invert* (2.2) in order to obtain the values of the unknown parameters x in this particular experiment. (We use $\overline{m}$ to distinguish the value of the random variable m in a particular realisation from the variable itself.)

Let us suppose that we have some *a priori* information about certain parameters. This may be represented by the a priori probability distribution $D_{pr}(x)$; *e.g.* it may be that N_e is known to be 10^{11} m^{-3} with a 1% standard error, obtained by some independent measurement like a simultaneous plasma line mesurement or from a similar experiment under similar conditions. The a posteriori knowledge may subsequently be used as a priori knowledge in another experiment. In practice, some a priori knowledge is always inserted into the analysis (programs), and it is certainly not at all clear what kind of a priori assumptions are built in into the existing IS analysis packages.

When both x and m are considered as random variables with distributions $D(x)$ and $D(m)$ and with a joint probability distribution $D(x,m)$, Bayes' theorem (Bayes, 1763; 1958) for conditional probabilities may be written as

$$D(x|m)D(m) = D(x,m) = D(m|x)D(x). \tag{2.3}$$

The a priori distribution $D_{pr}(x)$ is then defined as $D_{pr} = \int D(x,m)dm$, and the *a posteriori* probability distribution for the parameters x is the conditional probability distribution

$$D_p(x) = D(x|m) = C_0\, D_{pr}(x)D(m|x), \tag{2.4}$$

for the parameter vector x given the measurement results m. The normalisation constant $C = C_0 = 1/\int D(x,m)dx$ is independent of x; we shall use the same letter C for all normalisation constants in this paper, since their actual values are irrelevant.

The transition densities $D(m|x)$ are the conditional probability densities of the measurements supposing that the parameters x are known. In the classical Fisher terminology (Fisher, 1922), the transition densities may be called *likelihood functions*.

In order to calculate the a posteriori densities, one has to know the transition densities. Fixing x, we see that m is Gaussian with mean $f(x)$ and covariance matrix Σ_m. Hence the general expression for the transition density is the conditional probability density for obtaining a measurement result m given x:

$$D(m|x) = \frac{1}{(2\pi)^{M/2}|\Sigma_m|^{\frac{1}{2}}}\, exp\Big(-\frac{1}{2}(m-f(x))^{\mathrm{T}}\Sigma_m^{-1}(m-f(x))\Big) \tag{2.5}$$

where M is the number of components of the vector m and $|\Sigma_m|$ the determinant of the measurement error covariance operator Σ_m. Therefore, the unnormalised a posteriori distribution for having a parameter vector x, given measured values $\overline{m}$ is

$$D_p(x) = D(x|\overline{m}) = C\, D_{pr}(x) exp\Big(-\frac{1}{2}(f(x)-\overline{m})^{\mathrm{T}}\Sigma_m^{-1}(f(x)-\overline{m})\Big). \tag{2.6}$$

The a posteriori distribution as a function of x may in principle be calculated from (2.6). In the general nonlinear case, Σ_m is a function of x. If the a priori distribution is uniform, the a posteriori distribution is maximised when the function

$$(f(x)-\overline{m})^{\mathrm{T}}\Sigma_m^{-1}(f(x)-\overline{m}) \tag{2.7}$$

is minimised, and this point gives the the maximum likelihood estimate for the parameter values x, when ϵ does not depend on x. If ϵ depends on x, the covariance matrix and hence the determinant $|\Sigma_m|$ in (2.5) will depend on x, and the maximum likelihood point will not necessarily be the minimum of ((2.7)).

The second moments of the a posteriori distribution are the a posteriori variances, and represent the measurement errors. In practice, a straightforward calculation of the full a posteriori distribution to a satisfactory degree of precision is a prohibitive task. One can instead study marginal distributions, which are obtained by integrating the full distribution over a chosen set of parameters. The marginal distributions are then functions of the remaining variables. Conclusions about the behaviour of the distribution can be drawn also from these projections. Normally, only those points where the one-dimensional marginal distributions attain their maxima, are given. If these maxima are not steep, valuable information is easily lost by giving just one point instead of the confidence intervals, within which most of the probability mass is concentrated.

2.3. General linear theory

In the general inversion approach, a particularly useful case is linear theory with Gaussian errors, since the expression of the a posteriori distribution can be written analytically. Equation (2.2) is replaced by the matrix equation

$$m = Ax + \epsilon, \tag{2.8}$$

where the transition densities are obtained from (2.5) by just replacing $f(x)$ by Ax. The a priori distributions are supposed to be Gaussian with centre points x_0 and error covariance matrix Σ_0:

$$D_{pr}(x) = C\, exp\big(-\frac{1}{2}(x - x_0)^{\mathrm{T}}\Sigma_0^{-1}(x - x_0)\big). \tag{2.9}$$

Here the point x_0 represents the value of the variable x obtained by some independent (*e.g.* previous or simultaneous measurement) and the error covariance matrix Σ_0 contains the associated a priori variances.

The a priori errors may well be taken to be Gaussian. In many cases, this simplifies calculations, and serves at least as an approximation. If we have no a priori knowledge about the parameters, D_{pr} may be taken constant, in which case Σ_0^{-1} is a zero matrix. In a similar way, when the a priori distribution is non-informative, *i.e.*very broad compared with the second factor in (2.6), it affects the a posteriori distribution hardly at all. If it is very narrow, then it is essentially the a priori distribution which determines the accuracy, and the measurement does not add much information.

In the general Gaussian case, by combining the the two exponential factors in (2.6) into a perfect square, it is readily seen that the a posteriori distribution is of the standard Gaussian form

$$D_p(x) = C\, exp\big(-\frac{1}{2}(x - \overline{x}_0)^{\mathrm{T}}Q(x - \overline{x}_0)\big), \tag{2.10}$$

where the matrix Q is given by

$$Q = \Sigma_0^{-1} + A^{\mathrm{T}}\Sigma_m^{-1}A \tag{2.11}$$

and the centre point vector $\overline{x}_0$ is both the maximum likelihood point x_{ML} and the mean value of the posteriori distribution. The centre point is defined by the equation

$$Q\overline{x}_0 = \Sigma_0^{-1}x_0 + A^{\mathrm{T}}\Sigma_m^{-1}\overline{m}. \tag{2.12}$$

We now define the misfit function to be the the sum of the two exponential factors considered above. The sum may be written into the form

$$S(x) = (x - \overline{x}_0)^{\mathrm{T}}Q(x - \overline{x}_0) + S(\overline{x}_0) \tag{2.13}$$

where we have used the fact that the term linear in x vanishes at the maximum likelihood point $\overline{x}_0$. The factor $exp(-\frac{1}{2}S(\overline{x}_0))$ was absorbed into the constant C in (2.10).

The vector $\overline{m}$ represents the measured values. The matrix Q is usually called the Fisher information matrix, if the a priori distribution is uniform. If the measurement errors are uncorrelated, Σ_m is diagonal, and the misfit function is in this case the familiar χ^2 function, which is proportional to the residual of the fit. The minimum of the χ^2 function gives the maximum likelihood point. If the a priori probability distribution is uniform, then in the case of uncorrelated Gaussian errors the point $\overline{x}_0$ may be obtained simply by multiplying the measured values by the pseudoinverse of the matrix A, and will therefore be independent of the actual value of the measurement error variances.

If there are more independent measurements m_i, then each measurement adds a contribution $A_i^{\mathrm{T}}\Sigma_i^{-1}A_i$ to the information matrix, and the position of the maximum likelihood point will in general be shifted.

The second moments of the a posteriori distribution representing the uncertainties in the parameter estimates x_k are also variances σ_k^2 of the a posteriori distribution. In the linear theory, these variances are simply the diagonal elements of the inverse of the Fisher information matrix Q from (2.11), $\sigma_k^2 = (Q^{-1})_{kk}$. Since Q does not depend on the measured values $\overline{m}$, this means that in a linear theory, one can *predict* the errors in an experiment, when one only knows the direct theory (represented by the matrix A), the measurement variances and the a priori distribution D_{pr}. We thus see that while the *position* of the maximum likelihood point may turn out to be independent of the actual values of the measurement accuracy, its uncertainty will not.

One arrives at a linear theory quite naturally by linearising a nonlinear theory. One has to assume that the behaviour of the theory function f is sufficiently well approximated in the neighbourhood of some point x_0 by using the first partial derivatives only:

$$f(x_0 + \Delta x) = f(x_0) + \sum_i \frac{\partial f}{\partial x_i}\Delta x_i. \tag{2.14}$$

The columns A_i of the matrix A in (2.8) are therefore in the linear theory the partial derivatives of f with respect to the varying parameters x_i:

$$A_i = \frac{\partial f}{\partial x_i}. \tag{2.15}$$

The point x_0 is taken to be the fitted solution of the measurement, and in order to estimate the errors associated with this measurement, we have to develop the function f around this point.

In the direct theory, the uncertainties in the parameter values x will be reflected in the uncertainties of the measured data as $\Delta f = \sum_i (\partial f/\partial x_i)\, \Delta x_i$. In the inverse theory, we want to study the propagation of the measurement uncertainties due to the noise ϵ to the parameter values x inferred. This is the content of (2.10). We shall illustrate this in a trivial case. When we take both x and m to be one-dimensional, we see from (2.14) that when the measurement variance is σ^2, the uncertainty in x will be given by the inverse relation

$$\Delta x = \frac{\sigma}{\left|\frac{\partial f}{\partial x}\right|}. \tag{2.16}$$

In the case of no prior information, $\sigma_x^{-2} = (\Delta x)^{-2}$ equals $Q = A^{\mathrm{T}} \Sigma_m^{-1} A$ for one variable with A given by (2.15) gives just (2.16). In a more general case, the full matrix equations (2.11) have to be used.

We may also note the intuitively obvious fact that when the theory function f is insensitive to changes in some parameter, it is to be expected that the corresponding error in the parameter estimate is not small.

2.4. The linear theory of IS measurements

The results presented in the previous sections will be applied to the linearised classical theory of incoherent scatter. For simplicity, we exclude the drifts and the effects of the magnetic field. Therefore, we have a symmetric spectrum and a purely real ACF.

We suppose that the analysis is made by fitting a theoretical ACF, depending on a number of parameters, to the measured set of lags which may or may not contain the zero lag. By the ACF we mean the Fourier transform of the spectrum consisting of the ion component. We also suppose that we are considering an essentially collision-free plasma, so that in a single-ion plasma, the ACF is a smooth function having a first zero at about the time corresponding to the ion-acoustic velocity.

In order to facilitate data handling in the subsequent calculations, we shall use dimensionles variables of the order of unity only. This is possible, since the

theory is linear. Suppose first that there is only one ion species with mass m_1 and temperature T_i present. We *fix* the frequency scale ω_0 to be the ion angular Doppler frequency

$$\omega_0 = k\sqrt{\frac{2k_B T_i}{m_1}}, \tag{2.17}$$

where k is the scattering wave number and k_B the Boltzmann constant. For $T_e = T_i$, this is the ion–acoustic frequency. The collision frequencies are to be expressed in units of $\nu_0 = \omega_0/2\pi$.

Let us denote the natural time scale of the ACF by τ_0. This should be approximately equal to $\pi/2\omega_0$, which is the position of the first zero in the ACF for a spectrum where there are only two delta peaks at the frequencies $\omega = \pm\omega_0$, corresponding to a pure one-ion case with no Landau damping. Therefore, we *define* the time scale by using the value

$$\tau_0 = \pi/2\omega_0 \tag{2.18}$$

even for $T_e \neq T_i$ and the cases where there are more ion species present. We may require that the ion with mass m_1 dominates in order to have a truly natural fixed time scale.

We normalise the electron density scale N_0 and the ion temperature scale T_0 to the fitted values of the variables. Then the parameter vector x in (2.1) will actually be

$$x = (N_e/N_0, T_i/T_0, T_e/T_i, \nu_{in}/\nu_0, p)^{\mathrm{T}}, \tag{2.19}$$

where we now have fixed m_1 to be the molecular ion with mass 30.5 u (a mixture of 25% O_2^+ and 75% NO^+), and the second ion to be O^+ with mass 16 u, so that p is the fractional oxygen content $N(O^+)/N_e$ of the plasma.

We scale the spectrum $S(\mathbf{k},\omega)$ to the expression

$$\frac{S(\mathbf{k},\omega)\pi\omega_0}{N_0\sigma_0}, \tag{2.20}$$

where $\sigma_0 = 4\pi r_e^2$ was defined in the previous Chapter. Similarly, we know that for $T_e/T_i = 1$ the integral over ω of this expression over the whole spectrum (including the electron component with the plasma lines) should be equal to $\frac{1}{2}(1+k^2\lambda_D^2)/(2+k^2\lambda_D^2)$, independently of the collision frequency and the magnetic field. This integral also represents the total scattered power. Therefore, we *fix* the ACF scale $a(0)$, the value at zero lag to the constant value $\frac{1}{2}$.

Our aim is to study the fitted parameter uncertainties, resulting from the noise in the measured lag values. Our interest is in those cases where the signal is weak and the dominant noise is white noise; for a strong signal, the following assumptions may not be valid. We suppose that different lags are uncorrelated and that they have equal variances, although the zero lag estimate has in principle twice the variance of other lagged product estimates. In multipulse experiments, the zero

lag only has a variance usually different from the other lags and is often not used at all (see also Lehtinen and Huuskonen, 1986). Nevertheless, we assume for simplicity that measurement error covariance matrix Σ_m will be diagonal and a multiple of the unit matrix, $\Sigma_m = \sigma^2 1$. This is more or less true in multipulse or alternating code (Lehtinen and Häggström, 1987) experiments. The use of a single parameter for the measurement variances is not a serious limitation. It is at least intuitively clear that the final error estimates will be between the two obtained by using the smallest and largest values for the variances. When there is no prior information, one has to calculate the error table only once and scale the results according to the true value of σ.

When $\Sigma_m = \sigma^2 \mathbf{1}$, the elements of the Fisher information matrix Q are given by the expression

$$Q_{ij} = \frac{1}{\sigma^2} \sum_{lags} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \tag{2.21}$$

and it is easily seen that σ^2 is a function of the lag length $d\tau$. From a physical point of view, it is clear that after some point, decreasing the lag length does not add any new information, and so the information matrix Q will be essentially the same. Writing the defining sum in (2.21) in the form

$$\frac{1}{\sigma^2 d\tau} \sum_{lags} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} d\tau \tag{2.22}$$

we see immediately that $\sigma^2 d\tau$ *will have to be a constant*, since the sums are Riemann's sums approximating the integrals

$$\int_0^\infty \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} d\tau. \tag{2.23}$$

Since the ACF is a symmetric function, we can extend the integrals over all times from $-\infty$ to ∞, provided the the variance of the zero lag is *twice* the variance of the other lags, as noted in the previous paragraph.

The fact that $\sigma^2 \sim 1/d\tau$ may be understood in the following way. In order to obtain the same information matrix with different (but sufficiently small) lag lengths $d\tau$, let us consider $d\tau = nd\tau'$, where n is an integer. Then each term in (2.22) will be replaced by a sum of n terms, which by continuity are all approximately equal and give n times the original term. Therefore, the new variance σ'^2 has to be $n\sigma^2$. This amounts to replacing one term by a sum of terms and taking their average. On the other hand, the variance of this sample average over the original lag length $d\tau$ is of course $\sigma'^2/n = \sigma^2$ and hence remains constant. In a particular experiment, *i.e.*, for a given fixed $d\tau$, this constant can be identified as half the variance of the ACF zero lag estimate, or half the variance of the power profile estimates.

When we denote the number of lags to the first zero by $N_\tau = \tau_0/d\tau$, we have the result that $\sigma^2 \sim N_\tau$. Since the maximum value of the ACF is $a(0)$, proportional

to the total scattered power of the plasma, a natural definition for *the ACF noise level* γ, square root of the average relative variance per unit time delay, is given by

$$\gamma^2 = \frac{\sigma^2}{a(0)^2 N_r}, \tag{2.24}$$

and it must be emphasized that this definition is valid only if the lag measurement errors are independent. We use $a(0) = \frac{1}{2}$ and fix γ to be γ_0=0.01, and use the resulting value for σ throughout. This choice of γ_0 will ensure that the uncertainty in the electron density when fitted alone will be conveniently *about* one per cent.

The true measurement noise σ is often derived from the signal–to–noise ratio (SNR) of the measurement, and will be different in each individual experiment. When the SNR$\ll$ 1, the variance of a simple short-pulse power profile estimate is equal to (Lehtinen, 1986)

$$\sigma^2 = a(0)^2 (\mathrm{SNR})^{-2} / N_R, \tag{2.25}$$

where N_R is the number of repetitions of the radar pulse in the experiment. Increasing N_R, which is often called integration, improves the statistical accuracy of the estimate. If the SNR is not small, the effects of clutter would have to be included.

Similarly, in a multipulse experiment, the ACF lag estimates have the variances

$$\sigma^2 = \frac{1}{2} a(0)^2 (\mathrm{SNR})^{-2} / N_R \tag{2.26}$$

where SNR now refers to a single-pulse SNR$\ll$ 1.

2.5. Results of calculations

In this Section, we present results of calculations to illustrate the power of the statistical inversion methods. However, the presentation will be complete enough to allow the use of these results in the planning of those E– and F–region experiments where the collision frequencies can be neglected, and in the analysis of the error bars obtained in the experiments. The calculations are documented in Appendix. In the evaluation of an actual experiment, one may use the true measurement variances obtained as the diagonal elements of the measurement error covariance matrix Σ_m in order to obtain the true parameter estimate errors.

We study the cases where the time scale is $\tau_0 = 99.3\,\mu\mathrm{s} \approx 100\,\mu\mathrm{s}$, and the value of $(k\lambda_D)^2$, where λ_D is the Debye length, is 0.022 or 0.044, the collision frequency zero and O^+ content 30%. This is the case for the EISCAT UHF radar backscatter with frequency 933.5 MHz, electron density 10^{11} m^{-3} and ion temperature 300 K. (This may not be the most general realistic case, but we use it here as an example.)

The temperature ratio T_e/T_i is taken to be either 1 or 2. (The position of the true zero of the ACF is about $1.1\,\tau_0$ for $T_e/T_i = 1$ and $0.9\,\tau_0$ for $T_e/T_i = 2$.) In the calculations, we have taken $N_\tau = 10$, which means that $d\tau \approx 10\,\mu$s, and used 60 lags for the calculation of the ACFs. The complete error tables are presented for the case $T_e/T_i = 1$ only.

The positive parts of the spectra are shown in Fig. 3 for $T_e/T_i = 1$ and $T_e/T_i = 2$. The corresponding ACFs are shown in Fig. 4, and the partial derivatives with respect to the five variables included in the theory in Figs. 5a and 5b.

As the errors of a parameter x_k will be estimated to be equal to the square roots of the a posteriori variances σ_k^2, the remaining task is to calculate the inverse of Q from (2.11). Let the number of measured lags be M and let the number of variable parameters be n. Then all we need is the partial derivative matrix A, an $M\times n$ matrix from (2.15); the matrices $A^{\mathrm{T}}\Sigma_m^{-1}A$ and Q will be $n\times n$ –matrices. The dimensions of Q used will vary from a minimum 1×1 (one-parameter fits) to the maximum 5×5 (five-parameter fit), but we need to calculate the matrix A only once. For the vector $x = (1,1,1,0,0.3)^{\mathrm{T}}$ considered here, the full 5×5 matrix $A^{\mathrm{T}}\Sigma_m^{-1}A$ is equal to

$$10^3 \begin{pmatrix} 5.3707 & -1.3173 & -3.1882 & 0.0185 & -0.8660 \\ -1.3173 & 1.8637 & 1.7567 & -0.1807 & 1.2644 \\ -3.1882 & 1.7567 & 2.7007 & -0.1852 & 1.1218 \\ 0.0185 & -0.1807 & -0.1852 & 0.0411 & -0.1040 \\ -0.8660 & 1.2644 & 1.1218 & -0.1040 & 0.8768 \end{pmatrix}, \tag{2.27}$$

and all the other matrices needed will be submatrices of (2.27). The a posteriori covariance matrices are inverses of the submatrices of (2.27) when there is no a priori information; when *e.g.* T_i is known with a 10% accuracy, one has to add the number $(0.1)^{-2} = 100$ to the (2,2) element of (2.27) in order to obtain Q. The a posteriori covariance matrix of the 3–parameter fit of $N_e, T_i, T_e/T_i$ is equal to

$$\begin{pmatrix} 0.0012 & -0.0013 & 0.0023 \\ -0.0013 & 0.0028 & -0.0033 \\ 0.0023 & -0.0033 & 0.0052 \end{pmatrix}, \tag{2.28}$$

and the full 5–parameter fit a posteriori covariance matrix equal to

$$\begin{pmatrix} 0.0863 & 0.0002 & 0.1645 & 0.5506 & -0.0602 \\ 0.0002 & 0.6482 & -0.0457 & 0.6113 & -0.8036 \\ 0.1645 & -0.0457 & 0.3177 & 1.0081 & -0.0585 \\ 0.5506 & 0.6113 & 1.0081 & 4.1270 & -1.1380 \\ -0.0602 & -0.8036 & -0.0585 & -1.1380 & 1.0404 \end{pmatrix} \tag{2.29}$$

with the variances σ_k^2 on the diagonals and parameter estimate error correlations ρ_{ij} included in the off-diagonal elements as $\sigma_i\sigma_j\rho_{ij}$. For example, the correlation coefficient between T_i and p is noticeably high, about -0.98, as expected. The full

correlation matrix constructed from (2.29) is

$$\begin{pmatrix} 1.0000 & 0.0007 & 0.9935 & 0.9229 & -0.2008 \\ 0.0007 & 1.0000 & -0.1008 & 0.3738 & -0.9785 \\ 0.9935 & -0.1008 & 1.0000 & 0.8803 & -0.1018 \\ 0.9229 & 0.3738 & 0.8803 & 1.0000 & -0.5492 \\ -0.2008 & -0.9785 & -0.1018 & -0.5492 & 1.0000 \end{pmatrix}. \tag{2.30}$$

The errors obtained are shown in Tables 1 and 2 for all possible parameter combinations of the 5 parameters, keeping a subset fixed, denoted by the dashes. In each table, the first five columns give the predicted absolute errors of the scaled variables. In Table 1, there is no a priori information on the parameters. As is seen, fixing T_i in the 5-parameter fit makes the resulting 4-parameter composition fit error much smaller while leaving the other errors practically speaking the same. In Table 2, it is supposed that the ion temperature is known to be T_0 with a 10% accuracy (also those cases where it is supposed that T_i is known exactly are included; they of course duplicate the corresponding lines in Table 1). The result is almost equal to the previous case: knowing the ion temperature to less than 10% leads to about the same errors in the composition multiparameter fits.

The last column gives for each case the logarithm of the condition number of Fisher information matrix Q. The condition number may be defined as the ratio of the square roots of the largest and smallest eigenvalues of the matrix Q (Golub and Van Loan, 1983). It is a measure of the stability of the solutions of linear equations $Ax = y$; an upper bound on the relative solution error $\Delta x/x$ in solving for x by inverting A is equal to the condition number times the relative data error $\Delta y/y$. Thus the larger the condition number is, the larger the relative solution error may be. In the decomposition $Q = O\Lambda^2 O^{\mathrm{T}}$, where O is an orthogonal matrix and Λ^2 a diagonal matrix with non-negative diagonal elements $\lambda_1^2, \ldots, \lambda_n^2$, where the values λ_i^2 are the eigenvalues of Q and the rows of O its orthogonal eigenvectors, the condition number is the ratio $\lambda_{max}/\lambda_{min}$. In this orthogonal basis, the equiprobability surfaces of the a posteriori distribution (2.10) are ellipsoids whose semiaxes are proportional to the inverses of λ_i. Therefore, the condition number also measures the degree of oblateness of the a posteriori ellipsoids, and the larger the condition number, the more oblate (degenerate) the ellipsoid is and the closer Q is to a singular matrix.

Table 1 may be used in an actual experiment planning and evaluation as a quick-look table by multiplying the entries of the table by the ratio γ/γ_0, where γ is the true value of the ACF noise level calculated from (2.24). Table 1 was calculated assuming the noise level to be equal to 1%. The resulting error in the electron density estimate is then about 1.4%, and this is the statistical accuracy scale of the power profile measurements used here. The usual thumb rule is that the statistical accuracy ot the power profile measurements is the same as the accuracy of the electron density estimate and half of the accuracy of the estimates for T_i and T_e (Rishbeth and Williams, 1985).

Fitting the composition and the temperature at the same time in any of the

multiparameter fits leads to relative errors which are of the order of 100% or more. Therefore, the power profiles have to be known to an accuracy which is an order of magnitude better than about one per cent used in the previous calculations in order to be able to fit the composition to a satisfactory degree of accuracy. This may be achieved by integrating the data long enough. Alternatively, the ion temperature T_i may be known with an accuracy of at least 10%. The result of knowing T_i with this accuracy may be seen in Table 2. We see that the relative errors are still large, up to about 50%, except in the 5–parameter fit, where the error is almost 100%. It is important to note that in this case only a modest integration in time is necessary to bring the composition fit errors down to reasonable values. All of this is qualitatively well known among the pundits of incoherent scatter measurements. However, we have shown the existence of a quantitative tool to calculate and compare the errors for all possible parameter combinations in a given experiment, and we can say that these calculations show that a multiparameter composition fit is possible at least when the ionosphere is reasonably stable.

A comparison with real experiments shows that the errors reported by one of the standard analysis packages tend to be about 25% higher than the values obtained by the inversion method estimate (Huuskonen, 1987).

There are two questions of practical importance which will have to be answered separately by studying the behaviours of the theoretical ACFs for the given parameter vector x described above. The first one is: what is the poorest lag resolution $d\tau_{max}$ which still gives the same error bars as an infinitely good lag resolution (when the the lag extent $\Delta\tau$ is kept fixed)? The second question is: what is the shortest lag extent $\Delta\tau_{min}$ which still gives us the same error bars as an infinitely long lag extent (when the lag resolution $d\tau$ is kept fixed)? It is known that already in the case of an infinitely long and dense sampling of the ACF, the inversion problem for the plasma parameters is rather unstable. It is therefore evident that by choosing in a specific experiment either too coarse a lag resolution or too short a lag extent, this instability may become worse and lead to a radical increase in the errors for some lag resolution and lag extent values. The absolute values of the errors are not important, but rather the stability intervals of the fits. In the former case, the sampling interval for which the errors become unstable is the analogue of the Nyqvist frequency for the particular fit considered. We can therefore expect that the instability occurs at the latest when the sampling interval $d\tau$ approximately equals τ_0.

We form the error estimates from our basic ACF by sampling it at multiples of the basic sampling interval $0.1\,\tau_0$ up to lag $6\,\tau_0$; for practical purposes, this corresponds to an infinitely dense and long sampling of the ACF. We first plot the behaviour of the errors as a function of the lag resolution, but keeping the maximum lag fixed at about $6\,\tau_0$. Adding more lags only adds very small terms to the information matrix, but has no visible effect on the results.

Figure 5b shows that without the zero lag data, even sampling at $0.7-0.8\,\tau_0$ is adequate from the point of view of the stability of the errors. In Figures 6a and 6b, $T_e/T_i = 2$. This has the effect of requiring a lag resolution of at least $0.6 - 0.7\,\tau_0$.

In Figures 7a and 7b, and similarly in Figures 8a and 8b, the 5–parameter fit is considered. The effect of using the zero lag data is clearly seen here. However, we may note that it is the collision frequency which is most sensitive to zero lag data. All the other parameter errors start growing at about $0.4\,\tau_0$. In the composition fits, one may in most cases set the collision frequency equal to zero and fit the four remaining parameters. If one does use the zero lag data, then sampling at about $0.7\,\tau_0$ will be sufficient.

The effect of varying the lag extent can be studied in a similar way by starting from the same ACFs. We keep the lag resolution fixed at $0.1\,\tau_0$, vary the longest lag and calculate the a posteriori errors. Since for positive definite symmetric matrices it is generally true that $(Q^{-1})_{kk} \geq (Q_{kk})^{-1}$, we know that adding more lags will increase Q_{kk} and therefore decrease the lower bound of the errors, and probably also the errors themselves. This is seen in Figures 9a and and 9b, where we present the two cases of a 3–parameter fit for $T_e/T_i = 1$ and $T_e/T_i = 2$, respectively. It is seen that the errors do not change appreciably after about $1.8\,\tau_0$. The zero lag has practically no effect on the curves. For the 5–parameter fits in Figures 10a and 10b, stabilisation occurs at about $2.8\,\tau_0$ in both cases.

In order to see how well these conditions are satisfied in an individual experiment, we consider a multipulse experiment known with the code type EFOR-V4 (code F in Turunen and Silén, 1984; Lehtinen and Huuskonen, 1986). We only use the lag resolution and the lag extent of the experiment in these considerations; the zero lag is measured, but its variance is four times that of the other lags, meaning that the errors using this variance will be twice as large. The actual result will be between the bounds calculated with and without the zero lag. The code has the lag resolution $d\tau = 40\,\mu$s, and 7 lags are used, resulting in a lag extent $\Delta\tau = 240\,\mu$s, meaning $d\tau = 0.4\,\tau_0$ and $\Delta\tau = 2.4\,\tau_0$. For 3– and 5–parameter fits, the lag resolution is certainly sufficient; the lag extent is sufficient for the 3–parameter fit, and almost sufficient for the 5–parameter fit. This is true for $T_i = 300$ K.

We may also consider for which ion temperature intervals (T_{min}, T_{max}) this code is sufficient in the whole interval $1 \leq T_e/T_i \leq 2$. A change in T_i or m_i scales the τ-axis by $\sqrt{T_i/m_i}$. For increasing T_i, it moves the zero-crossing closer to the origin. Hence T_{max} determines the lag resolution requirement, and T_{min} sets the scale for the lag extent. Let $d\tau_{max}$ be the coarsest lag resolution for which this code is sufficient, and $\Delta\tau_{min}$ similarly the smallest lag extent which is sufficient for this code. Then

$$\begin{aligned} T_{max} &= \left(\frac{d\tau_{max}}{d\tau}\right)^2 T_i \\ T_{min} &= \left(\frac{\Delta\tau_{min}}{\Delta\tau}\right)^2 T_i. \end{aligned} \tag{2.31}$$

As mentioned earlier, the zero lag is measured in this code, but it has fourfold variance. Since we have used only one value for σ in calculating the behaviour of errors as a function of lag resolution and lag extent, a conservative estimate would be that the true validity intervals will be somewhere between those obtained by using the same σ also for the zero lag and not using the zero lag data at all, which

is equivalent to putting the variance equal to infinity. An educated guess is that the true answer is quite close to the former one. (It would not be difficult to do the exact calculation using the correct Σ_m, but we shall refrain from doing it here.)

With the zero lag data and for $T_e/T_i = 1$, the sampling interval $1.1\,\tau_0$ is in the the 3–parameter fit case sufficient for $T_i = 300$ K, $d\tau = 0.4\,\tau_0$ means $T_{max} = 2300$ K and $1.8\,\tau_0$ means $T_{min} = 170$ K, resulting in the validity interval 170 K $< T_i <$ 2300 K. For the 5–parameter fit, we get similarly roughly the interval 400 K $< T_i <$ 900 K. For $T_e/T_i = 2$, the corresponding intervals are 130 K $< T_i <$ 1500 K for the 3–parameter fit and $350 < T_i <$ 900 K for the 5–parameter fit. This means that for T_e/T_i between 1 and 2, this code is sufficient for 3–parameter fits when the ion temperature is between 170 and 1500 K, and for 5–parameter fits between 400 and 900 K.

Without the zero lag data, we would get correspondingly a stable region for the 3–parameter fits between 170 and 900 K, but nothing for the 5-parameter fits. Therefore, we can only state that the 3-parameter fit is certainly valid for $170 < T_i <$ 900 K and that the 5–parameter fits will be unstable except possibly in a narrow interval starting from 400 K.

2.6. Conclusions

We have presented a general method to estimate the probable errors in IS measurements under given ionospheric conditions. We have shown that it is possible to determine a sufficient lag resolution and lag extent for a stable solution. We have also shown that the error estimates have, due to the scaling properties of the ACFs, a larger range of validity than just the neighbourhood of the parameter point considered. The ACF scale is τ_0, which depends on the ion thermal velocity only. The values of the minimum lag resolution and lag extent in an actual experiment may be obtained from the curves presented here without doing the calculations explicitly, and these results can be effectively used in assessing the feasibility of a particular experiment when one wants to determine the values of ionospheric parameters.

One of the assumptions in the discussion of the previous Sections is that the plasma (or parts of it) is not drifting. If the drift is not negligible, but of the order of ω_0 in frequency space, one may in fact have to sample with a much better lag resolution than was deduced, since a slower sampling may give a wrong ACF through aliasing. However, it may be shown that the full a posteriori distribution will separate into a product of a velocity distribution and a distribution for the other variables. This means that the errors will be independent, and thus our conclusions will not be seriously affected.

Table 1

ERROR ESTIMATES OF THE PLASMA PARAMETERS WITH NO PRIOR INFORMATION ON THE PARAMETERS

N_e/N_0	T_i/T_0	T_e/T_i	ν/ν_0	p	$\log_{10}$
1.000	1.000	1.000	0	0.300	cond(Q)
–	–	–	–	0.034	0.000
–	–	–	0.156	–	0.000
–	–	–	0.186	0.040	1.496
–	–	0.019	–	–	0.000
–	–	0.028	–	0.049	0.974
–	–	0.023	0.188	–	1.982
–	–	0.029	0.194	0.051	2.089
–	0.023	–	–	–	0.000
–	0.157	–	–	0.229	2.322
–	0.031	–	0.206	–	1.906
–	0.366	–	0.434	0.483	3.161
–	0.037	0.031	–	–	0.934
–	0.199	0.036	–	0.264	2.719
–	0.041	0.031	0.207	–	2.248
–	0.805	0.064	0.782	0.999	4.031
0.014	–	–	–	–	0.000
0.015	–	–	–	0.037	0.888
0.014	–	–	0.156	–	2.117
0.015	–	–	0.192	0.045	2.317
0.025	–	0.035	–	–	1.112
0.030	–	0.057	–	0.060	1.712
0.071	–	0.121	0.535	–	3.358
0.294	–	0.561	1.884	0.210	4.491
0.015	0.025	–	–	–	0.611
0.015	0.159	–	–	0.230	2.678
0.016	0.036	–	0.219	–	2.450
0.033	0.801	–	0.964	1.015	4.196
0.035	0.052	0.072	–	–	1.840
0.113	0.747	0.267	–	0.852	4.066
0.288	0.166	0.561	1.698	–	4.431
0.294	0.805	0.564	2.031	1.020	4.626

Table 2

ERROR ESTIMATES OF THE PLASMA PARAMETERS WITH 10% A PRIORI –ACCURACY FOR T_i

N_e/N_0 1.000	T_i/T_0 1.000	T_e/T_i 1.000	ν/ν_0 0	p 0.300	$\log_{10}$ cond(Q)
–	–	–	–	0.034	0.000
–	–	–	0.156	–	0.000
–	–	–	0.186	0.040	1.496
–	–	0.019	–	–	0.000
–	–	0.028	–	0.049	0.974
–	–	0.023	0.188	–	1.982
–	–	0.029	0.194	0.051	2.089
–	0.023	–	–	–	0.000
–	0.084	–	–	0.126	1.803
–	0.029	–	0.202	–	1.912
–	0.096	–	0.213	0.133	2.167
–	0.035	0.030	–	–	0.885
–	0.089	0.030	–	0.126	2.049
–	0.038	0.030	0.204	–	2.240
–	0.099	0.030	0.215	0.133	2.407
0.014	–	–	–	–	0.000
0.015	–	–	–	0.037	0.888
0.014	–	–	0.156	–	2.117
0.015	–	–	0.192	0.045	2.317
0.025	–	0.035	–	–	1.112
0.030	–	0.057	–	0.060	1.712
0.071	–	0.121	0.535	–	3.358
0.294	–	0.561	1.884	0.210	4.491
0.015	0.025	–	–	–	0.585
0.015	0.085	–	–	0.126	2.143
0.016	0.034	–	0.213	–	2.426
0.016	0.099	–	0.225	0.133	2.548
0.033	0.046	0.066	–	–	1.752
0.033	0.099	0.067	–	0.128	2.335
0.160	0.086	0.307	0.988	–	3.954
0.294	0.099	0.561	1.887	0.244	4.535

3. The effect of a priori accuracies

3.1. Introduction

The dependence of an error in a multiparameter fit on the a priori width of another parameter is studied and an exact formula is given. This formula makes it possible to estimate how accurate the a priori information on a given parameter must be in order to obtain the other parameter with a prescribed accuracy. These estimates are useful in those cases where some errors are strongly correlated, like the ion temperature and the O^+ content in incoherent scatter measurements with the EISCAT UHF radar. The dependence of the solution on the accuracy of the power profile estimates is given. Likewise, the behaviour of the error when a linear combination of the parameters is known exactly is given. The formulae given have a very simple geometrical meaning, and together the two sets of formulae give a complete solution to the problem of estimating the accuracies in a multiparameter fit when a linear combination of the variables is known with a prescribed accuracy.

3.2. General results

In a typical multiparameter fit to a physical model, one of the main difficulties is the fact that when the number of parameters to be fitted grows, the parameter estimation errors grow more or less explosively with the number of correlated parameters. There are two cases of practical importance which effectively reduce the number of parameters in the fit: *i)* one parameter is known with a given accuracy or *ii)* a linear combination of some of the parameters is known exactly. The third case, when a linear combination is known with a prescribed accuracy, is a combination of the two.

As in the previous Chapter, we suppose that the measurement vector $m = (m_1, \ldots, m_M)^{\mathrm{T}}$ depends on the parameter vector $x = (x_1, \ldots, x_N)^{\mathrm{T}}$ through the equation

$$m = Ax + \epsilon, \tag{3.1}$$

where ϵ is a zero-mean Gaussian measurement noise vector with covariance matrix Σ_m. The a priori distributions are supposed to be Gaussian with centre points x_0 and error covariance matrix Σ_0:

$$D_{pr}(x) = C\, exp\big(-\frac{1}{2}(x - x_0)^{\mathrm{T}}\Sigma_0^{-1}(x - x_0)\big). \tag{3.2}$$

It was shown in the previous Chapter that the a posteriori distribution is of the standard Gaussian form

$$D_p(x) = C\,exp\big(-\frac{1}{2}(x-\overline{x}_0)^{\mathrm{T}}\Sigma_p^{-1}(x-\overline{x}_0)\big), \tag{3.3}$$

where Σ_p is the full a posteriori covariance matrix given by the inverse of the Fisher information matrix: $\Sigma_p = Q^{-1}$ and

$$Q = \Sigma_0^{-1} + A^{\mathrm{T}}\Sigma_m^{-1}A. \tag{3.4}$$

The centre point vector $\overline{x}_0$ is both the maximum likelihood point x_{ML} and the mean value of the posteriori distribution. The centre point is defined by

$$\overline{x}_0 = Q^{-1}(\Sigma_0^{-1}x_0 + A^{\mathrm{T}}\Sigma_m^{-1}\overline{m}), \tag{3.5}$$

where the vector $\overline{m}$ represents the measured values, and the a posteriori distributions are simply negative exponentials of half the positive definite quadratic forms

$$S(x) = (x-\overline{x}_0)^{\mathrm{T}}Q(x-\overline{x}_0) + S(\overline{x}_0). \tag{3.6}$$

The matrices A are the derivative matrices of the full theory function f and the parameters x are random variables with distribution functions which are multivariate normal (Gaussian) with mean $\overline{x}_0$ and covariance Σ_p in the usual statistical terminology (Rao, 1973; Tarantola, 1987). It follows that $\frac{1}{2}S$ has a χ^2 distribution with $\dim(M)-N$ degrees of freedom. In the following, the distributions being Gaussian leads to several useful simplifications.

3.3. The parameter blocks

The estimation errors, being equal to the second moments of the a posteriori distribution, are given by the square roots of the a posteriori covariance matrix Σ_p. It is not necessary to consider the effect of a priori information on all parameters simultaneously. In fact, we only need to consider the effect of the a priori knowledge on one parameter at a time; the equations for the rest of the parameters are all similar. This is seen as follows. When we have prior knowledge on a block of parameters x, the bare a posteriori covariance matrix $\Sigma_b = (A^{\mathrm{T}}\Sigma_m^{-1}A)^{-1}$ is written in the block form

$$\Sigma_b = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}, \tag{3.7}$$

where y denotes the block of the rest of the variables and the parameter vector has the block form $(x;y)$. From the properties of the Gaussian distributions, it follows that the matrix Σ_{xx} is the marginal a posteriori covariance matrix of the x–variables (likewise for Σ_{yy}) and Σ_{xy} the covariance matrix of the variables x, y (see *e.g.* Rao, 1973). The corresponding information matrix Q_b is similarly written in the block form

$$Q_b = \begin{pmatrix} Q_{xx} & Q_{xy} \\ Q_{yx} & Q_{yy} \end{pmatrix}. \tag{3.8}$$

When the a priori matrix is

$$\Sigma_0 = \begin{pmatrix} \Sigma_{0xx} & 0 \\ 0 & 1\infty \end{pmatrix}, \tag{3.9}$$

with an inverse

$$Q_0 = \begin{pmatrix} Q_{0xx} & 0 \\ 0 & 0 \end{pmatrix}, \tag{3.10}$$

where $Q_{0xx} = \Sigma_{0xx}^{-1}$, then the full a posteriori covariance matrix Σ_p is given by

$$\Sigma_p = (Q_b + Q_0)^{-1} \tag{3.11}$$

and hence the full a posteriori distribution by

$$D_p(x,y) = C\, exp(-\frac{1}{2}(x;y)^{\mathrm{T}}(Q_b + Q_0)(x;y)^{\mathrm{T}}). \tag{3.12}$$

Here we have neglected the dependence on $\overline{x}_0$, since we are not interested the position of the maximum likelihood point, but only in the behaviour of the diagonal elements of the a posteriori covariance matrix. Therefore, we only keep the quadratic coefficients in the variable $(x;y)$ The marginal distribution of x, when we are not interested in the rest of the variables y, is

$$\begin{aligned} D_p(x) =& C \int D_p(x,y)dy \\ =& C \int exp(-\frac{1}{2}(x;y)^{\mathrm{T}}(Q_b + Q_0)(x;y))dy \\ =& C\, exp(-\frac{1}{2}x^{\mathrm{T}}Q_{0xx}x) \int exp(-\frac{1}{2}(x;y)^{\mathrm{T}}Q_b(x;y))dy, \end{aligned} \tag{3.13}$$

and since the integral on the right is by definition just the marginal distribution of x, we have

$$\begin{aligned} D_p(x) =& C\, exp(-\frac{1}{2}x^{\mathrm{T}}Q_{0xx}x)exp(-\frac{1}{2}x^{\mathrm{T}}\Sigma_{xx}^{-1}x) \\ =& C\, exp(-\frac{1}{2}x^{\mathrm{T}}((\Sigma_{0xx})^{-1} + (\Sigma_{xx})^{-1})x). \end{aligned} \tag{3.14}$$

Since the result does not depend on y, we see that the uninteresting variables can be neglected, and *the marginal distribution in x only depends on the parameters in the x-block.*

3.4. One variable is known with a prescribed accuracy

We start by studying the simple case of the effect of a finite width of the a priori distribution in one variable on the width of the marginal distribution of any other variable. In practical terms, how the accuracy in the some parameter depends on the a priori accuracy of a given parameter.

Let the a priori distribution of x_1 have a width p; then the expression for Σ_0 is

$$\Sigma_0 = \begin{pmatrix} p^2 & 0 \\ 0 & \infty \end{pmatrix}, \tag{3.15}$$

and for Σ_b

$$\Sigma_b = \begin{pmatrix} a^2 & ab\delta \\ ab\delta & b^2 \end{pmatrix}, \tag{3.16}$$

where a in turn is the width of the marginal distribution of the first parameter x_1, b that of the second parameter x_2 and δ their correlation coefficient, which may have any sign. A two-dimensional a posteriori ellipsoid and the a priori on x_1 are represented in Fig. 12.

Q and hence Σ_p may be calculated easily; the result is

$$\Sigma_p = \frac{1}{1+(\frac{p}{a})^2} \begin{pmatrix} p^2 & ab\delta(\frac{p}{a})^2 \\ ab\delta(\frac{p}{a})^2 & b^2(\epsilon^2+(\frac{p}{a})^2) \end{pmatrix}, \tag{3.17}$$

where we have put for simplicity $\epsilon^2 = 1-\delta^2$.

Let us consider the behaviour of the errors as a function of the a priori uncertainty p for small values of p (compared with a). For x_1 the result is

$$(\Delta x_1)^2 = \frac{p^2}{1+(\frac{p}{a})^2} \approx p^2 \tag{3.18}$$

so that $\Delta x_1 \approx p$ for small p and

$$\begin{aligned} (\Delta x_2)^2 &= b^2 \frac{\epsilon^2+(\frac{p}{a})^2}{1+(\frac{p}{a})^2} \\ &\approx b^2(\epsilon^2+(\frac{p}{a})^2) \end{aligned} \tag{3.19}$$

so that $\Delta x_2 \approx b\sqrt{\epsilon^2+(\frac{p}{a})^2}$. This may be summarised as follows: for small p, the uncertainty (error) in the first parameter x_1 is directly equal to the a priori uncertainty, while for very large p it saturates to the value a of the marginal distribution; in the case of the second parameter x_2, for *very* small p, the error is the constant $c = b\epsilon$ obtainable from the value of the *conditional* a posteriori distribution supposing that x_1 is known exactly, then it starts growing roughly

linearly in p, until it for very large p saturates to the value b, the width of the marginal distribution. When the errors are strongly correlated, $\epsilon \approx 0$, and we see from (3.19) that the uncertainty in the second parameter is reduced by the factor p/a for small p. If the errors are not correlated, $\epsilon \approx 1$, and the a priori knowledge does not help to resolve the second parameter.

3.5. A linear combination of the variables is known exactly

As in the previous Section, we can study a limited problem of originally 3 variables x, y and z with a bare a posteriori covariance matrix Σ_b:

$$\Sigma_b = \begin{pmatrix} a^2 & ab\alpha & ac\beta \\ ab\alpha & b^2 & bc\gamma \\ ac\beta & bc\gamma & c^2 \end{pmatrix}, \tag{3.20}$$

where the diagonal elements a^2, b^2 and c^2 are variances of the variables and α, β, γ their correlation coefficients. We suppose that there is an independent measurement of x and y, so that $y = \lambda x$ exactly. The quadratic form described by (3.20) will now depend on only two parameters, x and z. It is a simple matter to substitute λx for y and to calculate the resulting covariance matrix for the two remaining independent variables. The result is

$$\begin{aligned} \tilde{\Sigma}_{xz} = \tilde{\Sigma}_{zx} =& (ab^2c(\beta - \gamma\alpha) + \lambda a^2 bc(\gamma - \alpha\beta))/\Delta \\ \tilde{\Sigma}_{xx} =& a^2b^2(1-\alpha^2)/\Delta \\ \tilde{\Sigma}_{zz} =& (b^2c^2(1-\gamma^2) + 2\lambda abc^2(\gamma\beta - \alpha) + \lambda^2 a^2c^2(1-\beta^2))/\Delta, \end{aligned} \tag{3.21}$$

where Δ is equal to

$$\Delta = a^2\lambda^2 + b^2 - 2\alpha\lambda ab. \tag{3.22}$$

These formulas may be very simply written in a geometric form by using the fact that all matrices involved are positive definite. The a posteriori covariance matrix Σ_b in (3.20) may be represented *e.g.* in the form $\Sigma_b = GG^{\mathrm{T}}$ where G consists of row vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ in some orthonormal basis. Their squared lengths are equal to the variances a^2, b^2 and c^2, and the correlation coefficients α, β, γ are cosines of their relative angles. When we further denote the vector $\mathbf{W}$ by

$$\mathbf{W} = \lambda\mathbf{a} - \mathbf{b}, \tag{3.23}$$

then we have that the squared length of the new variable $\mathbf{W}$ is equal to $\mathbf{W}^2 = \lambda^2\mathbf{a}^2 + \mathbf{b}^2 - 2\alpha\lambda ab = \Delta$. Also the equation (3.21) will have a simple geometrical meaning, when we use vector algebraic notation. The most important equation is the one for the variance of the independent variable z:

$$\begin{aligned} \tilde{\Sigma}_{zz} =& c^2(1 - (\lambda ab - b\gamma)^2/\Delta) \\ =& c^2(1 - (\mathbf{c}\cdot\mathbf{W})^2)/\mathbf{W}^2 \\ =& (\mathbf{c}\times\mathbf{W})^2/\mathbf{W}^2. \end{aligned} \tag{3.24}$$

We see immediately that when $\mathbf{c}$ is parallel with the new vector $\mathbf{W}$, the variance may become zero. If the vectors are perpendicular, knowing $\mathbf{W}$ has no effect on the variance of z. This is possible, when the variables are suitably correlated. The knowledge of the linear combination also affects the first variable:

$$\begin{aligned}\tilde{\Sigma}_{xx} =& \mathbf{a}^2(1-(\mathbf{a}\cdot\mathbf{W})^2)/\mathbf{W}^2 \\ =& (\mathbf{a}\times\mathbf{W})^2/\mathbf{W}^2,\end{aligned} \tag{3.25}$$

a completely analogous formula, and also the correlation coefficients:

$$\tilde{\Sigma}_{\mathbf{xz}} = (\mathbf{a}\times\mathbf{W})\cdot(\mathbf{c}\times\mathbf{W})/\mathbf{W}^2, \tag{3.26}$$

in a most natural way in geometric terms.

3.6. Applications to incoherent scatter

The results of Section 4 mean that Table 2 in Chapter 2 is superfluous. It may be calculated from Table 1 by using $p = 0.1$ and Eq.(3.17). We apply now these results to the EISCAT UHF radar operating at about 933 MHz. In Eq. (3.19) the values a and b may be read off as the errors in the particular fit considered, (*i.e.*, the widths of the marginal distributions), and the constant c as that value which corresponds to the same fit but with the parameter x_1 kept fixed (*i.e.*, as the width of the conditional distribution supposing that x_1 is known).

As was pointed out in Section 3, the a priori knowledge on some parameter is only then useful when the errors are highly correlated, meaning that ϵ is small. If ϵ is not small, a priori knowledge is of little use. In incoherent scatter, it is well known that the errors in T_i and ion content are highly correlated (with a negative correlation coefficient), so that ϵ is very small indeed. Then we see that for the cases of fitting T_i and O^+ content simultaneously, *the error in either parameter is linear in the a priori width of the other one* provided the a priori width (to be of any use) is comfortably smaller than the width of the marginal distribution. Taking the numbers from Table 1, we plot the behaviour of the error in O^+ content supposing that ν_{in} is zero (usually the case where composition is important), but fitting the remaining four parameters as a function of the a priori accuracy of T_i in Fig. 13.

The tables in the previous Chapter have been calculated with the assumption that the ACF noise level is $\gamma_0 = 0.01$, which means that the power profile estimates have a statistical accuracy 0.014 or 1.4%, which is a very common number in practice. It is easy to understand that integrating long enough reduces the amount of noise in the data, and hence any desired accuracy in the fitted parameters may be obtained. We pose the following question: between which values of the noise level (or SNR) is the a priori information useful? This question may be answered without doing an explicit calculation. The best case is when the 4–parameter fit $(N_e, T_i, T_e/T_i, p(O^+))$ itself gives the desired accuracy, meaning that any width of

the a priori distribution will do (the width of the marginal distribution equals the desired accuracy); the worst case is when the 3–parameter fit $(N_e, T_e/T_i, p(O^+))$ itself gives the desired accuracy, meaning that the a priori width is zero or that the first parameter is known exactly (the width of the conditional distribution is equal to the required accuracy).

When the noise level γ is increased by a factor $\xi = \gamma/\gamma_0$, the numbers a and b are scaled by the same number, but the error correlation coefficient δ and hence ϵ remain the same. Then we have that the error Δx_2 becomes

$$(\Delta x_2)^2 = \frac{(\xi\epsilon b)^2 + (\frac{pb}{a})^2}{1 + (\frac{p}{\xi a})^2}, \tag{3.27}$$

so that given a wanted accuracy $\Delta = \Delta x_2$, the required accuracy p for the a priori on x_1 becomes

$$(\frac{p}{a})^2 = \xi^2 \frac{\Delta^2 - \xi^2\epsilon^2 b^2}{\xi^2 b^2 - \Delta^2}. \tag{3.28}$$

As noted earlier, it is easily seen that a solution for p only exists for

$$\frac{\Delta}{b} < \xi < \frac{\Delta}{b\epsilon}. \tag{3.29}$$

Therefore, the noise level relative to γ_0 can be at most $\Delta/b\epsilon$, but it has to be at least Δ/b; the numbers b and $b\epsilon = c$ are the widths of the marginal and conditional distributions, respectively.

By taking again the numbers from the 4–parameter fit, we see that a 10% accuracy for the composition fit is obtainable with an a priori knowledge on T_i/T_0 for $0.117 < \xi < 1.667$. This is equivalent to knowing the power profiles with an accuracy between 11.7% and 0.8%. The behaviour of p/a is illustrated in Fig. 14.

The following question is important in practice in the design of experiments: how well should one know the ion temperature in order to obtain the ion composition with an accuracy of, say, $\pm 10\%$, for a given accuracy of the power profile estimate?

In the case considered in Table 1, the answer for the 2–parameter $(T_i, p(O^+))$, 3–parameter $(N_e, T_i, p(O^+))$, 4–parameter $(N_e, T_i, T_e/T_i, p(O^+))$ and 5–parameter $(N_e, T_i, T_e/T_i, \nu_{in}, p(O^+))$ fits is the following (the oxygen content was 30%, so we are asking for a relative accuracy of 33% only): for the 5–parameter fit, there is no solution: even if the ion temperature is known exactly, the error in composition is still 21%; for the 4–parameter fit, we have p=7.1%; for the 3– and 2–parameter fits, we get 7.1% and 7.2%. These numbers are remarkably similar, and a general answer is therefore that knowing T_i to better than 7% leads to composition errors that are 10% or better. This is obviously a consequence of the high correlation of the ion temperature and composition errors.

3.7. Conclusions

The error tables presented in the previous Chapter may be used effectively in studying the dependence of an error in a multiparameter fit on the a priori width in another parameter. This method makes it possible to estimate how accurate the a priori information must be in order to obtain the other parameter with a prescribed accuracy. These estimates are useful in the cases where the errors in some parameters are strongly correlated, like the ion temperature and the O^+ content in incoherent scatter measurements with the EISCAT UHF radar. The dependence of the solution on the accuracy of the power profile estimates is given.

4. Effect of parameter mixing within a measuring volume

4.1. Introduction

The effect of having parameters varying over a volume under consideration rather than being constant is studied in the framework of statistical inversion theory. It is shown that using parameter averages is equivalent to have the theory corrected by the covariances of the variables coupled with the second derivatives of the theory function. If the parameter distributions were known exactly, this would only introduce a bias in the theory and hence a systematic error in the parameter centre point estimates. When the distributions are not known exactly, there is another source of error consisting of the uncertainties in the parameter variation estimates. This leads to new error bounds on the allowed parameter variability within the volume under consideration if some prescribed accuracy in the parameter estimates is wanted. These considerations are applied to IS measurements, where it is important to be able to estimate the effect of integrating both in space and time over a volume with varying parameters in order to obtain spectra or autocorrelation functions. Numerical examples are given in the case of the O^+ content estimates in measurements with the EISCAT UHF radar, when *e.g.* the ion temperature varies over the integration ranges. The results obtained may be used in the design of experiments when high resolution composition measurements are required.

4.2. General results from traditional inversion theory

In the framework of general statistical inversion theory discussed in the previous Chapters, we study the modifications to the first-order linear theory when the measurements are made over volumes where the parameters to be estimated are allowed to vary. As earlier, it is assumed that the measurement vector $m = (m_1, \ldots, m_M)^{\mathrm{T}}$ and the unknown parameter vector $x = (x_1, \ldots, x_N)^{\mathrm{T}}$ are connected by the equation

$$m = f(x) + \epsilon, \tag{4.1}$$

where ϵ is a Gaussian random variable with zero mean and covariance Σ_m. In incoherent scatter measurements, f is the nonlinear function describing the dependence of the theoretical ACF values on the unknown plasma parameters x within some measuring volume V.

In statistical inversion theory, one tries to estimate the parameter values x when we have the measurements $\overline{m}$ and prior information on the values and accuracies of the parameters x. It is normally assumed that the parameters x are

constant within V. However, in many practical cases, *e.g.* when a measurement consists of receiving scattered signals from a random medium, the parameters vary over the volume where the signal comes from, according to the volume resolution set by the experimental setup and by the hardware requirements, and it is also necessary to integrate, *i.e.*, use time averages of the samples obtained in order to have a tractable signal. Then it is important to ask the question when the variation of the parameters over space and time is significant enough to affect conclusions deduced on the basis of the simple linear theory.

In the linear theory, instead of the previous equation, one has

$$m = Ax + \epsilon, \tag{4.2}$$

where the matrix A is most often a linear approximation consisting of the first derivatives of the full nonlinear theory function f with respect to the parameters of the theory. In the linear theory, the parameter error estimates are obtained as the square roots of the variances given by the diagonal of the a posteriori covariance matrix Σ_p which is the inverse of the Fisher information matrix Q:

$$Q = \Sigma_0^{-1} + A^{\mathrm{T}}\Sigma_m^{-1}A. \tag{4.3}$$

where Σ_0 is the a priori accuracy covariance matrix. If there are more independent measurements with Gaussian errors ϵ_i, then each measurement adds a contribution of the form $A^{\mathrm{T}}\Sigma_i^{-1}A$ to the information matrix.

The centre point vector $\overline{x}_0$ of the a posteriori distribution is in the linear and Gaussian case both the maximum likelihood point and the mean value of the a posteriori distribution, and is given by

$$\overline{x}_0 = Q^{-1}(\Sigma_0^{-1}x_0 + A^{\mathrm{T}}\Sigma_m^{-1}\overline{m}) \tag{4.4}$$

where x_0 is the centre point of the a priori distribution and $\overline{m}$ denotes the measured values in the experiment. It is seen that each independent measurement also modifies the a posteriori centre point value.

4.3. Linear inversion theory revisited with incoherent scatter in mind

Real IS measurement values are averages of the sort described in the previous Section, since the responses come from a finite scattering volume and also because the response is a result of time integration that may range typically from fractions of a second to several hours. Depending on the situation, one may be interested in the variation of the plasma parameters as a function of range, across the radar beam, as a function of integration time or any combination of these factors. Nevertheless, it is possible to formulate all this by saying that the resulting measurement value

corresponds to an average over a probability measure μ on the space of the plasma parameters X. Then the previous equation will be modified to

$$m = \int_X f(x)\,\mu(dx) + \epsilon. \tag{4.5}$$

Let us also suppose that the variation of the parameters within the measuring volume both in space and in time is sufficiently small so that the measure μ is concentrated in a small region in the space of the unknowns X around a centre of gravity $\overline{x}_0$ defined by

$$\overline{x} = \int_X x\,\mu(dx). \tag{4.6}$$

We may then write the integral ((4.5)) as

$$\int_X \left(f(\overline{x}) + \sum_i \left(\frac{\partial f}{\partial x_i}\right)_{x=\overline{x}} (x-\overline{x})_i + \frac{1}{2}\sum_{i,j}\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{x=\overline{x}} (x-\overline{x})_i (x-\overline{x})_j \right) \mu(dx). \tag{4.7}$$

Defining the central covariance matrix of the parameter variations within the measurement volume by

$$\mathrm{Cov}_V[x]_{ij} = \int_X (x-\overline{x})_i (x-\overline{x})_j\,\mu(dx), \tag{4.8}$$

we arrive at the modified relationship

$$m = f(\overline{x}) + \sum_{i,j} C_{ij}\mathrm{Cov}_V[x]_{ij} + \epsilon, \tag{4.9}$$

where we have denoted by C the curvature matrix $\frac{1}{2}(\partial^2 f/\partial x_i \partial x_j)_{x=\overline{x}}$.

Now our set of unknowns consists of the centre point and the second central moments of the parameter distribution in the integration volume. That is, we might define a new vector of unknowns x' by

$$\begin{aligned} x' = ((\overline{x})_1, \ldots, (\overline{x})_N, &\mathrm{Cov}_V[x]_{11}, \ldots, \mathrm{Cov}_V[x]_{1N}, \\ &\mathrm{Cov}_V[x]_{22}, \ldots, \mathrm{Cov}_V[x]_{2N}, \ldots, \mathrm{Cov}_V[x]_{NN}). \end{aligned} \tag{4.10}$$

The standard way of doing analysis therefore corresponds to the introduction of very definite a priori knowledge about the second central moments $\mathrm{Cov}_V[x]_{ij}$: they are neglected, *i.e.*, supposed to vanish. This simply means that the parameters are considered to be constant within V. It would certainly be much better if we used estimates of the parameter *gradients* from outside the measurement volume to infer the changes of the parameters inside the measurement volume and thus to arrive at some estimates of their values. Still, the knowledge about the values of $\mathrm{Cov}_V[x]_{ij}$ will be inaccurate and in the following, we shall estimate the influence of this on

the estimation results of the first part of our new unknown vector x', that is, $\overline{x}$, the average of the parameter distribution in the measurement volume. There is, however, some definite a priori knowledge about the covariance matrix: since it is positive definite, the squares of off-diagonal elements can at most be equal to the product of the corresponding diagonal elements.

The approach described here is completely new; a first version has been presented at the URSI conference in Tel-Aviv 1988 by Markku Lehtinen in an invited talk. As such, it is *not* the same as having an inexact theory as discussed in Tarantola and Valette (1982a). The difference is that we suppose that we have an exact theory, but our measurement values are averages over volumes where the parameters vary. It is of course possible that we may eventually end up with the same sort of answers to the parameter estimate uncertainties.

4.4. Applications to incoherent scatter: linear results revisited

We compare corrections obtained from the covariance estimates with the results derived in the previous Chapters. We shall calculate tables for the a posteriori errors in several important cases, where the second moments are supposed to be either fixed or completely unknown; the question of a priori knowledge will be important in inferring how accurately the second moments must be known so that their uncertainties do not affect the final parameter estimation errors.

We estimate the errors by starting again from the linear theory obtained from (4.9) by linearisation:

$$m = f(x_0) + \sum_i D_i(\overline{x} - x_0)_i + \sum_{i,j} C_{ij}\mathrm{Cov}_V[x]_{ij} + \epsilon. \tag{4.11}$$

where we have denoted by $D_i = (\partial f/\partial x_i)_{x=\overline{x}}$ the matrix A of the traditional linear theory. The point x_0 is the result of the fit. The uncertainties in the position of x_0 are also uncertainties of the position of the mean value $\overline{x}$ of the parameter x. It is clear that these uncertainties are not affected by the fact that the parameter x has been varying over the integration volume, since they only depend on the values of the derivative matrix A at the fit point x_0. One may consider this also as an effect of averaging a first–order, linear theory. The moral is that in the linear theory, the position of the centre point estimate is affected by the bias introduced by the second moments, but the accuracy of the centre point estimates is not affected by the values of the covariances, but rather by the accuracies with which the covariances are known.

The first–order errors for the multiparameter fits for the value $x_0 = (1, 1, 1, 0, 0.3)^{\mathrm{T}}$, taken from Chapter 2, are for the same value of the SNR, resulting in an uncertainty of the electron density when fitted alone of 1.4%:

N_e/N_0	T_i/T_0	T_e/T_i	ν/ν_0	$p(O^+)$
1.000	1.000	1.000	0.000	0.300
0.035	0.052	0.072	–	–
0.113	0.747	0.267	–	0.852
0.288	0.166	0.561	1.698	–
0.284	0.805	0.564	2.031	1.020

If we suppose that we know the parameter covariances exactly, then the measurement vector m will be modified to m' where

$$m' = m - \sum_{i,j} C_{ij} \mathrm{Cov}_V[x]_{ij}. \tag{4.12}$$

This means that we have introduced a bias into the measurements. The bias correction to the parameter estimates in the linear theory, supposing no a priori knowledge, is from (4.4)

$$b = -(A^{\mathrm{T}} \Sigma_m^{-1} A)^{-1} A^{\mathrm{T}} \Sigma_m^{-1} \sum_{i,j} C_{ij} \mathrm{Cov}_V[x]_{ij}. \tag{4.13}$$

We may also rewrite this for diagonal Σ_m as

$$b = -\mathrm{pinv}(A) \sum_{i,j} C_{ij} \mathrm{Cov}_V[x]_{ij}, \tag{4.14}$$

where $\mathrm{pinv}(A)$ denotes the pseudoinverse of A.

The simplest case to consider is a situation where the measurement volume extends in only one direction and where we can suppose that the plasma parameters vary linearly. This means that the parameters remain constant in other directions. If we denote the direction in which the parameters may change by h, the situation is easily modelled by stating that $h \epsilon V = (-1, 1)$ and $x = x_0 + ah$. Then the covariances will be given by $\mathrm{Cov}_V[x]_{ij} = \frac{1}{3} a_i a_j$. We now consider the products $a_i a_j$ as our new variables. Although these variables do not change the value of the spectrum or ACF, they affect the value of the centre point x_0: the bias correction in terms of the fifteen independent products $c_k = a_i a_j (i \leq j)$ is given by the product of a 5×15 matrix B and the column vector of the covariances,

$$b = -Bc. \tag{4.15}$$

Here *e.g.* $c_1 = a_1^2, c_2 = a_1 \cdot a_2, \ldots, c_6 = a_2^2, \ldots, c_{15} = a_5^2$. The calculations are

detailed in the Appendix. The matrix B is given (for lack of space) by its transpose

$$B^{\mathrm{T}} = \begin{pmatrix} 0.000 & 0.000 & 0.000 & 0.000 & -0.000 \\ 0.003 & 0.317 & -0.013 & -0.047 & 0.052 \\ -0.005 & 0.354 & 0.296 & -0.117 & 0.028 \\ 0.001 & 0.340 & 0.310 & 0.252 & 0.039 \\ -0.002 & 0.305 & 0.307 & 0.238 & 0.437 \\ 0.924 & -3.645 & 2.046 & 4.730 & 3.853 \\ -0.572 & -1.340 & -1.032 & -2.470 & 2.348 \\ -0.206 & -1.352 & -0.311 & -0.529 & 1.957 \\ 1.751 & -6.811 & 3.875 & 9.684 & 7.520 \\ -0.515 & 0.868 & -1.239 & -1.716 & -0.575 \\ 0.292 & -0.639 & 0.615 & 1.005 & 0.518 \\ 0.126 & -2.261 & 0.392 & 0.028 & 2.817 \\ -0.039 & 0.111 & -0.084 & -0.158 & -0.105 \\ 0.223 & 0.385 & 0.413 & 1.441 & -0.749 \\ 0.426 & -1.334 & 0.922 & 1.716 & 1.310 \end{pmatrix}. \tag{4.16}$$

The effect of the diagonal elements (variances) is visible in the rows 1,6,10,13 and 15; it is seen that the most important single variance is that of the temperature and that the effect of the variability of the electron density is negligible. The diagonal elements are given by

$$d(B) = \begin{pmatrix} 0.0003 & 0.9241 & -0.5147 & -0.0391 & 0.4262 \\ 0.0005 & -3.6446 & 0.8682 & 0.1114 & 1.3341 \\ 0.0001 & 2.0463 & -1.2385 & -0.0840 & 0.9219 \\ 0.0000 & 4.7301 & -1.7159 & -0.1581 & 1.7161 \\ -0.0004 & 3.8533 & -0.5747 & -0.1048 & 1.3103 \end{pmatrix}. \tag{4.17}$$

The bias introduced by *e.g.* T_i varying 10% around its central value leads to the bias vector $b = (0.1)^2(-0.924, 3.645, -2.046, -4.730, -3.853)^{\mathrm{T}}$ correction to the values $x_0 = (1, 1, 1, 0, 0.3)$, that is, the 'correct' x_0 is

$$x_0 = (0.99, 1.03, 0.98, -0.05, 0.26)^{\mathrm{T}}. \tag{4.18}$$

The thumb rule emerging from (4.16) is that variabilities of the order of 10% lead to bias corrections that are of the order of 1%, except for the electron density, which has no effect at all. Since the bias correction is quadratic in the variabilities a_i, the bias corrections will be unimportant for smaller values. If, on the other hand, the variability of the ion temperature is known to be of the order of 1%, all bias corrections may be neglected.

When the parameter variabilities are not known exactly, or may not be estimated sufficiently accurately, the matrix A of the linear theory will have extra columns resulting from the sums in (4.14); these are (up to a factor of 1/3) either the second derivatives (diagonal elements) or twice the mixed derivatives multipying the ordered products $a_i a_j$.

For simplicity, we give only the results of the single variances of all parameters.

The effect of the parameter variations is seen in the following tables, where again we denote by a dash a parameter which is kept fixed. The last column gives the uncertainty of the parameter a_i. There is a marked effect in the parameter estimate uncertainties. The mere fact that we do not know *how much* a given parameter varies over the volume increases the estimated errors from the simple linear theory by a varying factor. The results are seen to split into two groups by the simple fact that the second derivatives with respect to T_i, T_e/T_i and $p(O^+)$ are of the same order of magnitude as the first derivatives, while the second derivatives with respect to N_e and ν_{in} are very small indeed; the derivatives are represented in Fig. 15. This means, as pointed out in Chapter 2, that the errors corresponding to these variables will be large. The multiparameter fit errors of the first group are in the following three tables:

Ion temperature variation accuracy unknown

N_e/N_0	T_i/T_0	T_e/T_i	ν/ν_0	$p(O^+)$	Var
1.000	1.000	1.000	0.000	0.300	*
0.035	0.052	0.072	–	–	–
0.065	0.058	0.125	–	–	0.391
0.113	0.747	0.267	–	0.852	–
0.164	1.920	0.451	–	2.266	0.638
0.288	0.166	0.561	1.698	–	–
0.414	0.240	0.811	2.874	–	0.509
0.294	0.805	0.564	2.031	1.020	–
1.216	4.722	2.673	6.372	5.024	1.130

For the temperature ratio, we have similarly the table

Temperature ratio variation accuracy unknown

N_e/N_0	T_i/T_0	T_e/T_i	ν/ν_0	$p(O^+)$	Var
1.000	1.000	1.000	0.000	0.300	*
0.035	0.052	0.072	–	–	–
0.104	0.142	0.330	–	–	0.800
0.113	0.747	0.267	–	0.852	–
0.317	1.383	0.891	–	1.382	1.019
0.288	0.166	0.561	1.698	–	–
0.804	0.604	1.829	3.632	–	1.170
0.294	0.805	0.564	2.031	1.020	–
1.147	2.036	2.726	4.217	1.604	1.467

and for the composition variations,

Ion composition variation accuracy unknown

N_e/N_0	T_i/T_0	T_e/T_i	ν/ν_0	$p(O^+)$	Var
1.000	1.000	1.000	0.000	0.300	*
0.035	0.052	0.072	–	–	–
0.037	0.054	0.074	–	–	0.654
0.113	0.747	0.267	–	0.852	–
0.220	1.696	0.551	–	1.911	0.979
0.288	0.166	0.561	1.698	–	–
0.455	0.281	0.897	2.790	–	0.838
0.294	0.805	0.564	2.031	1.020	–
0.756	2.326	1.610	3.465	2.374	1.279

The effects of the second group, the electron density variations and the collision frequency variations are in the following two tables:

The electron density variation accuracy unknown

N_e/N_0	T_i/T_0	T_e/T_i	ν/ν_0	$p(O^+)$	Var
1.000	1.000	1.000	0.000	0.300	*
0.035	0.052	0.072	–	–	–
0.405	0.137	0.235	–	–	32.707
0.113	0.747	0.267	–	0.852	–
0.425	0.975	0.286	–	1.012	35.638
0.288	0.166	0.561	1.698	–	–
0.442	0.244	0.565	1.773	–	33.425
0.294	0.805	0.564	2.031	1.020	–
0.507	1.022	0.574	2.032	1.160	35.639

and

Ion-neutral collision frequency variation accuracy unknown

N_e/N_0	T_i/T_0	T_e/T_i	ν/ν_0	$p(O^+)$	Var
1.000	1.000	1.000	0.000	0.300	*
0.035	0.052	0.072	–	–	–
0.035	0.059	0.072	–	–	2.283
0.113	0.747	0.267	–	0.852	–
0.250	1.828	0.622	–	2.025	3.518
0.288	0.166	0.561	1.698	–	–
0.517	0.334	1.023	3.103	–	3.086
0.294	0.805	0.564	2.031	1.020	–
1.828	5.203	3.916	7.574	4.943	6.794

It is seen that with these two last parameters, the errors in the multiparameter fits are not necessarily smaller than the errors for the other parameters. However, we know in most cases a priori that their variation uncertainty cannot be larger than one in our units, and from the formulae in the previous Chapter, the effect of these a priori bounds will make the errors of other parameters in practice equal to the errors when the parameter variation uncertainty was zero. In terms of the previous Chapter, these a priori widths are so small that the widths of the marginal distributions will be equal to the widths of the conditional distributions supposing that the variations Δa_i were exactly known. When we suppose *e.g.* that the variation of the electron density is known to an accuracy of 1 (in units of N_0), which is always realised in practice, we have the result that electron density variations do not matter at all.

Similarly, giving a generous bound to the accuracy that the ion-neutral collision frequency variation is known: when we know that the collision frequency variation is known to an accuracy 1 (in units of ν_0), we have

The collision frequency variation known to an accuracy ν_0

N_e/N_0	T_i/T_0	T_e/T_i	ν/ν_0	$p(O^+)$	Var
1.000	1.000	1.000	0.000	0.300	*
0.035	0.052	0.072	–	–	–
0.035	0.053	0.072	–	–	0.991
0.113	0.747	0.267	–	0.852	–
0.115	0.759	0.271	–	0.865	0.998
0.288	0.166	0.561	1.698	–	–
0.291	0.169	0.568	1.719	–	0.997
0.294	0.805	0.564	2.031	1.020	–
0.296	0.813	0.570	2.038	1.025	1.000

In the same vein, we can say that the collision frequency does not matter in practice.

The most important parameter is in practice the ion temperature T_i. When the ion temperature variation is known to 40% (meaning a 5% accuracy in the estimate of the corresponding second moment or covariance matrix element), we have the following table:

Ion temperature variation known to 40%

N_e/N_0	T_i/T_0	T_e/T_i	ν/ν_0	$p(O^+)$	Var
1.000	1.000	1.000	0.000	0.300	*
0.035	0.052	0.072	–	–	–
0.053	0.055	0.103	–	–	0.333
0.113	0.747	0.267	–	0.852	–
0.121	0.988	0.298	–	1.147	0.386
0.288	0.166	0.561	1.698	–	–
0.327	0.189	0.640	2.089	–	0.369
0.294	0.805	0.564	2.031	1.020	–
0.328	0.991	0.651	2.166	1.189	0.398

This leads to 15-20% increase in the composition fit errors, when $T_e/T_i = 1$. When we let the temperature ratio vary, the errors in the composition fits become large first when the uncertainty reaches 60%, which is seldom realised in practice, but possible. This means a 12% accuracy in the second moment estimates. We then have the table

Temperature ratio variation known to 60%

N_e/N_0	T_i/T_0	T_e/T_i	ν/ν_0	$p(O^+)$	Var
1.000	1.000	1.000	0.000	0.300	*
0.035	0.052	0.072	–	–	–
0.060	0.083	0.174	–	–	0.560
0.113	0.747	0.267	–	0.852	–
0.149	0.839	0.386	–	0.924	0.583
0.288	0.166	0.561	1.698	–	–
0.345	0.222	0.714	1.884	–	0.590
0.294	0.805	0.564	2.031	1.020	–
0.346	0.862	0.715	2.121	1.040	0.596

In the case of the 4–parameter fit, the error does become large, about 10%; the 5–parameter fit errors have increased only slightly.

A final remark is that if we now let the composition vary, the increase in the errors become noticeable first at impractically high uncertainties, of the order of 100%. This table is not shown here.

4.5. Discussion

We have learned from the results shown above that in those cases where the collision frequency is negligible, the only parameter whose variation one has to know to better that 50%, especially in the composition fits, is the ion temperature. It has to be stressed that this means only 5% in the second moments to be estimated somehow from the lag profiles. It may be suspected that there are situations when this estimate will turn out to be more inaccurate. It is also possible that the variations in the temperature ratio may in some cases cause trouble.

There has been no systematic effort to attack the problem of parameter mixing; *e.g.* Lathuillère *et al.* (1983a) only discussed the bias problem in fitting ion temperature and collision frequencies; however, the authors specifically excluded noise in their numerical simulations (which is just the opposite of the approach presented here), but nevertheless were able to note that the electron density profiles had little effect.

5. Posteriori distributions for semilinear theories

5.1. Introduction

In this Chapter we study a question which is very important in practice, namely that of using results of fits from subsets of some larger measuring volume V. We are not integrating the *data*, but rather (post)integrating or averaging the *results* over V. In order that the integration of these results be physically meaningful, we must be able to assume that the parameter variations are sufficiently small in some sense. What we are in fact doing is multiplying in a definite way a posteriori marginal distributions in subsets of V in order to obtain another marginal distribution describing the whole of V with better resolution. Most often this is done in order to make the a posteriori distribution of some single parameter within a larger volume such that the parameter gets resolved. How much integration in time or height is allowed is answered by the considerations from the previous Chapter. In the case of postintegration, it will turn out that the *residuals* of the fits play the key role.

5.2. Introductory considerations: semilinear theories

Let us suppose that the measurement is described by

$$m = f(x;y) + \epsilon, \tag{5.1}$$

where $x \mapsto f(x;y)$ is a linear function in x for each fixed y. In this case we can rewrite the measurement equation as a matrix equation

$$m = f(y)x + \epsilon, \qquad \langle \epsilon\epsilon^{\mathrm{T}} \rangle = \Sigma_m \tag{5.2}$$

where $f(y)$ is a matrix defining a linear mapping between the parameter space X and the space of measurements M.

The a posteriori distribution is then given (up to constants) by the identity

$$\begin{aligned} D_p(x;y) &= \exp\left(-\frac{1}{2}(f(y)x-m)^{\mathrm{T}}\Sigma_m^{-1}(f(y)x-m)\right) \\ &= |Q(y)|^{-1/2}\exp\left(-\frac{1}{2}(f(y)\overline{x}(y)-m)^{\mathrm{T}}\Sigma_m^{-1}(f(y)\overline{x}(y)-m)\right) \\ &\quad \cdot |Q(y)|^{1/2}\exp\left(-\frac{1}{2}(x-\overline{x}(y))^{\mathrm{T}}Q(y)(x-\overline{x}(y))\right) \\ &= D_p(y)D_p(x|y), \end{aligned} \tag{5.3}$$

where the inversion solution for the linear variable is as in Chapter 2,

$$\begin{aligned} Q(y) &= f(y)^{\mathrm{T}}\Sigma_m^{-1}f(y) \\ \overline{x}(y) &= Q(y)^{-1}f(y)^{\mathrm{T}}\Sigma_m^{-1}m. \end{aligned} \tag{5.4}$$

By the Bayes formula, the first factor of the solution is the marginal distribution of the a posteriori distribution for variable y, while the second factor is the conditional a posteriori distribution for x when y is supposed to be known. This solution corresponds to a practical analysis method that is often used: one fixes the nonlinear parameter and fits the remaining linear parameters. Then one changes the nonlinear parameter, until a best fit for the linear parameters is found. The quadratic form in the first term is proportional to the residual of the linear fit as a function of the nonlinear parameter. We have thus obtained a statistical interpretation of these kinds of methods: if the residual of the fit as a function of the nonlinear parameters is defined by

$$S(y) = (f(y)\overline{x}(y) - m)^{\mathrm{T}}\Sigma_m^{-1}(f(y)\overline{x}(y) - m), \tag{5.5}$$

the marginal distribution for the nonlinear parameters of the a posteriori distribution is given by

$$D_p(y) = |Q(y)|^{-1/2}\exp\left(-\frac{1}{2}S(y)\right). \tag{5.6}$$

Moreover, we see that the total a posteriori distribution is Gaussian in the directions of the linear parameter x and that it is possible to solve numerically its centre point $\overline{x}(y)$ and Fisher information matrix $Q(y)$ as a function of the nonlinear parameter y. This is represented in Fig. 16.

5.3. Partially linearisable case

It is not necessary that some part of the theory be exactly linear for the above discussion to be useful. It may be that while the total a posteriori distribution is so wide or complicated that it is impossible to base its study on a linearisation of the nonlinear theory around the best fit point, its conditional distributions with some parameter(s) fixed may well be narrow enough so that they may be approximated by Gaussians. This is due to the fact that the theory function itself is well represented by the first term in its Taylor expansion in the region where most of the conditional posteriori probability mass is situated. In this case we arrive at a similar situation as above.

Let us suppose again that

$$m = f(x;y) + \epsilon, \qquad \langle\epsilon\epsilon^{\mathrm{T}}\rangle = \Sigma_m, \tag{5.7}$$

with ϵ a zero-mean Gaussian random variable. Here a typical situation could be that x is a set of ordinary ionospheric variables, for example $x = (N_e, T_i, T_e/T_i)$ and y is a difficult parameter, like the ion composition. In this case it is seen from the tables of estimation errors in Chapter 2 that *the conditional a posteriori distributions for the ordinary parameters with composition supposed to be known are an order of magnitude narrower than their marginal a posteriori distributions.* If it is possible to analyse composition (with the collision frequency supposed to be zero) with error bars of the order of 10%, then it follows from Table 1 in Chapter 2 by scaling with the ratio 0.1/0.852 that the *conditional* error bars (supposing that the composition is known exactly) in the 3–parameter fit for N_e, T_i and T_e/T_i will be about 0.41%, 0.61% and 0.85%, respectively. Thus the marginal a posteriori distributions for these parameters would be very narrow indeed. This means also that the measurement errors would have to be very small; it would correspond to an accuracy of 0.16% in the power profiles. This kind of statistical accuracy means under normal circumstances such long integration periods (about one hour) and integrations over tens of kilometers that the question of the stability of the ionosphere or of the parameter mixing errors discussed in the previous Chapter becomes important.

Let the a posteriori distribution be again given by

$$\begin{aligned} D_p(x;y) &= \exp\left(-\frac{1}{2}(f(x;y)-m)^{\mathrm{T}}\Sigma_m^{-1}(f(x;y)-m)\right) \\ &= \exp\left(-\frac{1}{2}S(x;y)\right), \end{aligned} \tag{5.8}$$

and let us define the residual of the fit, $\overline{S}(y)$, by

$$\overline{S}(y) = \min_{x\in X} S(x;y) \tag{5.9}$$

for each fixed y and denote the point in X at which the minimum is attained by $\overline{x}(y)$:

$$\overline{S}(y) = S(\overline{x}(y);y). \tag{5.10}$$

We must now suppose here that there is a unique minimum point $\overline{x}(y)$ for each y or at least that the possible secondary minima lead to posteriori probabilities that are so small that they can be neglected. If $\Sigma_m \to 0$, this will always happen, if the minima are not exactly equal. The conditional posteriori distribution $D_p(x;y)$ can then be approximated by using the first three terms of the Taylor series expansion of $S(x;y)$ with respect to x at the point $\overline{x}(y)$:

$$S(x;y)\approx S(\overline{x}(y);y) + \frac{1}{2}\sum_{i,j}\frac{\partial^2 S(x;y)}{\partial x_i \partial x_j}(x_i-\overline{x}_i(y))(x_j-\overline{x}_j(y)), \tag{5.11}$$

where the derivatives have been evaluated at $x = \overline{x}(y)$ and we have used the fact that the gradient term vanishes at the minimum point. The second derivatives are given by

$$\frac{\partial^2 S(x;y)}{\partial x_i \partial x_j} = 2\frac{\partial f}{\partial x_i}^{\mathrm{T}}\Sigma_m^{-1}\frac{\partial f}{\partial x_j} + 2(f-m)^{\mathrm{T}}\Sigma_m^{-1}\frac{\partial^2 f}{\partial x_i \partial x_j}. \tag{5.12}$$

Since $f(\overline{x}(y), y) - m$ may often be considered small (it is the residual vector of the partial fit), we may approximate $S(x; y)$ by

$$S(x; y) \approx S(\overline{x}(y); y) + (x - \overline{x}(y))^{\mathrm{T}} Q(y)(x - \overline{x}(y)), \tag{5.13}$$

where we have defined the conditional Fisher information matrix $Q_{ij}(y)$ by

$$Q_{ij}(y) = \frac{\partial f}{\partial x_i}^{\mathrm{T}} \Sigma_m^{-1} \frac{\partial f}{\partial x_j}. \tag{5.14}$$

It now follows with these definitions for $Q(y)$ and $\overline{x}(y)$ that simply

$$\begin{aligned} D_p(x; y) &= \exp\left(-\frac{1}{2}(f(x; y) - m)^{\mathrm{T}} \Sigma_m^{-1} (f(x; y) - m)\right) \\ &= |Q(y)|^{-1/2} \exp\left(-\frac{1}{2}(f(\overline{x}(y), y) - m)^{\mathrm{T}} \Sigma_m^{-1} (f(\overline{x}(y), y) - m)\right) \\ &\cdot |Q(y)|^{1/2} \exp\left(-\frac{1}{2}(x - \overline{x}(y))^{\mathrm{T}} Q(y)(x - \overline{x}(y))\right), \end{aligned} \tag{5.15}$$

formally similarly to the semilinear case.

5.4. Postintegration of calculated a posteriori distributions

In this Chapter we consider a situation where we have a posteriori distributions $D_p^i(x)$ available from a number of independent measurements. These a posteriori distributions may be analysis results of several consecutive heights or from many consecutive integration periods. If the a priori distributions can be taken to be constant and if it is supposed that the parameter values for the different measurements are the same, it is clear that the combined a posteriori distribution for the independent measurements is just the product of the a posteriori distributions of these readily calculated measurements.

A more complicated situation arises, if we suppose that the correct parameter values depend on the index i in some special way. Let us suppose that the values of the parameters are given by x^i and y^i, where the first ones are the ordinary or linear parameters and the second ones are the difficult or non-linear parameters. Then the measurements are described by

$$m^i = f(x^i; y^i) + \epsilon^i \tag{5.16}$$

with ϵ^i independent from ϵ^j, if $i \neq j$, and ϵ^i is a Gaussian zero-mean random variable with covariance matrix Σ_m^i.

The a posteriori distribution is then given by the product (supposing constant a prioris)

$$\begin{aligned} & D_p(x^i; y^i) \\ & = \prod_i D_p^i(x^i; y^i) \\ & = \prod_i D_p^i(y^i) D_p^i(x^i|y^i) \\ & = \prod_i \exp\left(-\frac{1}{2}(f(x^i; y^i) - m^i)^{\mathrm{T}} (\Sigma_m^i)^{-1} (f(x^i; y^i) - m^i)\right), \end{aligned} \tag{5.17}$$

that is, by using the conditional information matrices $Q(y)$ defined above,

$$\begin{aligned} & D_p(x^i; y^i) \\ & = \prod_i |Q^i(y^i)|^{-1/2} \exp\left(-\frac{1}{2}(f(\overline{x}^i(y^i), y^i) - m^i)^{\mathrm{T}} (\Sigma_m^i)^{-1} (f(\overline{x}^i(y^i), y^i) - m^i)\right) \\ & \cdot \prod_i |Q^i(y^i)|^{1/2} \exp\left(-\frac{1}{2}(x^i - \overline{x}^i(y^i))^{\mathrm{T}} Q^i(y^i)(x^i - \overline{x}^i(y^i))\right), \\ & = \prod_i \exp\left(-\frac{1}{2}\overline{S}^i(y^i)\right) \exp\left(-\frac{1}{2}(x^i - \overline{x}^i(y^i))^{\mathrm{T}} Q^i(y^i)(x^i - \overline{x}^i(y^i))\right). \end{aligned} \tag{5.18}$$

The a posteriori distribution is thus decomposed into products of residuals of the fits with y^i kept fixed, and it is seen to be Gaussian in the linear parameters x^i. Here the only assumption is that the points $\overline{x}^i(y^i)$ are the points where the a posteriori distributions for fixed y^i attain their minima, so that the conditional distributions may be approximated by Gaussians. This is represented schematically in Fig. 17. Let us suppose next that the indexed parameters x^i are related to some basic parameters x by

$$x^i = A^i x \quad \text{an} \quad y^i = y. \tag{5.19}$$

For example, if i represents the radar range, it is natural to introduce a linear height gradient to consecutive range gates by specifying $x^i = A^i x = x + i\Delta x$, the set of new unknowns being $(x, \Delta x)$. Let us in addition suppose that we are interested in the marginal distribution for the difficult parameter y. This is given by the integral

$$\begin{aligned} D_p(y) &= \int_X \prod_i D_p^i(x^i; y)\, dx \\ &= \prod_i \exp\left(-\frac{1}{2}\overline{S}^i(y^i)\right) \int_X \prod_i \exp\left(-\frac{1}{2}(A^i x - \overline{x}^i(y))^{\mathrm{T}} Q^i(y)(A^i x - \overline{x}^i(y))\right). \end{aligned} \tag{5.20}$$

Here, the last product is the same expression as in the linear problem described by

$$m_i = A^i x + \epsilon_i, \quad < \epsilon_i \epsilon_i^{\mathrm{T}} >= Q_i^{-1}(y), \quad m_i = \overline{x}^i(y). \tag{5.21}$$

As mentioned in Chapter 2, this inverse problem has the solution given by the information matrix

$$Q_V(y) = \sum_i A^{i\mathrm{T}} Q^i(y) A^i \tag{5.22}$$

and the centre point $\overline{x}(y)$ given by

$$\overline{x}(y) = Q_V^{-1} \sum_i A^{i\mathrm{T}} Q^i(y) \overline{x}^i(y). \tag{5.23}$$

The information matrices $Q^i(y)$ of the linear problem (5.21) are outputs of the fits of $\overline{x}^i(y^i)$ for fixed y. The height gradient matrices A^i are obtained by assuming or knowing some functional form as a function of height in (5.19), and $\overline{x}(y)$ can be calculated. The value of the integral in (5.20) may be calculated by noting that by virtue of (5.21)–(5.23), the product of the exponentials may be written as the exponential of the sum

$$S(x;y) = (x - \overline{x}(y))^{\mathrm{T}} Q_V (x - \overline{x}(y)) + S(\overline{x}(y);y) \tag{5.24}$$

as in Chapter 2. The value of the Gaussian integral is then (up to constants) equal to the square root of determinant of the coefficient matrix Q_V, and the factor $S(\overline{x}(y);y))$ is by definition the residual of the gradient fit. Substituting $x = \overline{x}(y)$ into the integrand in (5.20) we have the final result

$$\begin{aligned} D_p(y) = \left| \sum_i A^{i\mathrm{T}} Q^i(y) A^i \right|^{1/2} \prod_i \exp\left(-\frac{1}{2} \overline{S}^i(y^i) \right) \\ \cdot \prod_i \exp\left(-\frac{1}{2} \left(\overline{x}(y) - \overline{x}^i(y) \right)^{\mathrm{T}} Q^i(y) \left(\overline{x}(y) - \overline{x}^i(y) \right) \right). \end{aligned} \tag{5.25}$$

In practice, one therefore has to calculate both $\overline{x}(y)$ as the sum given above and the corresponding residuals for all intermediate fits for different values of y. The information matrix Q_V may depend on y, and therefore, its determinant has to be calculated, too.

5.5. Discussion

We have derived the postintegrated a posteriori distribution of the difficult parameter y in a larger volume V by combining several previously calculated a posteriori distributions in subsets of V for all parameters. We have seen that the problem can be reduced to several simpler fits and that the residuals of these fits play a key role in the final solution. Being matrix operations, these fits are relatively easy to bring to a practical form, once we have the results of the intermediate fits *with the information or covariance matrices* available.

6. Determination of ion composition from alternating code data

6.1. Introduction

In this Chapter we describe the design, realisation and analysis of the first alternating code experiments ever performed. Alternating code is a new method first described in Lehtinen (1986) and Lehtinen and Häggström (1987). Alternating codes consist of series of long coded pulses repeated over and over again. The collected lag profiles can be specially decoded to produce ACF estimates with ambiguity functions similar to those obtained from ordinary double-pulse codes with all possible pulse separations. After this decoding, the data can be analysed by standard analysis programs. Alternating codes can be used instead of multipulse experiments, and they lead to much better ACF accuracy than any other methods with corresponding resolutions. In fact, it was proven in Lehtinen (1986) that the accuracy is close to the theoretical optimum in the case where the signal-to-noise ratio (SNR) is small. It will be seen that the analysis of first-phase results shows that the composition can indeed be determined with a good resolution.

6.2. The design requirements of the experiment

The primary goal of the analysis was the determination of the ion composition $p(O^+)$ in a population consisting of O^+ (mass 16 u) and a mixture (mass 30.5 u) of the heavier molecular ions NO^+ (75%) and O_2^+ (25%) with the best possible height and time resolution.

As the composition is a 'difficult' parameter in the sense of Chapter 5, we chose to do the analysis in the way outlined there; the O^+ content was fixed at different values and the other 3 standard parameters ($N_e, T_i, T_e/T_i$) were then fitted to the data together with the ion velocity; the collision frequency was supposed to be zero. The analysis outputs were the best fit values of the other parameters, their error estimates and the fit residuals given as a function of composition. As it may be shown that the total a posteriori distribution is a product of the ion velocity a posteriori distribution and that of the rest of the variables, we can use the results of Chapter 2 as if the velocity was known to be zero. The true number of fitted parameters is nevertheless 5.

The fits were performed on data postintegrated to roughly 20 minute dumps. The basic height resolution was about 4 km, and the separation between the gates was 1.5 km. The final composition results were then derived using the results of

Chapter 5 to combine calculated a posteriori densities of ion composition for further height and time integration in order to achieve better resolution for the composition.

This work is a co-operation of Ingemar Häggström, Tauno Turunen, Markku Lehtinen and Matti Vallinkoski. The basic specification of the sampling rates, pulse lengths, code lengths and other technical data was based on the methods described in Chapter 2. The practical experiment design was limited by the existing EISCAT hardware, which has certainly not been designed with alternating codes in mind. A possible practical solution fulfilling the calculated lag resolution and lag extent requirements was found by Tauno Turunen and Markku Lehtinen. The actual experiment programming for the EISCAT radar was done by Tauno Turunen and Ingemar Häggström; Ingemar Häggström performed the decoding and the first phase of the analysis. The second phase of the analysis combining the calculated a posteriori densities was then performed by Matti Vallinkoski and Markku Lehtinen. The results of this second phase will be presented in the following Sections.

6.3. Design of the experiment

The design goals of the experiment were the following:

1) The lowest gates taken into account should be in the region where one can safely neglect the collision frequency.

2) The positions of the first zeros of the ACFs can vary between about 100μs and about 20μs, equivalent to ion temperatures anticipated to lie between 300 K for purely molecular ions and about 3000 K for pure O^+ with temperature ratios in the range 1...2.

3) The experiment should have a resolution of the same order as ordinary pulse codes or even better, so that the parameter mixing errors will not become too serious inside the basic scattering volume.

The collision frequency can be set to zero above 120 km, which is therefore about the lowest altitude we can use. The maximum lag extent must be smaller than the total code length, which is limited by the fact that the transmission must be well off before the front end of the scattered signal reaches the radar, about 400–500μs. The minimum required lag extent may be obtained by repeating the calculations of Chapter 2 for a 4-parameter fit using as the basic temperature $T_0 = 300$ K. There it was shown that the minimum possible temperature T_{min} for which the errors are still in the stable region is determined by the lag extent, *i.e.*, the behaviour of the ACF derivatives for large lags. If we want $T_{min} = 300$ K for purely molecular ions, $p(O^+) = 0\%$, the lag extent requirement is easily satisfied for the worst case $T_e/T_i = 1$. If we add some admixture of O^+, the shoulder due to the lighter O^+ moves to higher lag extent values. Fig. 18 (upper panel) shows the behaviour of the errors as a function of the lag extent for $p(O^+) = 30\%$, which is

a plausible estimate for the maximum amount of O^+ for this temperature. We see that the errors are reasonably stable down to a lag extent of about $\Delta\tau \approx 3\,\tau_0$ where $\tau_0 \approx 100\mu$s, which gives the minimum temperature $T_{min} \leq T_0$ for all lag extents $\Delta\tau \geq 3\,\tau_0$. On the other hand, the high temperature end means 6000 K for purely molecular ions. The lowest value for T_{max} is obtained for $T_e/T_i = 2$ (Fig. 18, lower panel). As shown in Chapter 2, the highest possible temperature for which the fit errors are still in the stable region is determined by the lag resolution, *i.e.*, the behaviour of the ACF derivatives for small lags. The lag resolution requirement is seen to be reasonably well satisfied up to about $d\tau \approx 0.5\tau_0$, which gives the desired maximum lag resolution $d\tau_{max} = 0.5\,\tau_0\sqrt{T_0/T_{max}} \approx 11\mu$s. The values satisfying the above requirements and used in the experiment were $d\tau = 10\mu$s and $\Delta\tau = 390\mu$s (denoted by the arrows in Fig. 18). The lowest gate was chosen to be about 125 km. Thus requirements 1) and 2) are satisfied.

The simplest solution for requirement 3) would then be to have an alternating code consisting of 10μs pulses, which is a typical number in ordinary pulse codes (the smallest possible pulse length in the EISCAT UHF radar is 2μs). As it was not possible to use 32-bit alternating codes with 10μs pulses with the present EISCAT hardware, we solved the problem by using a 16-bit alternating code with 30μs pulses but with 10μs sampling interval, *i.e.*, the lag increment is one third of the bit length 30μs. We may thus define *a fractional lag* to be a lag whose length is a fraction of the bit length, and *a main lag* to be a lag whose length is an integer multiple of the bit length. The main lags are therefore in this case multiples of 3 (lengths multiples of 30μs), and all other lags fractional lags, like lags $1, 2, 4, 5, 7, 8, \ldots$ (lengths $10, 20, 40, 50, 70, 80, \ldots\,\mu$s). The total code length thus became 480μs, but the ACFs were sampled only up to 390μs; the lag statistics for 390μs also became better than in the more straightforward solution, as the number of independent lag estimates from an alternating code goes to zero as the lag extent approaches the total pulse length, and we now get four independent estimates for the last lag instead of just one. The use of fractional lags and a longer pulse also results in larger scattering volumes giving better statistics, but the possibility of using fractional lags with their narrower range ambiguities still makes high resolution possible in cases where one can otherwise loosen the requirements of accuracy. Thus, this experiment is a 'multiple-resolution' experiment.

In Fig. 19 we have calculated the resulting errors of the 4–parameter fits for the actual lag distribution used in the experiments. The calculation is documented in the Appendix. In reality, the number of zero lag estimates is two, but with the variance being twice that of the other lag estimates and with the higher statistical weights for the other lags $(45, 30, 15, 14, 14, 14, 13, \ldots)$ we get a good upper limit to the error estimates by simply neglecting the zero lag. It is seen from the curves that the temperature limits estimated above conform well with this more exact calculation. The calculation was done with $p(O^+) = 0\%$ and for the two extreme values of the temperature ratio, and it was assumed that the zero lag data was not used at all. The large errors in the low temperature end are due to the smaller statistics at large lags; in the last seven 30μs pulses, only the last 10μs were sampled.

In Fig. 20 we have calculated in a similar way the resulting parameter fit

errors as a function of composition. Also this calculation is documented in the Appendix. We see that around $p(O^+)$=35% the errors become very large. This fact may be seen in the a posteriori plots for actual data. Interestingly enough, a similar curve for fixed electron density has been presented in Lejeune (1980) in an almost unavailable report on the basis of a crude model just consisting of rectangles for the incoherent scatter spectrum.

6.4. Experiments

The first short experiment was run at the EISCAT incoherent scatter facility on 30 April 1987. It lasted for three hours, but only three 20 minute intervals proved to be free of technical errors due to exceptionally high data rates and correlator dump rates. After that, the experiment was repeated on 4-5 October 1987 using Swedish special programme time starting at 07.00 UT on 4 October 1987, and on 4-5 January 1988 using Finnish special programme time starting at 15.00 UT on 4 January 1988. These experiments succeeded well and produced over 20 hours of alternating code data each. The geomagnetic conditions for the experiments were rather different: the April run was the most quiet with a K_p index value of 1−, the October run more disturbed with an average K_p index value of 2+, and the January run was done during a weak polar cap absorption event with an average K_p index value of 3+.

Due to the limitations of the present EISCAT correlator (it is impossible to enter codes or make subtractions) each of the 32 codes was transmitted in a separate radar controller program which was changed and recorded to tape every 2 seconds so that the whole cycle was completed in 64 seconds. With a sampling interval of 10μs we obtained 19 lag profiles with 10μs lag intervals from 0 to 180μs, and in addition to that 7 lag profiles with a lag interval of 30μs from 210μs to 390μs. Using four frequencies it was possible to add two 30μs single pulses to be used as power profiles and zero lag estimation.

In all, from the lag profiles it was possible to form 114 ACFs ranging from 125 km up to 300 km altitude, each with a height resolution of about 4 km and a range increment of 1.5 km. The transmitting antenna was pointing along the local geomagnetic field line.

An example of a measured ACF with the corresponding measured and fitted theoretical spectrum is presented in Fig. 21. The integration time is 256 s and the height 206 km. The first minimum of the ACF is due to the fact that the scattering volumes of lag 1 and lag 2 are much smaller than those of the main lags (with ratios of $\approx$ 0.40 and 0.75) and that the scattering volumes of the other fractional lags are slightly larger (ratios $\approx$ 1.15).

6.5. Decoding of the data

The easiest way to understand how the decoding is performed is to start from the lagged profile or UNIPROG matrix (Ho *et al.*, 1983). The code pattern is set on the axes and the matrix is formed by cross-multiplying the pattern. Fig. 22a shows an example for one of the patterns used in the 4-bit code. If the data is stored in the same way, the decoding is done by putting the code matrix on top of the data matrix, and the decoding elements give the signs to the individual lagged products in the summation. The next gate is then decoded by sliding the code matrix one step along the diagonal of the data matrix, so that the code matrix works as a filter through which the lag profiles run (Fig. 22b). For fractional lags, we have to leave gaps in the code patterns, the length of the gaps being equal to the number of fractional lags between each main lag. Fig. 22c shows an example with the 4-bit code where the lag increment is one third of the bit length. Each one of the filter coefficients is in that case spread out in the usual way as for a double pulse code like a butterfly with the main lag coefficients in the middle. For the main lags, we know that the ambiguity functions are the same as for a single double-pulse code with the appropriate separation; in order to calculate some of the fractional lag coefficients, we have to use the ambiguity functions of the two different two-pulse codes corresponding to the two closest main lags. These can then be added or either one of them skipped according to the size and shape of the scattering volumes and the resulting statistical errors, which we want to minimise. The shape of the scattering volume of the fractional lags is different from that of the main ones, but their range coverage is the same.

In our case, we chose to add the overlapping contributions for the fractional lags, except for lag 1 and lag 2, where there is no overlap (the only contributing two-pulse code is the one with 30μs separation). The reason for doing this is that then the number of estimates for each of the scattering volume sizes is the same and this makes the analysis procedure easier.

6.6. First phase analysis

The basic analysis was performed by calculating theoretical ACFs with filter and waveform corrections and then fitting these values to the measured and decoded ACF values. In the fits the variances of the measured and decoded values were estimated in order to be able to give correct weights to the lags. For alternating codes all background noise and correlated clutter is automatically removed by the decoding process (their averages go to zero), but they still contribute to the variance of the ACF estimates. Zero lag data was only available from the two single pulses, since the zero lag deconvolution (see Lehtinen and Huuskonen, 1986) is not included in the standard analysis packages.

In our case, we decoded all ACFs before starting the analysis, and for the variance calculation we added the correlated clutter and the background noise to the decoded ACFs. The clutter was estimated from the neighbouring gates with a longer range for short lags.

The waveform correction was done in the time domain by the commonly used so-called 'pulse ACF' method. The pulse ACF for alternating codes is somewhat more complicated than for multipulse schemes and is shown in Fig. 23 for the 4-bit case. Each main lag has its own pulse ACF with the same form as if two pulses, with the same length as the bit length, were transmitted with a separation of that lag, except that the contribution from the correlation inside the pulses themselves is missing. This means that the standard way of correcting the theoretical ACFs (multiply with the pulse ACF, transform to the frequency domain, multiply with the receiver filter form and transform back to time domain) becomes very time consuming, as this has to be done for all the main lags separately. Since this whole operation is a linear mapping from spectra to measurement space, it could be done far more effectively by calculating first the matrix of the mapping, which depends only on the experiment, and then just a matrix product. This will be implemented in the analysis package currently under development.

In the integration process, the satellite echos in the background were removed. If a satellite was detected in the signal part of the data, the whole 2–second dump was skipped, as the alternating code technique would fail for too many gates in such a case and make the ion composition determination impossible. On the average about 4 satellite echos per hour were detected. The data was integrated to 20 minute dumps and in each gate the electron density, electron and ion temperatures and ion drift velocity were fitted for a fixed ion composition varying from 0 to 100% O^+ in steps of 5%.

6.7. Results

We have applied the methods developed in Chapter 5 to the first-phase analysis results described above. As the present first-phase analysis package does not output the whole estimated error covariance matrix of the fit results, but only the diagonal elements, we have reconstructed the off-diagonal elements by using the correlations calculated in Chapter 2. This may cause some error in the results and can only be corrected by generalising the standard analysis package used. However, we believe that these errors are so small that they do not change the overall behaviour of the results. We also have to assume that the information matrix Q_V of the whole integration volume V from (4.25) is constant.

In Fig. 24 we show the results of the second phase of the analysis of the ion composition for the three 20 minute dumps measured on 30 April 1987. It was supposed in all calculations that the dependence of the gradients on height was linear; assuming independence of height produced similar results. In panel a),

we have performed a height integration over 5 consecutive gates and in panel b), the height integration is over 10 gates. The time integration is over all available three dumps, that is, one hour. The time integration was necessary to make the distributions sharper. The basis for the time integration as well as for the range integrations is that the ionosphere is stable during this period, as could be seen in the colour panels showing the analysis results of the 256-second data. The a posteriori distributions in panel a) are rather broad, but the additional height integration performed in b) has made the distributions much sharper. We may say that in this case the integration is justified by the results. In panel b), the half-widths vary from $\pm 2.5\%$ to $\pm 6\%$, except for $p(O^+) > 25\%$, where they are larger (but still less than 10%) in accordance with the curves in Fig. 21.

In Fig. 25 results from the experiment on 4 October 1987 are shown. The a posteriori distributions are integrated over 6 dumps and 10 consecutive range gates each shown in the four panels. The electron densities are small in this experiment, especially after 16 UT, leading to a worse SNR and thus to broader a posteriori densities than in the previous experiment. As is seen from the colour panels (not reproduced here), the ionosphere remained reasonably constant during the periods shown in this experiment, so that the integration times during these periods may be justified by the data. The longest period is in panel d), and it is seen that there the electron density was very low, and the resulting a posteriori densities very broad.

In Fig. 26 results from the experiment run on 4–5 January 1988 are shown, integrated over 6 dumps and over 10 range gates each. The ionosphere may not be termed stable during this experiment, and this fact is apparent in the two panels shown. Some stability is seen in the higher gates, where also the electron density and thus the SNR are significantly higher.

6.8. Discussion

The colour panels (not reproduced here) with 256 second integration time and 1.5 km gate separation demonstrate clearly how powerful the alternating code method is in the the ordinary 3-parameter plus ion velocity fits. The difficulty of determining also the ion composition accurately is clearly seen in the previous examples, although the alternating codes are about the best measurements one can do. The difficulty is known from earlier experiments, *e.g.* Lathuillère *et al.* (1983b), where in a long-pulse (360μs) experiment the altitude resolution was 54 km and integration times one hour. The accuracy of their final composition fit results was not discussed. Our range resolution resulting from the integration of ten consecutive gates is about 20 km, and the integration time the same, and we know the accuracy of our results. However, the alternating code measurements described here were not made to optimise the height resolution, but to show the viability of alternating codes. If one wanted to aim directly at a final height resolution of about 15 km, then this would have been possible *e.g.* with 100μs pulses, a 4-bit alternating code, resulting in an improvement of $(100/30)^2$ in the measurement speed as defined in

Lehtinen (1986). The time integration eats up one power of this speed, so that the final improvement in integration time would be 10/3. This means then that the actual numbers to compare here would be 20 km vs. 54 km and 18 minutes vs. 60 minutes, and still the same very small errors vs. unknown estimates. Therefore, we believe that we have demonstrated that it is indeed possible to determine the composition to a high degree of accuracy, less than 10%, and if necessary, with a good height and time resolution with the present hardware and with a standard analysis package. In the analysis package currently under development, where also whole covariance matrices are output and the filter corrections are done correctly, we believe we shall even be able to improve the results shown above considerably.

References

Bayes, T. (Reverend): An essay towards solving a problem in the doctrine of chances. Phil. Trans. R. Soc. A, **53**, 370–418, 1763. Republished in Biometrika, **45**, 298–315, 1958

Beynon, W.J.G. and Williams, P.J.S.: Incoherent scatter of radio waves from the ionosphere. Rep. Prog. Phys., **41**, 909–956, 1978

Bhatnagar, P.L., Gross, E.P., Krook, M.: A model for collision processes in gases, 1, Small amplitude processes in charged and neutral one-component systems. Phys. Rev. **94**, 511–525, 1954

Bowles, K.L.: Observations of vertical incidence scatter from the ionosphere at 41 Mc/s. Phys. Rev. Lett., **1**, 454–455, 1958

Bowles, K.L.: Incoherent scattering by free electrons as a technique for studying the ionosphere and exosphere: Some observations and theoretical considerations. J. Research NBS, **65D**, 1–13, 1961

Box, G.E.P., and Tiao, G.C.: Bayesian inference in statistical analysis. Addison-Wesley, Massachusetts, U.S.A., 1973

Buneman, O.: Scattering of radiation by the fluctuations in a nonequilibrium plasma. J. Geophys. Res. **67**, 2050–2053, 1962

Callen, H.B. and Welton, T.A.: Irreversibility and generalized noise. Phys. Rev. **83**, 34–40, 1951

Dougherty, J.P. and Farley, D.T.: A theory of incoherent scattering of radio waves by a plasma. Proc. Roy. Soc. A, **259**, 79–99, 1960

Dougherty, J.P. and Farley, D.T.: A theory of incoherent scattering of radio waves by a plasma, 3. Scattering in a partly ionized gas. J. Geophys. Res. **68**, 5473–5486, 1963

Dougherty, J.P. and Farley, D.T.: A theory of incoherent scattering of radio waves by a plasma, 3. Scattering in a partly ionized gas. J. Geophys. Res. **68**, 5473–5486, 1963

Evans, J.V.: Theory and practice of ionosphere study by Thomson scatter radar. Proceedings of the IEEE, **57**, 496–530, 1969

Farley, D.T., Dougherty, J.P. and Baron, D.W.: A theory of incoherent scattering of radio waves by a plasma, II. Scattering in a magnetic field. Proc. Roy. Soc. A, **263**, 238–258, 1961

Farley, D.T.: The effect of Coulomb collisions on incoherent scattering of radio

waves by a plasma. J. Geophys. Res. **69**, 197–200, 1964; J. Geophys. Res. **69**, 2402, 1964

Farley, D.T.: A theory of incoherent scattering of radio waves by a plasma, 4, The effect of unequal ion and electron temperatures. J. Geophys. Res. **71**, 4091–4098, 1966

Fejer, J.A.: Scattering of radio waves by an ionized gas in thermal equlibrium. Can. J. Phys. **38**, 1114–1133, 1960

Fisher, R.A.: On the mathematical foundations of theoretical statistics. Phil. Trans. Roy. Soc., Ser. A. **222**, 309–, 1922

Fried, B.D., and Conte, S.D.: The plasma dispersion function. Academic Press, New York and London, 1961

Golub, G.H. and Van Loan, C. F.: Matrix Computations, Section 12. Johns Hopkins University Press, Baltimore, U.S.A., 1983

Gordon, W.E.: Incoherent scattering of radio waves by free electrons with applications to space exploration by radar. Proc. IRE, **46**, 1824–1829, November 1958

Grassmann, V.: Inkohärente Streuung von Radiowellen durch ein Plasma unter Berücksichtigung von Teilchen–Stößen. MPAE-W-100-88-07, Max Planck Institut für Aeronomie, Katlenburg-Lindau, BRD, 1988

Hagen, J.B. and Behnke, R.A.: Detection of the electron component of the spectrum in incoherent scatter of radio waves by the ionosphere. J. Geophys. Res. **81**, 3441-3443, 1976

Hagfors, T.: Density fluctuations in a plasma in magnetic field. J. Geophys. Res. **66**, 1699–1712, 1961

Ho, T., Turunen, T., Silén, J. and Lehtinen, M.: The lag profile routine and the universal program for the EISCAT digital correlators. EISCAT Technical Note 83/37, 1983

Huuskonen, A.: Private communication. 1987

Huuskonen, A., Pollari, P., Nygrén, T. and Lehtinen, M.: Range ambiguity effects in Barker-coded multipulse experiments with incoherent scatter radar. J. Atmos. Terr. Phys. **50**, 265–276, 1988

Häggström, I., Turunen, T., Lehtinen, M., Vallinkoski, M.: Incoherent scatter radar measurements of ion composition measurement using alternating codes. (submitted to Radio Sci., 1989)

Lathuillère, Ch., Lejeune, G. and W. Kofman: Direct measurements of ion composition with EISCAT in the high-latitude F_1–region. Radio Science, **18**, 887-893, 1983a

Lathuillère, Ch., Wickwar, V.B., Kofman, W.: Incoherent scatter measurements of ion–neutral collision frequencies and temperatures in the lower thermosphere of the auroral region. J. Geophys. Res. **88**, 10137–10144, 1983b

Lehtinen, M.S.: Statistical theory of incoherent scatter measurements. EISCAT Technical Note 86/45. ISBN 951-99743-2-6. (Thesis, University of Helsinki 1986)

Lehtinen, M. and Huuskonen, A.: The use of multipulse zero lag data to improve incoherent scatter radar power profile measurements. J. Atmos. Terr. Phys. **48**, 787–793, 1986

Lehtinen, M. S. and Häggström, I.: A new modulation principle for incoherent scatter measurements. Radio Sci., **22**, 625–634, 1987

Lehtinen, M.: On statistical inversion theory. In: Theory and applications of inverse problems, H. Haario (ed.), Longman Scientific & Technical, Longman Group UK Ltd., 1988

Lejeune, G.: Détermination des incertitudes sur les paramètres mesurés par diffusion incohérente. In: Le système EISCAT et l'étude du couplage magnetosphère–ionosphère. D. Alcaydé, ed., pp. 11–22. Centre d'Etudes Spatiales des Rayonnements, Toulouse, France, 1980

Menke, W.: Geophysical data analysis: discrete inverse theory. Academic Press, London, U.K, 1984

Nyquist, H.: Thermal agitation of electric charge in conductors. Phys. Rev. **32**, 110–113, 1928

Raman, R.S.V., St-Maurice, J.P., Ong, R.S.B.: Incoherent scattering of Radar waves in the auroral ionosphere. J. Geophys. Res. **86**, 4751–4762, 1981

Rao, C.R.: Linear statistical inference and its applications. John Wiley and Sons, New York, U.S.A., 1960

Rishbeth, H. and Williams, P.J.S.: The EISCAT ionospheric radar: the system and its early results. Q. Jl. astr. Soc. **26**, 478–512, 1985

Schlegel, K.: A program library for incoherent scatter calculations (version B). MPAE-W-05-79-13, Max Planck Institut für Aeronomie, Katlenburg-Lindau, BRD, 1979

Suvanto, K.: Auroral F-region ion velocity distributions in the presence of large flows and electrostatic waves. Planet. Space Sci. **35**, 1429–1435, 1987

Swartz, W.E. and Farley, D.T.: A theory of incoherent scattering of radio waves by a plasma, 5, The use of the Nyquist theorem in general quasi-equilibrium situations. J. Geophys. Res. **84**, 1930–1932, 1979

Sitenko, A.G.: Electromagnetic fluctuations in plasmas. Academic Press, New York, U.S.A., 1967

Tarantola, A.: Inverse Problem Theory. Elsevier, Netherlands, 1987

Tarantola, A. and Valette, B.: Inverse problems= quest for information. J. Geophys. **50**, 159–170, 1982a

Tarantola, A. and Valette, B.: Generalized nonlinear inverse problems solved using the least squares criterion. Rev. Geophys. Space Phys. **20**, 219–232, 1982b

Turunen, T. and Silén, J.: Modulation patterns for the EISCAT incoherent scatter. J. atmos. terr. Phys. **46**, 593–600, 1984

Vallinkoski, M.: Statistics of incoherent scatter multiparameter fits. J. Atmos. Terr. Phys. **50**, 839–851, 1988

Vallinkoski, M. and Lehtinen, M.: Behaviour of parameter estimation errors in linear statistical inversion theory with applications to incoherent scatter error estimates. (to be submitted, 1989a)

Vallinkoski, M. and Lehtinen, M.: Parameter mixing errors in statistical inversion theory with applications to incoherent scatter error estimates. (to be submitted, 1989b)

Appendix. IS spectrum, ACF, derivatives and error routines

A.1. Introduction

The purpose of this Appendix is to give the complete source codes of the standard MATLAB user routines with which the calculations of the present work have been carried out. The routines are documented, so that the potential user may tailor them to suit one's particular tastes. All the routines have been tested and used in PC–MATLAB; the routines including calculations of second derivatives may turn out to be so time-consuming that it is preferable to use the faster Mac II–MATLAB or 386–MATLAB or any equivalent version without severe memory limitations, like Pro–MATLAB in VAX computers, SUN workstations &c to calculate and store these derivatives for later use.

A.2. The calculation of the IS spectrum

The IS spectrum subroutines, the ACF routines, their first and second derivatives are given. The general spectral formula and its calculation are given for the non-magnetic case with any number of ions. All populations are assumed to be Maxwellian, with speeds and temperatures independent of each other.

$$\frac{S(\theta_0)\,\omega_0}{N_e\sigma_0}\,d\theta_0 = \frac{|y_e|^2\sum_i \eta_i \mathrm{Re}y_i + |\sum_i \mu_i y_i + ik^2D^2|^2\mathrm{Re}y_e}{|y_e+\sum_i \mu_i y_i + ik^2D^2|^2}\frac{d\theta_0}{\pi\theta_0}$$

where $\theta_0 = \omega/\omega_0$. In the following, we need also $\psi_0 = \nu/\omega_0$, where $\nu = \nu_{in}$ is the ion-neutral collision frequency. The definition of the collision frequencies follows Schlegel (1979).

The admittances are defined by

$$y_s(\theta_s - i\psi_s) = i + \theta_s\frac{J(\theta_s)}{1-\psi_s J(\theta_s)} \equiv i + \theta_s u_s, \quad \text{with } J = -iG \quad \text{and where}$$

$$\theta_s = \omega/\omega_s, \quad \omega_s = k\sqrt{\frac{2\kappa T_s}{m_s}}, \quad \psi_s = \frac{\nu_s}{\omega_s}, \quad \psi_i = \psi_0, \quad \psi_e = 0.35714\,\psi_0$$

$$G(z) = i\sqrt{\pi}e^{-z^2} + 2e^{-z^2}\int_0^z e^{t^2}dt = 2e^{-z^2}\int_{-i\infty}^{z} e^{t^2}dt$$

$$\eta_i = N_i/N_e, \quad \mu_i = \eta_i T_e/T_i = \frac{N_i}{N_e}\frac{T_e}{T_i}, \quad \gamma_i = \omega_0/\omega_i = \sqrt{\frac{T_0}{T_i}\frac{m_i}{m_0}}$$

$$D = \lambda_D = v_t/\omega_p = \sqrt{\frac{\kappa T_e}{m_e}}\sqrt{\frac{\epsilon_0 m_e}{N_e e^2}} = \sqrt{\frac{T_e}{T_0}}\sqrt{\frac{N_0}{N_e}}D_0, \quad D_0 = \sqrt{\frac{\epsilon_0\kappa T_0}{N_0 e^2}}, \quad \kappa = k_B.$$

The spectrum is developed to a form symmetric with respect to electron/ion populations. All parameters are relative to the reference parameters m_0, T_0 and N_0.

$$\frac{S(\theta_0)\,\omega_0}{N_0\sigma_0}\,d\theta_0 = \frac{N_e}{N_0}\frac{\left|y(\frac{\theta_e}{\theta_0}\theta_0)\right|^2\sum_i\frac{N_i}{N_e}\gamma_i\mathrm{Re}u(\frac{\theta_i}{\theta_0}\theta_0) + \left|\sum_i\frac{N_i}{N_e}\frac{T_e}{T_i}y(\frac{\theta_i}{\theta_0}\theta_0) + i\frac{T_e}{T_0}\frac{N_0}{N_e}(kD_0)^2\right|^2\gamma_e\mathrm{Re}u(\frac{\theta_e}{\theta_0}\theta_0)}{\left|y(\frac{\theta_e}{\theta_0}\theta_0) + \sum_i\frac{N_i}{N_e}\frac{T_e}{T_i}y_i + i(kD_0)^2\frac{T_e}{T_0}\frac{N_0}{N_e}\right|^2\pi}d\theta_0$$

$$= \frac{N_e}{N_0}\frac{\left|\frac{T_0}{T_e}\frac{N_e}{N_0}y(\gamma_e\theta_0)\right|^2\frac{N_0}{N_e}\sum_i\frac{N_i}{N_0}\gamma_i\mathrm{Re}u(\gamma_i\theta_0) + \left|\sum_i\frac{T_0}{T_i}\frac{N_i}{N_0}y(\gamma_i\theta_0) + i(kD_0)^2\right|^2\gamma_e\mathrm{Re}u(\gamma_e\theta_0)}{\left|\frac{T_0}{T_e}\frac{N_e}{N_0}y(\gamma_e\theta_0) + \sum_i\frac{T_0}{T_i}\frac{N_i}{N_0}y(\gamma_i\theta_0) + i(kD_0)^2\right|^2\pi}d\theta_0$$

For collisional cases and with Doppler velocities taken into account the arguments in the y and u functions should be replaced by

$$\gamma_i\theta_0 \to \gamma_i(\theta_0 - v_i) - j\psi_i \quad \text{and} \quad \gamma_e\theta_0 \to \gamma_e(\theta_0 - v_e) - j\psi_e$$

The spectrum calculation relies on the specification of a transformation function **transf(p)** whose output should be a series of MATLAB vectors specifying the different ion species. The ions are indexed starting from 1 and the last element in the vectors is for the electrons.

$$\begin{aligned} \texttt{nin0(i)} &= N_i/N_0 \\ \texttt{tit0(i)} &= T_i/T_0 \\ \texttt{mim0(i)} &= m_i/m_0 \\ \texttt{psi(i)} &= \psi_i \\ \texttt{kd2(i)} &= (k_{scatt} D_0)^2 \\ \texttt{vi(i)} &= v_i/v_0 \end{aligned}$$

The calculation of the spectrum for arbitrary drifts, temperatures or an arbitrary number of ions is the following vector operation allowing the calculation of the spectra for a set of input parameter vector values p. The argument om is the scaled variable $\theta_0 = \omega/\omega_0$ given as a column vector.

```
% spec.m
% The incoherent scatter spectrum without the magnetic field.
%
% s=spec(p,om,outflag);
%
  function s=spec(p,om,outflag);
%
  t0=clock; f0=flops;
  [ns,np]=size(p);
  tt0=clock; ff0=flops;
  for i=1:ns,
    [nin0(i,:),tit0(i,:),mim0(i,:),psi(i,:),kd2(i),vi(i,:)]=transf(p(i,:));
  end;
  fprintf('time used in transf= %g,  flops=%g\n',etime(clock,tt0),flops-ff0);
  if nargin>2, format compact; nin0, tit0, psi, mim0, kd2, vi, end;
  j=sqrt(-1);
  [ns,nion]=size(mim0); nion=nion-1;
  nom=length(om);
  oma=zeros(nom,(nion+1)*ns);
  ua=oma;
  ya=oma;
  gam=zeros(ns,nion+1);
  for is=1:ns, for i=1:nion+1,
    icol=is+(i-1)*ns;
    gam(is,i)=sqrt(mim0(is,i)/tit0(is,i));
    oma(:,icol)=(om-vi(is,i))*gam(is,i);
    ua(:,icol)=oma(:,icol)-j*psi(is,i);
  end; end;
  tt0=clock; ff0=flops;
  ua=pldfi(ua);
  fprintf('time used in pldf  =%g,  flops=%g\n',etime(clock,tt0),flops-ff0);
  for is=1:ns, for i=1:nion+1,
    icol=is+(i-1)*ns;
    ua(:,icol)=-j*ua(:,icol)./(1+psi(is,i)*j*ua(:,icol));
    ya(:,icol)=(j+oma(:,icol).*ua(:,icol))*(nin0(is,i)/tit0(is,i));
    ua(:,icol)=ua(:,icol)*gam(is,i)*nin0(is,i);
  end; end;
  for is=1:ns, ya(:,is)=ya(:,is)+j*kd2(is); end;
  yaa=zeros(nom*ns,nion+1);
  yaa(:)=ya; ya=yaa';
  yaa(:)=ua; ua=yaa';
  s=(abs(ya(nion+1,:)).^2).*sum(real(ua(1:nion,:)));
```

```
  s=s+(abs(sum(ya(1:nion,:))).^2).*real(ua(nion+1,:));
  s=(s./(abs(sum(ya)).^2))'/pi;
  yya=zeros(nom,ns); yya(:)=s; s=yya;
  fprintf('total time used    =%g,  flops=%g',etime(clock,t0),flops-f0);
  fprintf(',  points calculated=%g\n',nom*ns);
end;
```

The value $\theta = \pm 1$ with one ion and no drifts or collisions is exactly the ion line position.

Transformation function from user parameters to plasma parameters. A call **transf('names')** should return a MATLAB array of names to be used for all the parameters and **transf('names',n)** should return the name of the n'th parameter **p(n)**. It is the responsibility of the user to take care of this if he/she wants to change the parametrization. The name strings are used for all output and labelling purposes.

The following form is for two ion species with equal temperatures and drifts, namely the molecular ion with mass 30.5 u and the O^+ with mass 16 u; also, for example, H^+ with mass 1 u and metallic ions, like Fe^+ with mass 56 may be added, but this will make it difficult to perform the spectrum calculations vectorially. The basic parameters N_0, T_0 and m_0 are contained in the file **bp.mat**, and the physical constants in the file **pars.mat**. The radar frequency **f** is fixed to be the EISCAT UHF frequency 933.5 MHz, and the wave number **k** is for backscatter. The initialisation is done in the following separate script **m**-file.

```
% specinit.m
% the EISCAT UHF calculation initialisation of parameters.
% All relevant parameters will have to be declared global.
% The basic parameters are in the MATLAB files pars.mat
% and bp.mat on disk and can be easily modified. The plasma
% dispersion function is tabulated in pldfv.mat.
%
% The values used are
% N0=1e+11;T0=300;m0=30.5;
% lightspeed=2.99792458e+8;fradar=933.5e+6;epsilon0=8.85419e-12;
% elcharge=1.6021892e-19;kB=1.380662e-23;u=1.660565e-27;
% me=5.485804e-4 u;mp=1.007277 u;
% me=9.109535e-31 kg;mp=1.672649e-27 kg
%
  load bp;
  N0=bp(1);T0=bp(2);m0=bp(3);
%
  load pars;
  lightspeed=pars(1);fradar=pars(2);epsilon0=pars(3);
  elcharge=pars(4);kB=pars(5);u=pars(6);
%
  kscatt=4*pi*fradar/lightspeed;
  v0=sqrt(2*kB*T0/(m0*u));
  om0=kscatt*v0;nu0=om0/(2*pi);tau0=pi/(2*om0);
  D2=kB*epsilon0*T0/(N0*elcharge^2);
%
  [lenp,n]=size(transf('names'));lenp=lenp-1;
  clear lightspeed fradar epsilon0 elcharge kB u n
%
  load pldfv
  global pldfv lenp m0 N0 T0 kscatt v0 om0 nu0 tau0 D2;
```

The function itself

```
% transf.m
% Function to transform user parameters to spectrum parameters
%
% [nin0,tit0,mim0,psi,kd2,vi]=transf(p,np);
```

```
%
  function [nin0,tit0,mim0,psi,kd2,vi]=transf(p,np);
%
  if isstr(p),
    nin0=['Ne/N0 ',
          'Ti/T0 ',
          'Te/Ti ',
          'nu/nu0',
          'O+    ',
          'v/v0  '];
    if nargin>1, nin0=nin0(np,:); end;
  else,
    load bp;
    m1=bp(3);m2=bp(4);
%
    load pars;
    me=pars(7);
%
    kd2=kscatt^2*D2;                % (kD0)^2
    psi0=p(4)./(2*pi);              % psi0
%
    nin0(3)=p(1);                   % Ne/N0;
    nin0(1)=p(1)*(1-p(5));          % N1/Ne
    nin0(2)=p(1)*p(5);              % N2/Ne
    tit0(1)=p(2);                   % T1/T0
    tit0(2)=p(2);                   % T2=T1
    tit0(3)=p(2)*p(3);              % Te/T0
    mim0(1)=m1/m0;                  % m1/m0
    mim0(2)=m2/m0;                  % m2/m0
    mim0(3)=me/m0;                  % me/m0 from u
    vi(1)=p(6);                     % v1/v0; v0=phase velocity om0/kscatt
    vi(2)=p(6);                     % v2/v0
    vi(3)=p(6);                     % ve/v0
    psi(1)=psi0;                    % psi1
    psi(2)=psi0;                    % psi2=psi1 for all temperatures
    psi(3)=psi0*0.35714;            % psie independently of Te/Ti
  end;
```

The following function is used to interpolate the plasma dispersion function values. Lagrangian interpolation is used. If the function values y_i are given at argument values x_i, the Lagrange interpolation polynomial is given by

$$P(x) = A(x)\sum_{j=1}^{N}\frac{y_j}{(x-x_j)\prod_{i\neq j}(x_i-x_j)} + R_N \quad \text{with} \quad A(x) = \prod_{i=1}^{N}(x-x_i).$$

The error term is given in a number of cases by

$$R_2 \approx 0.125\ \Delta^2 \approx 8.15\,10^{-4}$$
$$R_4 \approx 0.024\ \Delta^4 \approx 1.41\,10^{-5}$$
$$R_6 \approx 0.0042\Delta^6 \approx 7.00\,10^{-8}$$
$$R_8 \approx 0.0011\Delta^8 \approx 2.96\,10^{-9},$$

where the finite differences are approximately given by $\Delta^n \approx f^{(n)}(\Delta x)^n$. The numerical values calculated correspond to the maximum value for the real part of the plasma dispersion function with $\Delta x = 0.075$.

Four-point interpolation is used. If the user wants to use six-point interpolation instead, the parameter accur must be set to 6 below.

```
% pldfi.m
% Plasma dispersion function interpolation
%
% res=pldfi(z)
%
  function res=pldfi(z),
%
  accur=4;
  res=z;
  z=z(:);
  f1=find( abs(real(z)) >=3.9 | imag(z) <= -3.5 );
  f2=find( abs(real(z)) < 3.9 & imag(z) > -3.5 & imag(z) <0.0375 );
  f3=find( imag(z) >= 0.0375);
  if max(size(f1))>0, z(f1)=pldfas(z(f1),accur+2); end;
  if max(size(f3))>0, z(f3)=pldf(z(f3)); end;
%
  if max(size(f2)) >0,
    x=abs(real(z(f2)))/0.075+4;
    f1=find(real(z(f2)) < 0);
    y=-imag(z(f2))/0.075;
    p=x-floor(x)-sqrt(-1)*(y-round(y));
    x=floor(x)+61*round(y);  % index for linear pldfv table lookup
    if accur==4,
      y=(p-1).*(p-2).*( (-p/3).*pldfv(x-1) + (p+1).*pldfv(x) );
      y=(y+p.*(p+1).*( (2-p).*pldfv(x+1) + ((p-1)/3).*pldfv(x+2)) )*.5;
    elseif accur==6,
      p(find(abs(p-round(p))<=1e-99))=p(find(abs(p-round(p))<=1e-99))+2e-99;
      y=(pldfv(x+3)./(p-3)-pldfv(x-2)./(p+2))/10+pldfv(x+1)./(p-1);
      y=y+(pldfv(x-1)./(p+1)-pldfv(x+2)./(p-2))/2-pldfv(x)./p;
      y=y.*(p-3).*(p-2).*(p-1).*p.*(p+1).*(p+2)/12;
    end;
    y(f1) = -conj( y(f1) );
    z(f2)=y;
    res(:)=z;
  end;            % if max(size(f2)) >0
```

The definition of the plasma dispersion function G used here is

$$\mathrm{pldf}(z) = 2e^{-z^2}\int_0^z e^{t^2} + i\sqrt{\pi}e^{-z^2} = 2i\int_0^\infty e^{-t^2-2iz}dt = 2ie^{-z^2}\int_{iz}^\infty e^{-t^2}dt,$$

and for the set of values specified by the input vector z, the function returns a set of values specified by the output vector.

```
% pldf.m
% complex plasma dispersion function / Markku Lehtinen 10.3.1979
%
% accuracy:  8 numbers in the whole complex plane
%            (main algorithm accurate to 1e-12, accuracy is
%            limited by subroutine cerfexp and numerical accuracy)
%
% method:    approximation described by Salzer in
%            Math. Tables Aids Comput., Vol. 5 (1951) pp. 67-70.
%            The same formula appears in
```

```
%           Abramovitch-Stegun p.299 formula 7.1.29.
%           (the exponentials in the formulas have been written
%           in a slightly modified way to get rid of unnecessary
%           overflows)
%
% res=pldf(z);
%
  function res=pldf(z);
%
    if max(size(z))==1, res=0; else, res=zeros(z); end;
    [nrow ncol]=size(z);
    j=sqrt(-1);
    for irow=1:nrow, for icol=1:ncol,
      z1=-j*z(irow,icol);
      x=real(z1);
      y=imag(z1);
      if(abs(2*x*y)<30000),
         cs=cos(2*x*y);
         sn=sin(2*x*y);
      else,
          cs=1; sn=0;
      end;
      pisqr=sqrt(pi);
      fn=max(1,floor(abs(2*y))-11):floor(abs(2*y))+11;
      term1=exp(-fn.*fn/4-y*y);
      if max(abs(term1))<1e-100, term1=zeros(fn); end;
      term2=exp(-(fn/2-y).*(fn/2-y))/2;
      term3=exp(-(fn/2+y).*(fn/2+y))/2;
      factor=ones(fn)./(fn.*fn+4*x*x);
      sume=sum(factor.*(term1*cs-term2-term3));
      sumf=sum(factor.*(term1*2*x*sn+fn.*(term2-term3)));
      if (abs(x*y)<1e-4),
            e1=(-exp(-y*y)*x*y*y/2+2*x*sume)/pisqr;
            f1=(exp(-y*y)*y/2+sumf)/pisqr;
      else,
            e1=(exp(-y*y)*(cs-1)/4/x+2*x*sume)/pisqr;
            f1=(exp(-y*y)*sn/4/x+sumf)/pisqr;
      end;
      if abs(e1)<1e-100, e1=0; end;
      z1=2*j*(e1+j*f1)+j*pisqr*exp(-y*y)*(cs+j*sn)*cerfexp(-x);
      res(irow,icol)=z1;
    end; end;
    end;
```

The Lehtinen function

```
% cerfexp.m
%   Approximation for the function  (1-erf(x))*exp(x**2) with
%   accuracy 1e-8 for all x.  Method: power series for
%   erf(x) for small x, asymptotic continued fractions expansion
%   for large x (Abramovitz-Stegun formula 7.1.14)
%   Markku Lehtinen 2.7.1979
%
%   res=cerfexp(x);
%
    function res=cerfexp(x);
%
```

```
        x1=abs(x);
        sqx=x*x;
        if (x1<1.),
              tail=0.;
              term=sqx*sqx/10.;
              erf=1.-sqx/3.+term;
              for n=3:12,
                 term=-term*sqx*(2*n-1.)/(2*n+1.)/n;
                 tail=tail+term;
              end;
              erf=(erf+tail)*x*2./sqrt(pi);
              res=(1-erf)*exp(sqx);
        else,
              dum=0;
              for n=(55:-1:1),
                dum=n/2/(dum+x1);
              end;
              res=1/(dum+x1)/sqrt(pi);
              if (x<0), res=2*exp(sqx)-res; end;
        end;
      end;
```

The asymptotic formula for the plasma dispersion function is

$$\text{pldf}(z) \approx z^{-1}(1 + 1/z^2 + 1 \cdot 3/(z^2)^2 + 1 \cdot 3 \cdot 5/(z^2)^2 + \ldots),$$

and the expansion is valid for z not close to positive y-axis. We must require that $|\arg z - \pi/2| > \pi/4$.

```
% pldfas.m
% asymptotic formula for the plasma dispersion function
% n is the order of the expansion
%
% res=pldfas(z,n)
%
 function res=pldfas(z,n);
%
  if max(size(z))==0, res=[]; return; end;
  term=1;
  res=1;
  for i=1:n,
    term=term*(2*i-1)./(2*z.^2);
    res=res+term;
  end;
  res=res./z;
```

The following function can be used to create different tabulations of the plasma dispersion function

```
% pldftab.m
% function to create the plasma dispersion function table
%
% res=pldftab(dx,dy,nx,ny,nx1,ny1);
%
  function res=pldftab(dx,dy,nx,ny,nx1,ny1);
%
  res=zeros(nx*ny,1);
  for i=1:nx, for j=1:ny,
    res(i+(j-1)*nx,1)=pldf( (nx1+i-1)*dx +sqrt(-1)*(ny1+j-1)*(-dy) );
  end;
```

The following **m-file** was used to create the present pldfv. The tabulation interval is 0.075 in both directions, three x-values below zero, first y-value zero, 61 x-values leading to last value being (61-4)*.075=4.275 and 48 y-values leading to last value being -47*0.075=-3.525.

```
% pldftabb.m
% script file to create and save the plasma dispersion function table
%
  pldfv=pldftab(0.075,0.075,61,48,-3,0);
  save pldfv pldfv;
```

A.3. The calculation of IS autocorrelation function and its derivatives

The autocorrelation function (ACF) subroutines are calculated in the non-magnetic case for two ions with zero drifts. All populations are assumed to be Maxwellian. We restrict our considerations to the case of the EISCAT UHF radar. For other radars or more ions, **specinit** has to be modified.

The ACF is the Fourier transform of the incoherent scatter spectrum. The present restrictions lead to a spectrum which is symmetric with respect to the origin (zero Doppler shift) in frequency. The Fourier transforms are calculated by the FFT algorithm in MATLAB. In practice, this means that one cannot specify arbitrarily *both* the spectral resolution $d\omega$ with which the spectrum is calculated *and* the lag resolution $d\tau$ of the ACF. They are connected by the formula

$$d\omega d\tau = \frac{2\pi}{N_{fft}}$$

where N_{fft} is the total number of points in the FFT, which in practice has to be a power of two. The spectral resolution may be defined as $d\omega = \omega_0/N_\omega$ and the lag resolution as $d\tau = \tau_0/N_\tau$ where ω_0 gives the scale of the basic ion line position and by $\tau_0 = \pi/2\omega_0$ the scale of the position of the first zero of the ACF (when the effect of the ion–neutral collision frequencies is small). Technically, when the positive part of the spectrum has been calculated at N points where N is a power of two, the spectrum has to folded to become symmetric around the origin. Since the the origin can appear only once, one needs an 'extra' point to have an FFT with a total of $2N$ points. When we use N instead of N_{fft}, then the above equation will become

$$N = 2N_\omega N_\tau .$$

Usually one wants to specify the lag resolution $d\tau$ first. Then both N_ω and N will have to be determined; since N has to be a power of two, it is useful first to specify the minimum spectral resolution $d\omega_{min}$ and then find the smallest possible N. The resulting spectral resolution will then be at least $d\omega_{min}$. For the specific values considered in Chapter 2, we have *e.g.* for $N_\tau \approx 2.5$, $d\tau \approx 40\mu$s, and $N = 32$ with $N_\omega \approx 6.5$; for the case $N_\tau = 10$, $d\tau \approx 10\mu$s, and $N = 128$ with $N_\omega \approx 6.5$; $d\tau \approx 10\mu$s already serves as an infinitely good lag resolution.

The calculation of the corresponding ACF without drifts is done by the following **m-file**. Note that since the frequencies ω are a column vector, the spectrum is given as a column vector. Otherwise the **flipx** routine would have no effect. Note that the frequencies ω are defined to be non-negative.

```
% acfcalc.m
% The ACFs of the spectra with N+1 points. The function returns
% the ACFs along with the lag axis specifications. The parameter
% values are all contained in the matrix p consisting of the
% parameter value row vectors (Ne/N0,Ti/T0,Te/Ti,nu/nu0,p(0+)).
% The zero drifts are appended to p within the routine.
%
% [res,lagx]=acfcalc(p,Ntau,N)
%
  function [res,lagx]=acfcalc(p,Ntau,N);
%
  pp=zdrft(p);
  dom0=2*Ntau/N;dtau=1/Ntau;
```

```
om=(0:N)'*dom0;
sp=spec(pp,om);
[m,n]=size(pp);
for i=1:m,
  ssp(:,i)=[sp(1:N,i);flipx(sp(2:N+1,i))];
end;
clear sp;
af=fft(ssp);
res=dom0*real(af(1:N,:));
lagx=(0:N-1)'*dtau;
```

The first derivatives needed in the linear inversion calculations in Section 2 are calculated by the following **m-file** for a single given parameter vector without drifts. The parameter values are given by the vector **pvec** consisting of the parameter value row vector $(N_e/N_0, T_i/T_0, T_e/T_i, \nu/\nu_0, p(O^+))$. The vector **h=incr** contains the values of the increments used in estimating derivatives at the given point **x0=pvec**:

$$\left(\frac{\partial f}{\partial x_i}\right)_{x=x_0} = \frac{f(x_0+h_i)-f(x_0-h_i)}{2h_i}.$$

Note that there is no increment for drifts and that the lag axis has to remain the same for all ACFs in the calculations (same scalars **N** and **Nfft**). The zero drifts are appended to **p** and **incr** within the routine. The following form is for those versions of MATLAB where there are in practice no severe memory limitations. If this routine produces out of memory –messages, the routine **der1acfc** must be used. The calls to routines **der0acfc** and **der1acfc** may otherwise be freely interchanged in all subsequent routines.

```
% der0acfc.m
% The derivatives of a given single ACF calculated as a vector operation.
% The function returns the derivatives with the lag axis specifications.
% The parameter values are given by the parameter row vector pvec.
% The row vector incr contains the values of the increments used
% in calculating the symmetric derivatives at the given point pvec.
% The zero drifts are appended to p and incr within the routine.
%
% [res,lagx]=der0acfc(p,Ntau,N,incr)
%
  function [res,lagx]=der0acfc(pvec,Ntau,N,incr);
%
  p=row(pvec);inc=row(incr);
  n=length(pvec);
  inc0=diag(inc);
  inc1=kron(inc,ones(N,1));
  p1=kron(pvec,ones(n,1))+inc0;
  p2=kron(pvec,ones(n,1))-inc0;
  p1=zdrft(p1);
  p2=zdrft(p2);
  [acf1,lagx]=acfcalc(p1,Ntau,N);
  [acf2,lagx]=acfcalc(p2,Ntau,N);
  dacf=(acf1-acf2)./(2*inc1);
  end;
  res=dacf;
```

and the following form is the poor man's version calculating one derivative at a time:

```
% der1acfc.m
% The derivatives of a given single ACF calculated one at a time.
% The function returns the derivatives with the lag axis specifications.
% The parameter values are given by the parameter row vector pvec.
```

```
% The vector incr contains the values of the increments used
% in calculating the symmetric derivative estimates at the given point pvec.
%
% [res,lagx]=der1acfc(p,Ntau,N,incr)
%
  function [res,lagx]=der1acfc(pvec,Ntau,N,incr);
%
  p=row(pvec);inc=row(incr);
  n=length(p);
  p0=kron(p,ones(2,1));
  for i=1:n,
    p1=p0;di=incr(i);
    p1(1,i)=p0(1,i)+di;
    p1(2,i)=p0(2,i)-di;
    p1=zdrft(p1);
    [acf,lagx]=acfcalc(p1,Ntau,N);
    res(:,i)=(acf(:,1)-acf(:,2))/(2*di);
  end;
```

The second derivatives are needed in the second moment calculations in Chapter 4.
The second derivative estimates are obtained from the symmetric formulae

$$\left(\frac{\partial^2 f}{\partial x_i^2}\right)_{x=x_0} = \frac{f(x_0+h_i) - 2f(x_0) + f(x_0-h_i)}{h_i^2},$$

for $i = j$, and

$$\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{x=x_0} = \frac{f(x_0+h_i+h_j) - f(x_0+h_i-h_j) - f(x_0-h_i+h_j) + f(x_0-h_i-h_j)}{4h_ih_j},$$

for $i \neq j$.
The following routine calculates the second derivatives of an ACF. Only the poor man's version is given.

```
% der2acfc.m
% The second derivatives of a given single ACF. The function returns
% the derivatives with the lag axis specifications. The parameter
% values are given by the parameter row vector pvec.
% The vector incr contains the values of the increments used
% in calculating the symmetric derivatives at the given point pvec.
%
% [res,lagx]=der2acfc(p,Ntau,N,incr)
%
  function [res,lagx]=der2acfc(pvec,Ntau,N,incr);
%
  p=row(pvec);inc=row(incr);
  n=length(p);
  p0=kron(p,ones(2,1));
  ij=0;
  for i=1:n
    p1=p0;p2=p0;
    di=inc(i);
%   the diagonal derivatives
    p1(1,i)=p0(1,i)+di;
    p1(2,i)=p0(1,i);
    p2(1,i)=p0(1,i)-di;
    p2(2,i)=p0(1,i);
    ij=ij+1;
    p1=zdrft(p1);
```

```
    p2=zdrft(p2);
    [acf1,lagx]=acfcalc(p1,Ntau,N);
    [acf2,lagx]=acfcalc(p2,Ntau,N);
    res(:,ij)=(acf1(:,1)-acf1(:,2)+acf2(:,1)-acf2(:,2))/(di*di);
%   the mixed derivatives
    p1(1,i)=p0(1,i)+di;
    p1(2,i)=p1(1,i);
    p2(1,i)=p0(1,i)-di;
    p2(2,i)=p2(1,i);
    for j=i+1:n,
      ij=ij+1;fprintf('round %3.0f\n',ij);
      dj=incr(j);
      p1(1,j)=p1(1,j)+dj;
      p1(2,j)=p1(2,j)-dj;
      p1=zdrft(p1);
      [acf1,lagx]=acfcalc(p1,Ntau,N);
      p2(1,j)=p2(1,j)+dj;
      p2(2,j)=p2(2,j)-dj;
      p2=zdrft(p2);
      [acf2,lagx]=acfcalc(p2,Ntau,N);
      res(:,ij)=(acf1(:,1)-acf1(:,2)-acf2(:,1)+acf2(:,2))/(4*di*dj);
    end;
  end;
```

A.4. The error calculation routines in Chapter 2

The error calculations require the calculation of the Fisher information matrix Q:

$$Q = \Sigma_0^{-1} + A^{\mathrm{T}}\Sigma_m^{-1}A,$$

where A is the derivative matrix of the ACF calculated in the previous Section. The ACF noise level is fixed to 1%, and the variance σ is calculated from Eq.(2.24). The error tables Table 1 and Table 2 have been calculated by the following routine, where also Q is output. The parameter **Nlags** is the number of lags taken into account, and the vector pa is the a priori information vector: when the element corresponding to a parameter is zero, the rows and columns in the information matrix are zero, and a nonzero element denotes the standard deviation or the square root of the variance (accuracy). So *e.g.* a 10% accuracy in T_i means `[0 0.1 0 0 0]`.

```
% errcalc1.m
% The routine calculates and ouputs the error table and the information
% matrix Q.
% The user must choose the values of Ntau,N and incr
% for the deriacfc routine. The vector pa is the vector specifying
% the a priori widths of specified parameters. Zero means infinite
% variance. Correlations are not taken into account.
%
% [err,Q]=errcalc1(pvec,Ntau,N,incr,Nlags,pa);
%
  function [err,Q]=errcalc1(pvec,Ntau,N,incr,Nlags,pa);
%
  Q=errq(pvec,Ntau,N,incr,Nlags,pa);
  p=row(pvec);
  n=length(p);
  num=2^n;
  err=zeros(num,n+1);
  f=zeros(1,n);
```

```
  mask=zeros(1,n);
%
  err(1,1:n)=p;
  for i=2:num,
    f=bits(f,mask);
    d=sqrt(diag(pcovar(Q,f))');
    err(i,1:n)=d;
    err(i,n+1)=log(cond(Q(f,f)))/log(10);
  end;
```

The actual values used were `N=128, Ntau=10, incr(1:5)=0.005 and Nlags=60`. The help routines `bits.m` and `pcovar.m` are listed at the end of this Appendix.

The routine calculating the full information matrix Q is

```
% errq.m
% The user must choose the values of N, Ntau, incr and Nlags
% as in the der1acfc routine. The vector pa is the vector specifying
% the a priori widths of specified parameters. Zero means infinite
% variance. Correlations are not taken into account.
%
% Q=errq(p,Ntau,N,incr,Nlags,pa);
%
  function Q=errq(p,Ntau,N,incr,Nlags,pa);
%
  der=der1acfc(p,Ntau,N,incr);
  der=der(1:Nlags,:);
  sigma=sqrt(Ntau)*0.5*0.01;
  scale=sigma^(-2);
  ata=der'*der;
  n=length(p);pi=ones(1,n);
  apr=aprinfo(pa,pi);
  Q=ata*scale+apr;
```

The routine calculating the errors as a function of the sampling interval $d\tau$ is the following. The input needed is the same as for the derivative calculations plus `Nlags`. The parameters of the fit are given by the mask vector `f`, where 1 denotes the parameter which is considered variable and 0 the parameter which is supposed to be known. For the 3-parameter fit $(N_e, T_i, T_e/T_i)$, the mask is thus `f=[1 1 1 0 0]`. The basic lag length $d\tau$ must be so chosen that it can represent the infinite sampling frequency. Since the matrices (may) become ill-conditioned for large lag lengths, the parameter `K` denoting the last multiple of the lag length taken into account is given interactively. The results are calculated with and without the zero lag.

```
% lagres.m
% The user must choose the values of Ntau,N and incr
% for the der1acfc routine.
% [err1,err0,lagx]=lagres(p,Ntau,N,incr,Nlags,f)
%
  function [err1,err0,lagx]=lagres(p,Ntau,N,incr,Nlags,f)
%
  der=der1acfc(p,Ntau,N,incr);
  der=der(1:Nlags,:);
  J=sum(f);K=input('longest lag used?');
  err1=zeros(K,J);
  err2=zeros(K,J);
  lagx=(1:K)/Ntau;
  for j=1:K
    der1=der(1:j:Nlags,:);
    a=der1(:,f);
```

```
  Neff=Ntau/j;
  sigma=sqrt(Neff)*0.5*0.01;
  scale=sigma^(-2);
  ata=a'*a;
  ata=ata*scale;
  err1(j,:)=sqrt(diag(inv(ata))');
end;
for j=1:K
  der0=der(1:j:Nlags,:);
  M=size(der0);
  a=der0(2:M,f);
  Neff=Ntau/j;
  sigma=sqrt(Neff)*0.5*0.01;
  scale=sigma^(-2);
  ata=a'*a;
  ata=ata*scale;
  err0(j,:)=sqrt(diag(inv(ata))');
end;
```

The routine calculating the errors as a function of the longest lag used $\Delta\tau$ is the following. The input needed is the same as for the derivative calculations plus **Nlags**. The parameters of the fit are given by the mask vector **f**, where 1 denotes the parameter which is considered variable and 0 the parameter which is supposed to be known. For the 3–parameter fit $(N_e, T_i, T_e/T_i)$, the mask is thus `f=[1 1 1 0 0]`. The basic lag length $d\tau$ must be so chosen that it can represent the infinite sampling frequency. Since the matrices (may) become ill-conditioned for small lag extents, the parameter `K` denoting the length of the first lag taken into account is given separately. The results are calculated with and without the zero lag.

```
% lagext.m
% The user must choose the values of Ntau,N and incr
% for the der1acfc routine.
%
% [err1,err0,lagx]=lagext(p,Ntau,N,incr,Nlags,f)
%
  function [err1,err0,lagx]=lagext(p,Ntau,N,incr,Nlags,f)
%
  der=der1acfc(p,Ntau,N,incr);
  der=der(1:Nlags,:);
  J=sum(f);K=input('shortest lag extent used?');
  lagx=(K:Nlags)/Ntau;
  M=length(lagx);
  err1=zeros(M,J);
  err2=zeros(M,J);
  Neff=Ntau;
  for j=K:Nlags
    der1=der(1:j,:);
    a=der1(:,f);
    sigma=sqrt(Neff)*0.5*0.01;
    scale=sigma^(-2);
    ata=a'*a;
    ata=ata*scale;
    err1(j-K+1,:)=sqrt(diag(inv(ata))');
  end;
  for j=K:Nlags
    der0=der(2:j,:);
    a=der0(:,f);
    sigma=sqrt(Neff)*0.5*0.01;
    scale=sigma^(-2);
```

```
  ata=a'*a;
  ata=ata*scale;
  err0(j-K+1,:)=sqrt(diag(inv(ata))');
end;
```

A.5. The error calculation routines in Chapter 4

The error routines are easily constructed from the routines in used in Chapter 2. If the first and second derivatives together with the full information matrix have been calculated separately previously into a `.mat` file, they may be loaded from the disk directly. In that case, the variable names must have been declared `der1` and `der2`. (The potential user's taste for names &c may be different. This is one possibility to use previously stored variables.) The formula containing the information matrix $Q = A^{\mathrm{T}}\Sigma_m^{-1}A$

$$b = -(A^{\mathrm{T}}\Sigma_m^{-1}A)^{-1}A^{\mathrm{T}}\Sigma_m^{-1}\sum_{i,j} C_{ij}\,\mathrm{Cov_V}[\mathrm{x}]_{\mathrm{ij}},$$

where the curvature C is defined in (4.9), may be written using the pseudoinverse (since Σ_m is diagonal) of A as

$$b = -\mathrm{pinv}(A)\sum_{i,j} C_{ij}\,\mathrm{Cov_V}[\mathrm{x}]_{\mathrm{ij}}$$

The following routine calculates the bias correction matrix B, and the corrections to the basic parameter vector `x0`=$(N_e/N_0, T_i/T_0, T_e/T_i, \nu/\nu_0, p)$ without drifts. The full matrix B is output, and the also the correction vector `dx` plus the result vector `x=x0+dx`. `vw` specifies the parameter variation ranges Δa.

```
% biascorr.m
% The user must specify as in errcalc1 the parameter vector p.
% The parameter vw is the 'half-widths' vector of the parameter variations.
% If the flag is on, the derivatives and the information matrix are loaded
% from disk. The derivative matrix names must be der1 and der2.
% If there is a priori information, one cannot directly use the
% pseudoinverse pinv(A). The routine must be modified accordingly.
% Since the second derivatives and the covariance matrix (up to
% the constant 1/6) elements vw(i)vw(j) are stored as a vector
% consisting of the upper triangular parts for each lag, the off-diagonal
% elements must be counted twice in the final sum.
%
% [x,dx,B]=biascorr(flag,vw,Nlags,p,Ntau,N,incr)
%
  function [x,dx,B]=biascorr(flag,vw,Nlags,p,Ntau,N,incr)
%
  if flag==0 & nargin==7,
    der1=der1acfc(p,Ntau,N,incr);
    der2=der2acfc(p,Ntau,N,incr);
  elseif flag==1 & nargin >3,
    fnam=input('derivatives file? ','s');
    eval(['load ' fnam]);
  else
    error('flag must either 0 or 1 and input 7 or at least 4 parameters');
  end;
%
  der1=der1(1:Nlags,:);
  der2=der2(1:Nlags,:);
%                                          The diagonal mask
  V=vectoris(vw);
  n=length(p);
  f=diagindv(n);
```

```
%                                                    The bias
  B=pinv(der1)*der2/3;
  B(:,f)=0.5*B(:,f);
  dx=-B*V;
  x=col(p)+dx;
```

The following routine calculates the effect of parameter variations on the multiparameter fits. The fact that the covariance matrix is a positive definite matrix is not used, but the errors calculated satisfy the positivity constraints. It is advisable to use the routine only for one-parameter variations until further notice. The a priori knowledge on parameter variations has to be input.

```
% errcalc2.m
% The input is as in errcalc1 without a priori information.
% The parameter v is now the flag vector of the parameter variations.
% The parameter va is the half-widths vector of the parameter variations.
% 3-parameter fits and higher are output. The last items in the output
% are square roots of the errors, ie. parameter variation errors
% (for the diagonal elements). Each line in the output is duplicated
% by the line containing the effect of unknown variations of the parameter.
%
% err=errcalc2(flag,v,va,Nlags,p,Ntau,N,incr,pa)
%
  function err=errcalc2(flag,v,va,Nlags,p,Ntau,N,incr,pa)
%
  if flag==0 & nargin >=8,
    der1=der1acfc(p,Ntau,N,incr);
    der2=der2acfc(p,Ntau,N,incr);
  elseif flag==1 & nargin >=6,
    fnam=input('derivatives file? ','s');
    eval(['load ' fnam]);
  else
    error('flag must either 0 or 1 and input 8,9 or at least 6 parameters');
  end;
%
  der1=der1(1:Nlags,:);
  der2=der2(1:Nlags,:);
%                                    The curvature natrix
  n=length(p);pi=ones(1,n);
  if length(v)~=n, error('incompatible v'), end;
  f=diagindv(n);
  der2(:,f)=der2(:,f)/2;
%                                    The variable covariances mask
  V=vectoris(v);
  lenv=length(V);
  g=ones(1,lenv) & V;
  J=sum(g);
%                                    The bare information matrix
  a=[der1 der2(:,g)/3];
  sigma=sqrt(Ntau)*0.5*0.01;
  scale=sigma^(-2);
  ata=scale*a'*a;
%                                    The a priori information matrix
  nJ=n+J;
  nj=n+1:nJ;
  apr=zeros(nJ);
  if nargin==9,
    aprp=aprinfo(pa,pi);
```

```
    apr(1:n,1:n)=aprp;
  end;
  vaa=va.^(2);
  apra=aprinfo(vaa,v);
  apr(nj,nj)=apra;
  Q=ata+apr;
%                                  The error matrix
  h=zeros(1,nJ);
  h(1:3)=[1 1 1];
  count=nJ-3;
  num=2^count+1;
  err=zeros(num,nJ);
%                                  The errors
  mask=h;
  err(1,1:n)=p;
  for i=(2:num),
    C=pcovar(Q,h);
    d=sqrt(diag(C)');
    err(i,1:n)=d(1:n);
    err(i,nj)=sqrt(d(nj));
    h=bits(h,mask);
  end;
```

A.6. The error calculation routines in Chapter 6

The routine calculating the errors in chosen fits specified by the mask `f` as a function of the composition p is the following. The mask consists of ones and zeros, one denoting a variable which is considered variable and zero a variable which is fixed. The output is `rerr` (relative errors for $N_e, T_i, T_e/T_i$), composition value axis `lc` and absolute errors `aerr`, if needed.

The user must specify the starting value, increment and the number of the composition values wanted in a vector `pp=[pmin dp Np]`. The rest of the parameters is as in `der1cfc.m`. The actual lag distribution is given as a mask `F` of length 40 with the zero lag data missing (`Z=0`).

```
% comperr.m
% for the parameter combination
% given by the mask f with the last bit set to one.
%
% [rerr,lc,aerr]=comperr(p,Ntau,N,incr,pp,f)
%
  function [rerr,lc,aerr]=comperr(p,Ntau,N,incr,pp,f)
%
  Z=0;
  F=[Z 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 ..
     0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1];
  M=length(F);                                          %M=40
  J=sum(f);
  pmin=pp(1);dp=pp(2);Np=pp(3);
  lc=pmin+(0:Np-1)'*dp;
  err=zeros(Np,J);
  sigma=sqrt(Ntau)*0.5*0.01;
  scale=sigma^(-2);
  t0=clock;
%
  p0=[p(1:3) 1 1];
  for j=1:Np
    fprintf('starting round %3.0f of %3.0f\n',j,Np);
```

```
    p(5)=lc(j);
    der=der0acfc(p,Ntau,N,incr);
    a=der(1:M,f);
    a=a(F,:);
    ata=a'*a*scale;
    aerr(j,:)=sqrt(diag(inv(ata))');
  end;
  pp=kron(ones(Np,1),p0);
  rerr=aerr./pp(:,f);
%
  dt=etime(clock,t0);
  fprintf('total time used %6.2f minutes\n',dt/60);
```

The routine calculating the errors in chosen fits specified by the mask `f` as a function of the ion temperature T_i is the following. It is supposed that the second bit of the mask is one. The errors will be given as relative errors for the first three parameters in the output vector **`rerr`**; **`lt`** specifies the temperature axis in Kelvins, and **`aerr`** (if needed) gives the absolute errors with constant **`Neff`**.

The user must specify the starting value, increment and the number of the composition values wanted in a vector **`pt=[tmin dt Np]`**. The rest of the parameters is as in **`der1cfc.m`**. The actual lag distribution is given as a mask `F` of length 40 with the zero lag data missing (`Z=0`). Due to the temperature dependence of the first zero of the ACF, the number **`Neff`**, which is inversely proportional to the square root of T_0 (the real power is slightly less), can no longer be held constant. Also the temperature dependence of the normalised collision frequency variable ν/ν_0 on T_0 has to be taken into account in calculating relative errors. The absolute errors have to be looked upon with great caution.

```
% temperr.m
% The routine calculates the result in a loop
% for the parameter combination
% given by the mask f with the second bit set to one.
% [rerr,lt,aerr]=temperr(p,Ntau,N,incr,pt,f)
%
  function [rerr,lt,aerr]=temperr(p,Ntau,N,incr,pt,f)
%
  Z=0;
  F=[Z 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 ..
     0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1];
  M=length(F);                                          %M=40
  np=length(p)-1;
  J=sum(f);
  tmin=pt(1);dt=pt(2);Np=pt(3);
  lt=tmin+(0:Np-1)'*dt;
  err=zeros(Np,J);
  sigma=sqrt(Ntau)*0.5*0.01;
  scale=sigma^(-2);
  t0=clock;
%
  p0=[p(1:3) 1 1];
  for j=1:Np
    fprintf('starting round %3.0f of %3.0f\n',j,Np);
    p(2)=lt(j);
    der=der0acfc(p,Ntau,N,incr);
    a=der(1:M,f);
    a=a(F,:);
    ata=a'*a*scale;
    aerr(j,:)=sqrt(diag(inv(ata))');
  end;
  pp=kron(ones(Np,1),p0);
```

```
  pp(:,2)=lt;
  pp(:,4)=sqrt(lt);
  aerr=aerr./kron(sqrt(lt),ones(1,J));
  rerr=aerr./pp(:,f);
  lt=lt*T0;
%
  dt=etime(clock,t0);
  fprintf('total time used %6.2f minutes\n',dt/60);
```

A.7. The help routines used in the preceding calculations

```
% bits.m
% The routine steps in binary form over a given binary vector using
% a mask m determining which bits are stepped. If the mask bit is 1,
% the corresponding bit in f is not acted upon. Bits 1 in f are
% conserved.
%
% [res,ifend]=bits(f,m),
%
  function [res,ifend]=bits(f,m),
%
  c=1;m=row(m);n=length(m);
  for i=n:-1:1,
    if m(i)==0 & c,
      if f(i),
         c=1;
         f(i)=0;
      else
         c=0;
         f(i)=1;
      end;
    end;
  end;
  res=f;
  ifend=c;


% pcovar.m
% The routine calculates the a posteriori covariances
% from the information matrix Q and the mask f consisting of zeros
% and ones specifying which parameters are fixed. The output
% is a maximum fixed length vector consisting of the variances of
% the variable parameters and zeros for fixed parameters.
%
% res=pcovar(Q,f)
%
  function res=pcovar(Q,f)
%
  n=length(f);
  res=zeros(n);
  res(f,f)=inv(Q(f,f));


% aprinfo.m
% The construction of the a priori partial information matrix from a vector
```

```
% va containing the the a priori widths of the parameters and the
% parameter vector mask specifying which parameters are to be taken
% into account. Correlations are neglected, so that Qa will be diagonal.
% Qa=aprinfo(va,v)
%
  function Qa=aprinfo(va,v);
%
  u=row(v);ua=row(va);
  n=length(u);d=zeros(1,n);
  f=ua&ua;
  d(f)=ua(f).^(-2);
  ui=find(u==1);
  Qa=diag(d(ui));
```

The row and column operations on a vector (works also on a matrix); the row operation

```
% row.m
% makes a vector or a matrix into a row vector
%
% res=row(A)
%
  function res=row(A)
%
  res=A(:)';
```

The column operation

```
% col.m
% makes a vector or a matrix into a column vector
%
% res=col(A)
%
  function res=col(A)
%
  res=A(:);
```

The parameter vector completed to include the zero drifts

```
% zdrft.m
% The input is nxlenp matrix of single spectral parameter vectors
% p. It is completed to pp to include the zero drifts.
%
% pp=zdrft(p)
%
  function pp=zdrft(p)
%
  n=size(p);
  pp=[p zeros(n,1)];


% vectoris.m
% calculates the independent elements an nx(n+1)/2 matrix of the
% products of a vector v stored as a vector
%
% V=vectoris(v)
%
  function V=vectoris(v)
%
  ij=0;
  n=length(v);
```

```
  for i=1:n,
    for j=i:n,
      ij=ij+1;
      V(ij)=v(i)*v(j);
    end;
  end;


% diagindv.m
% the index function of the diagonal elements of a symmetric nxn matrix
% put into a vector V as in vectoris.m
%
% f=diagindv(n)
%
  function f=diagindv(n)
%
  f=zeros(1,n);f(1)=1;
  l=(n:-1:2);
  for i=2:n,
    f(i)=f(i-1)+l(i-1);
  end;
```

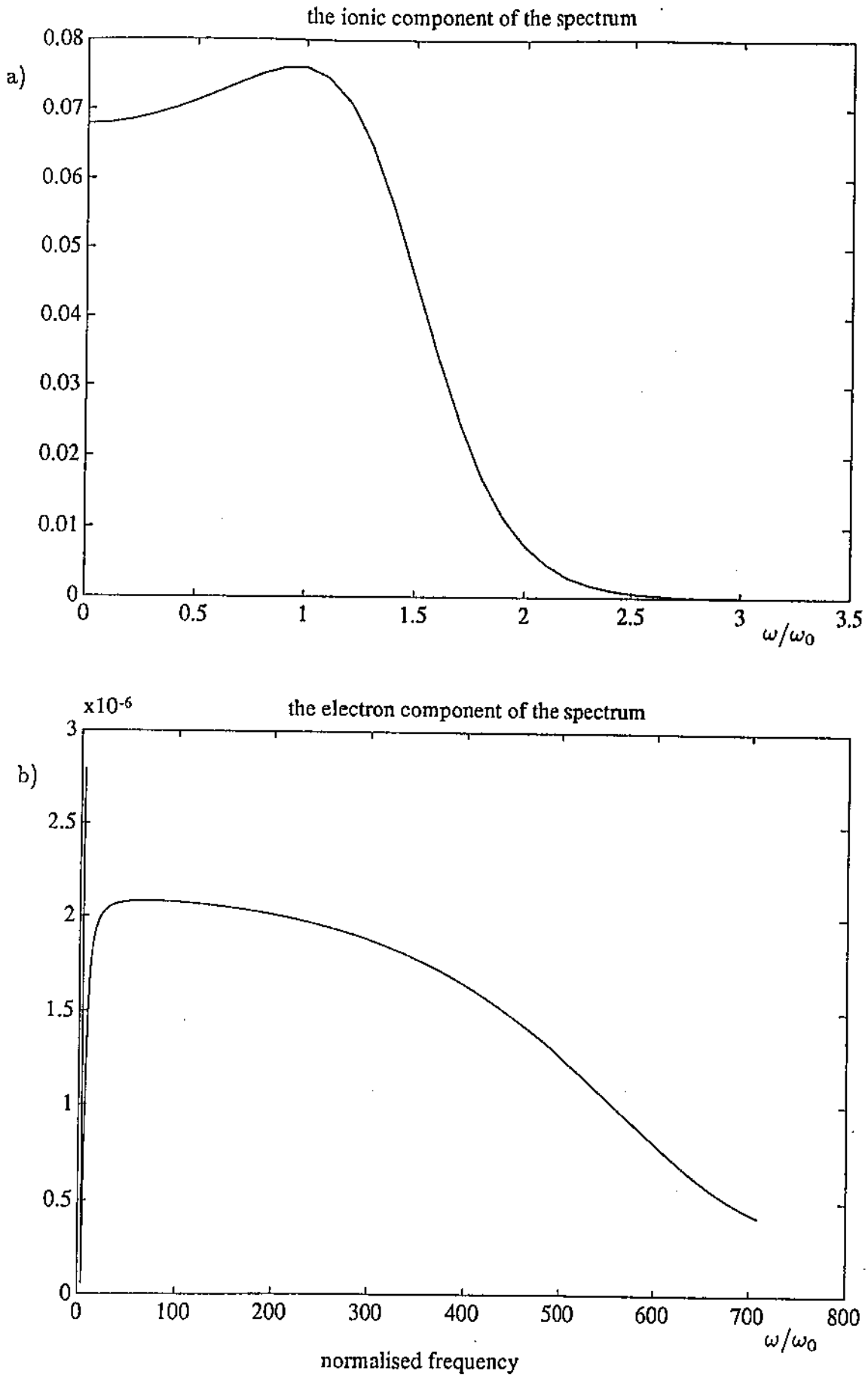


Fig. 1. A sample of a) the ionic component and b) the non-resonant part of the electron component of the IS spectrum. The normalisation and notation is as in Chapter 2. The values used are $N_e = 0.5{\cdot}10^{11}\ \mathrm{m}^{-3}$, $T_i = 300\ \mathrm{K}$, $T_e/T_i = 1$, $\nu_{in} = 0$ and $p(\mathrm{O}^+) = 0$.

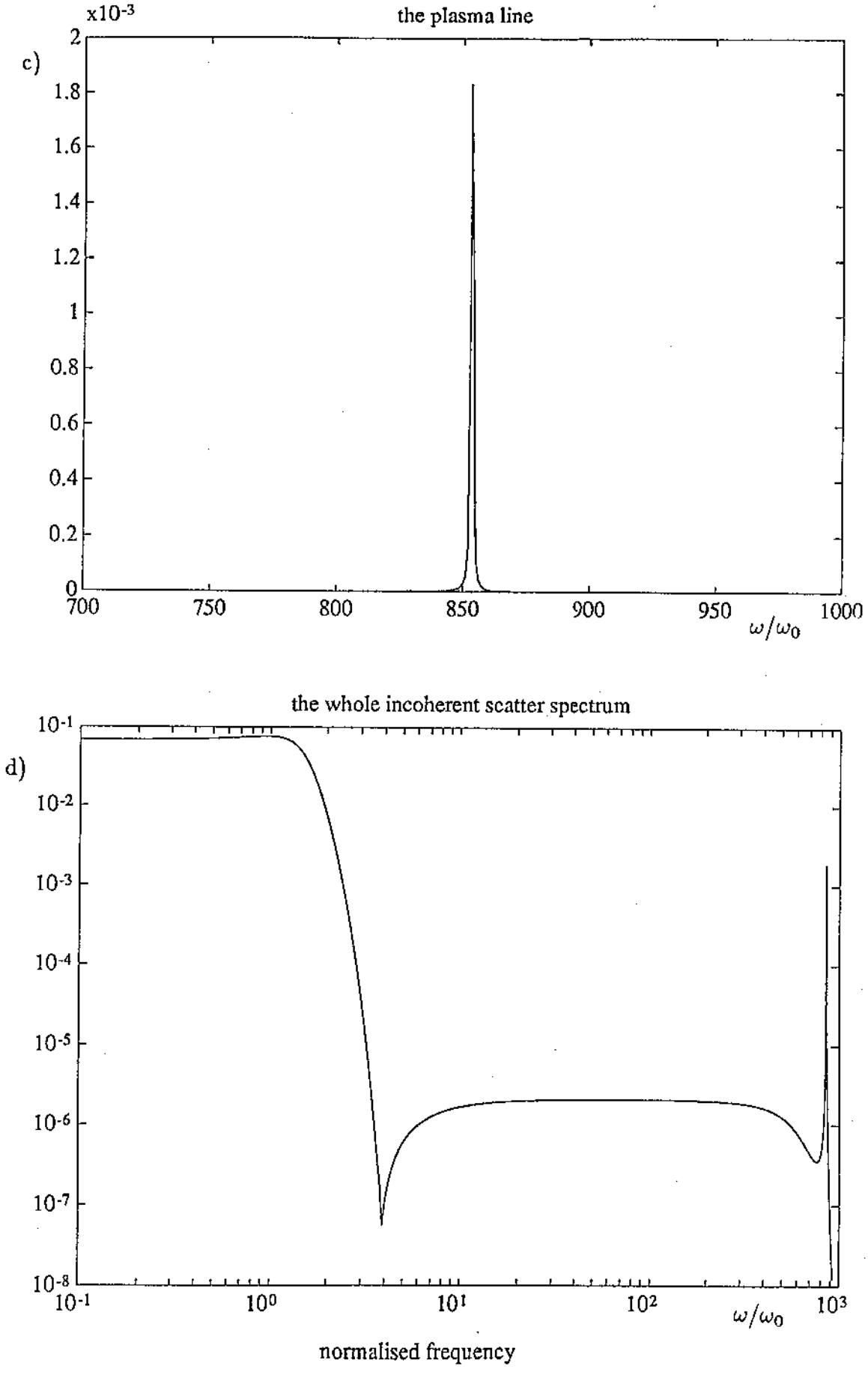


Fig. 1. A sample of c) the resonant plasma line in the electron component of the IS spectrum and d) the whole IS spectrum. The density value used was chosen to enhance the electron component somewhat.

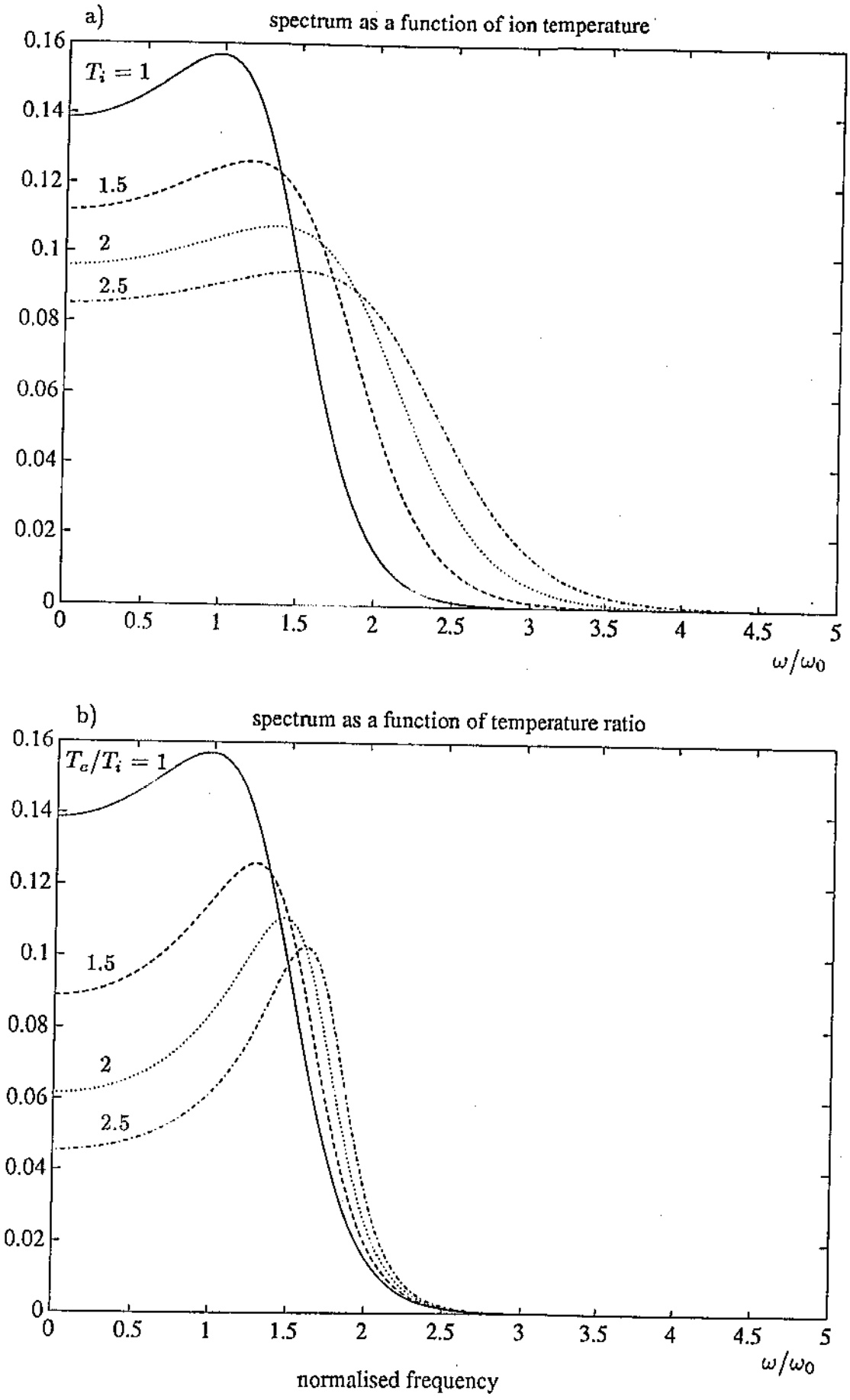
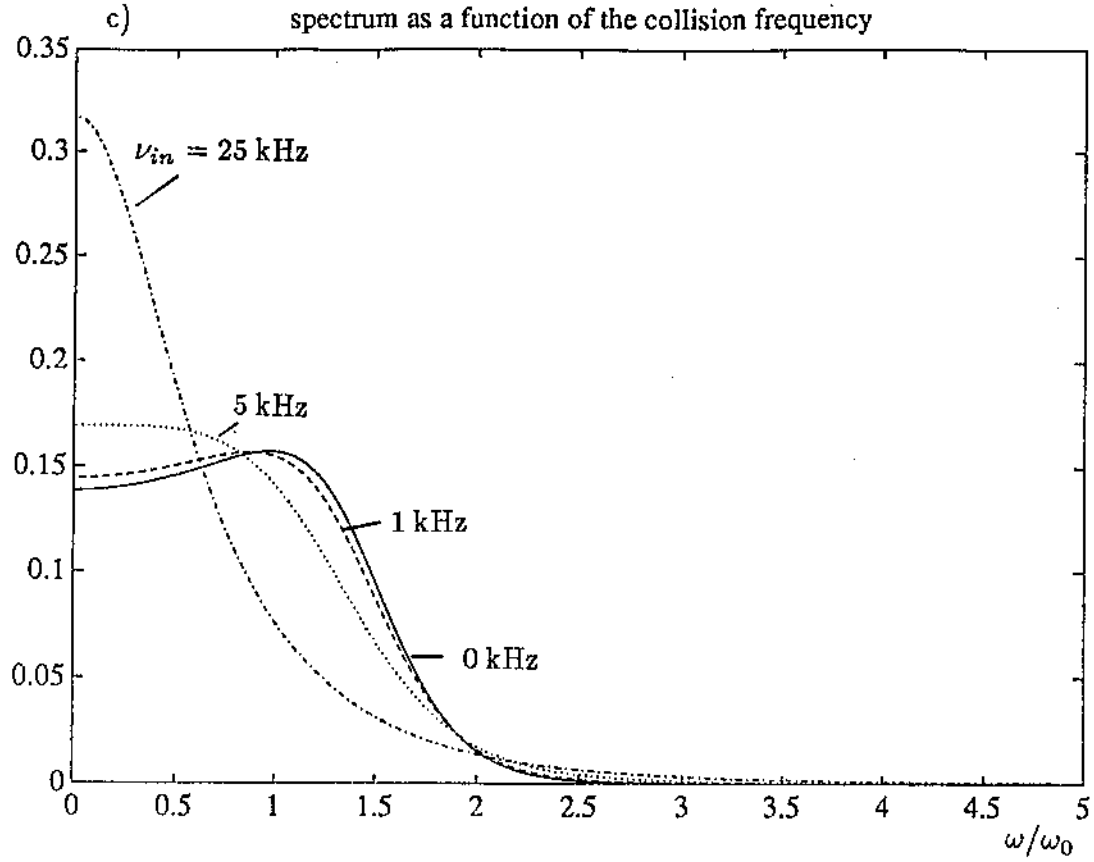


Fig. 2. The normalised spectrum as a function of a) the ion temperature T_i and b) the temperature ratio T_e/T_i. T_i is given in units of $T_0 = 300$ K.



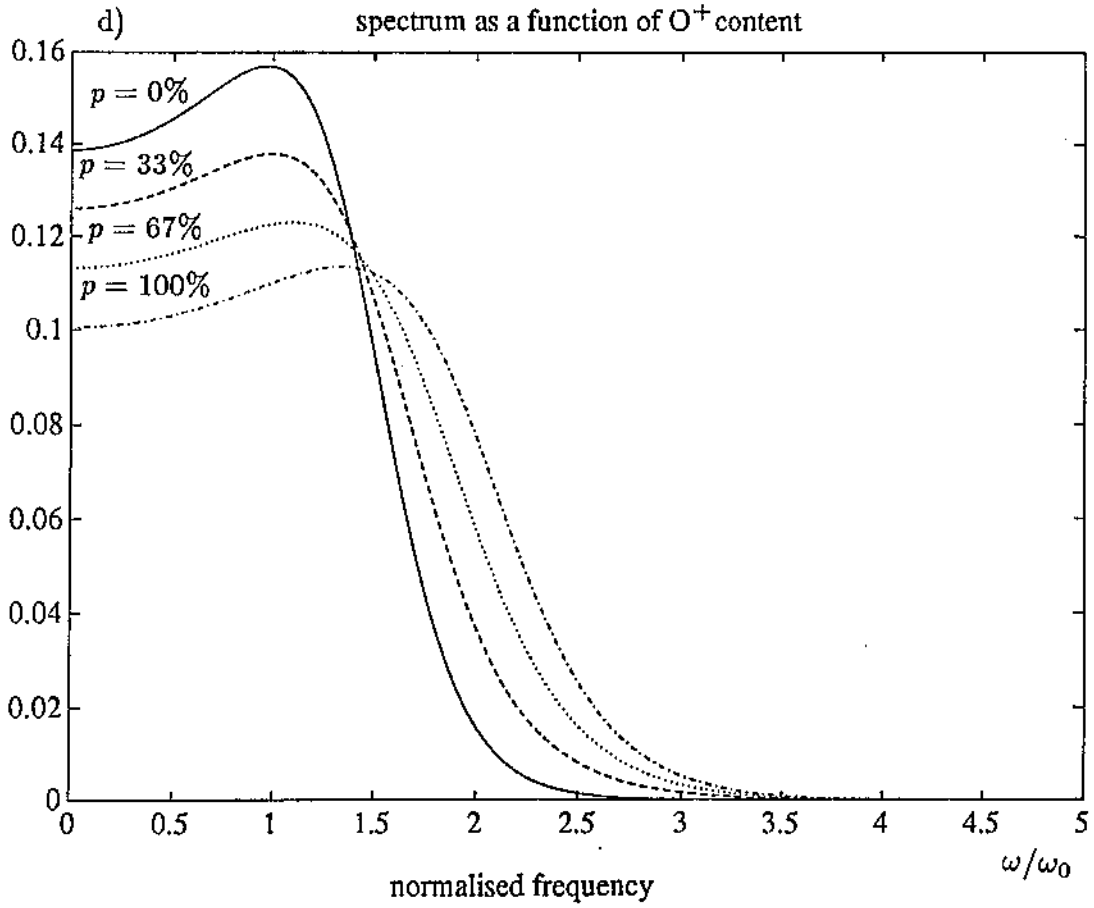


Fig. 2. The normalised spectrum as a function of c) the ion-neutral collision frequency ν_{in} and d) of the O^+ content $p(O^+)$.

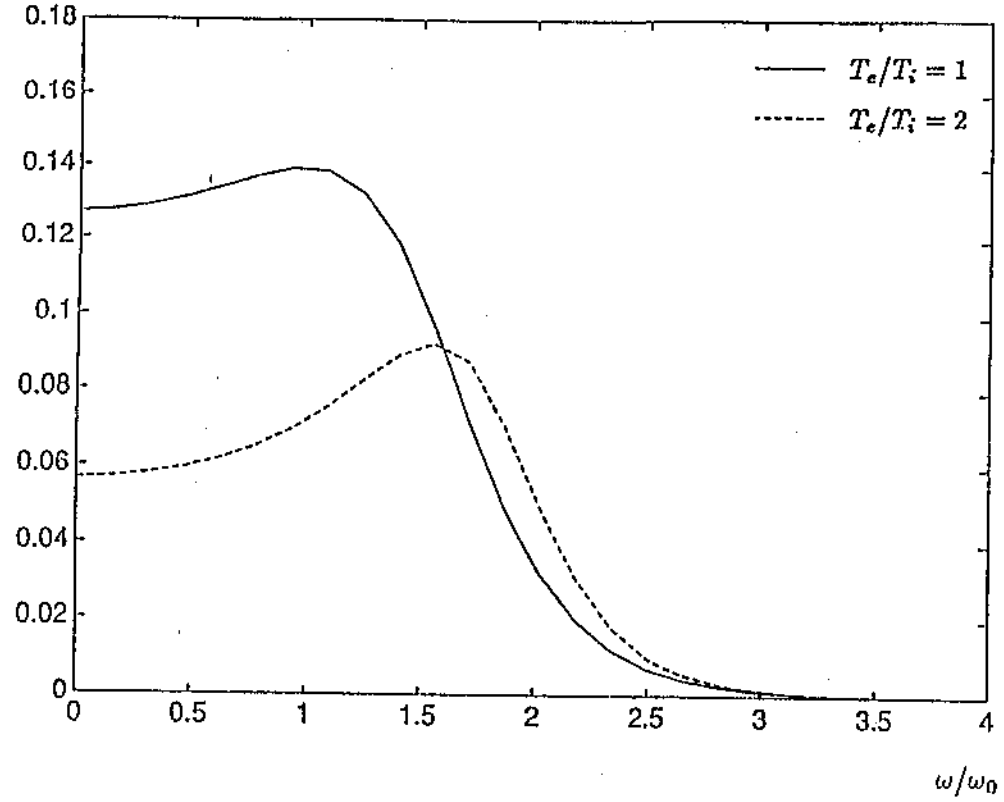


Fig. 3. The positive parts of the spectra.

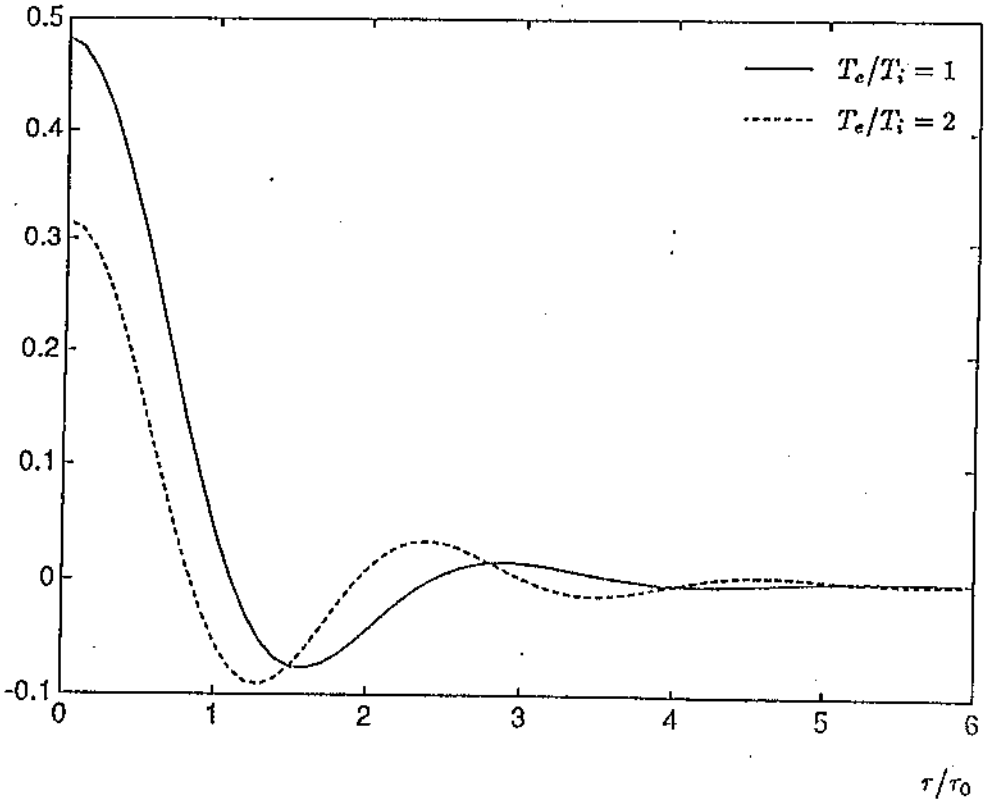


Fig. 4. The autocorrelation functions of the plasma with the same parameters as in Fig. 3.

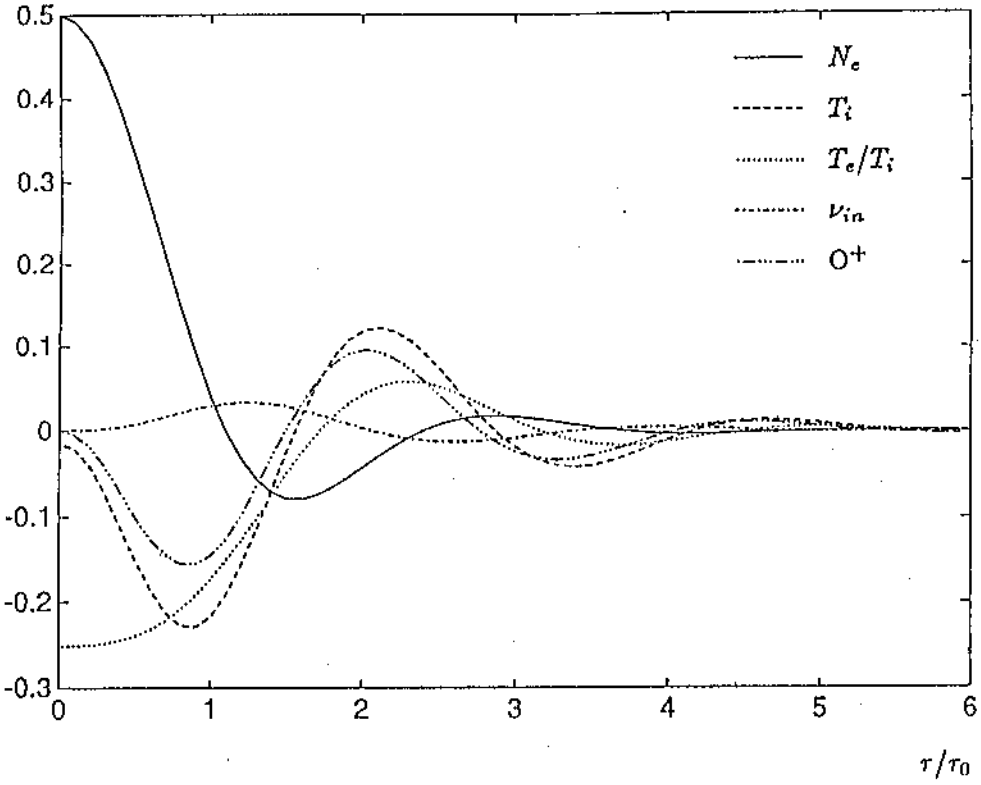


Fig. 5a. The partial derivatives of the ACF with respect to the plasma parameters for $T_e/T_i = 1$.

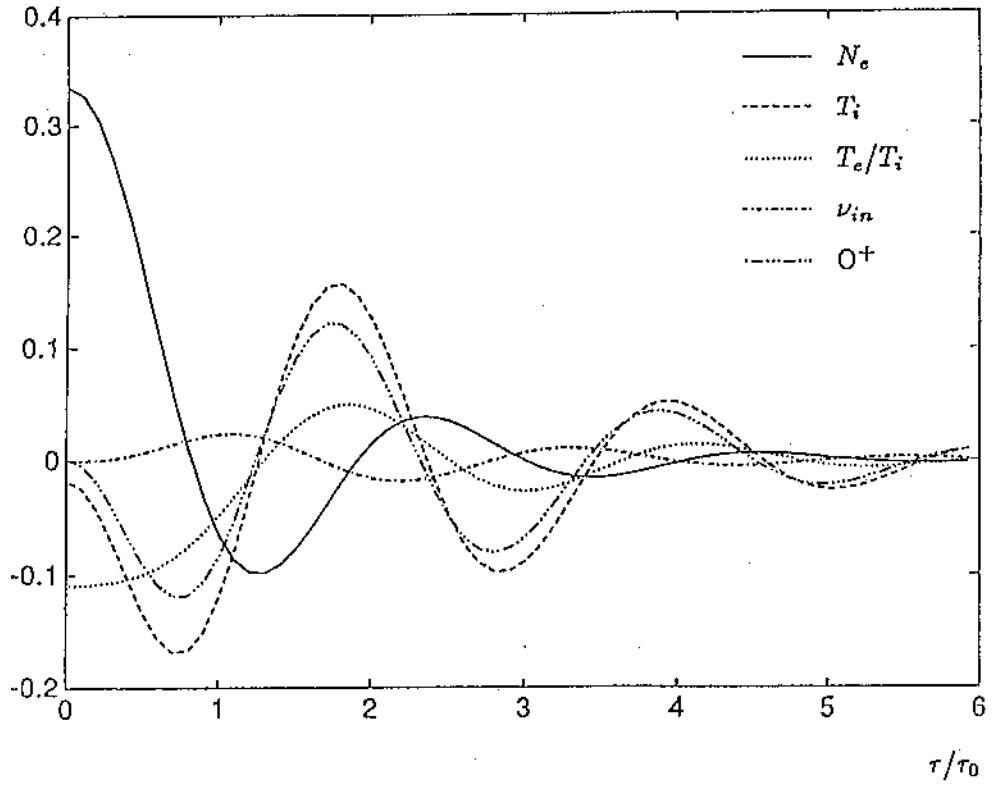


Fig. 5b. The partial derivatives of the ACF with respect to the plasma parameters for $T_e/T_i = 2$.

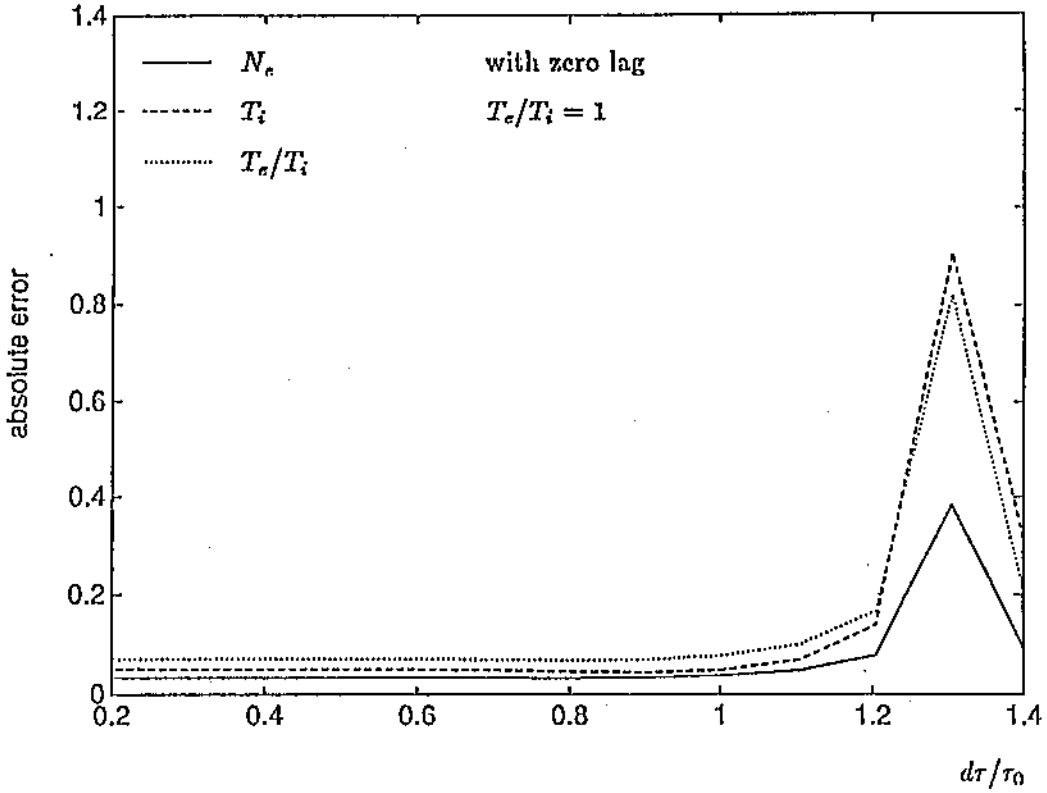


Fig. 6a. The behaviour of errors in 3-parameter fits as a function of the sampling interval for $T_e/T_i = 1$. The zero lag data is used in the inversion.

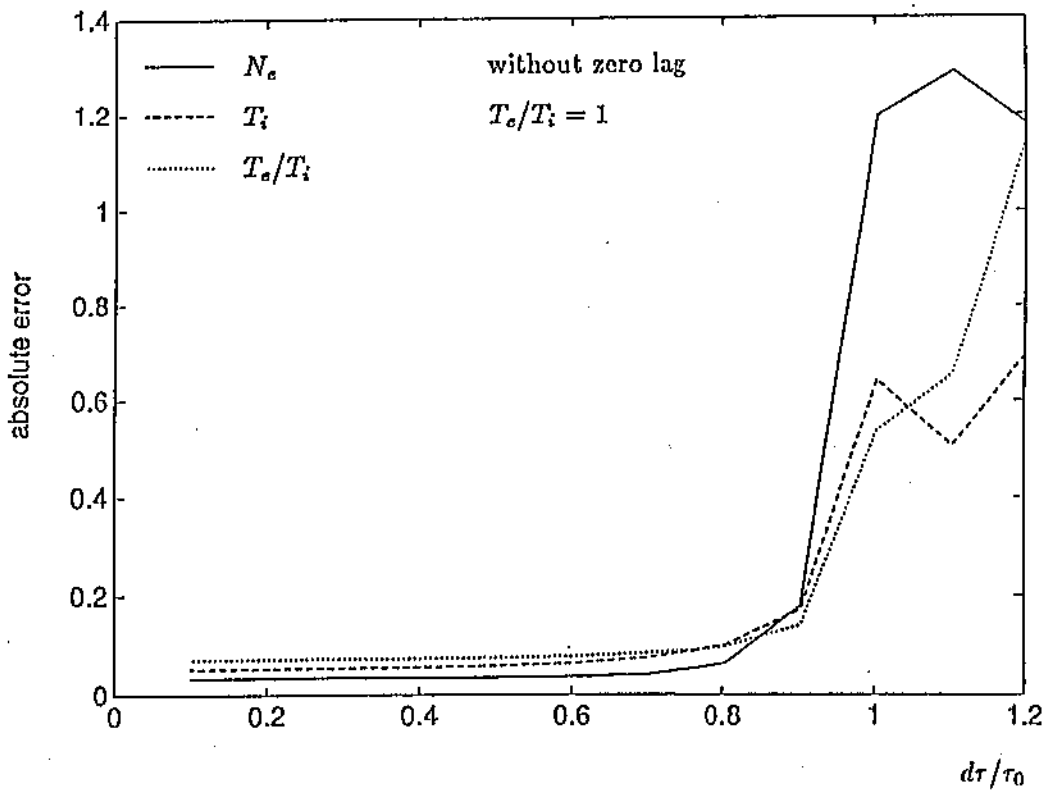


Fig. 6b. The behaviour of errors in 3-parameter fits as a function of the sampling interval for $T_e/T_i = 1$. The zero lag data is not used in the inversion.

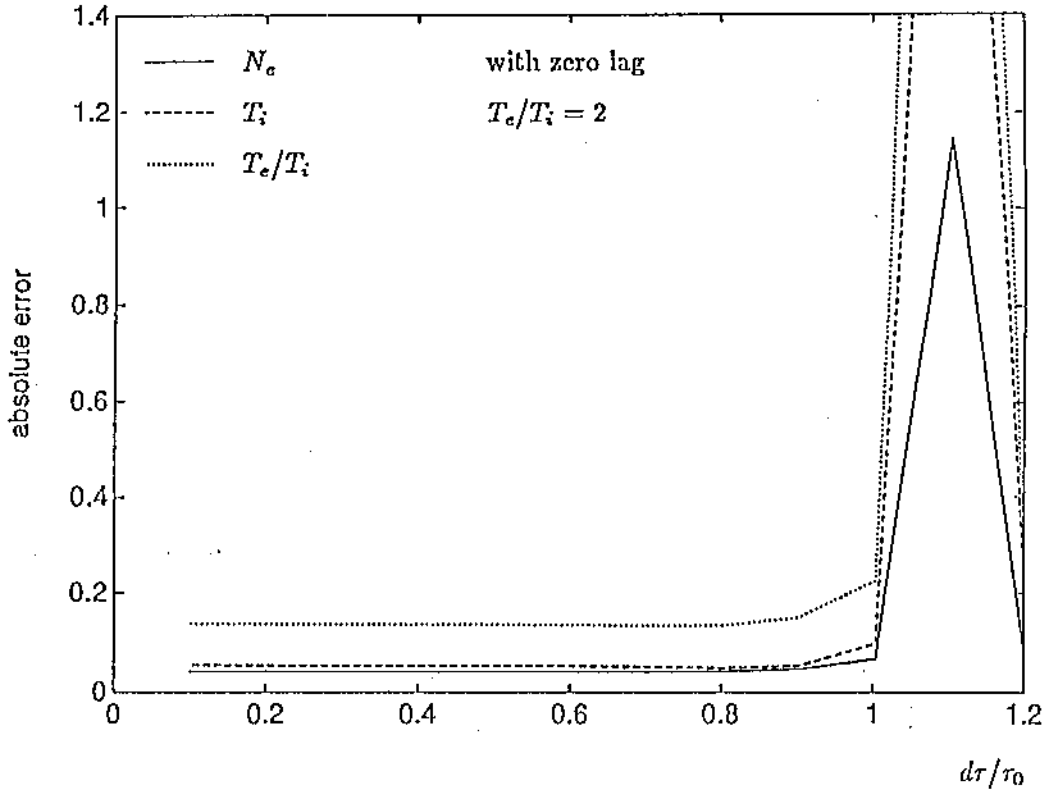


Fig. 7a. The behaviour of errors in 3-parameter fits as a function of the sampling interval for $T_e/T_i = 2$. The zero lag data is used in the inversion.

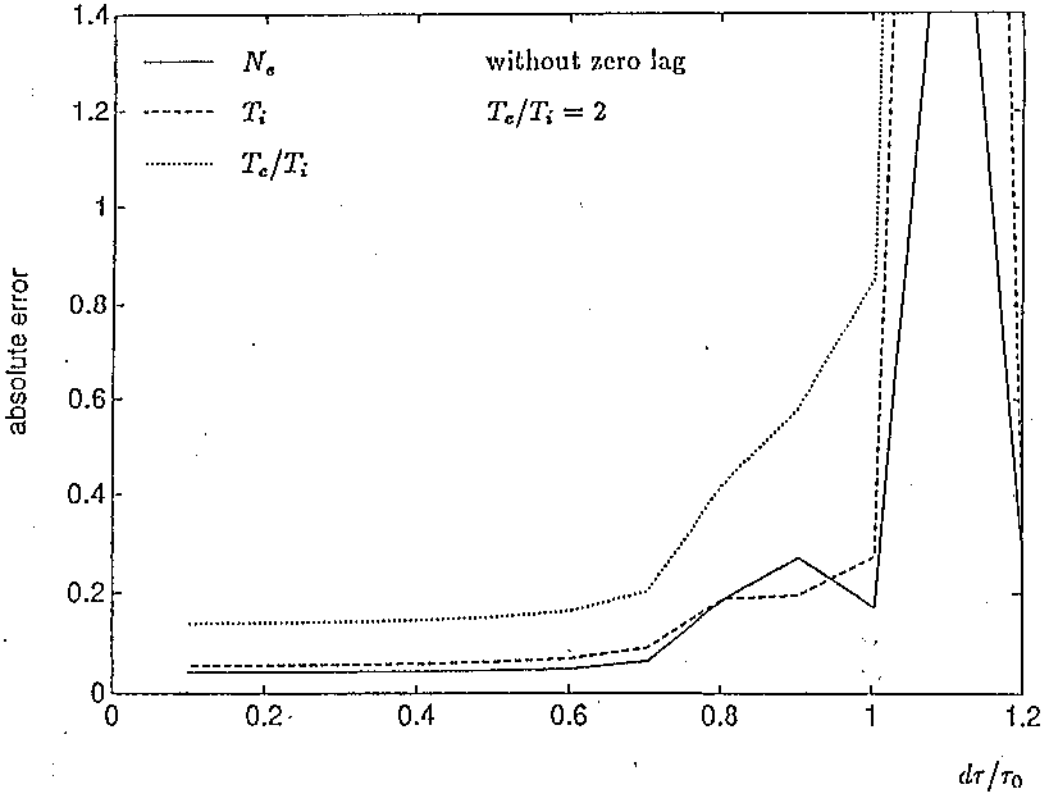


Fig. 7b. The behaviour of errors in 3-parameter fits as a function of the sampling interval for $T_e/T_i = 2$. The zero lag data is not used in the inversion.

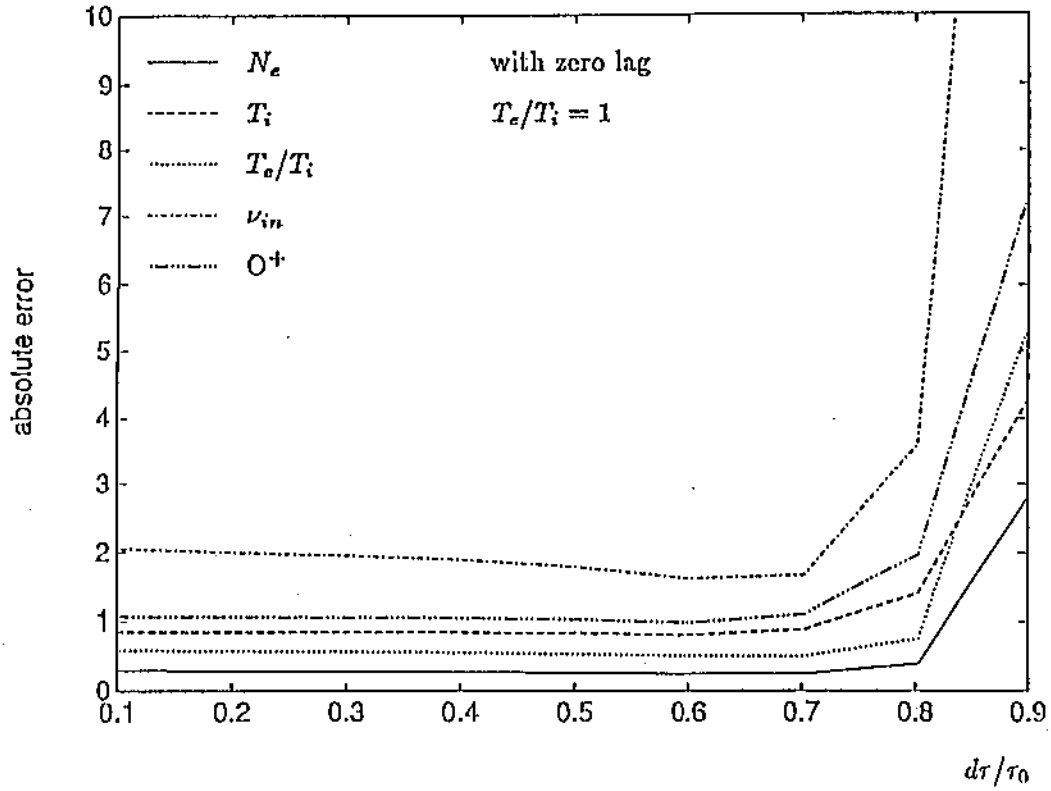


Fig. 8a. The behaviour of errors in 5-parameter fits as a function of the sampling interval for $T_e/T_i = 1$. The zero lag data is used in the inversion.

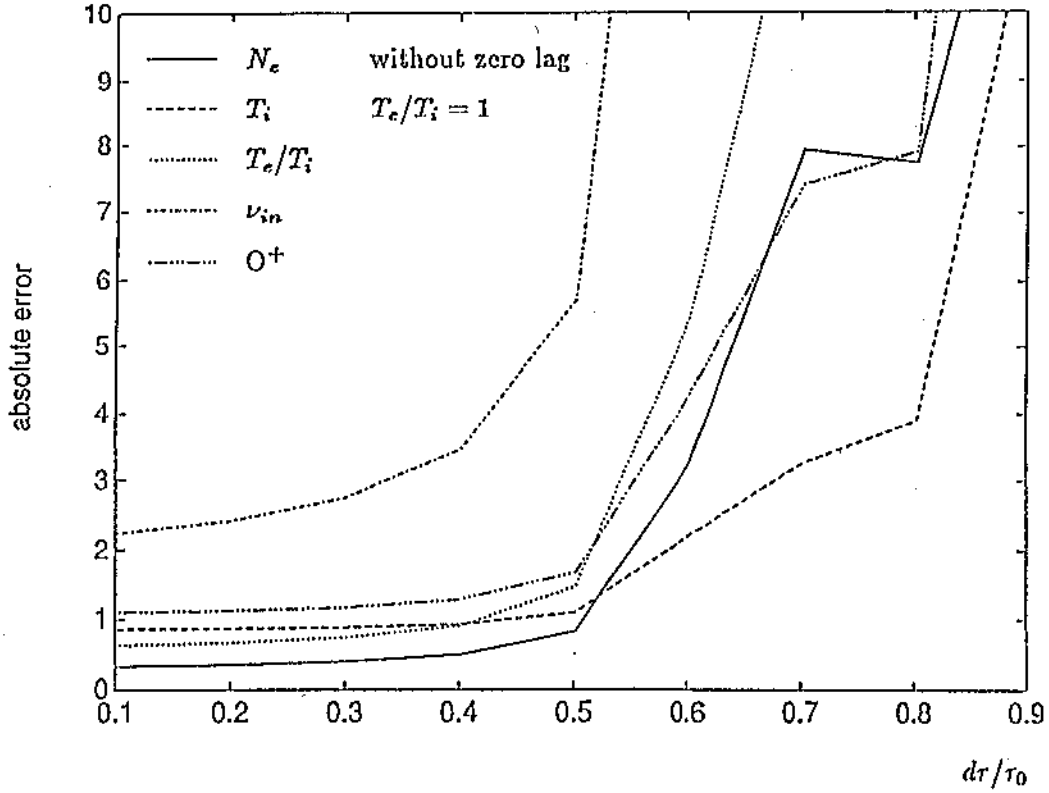


Fig. 8b. The behaviour of errors in 5-parameter fits as a function of the sampling interval for $T_e/T_i = 1$. The zero lag data is not used in the inversion.

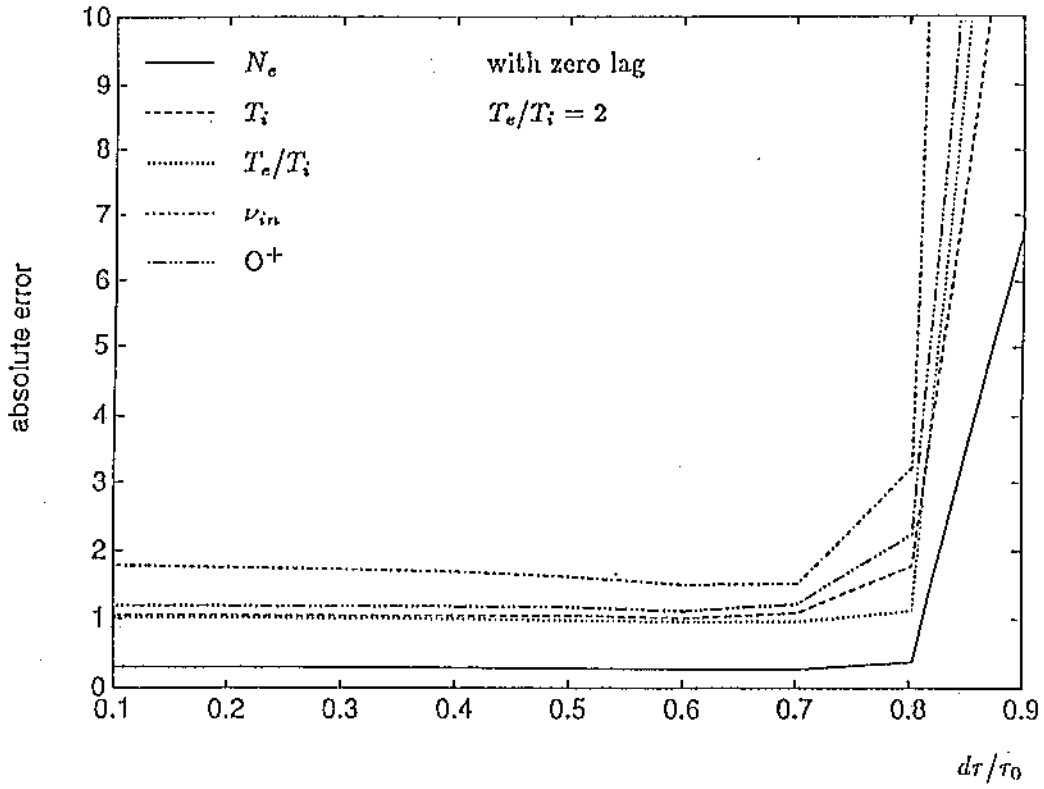


Fig. 9a. The behaviour of errors in 5-parameter fits as a function of the sampling interval for $T_e/T_i = 2$. The zero lag data is used in the inversion.

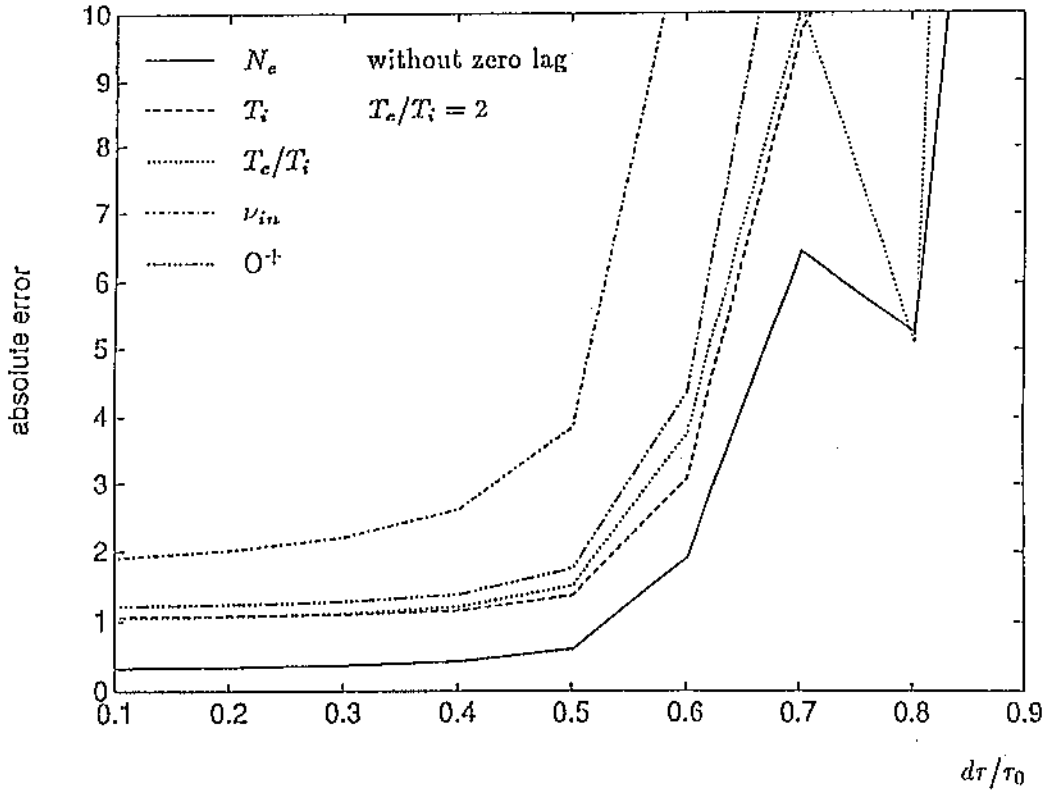


Fig. 9b. The behaviour of errors in 5-parameter fits as a function of the sampling interval for $T_e/T_i = 2$. The zero lag data is not used in the inversion.

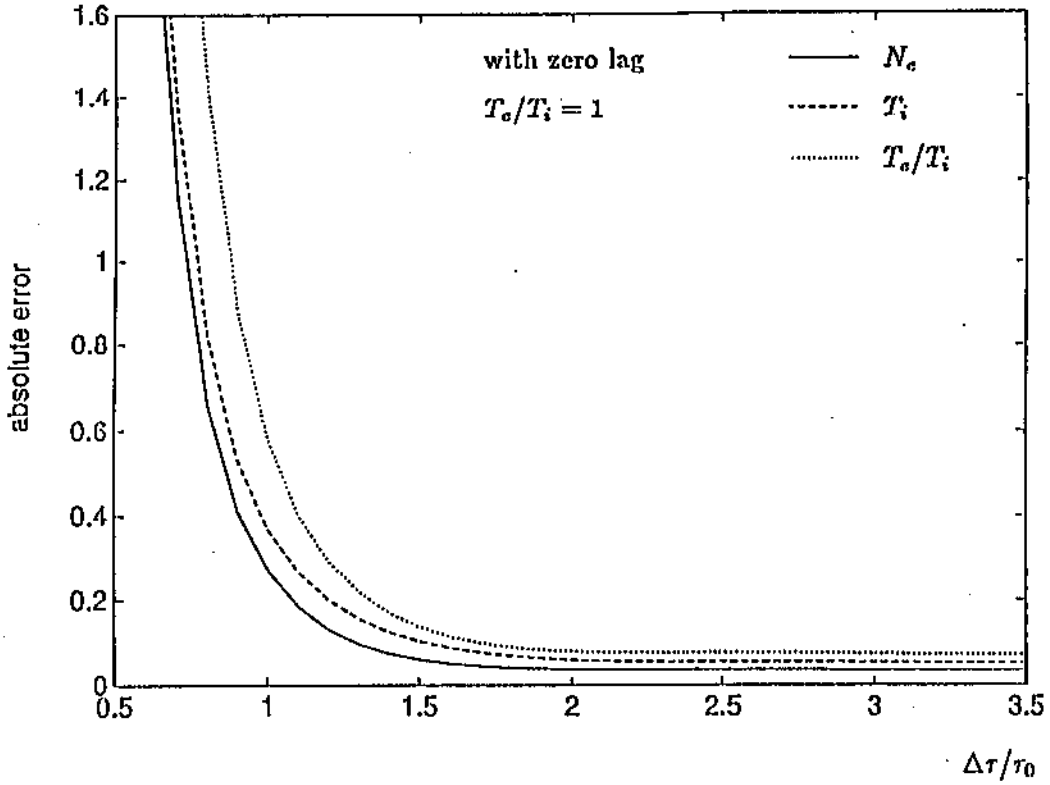


Fig.10a. The behaviour of errors in 3-parameter fits as a function of the longest lag used for $T_e/T_i = 1$.

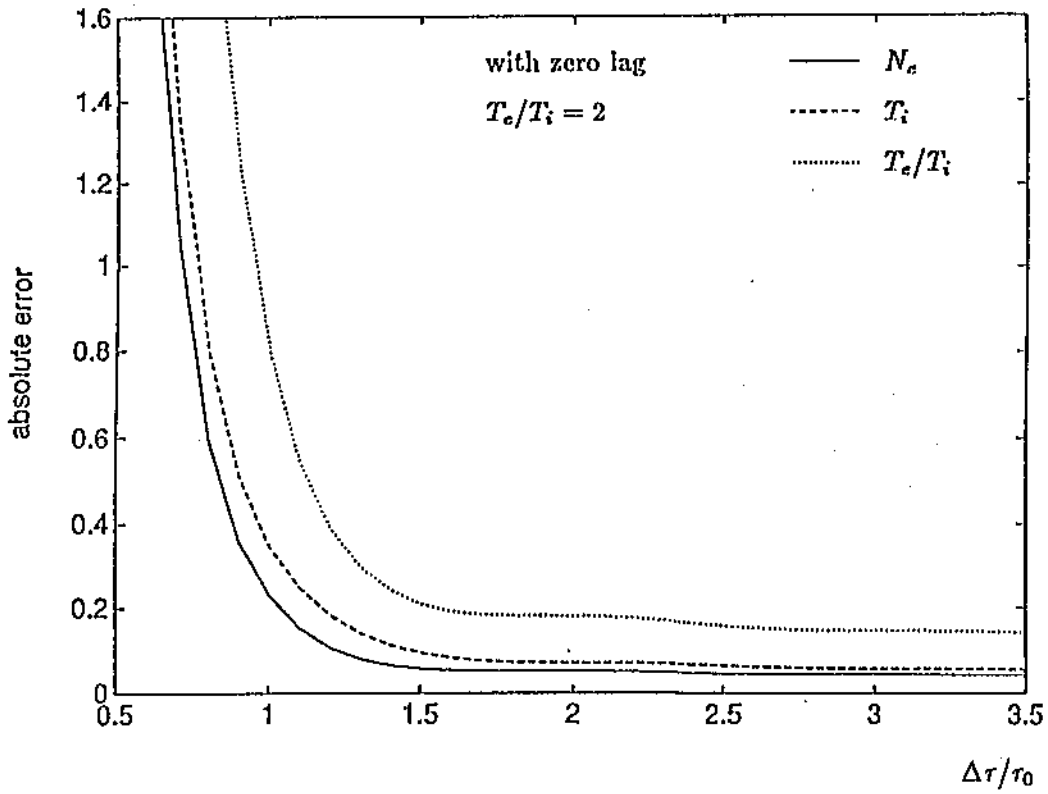


Fig.10b. The behaviour of errors in 3-parameter fits as a function of the longest lag used for $T_e/T_i = 2$.

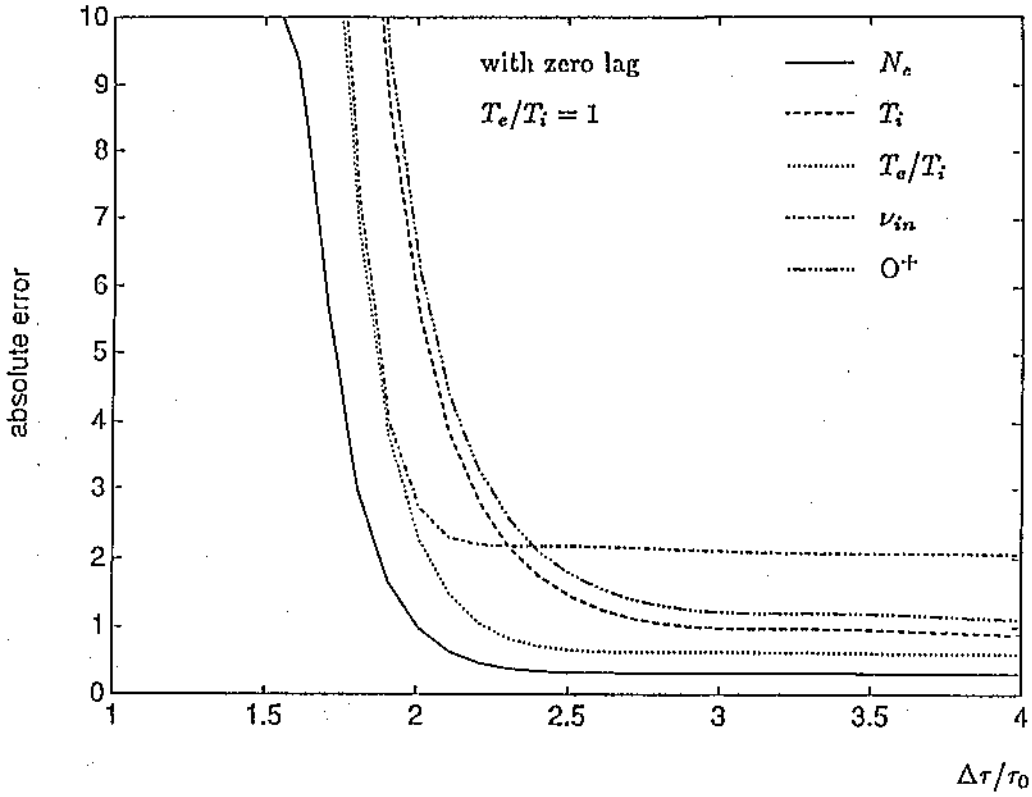


Fig.11a. The behaviour of errors in 5-parameter fits as a function of the longest lag used for $T_e/T_i = 1$.

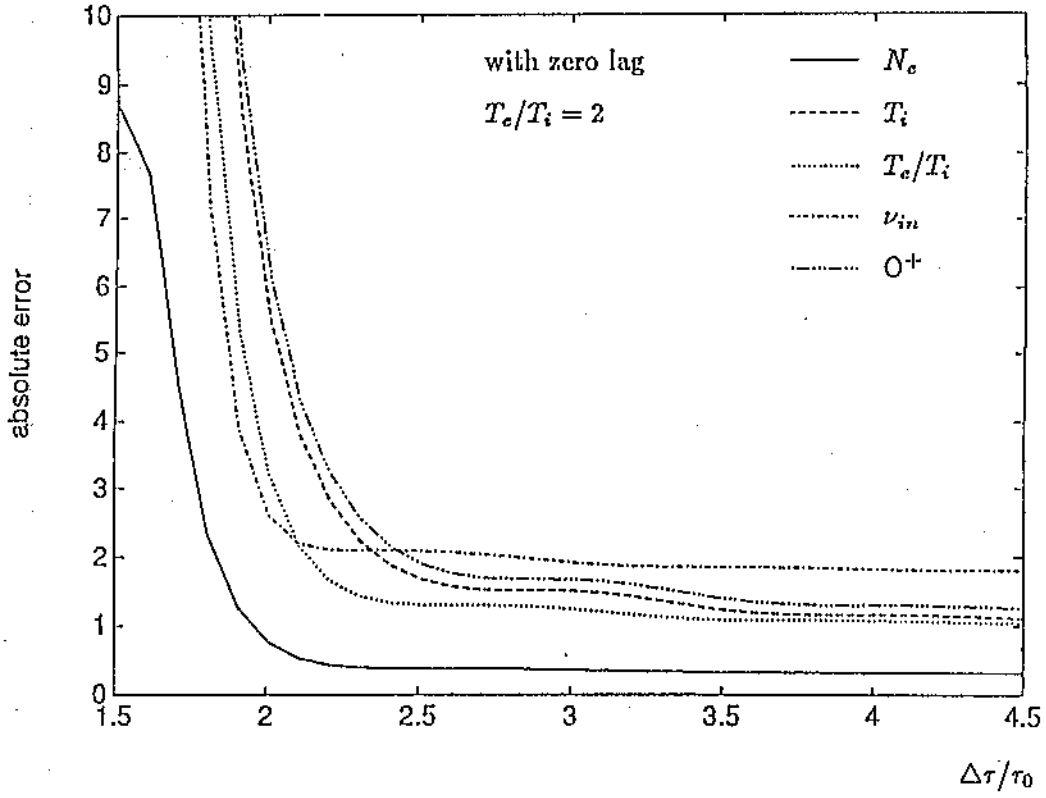


Fig.11b. The behaviour of errors in 5-parameter fits as a function of the longest lag used for $T_e/T_i = 2$.

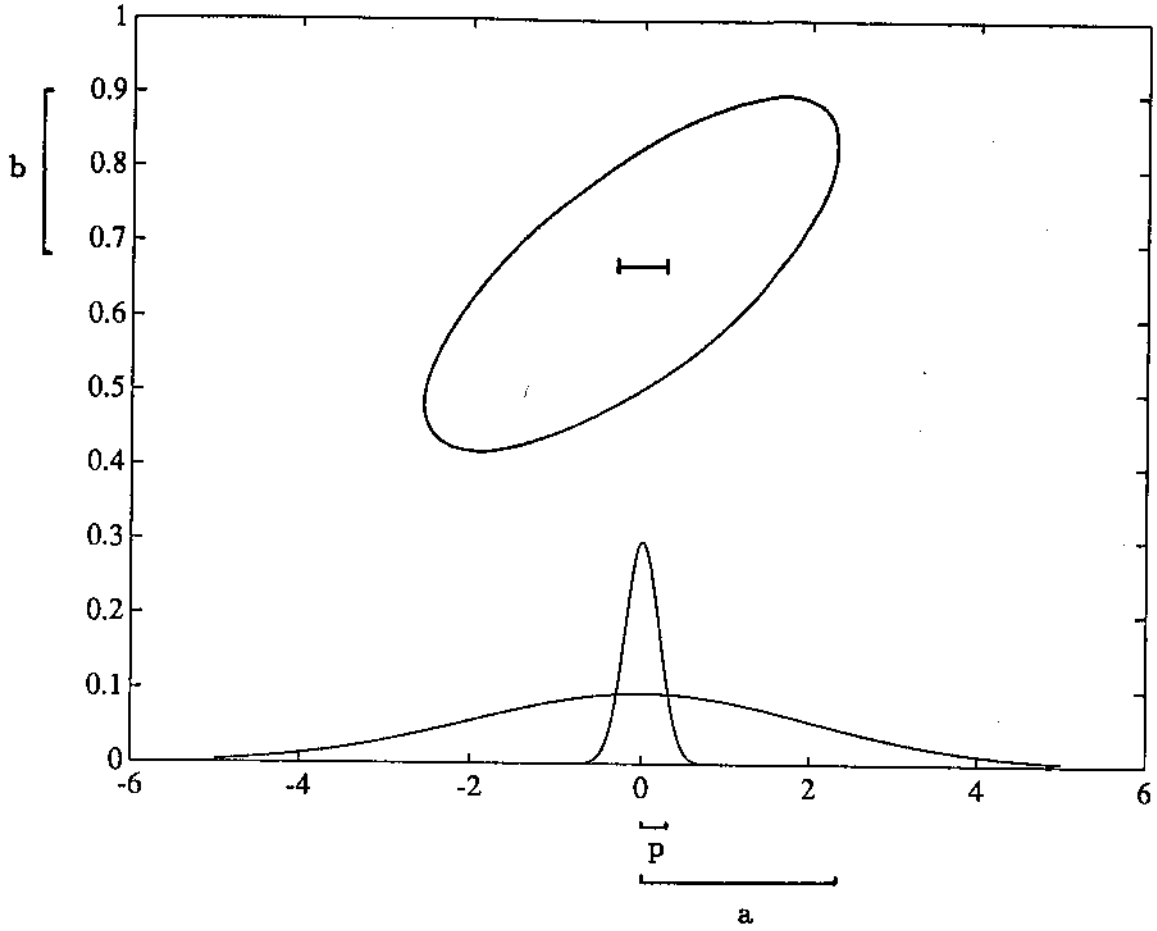


Fig. 12. The bare a posteriori and a priori distributions.

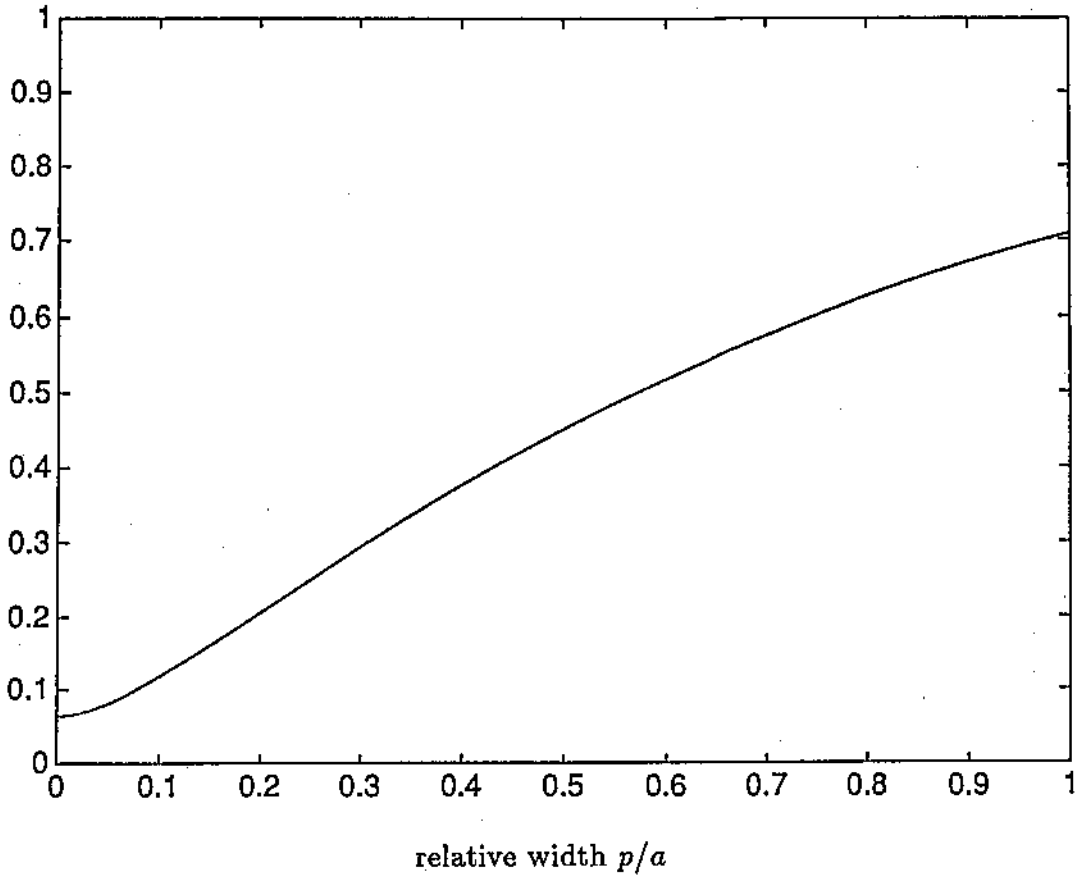


Fig. 13. The behaviour of the a posteriori error in O^+ content as a function of the a priori accuracy on T_i/T_0.

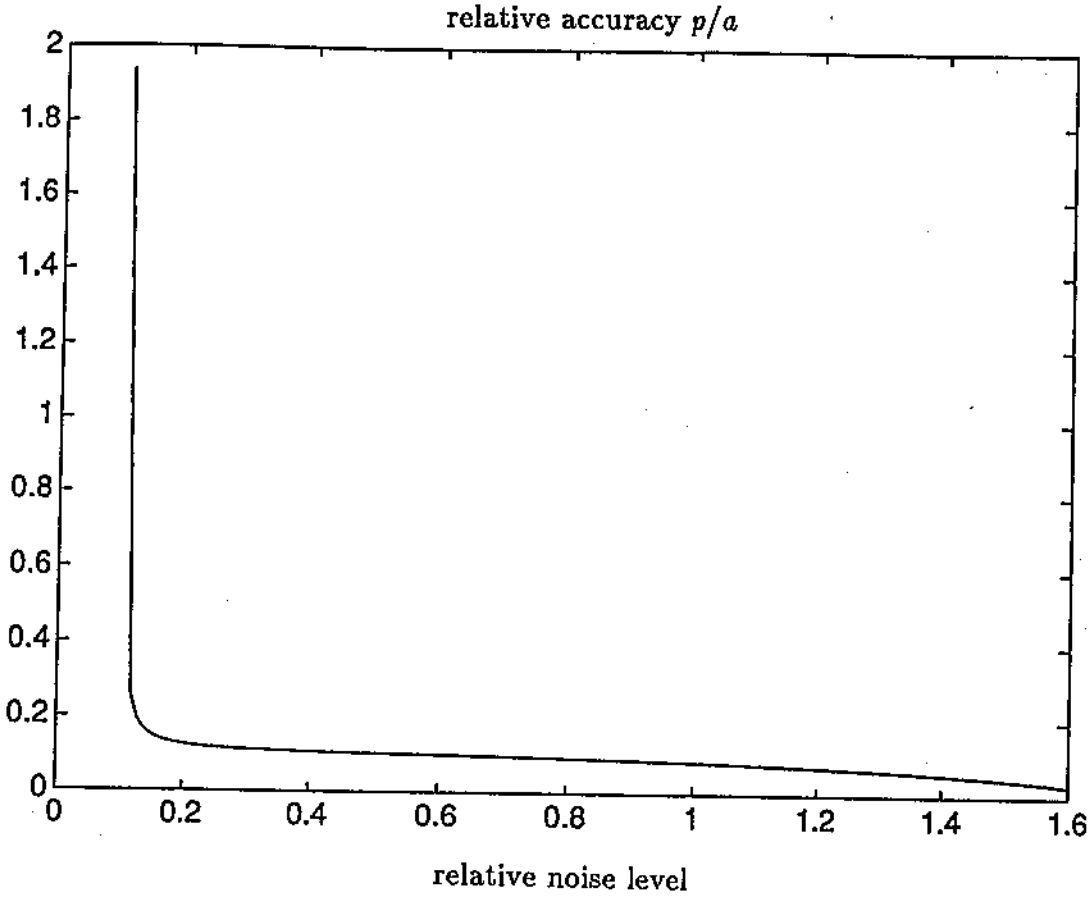
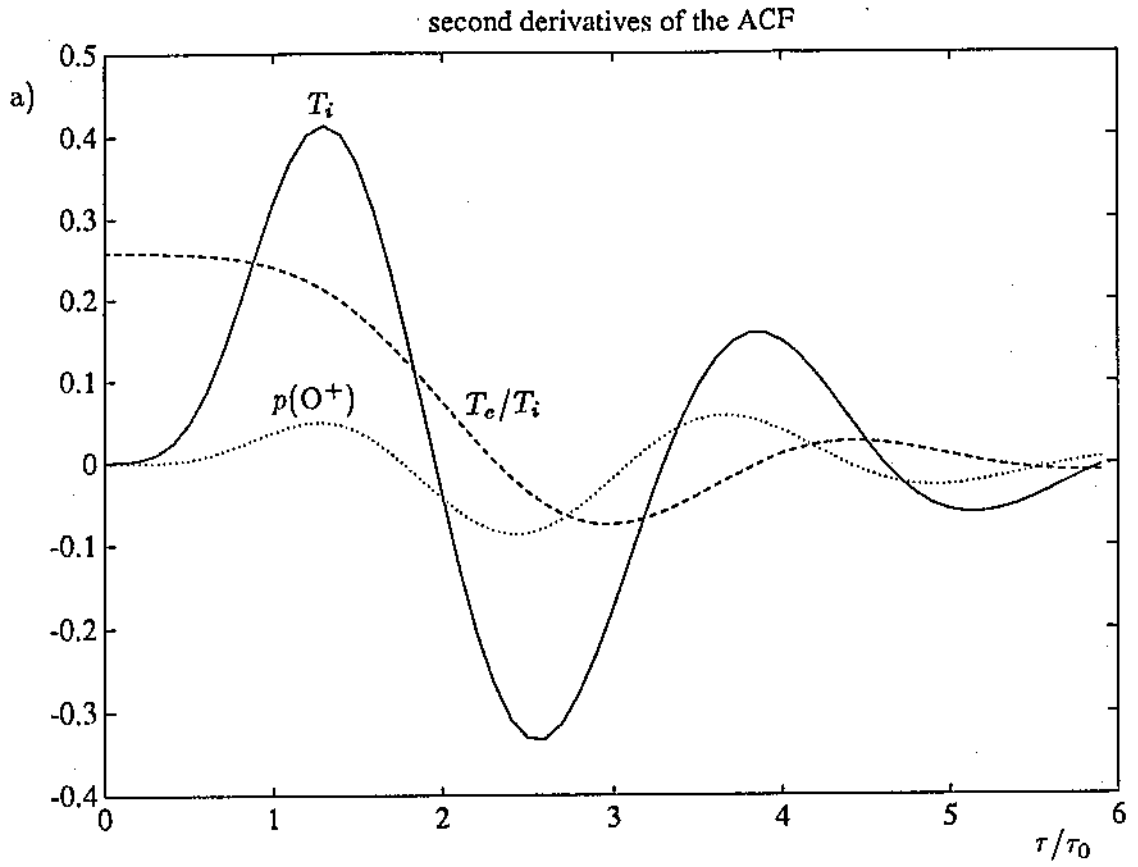


Fig. 14. The behaviour of the a priori accuracy in the first parameter required to obtain a given accuracy in the second parameter as a function of the noise level γ/γ_0.



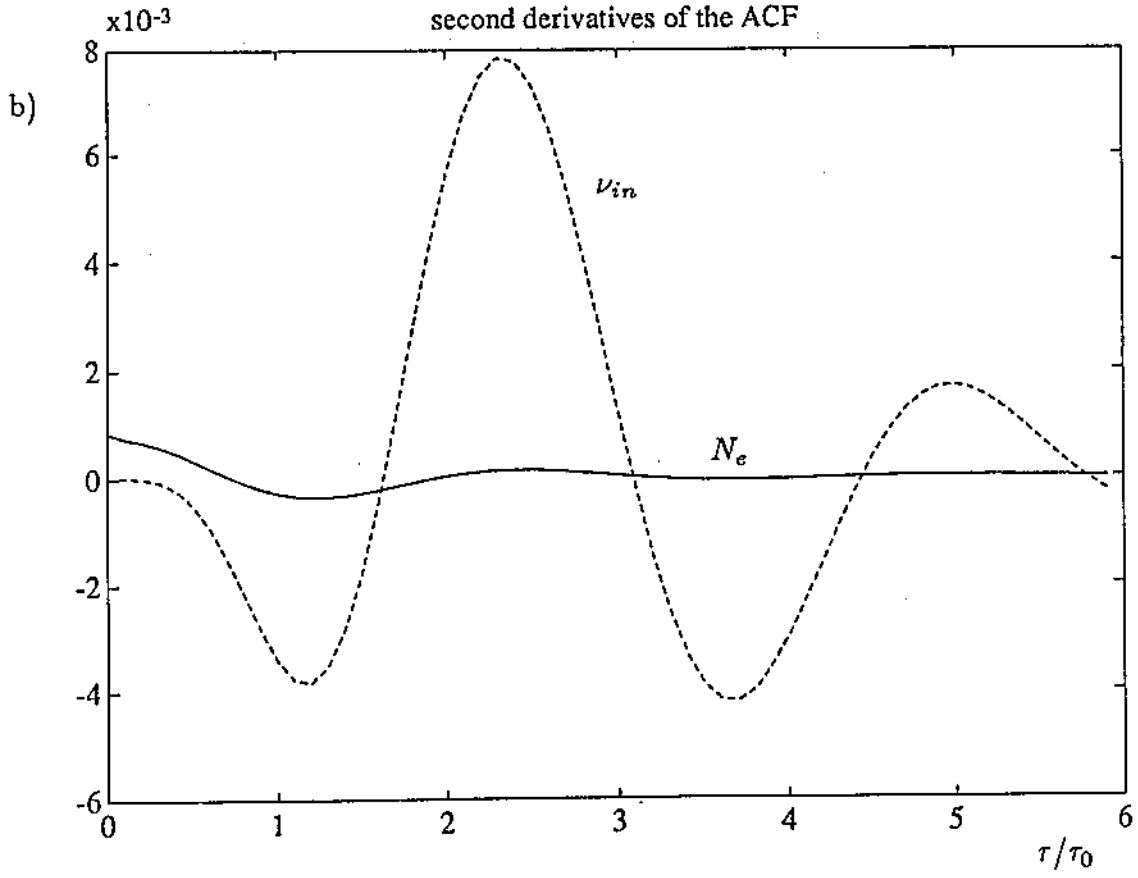


Fig. 15. The second derivatives of the plasma autocorrelation function with respect to the plasma parameters.

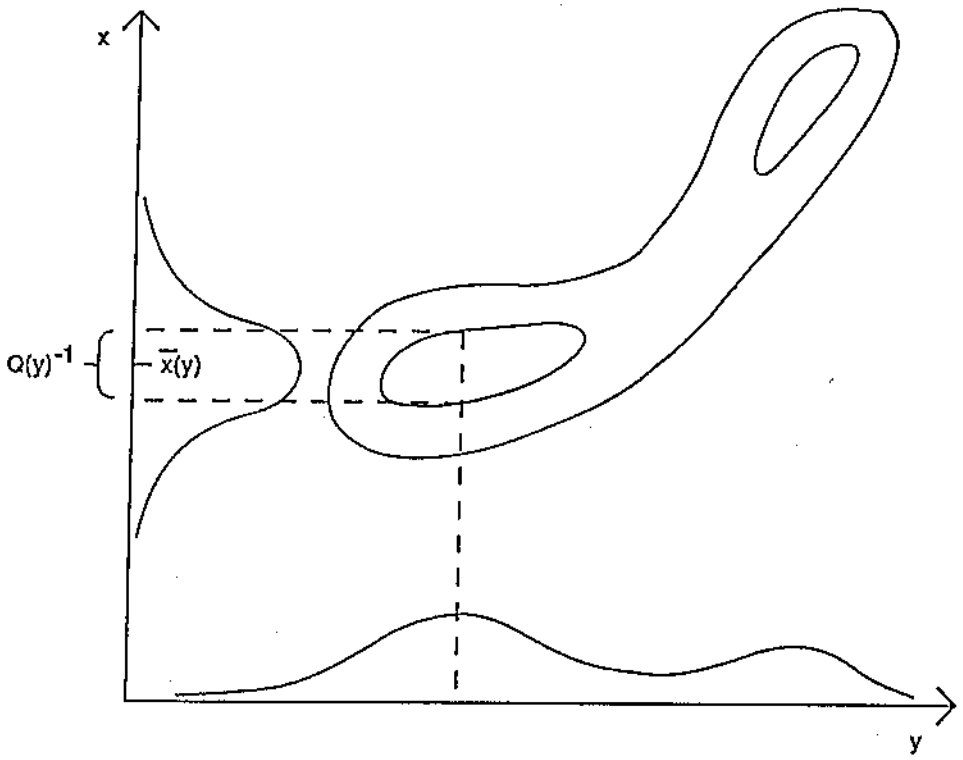


Fig. 16. The a posteriori distribution as a function of the nonlinear 'difficult' parameter y and of the linear 'ordinary' parameter x.

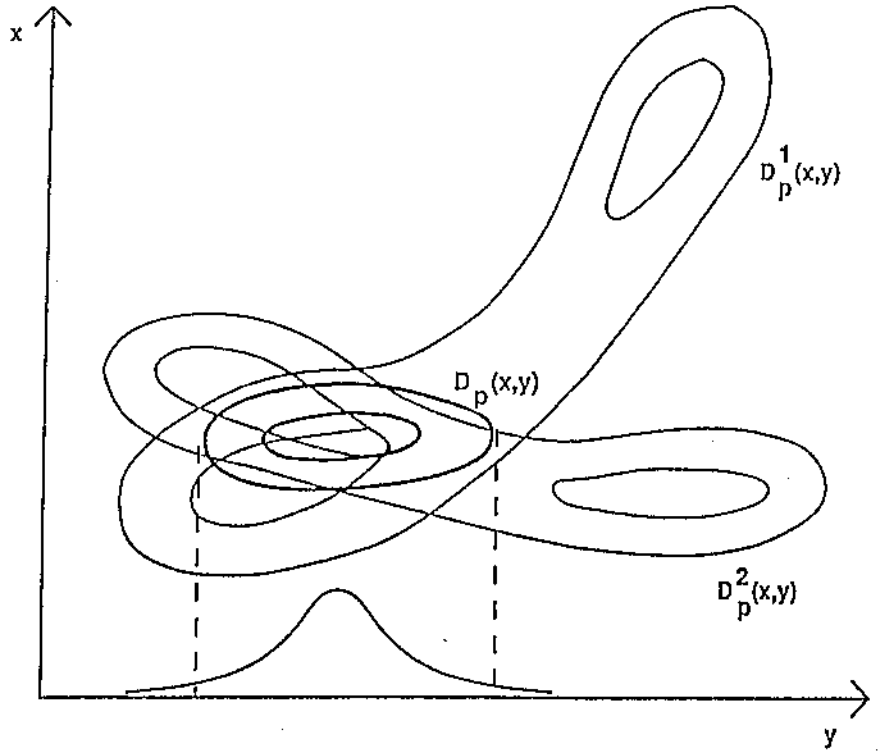


Fig. 17. The combination of two calculated a posteriori distributions into a single one to increase the resolution in the 'difficult' parameter y.

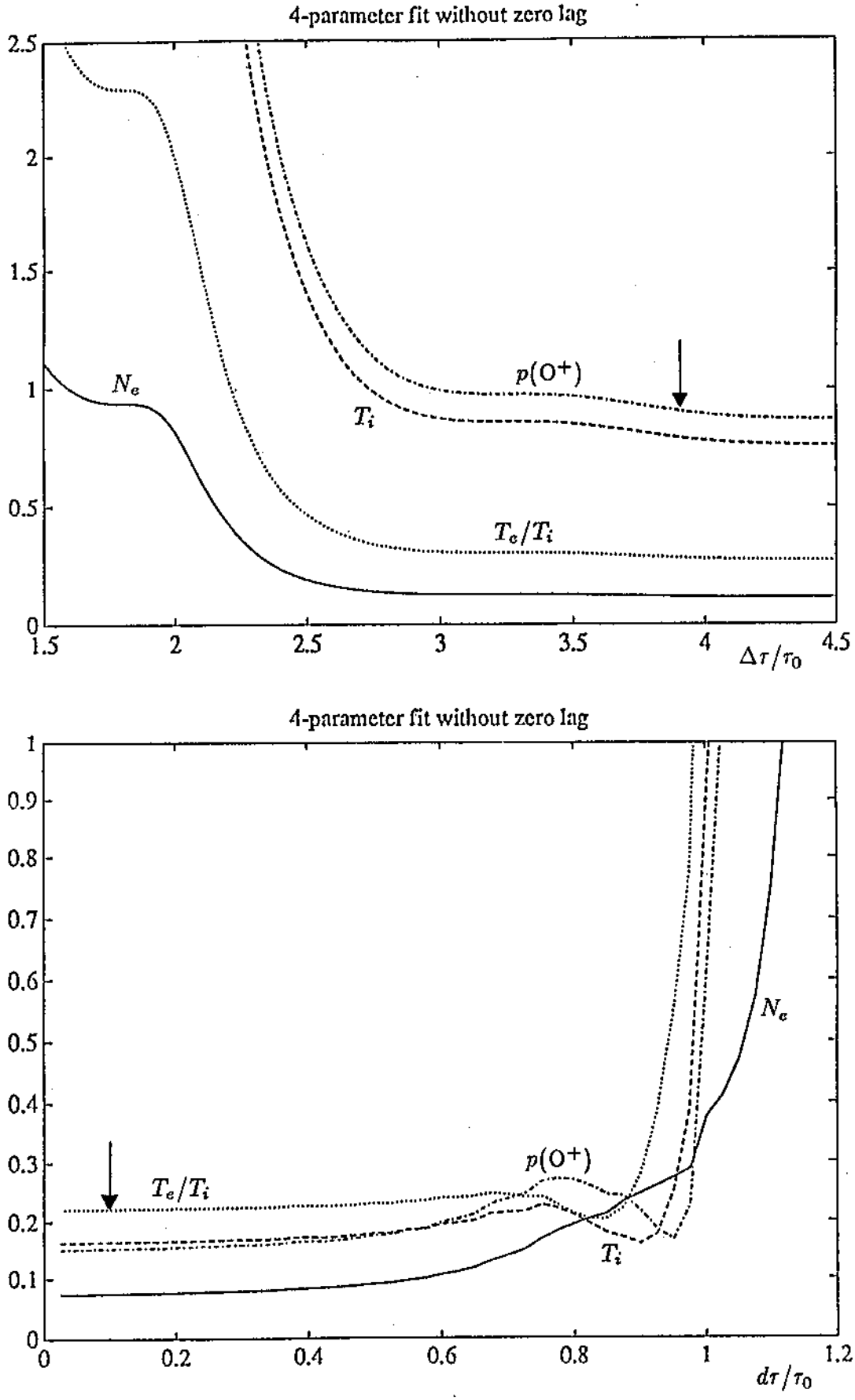
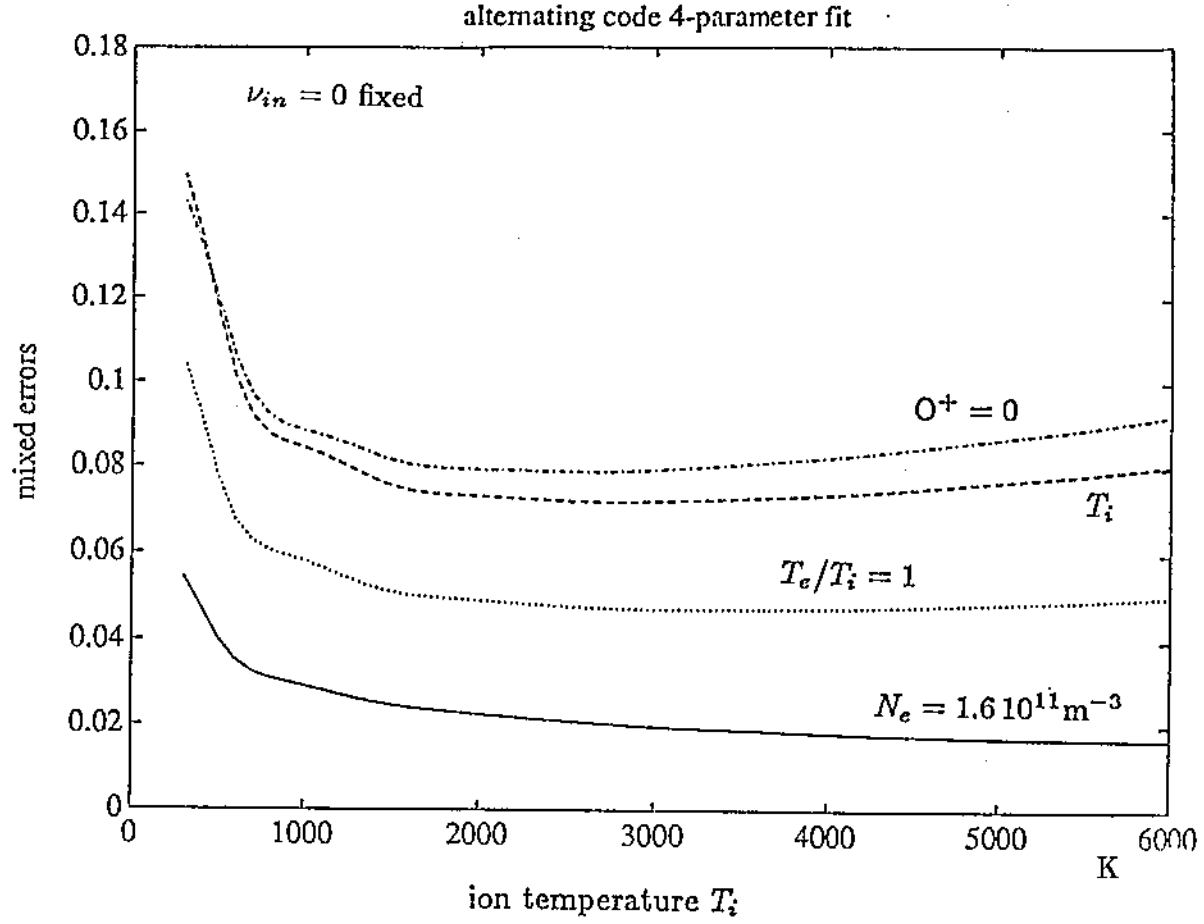


Fig. 18. The behaviour of expected errors in the 4–parameter fits of the alternating code experiments. The actual lag resolution and lag extent values are denoted by the arrows. The parameter values used in calculating the curves are $T_i = 300$ K, $T_e/T_i = 1$ and $p(O^+) = 0.3$ in the upper panel and $T_i = 300$ K, $T_e/T_i = 2$ and $p(O^+) = 0$ in the lower panel.



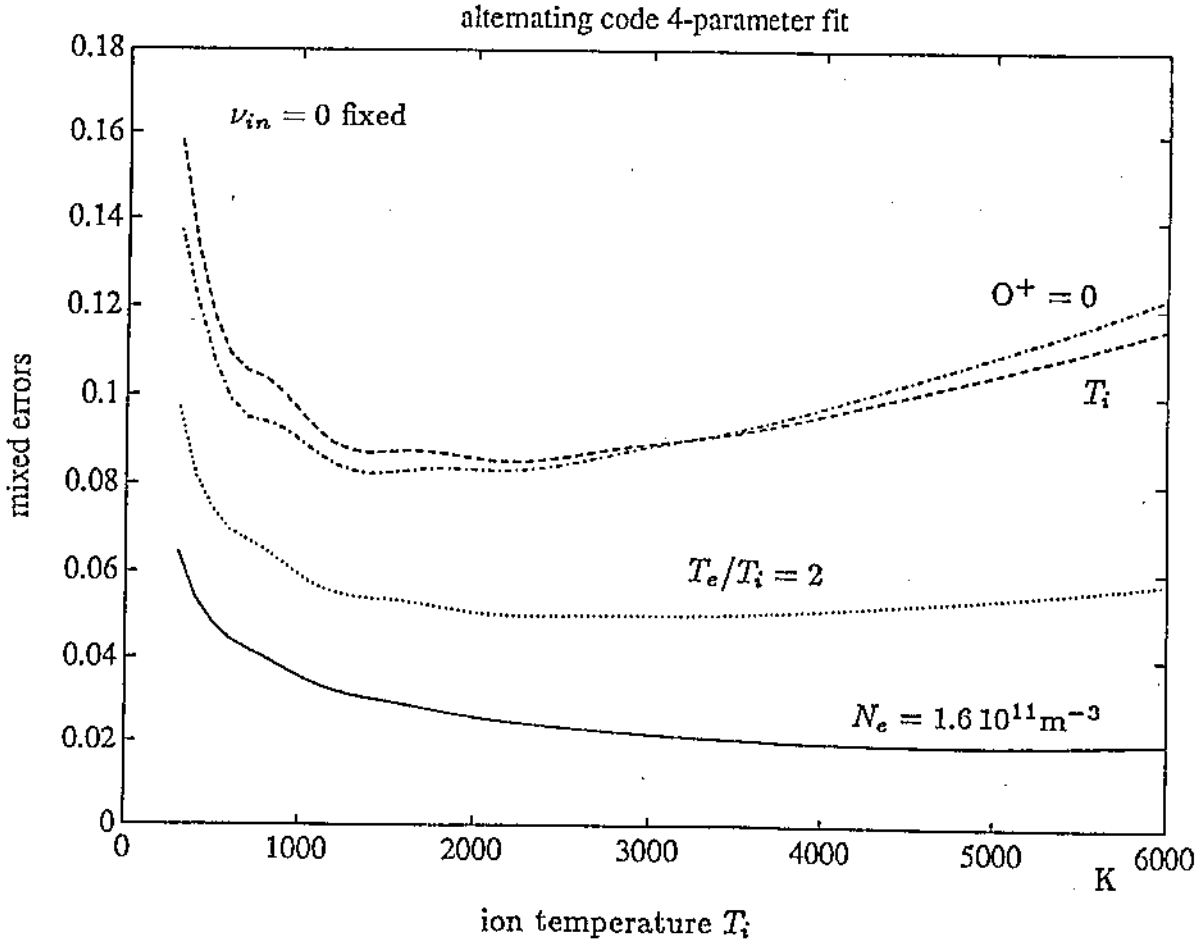
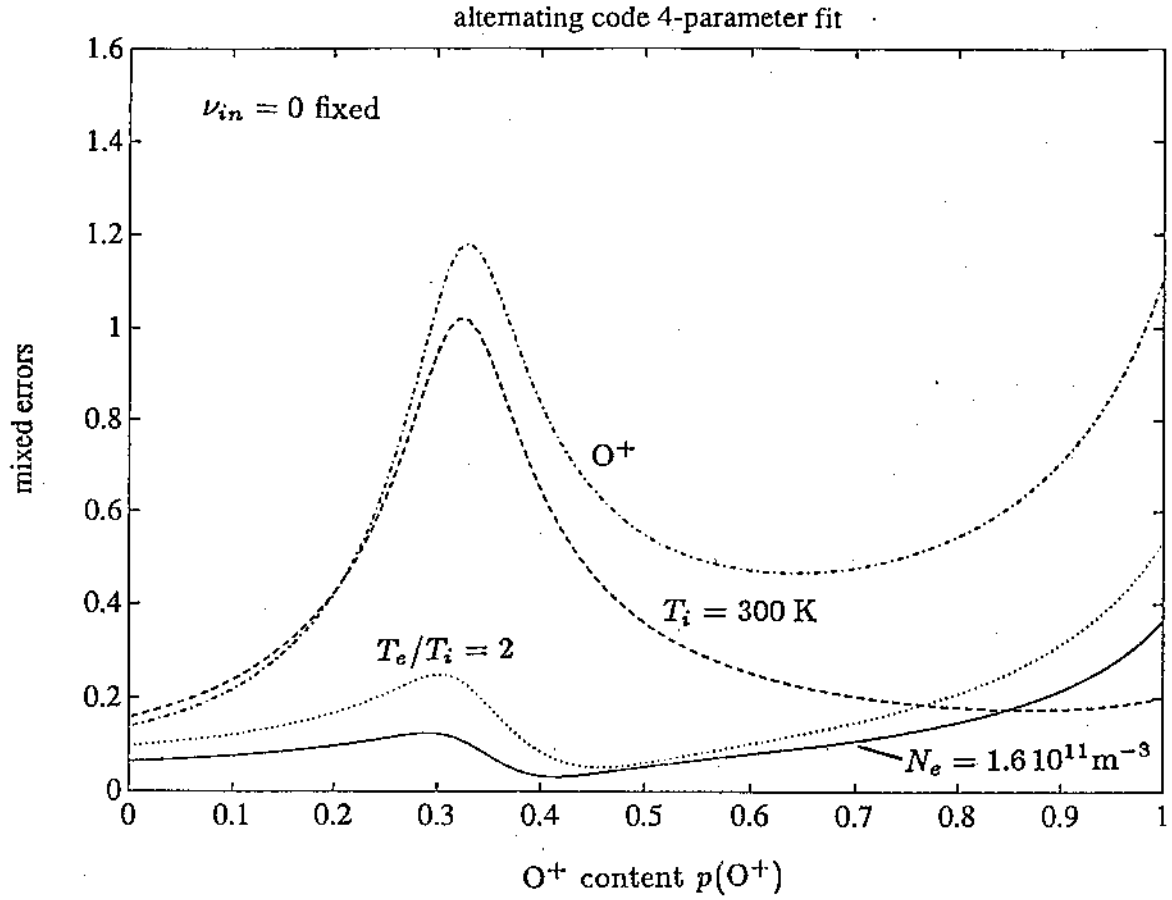


Fig. 19. The expected fit errors as a function of ion temperature T_i in the alternating code experiments.



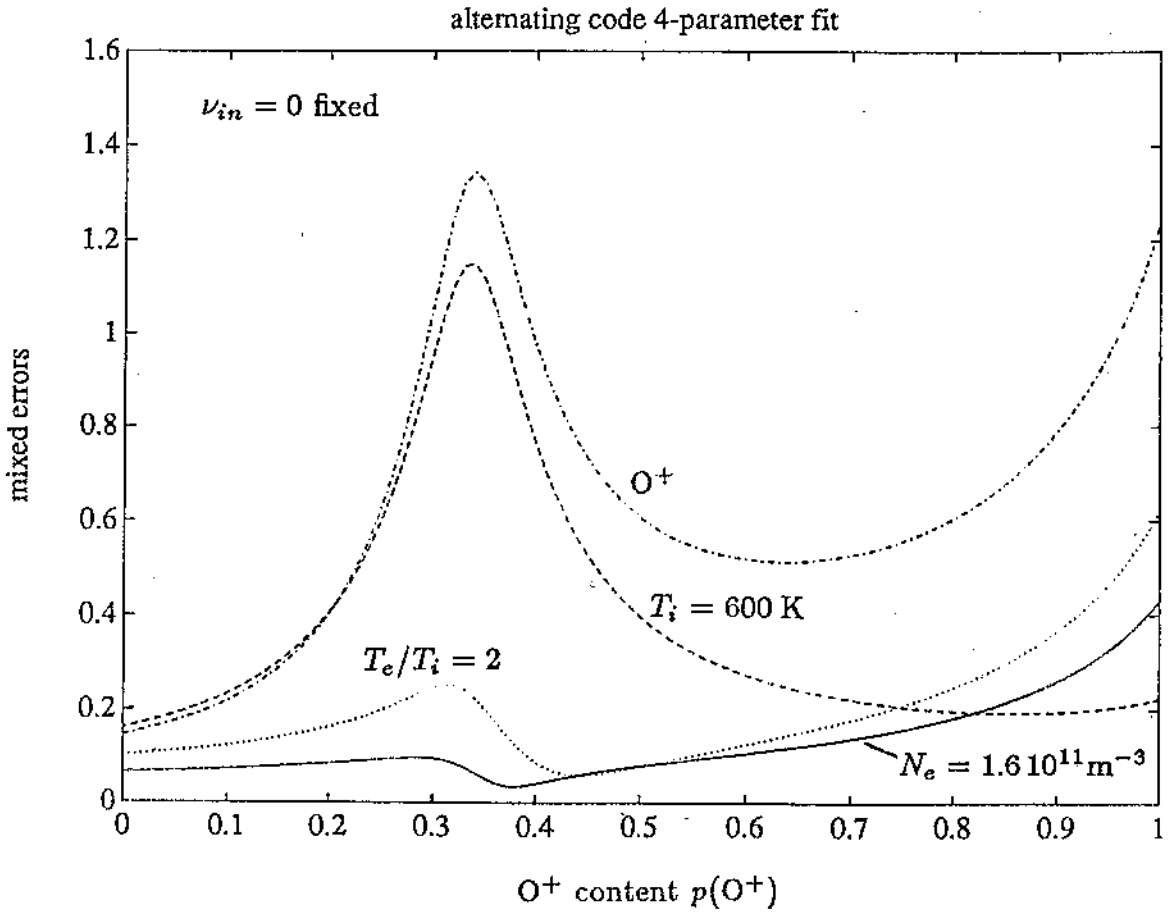


Fig. 20. The fit errors as a function of composition $p(O^+)$ in the alternating code experiments.

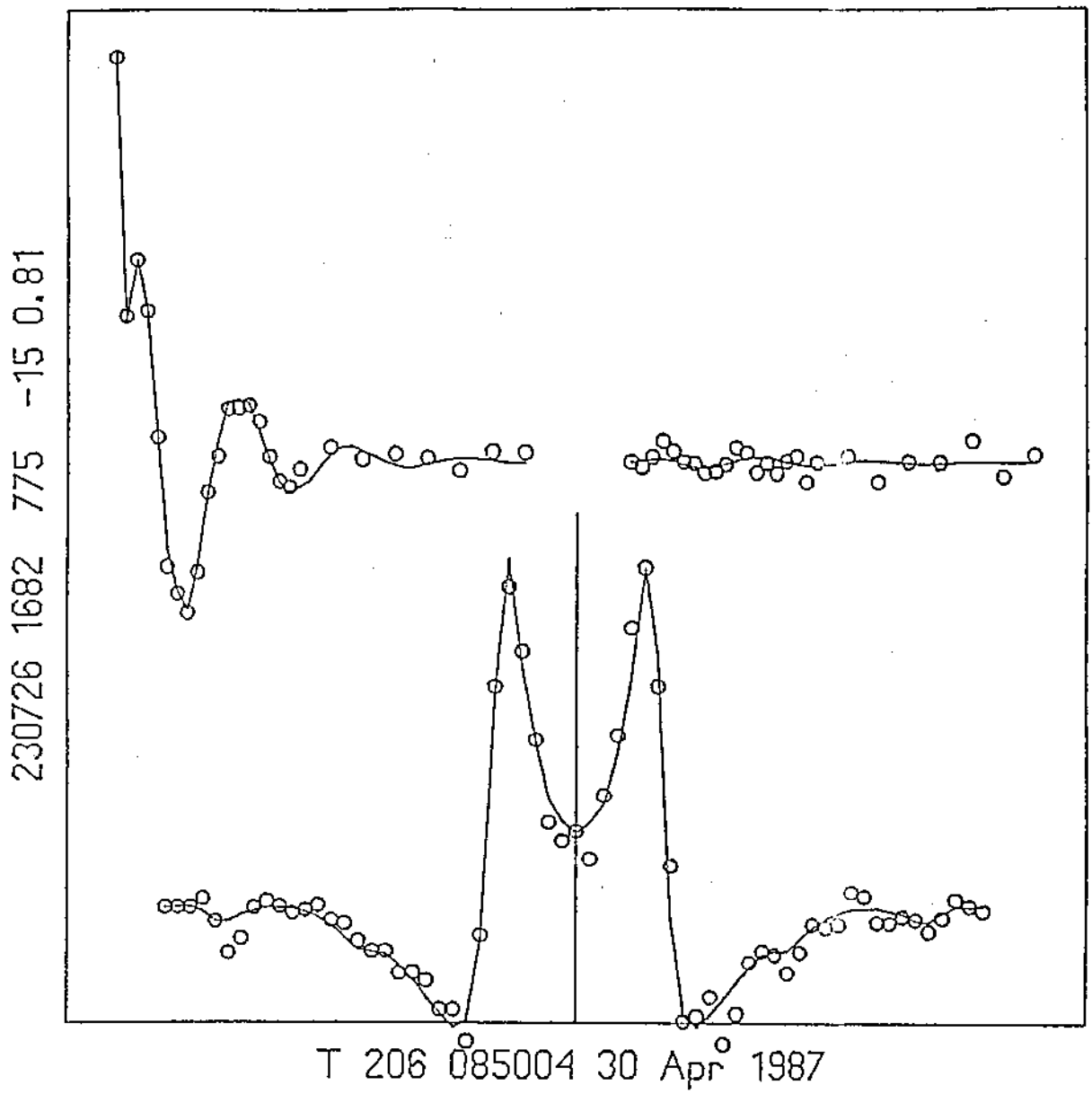


Fig. 21. A sample ACF and spectrum for the alternating code experiment.

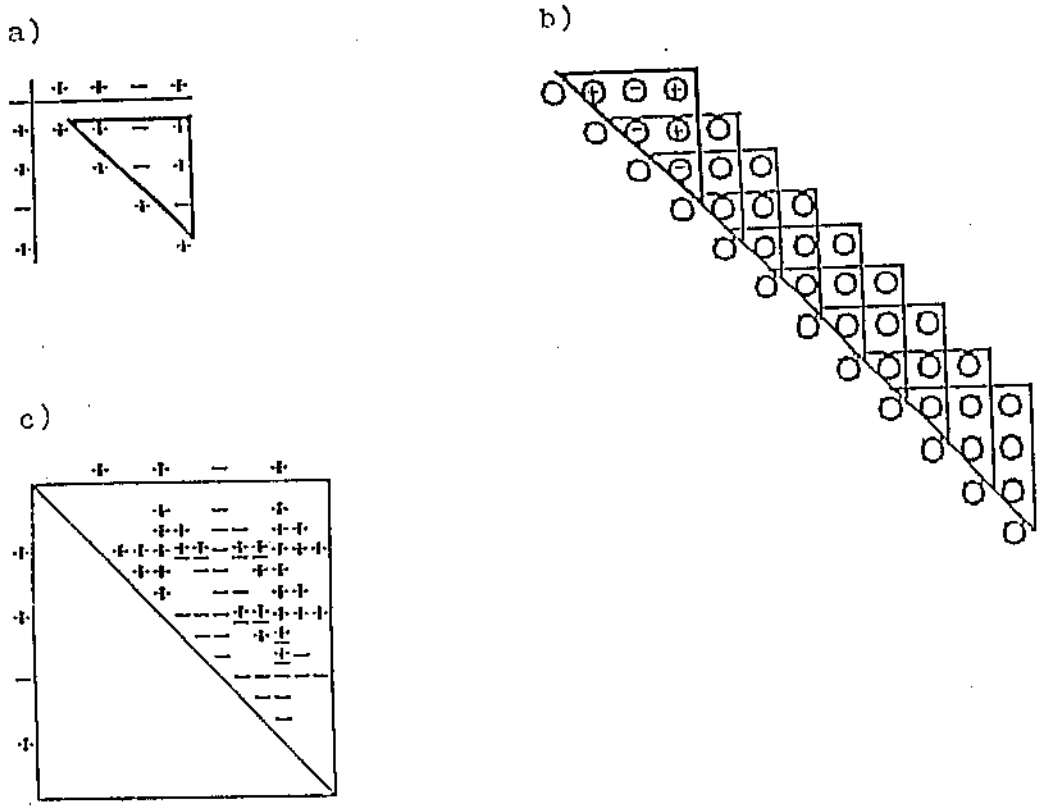


Fig. 22. The filter decoding scheme for lag profile filtering of the 4–pulse alternating code.

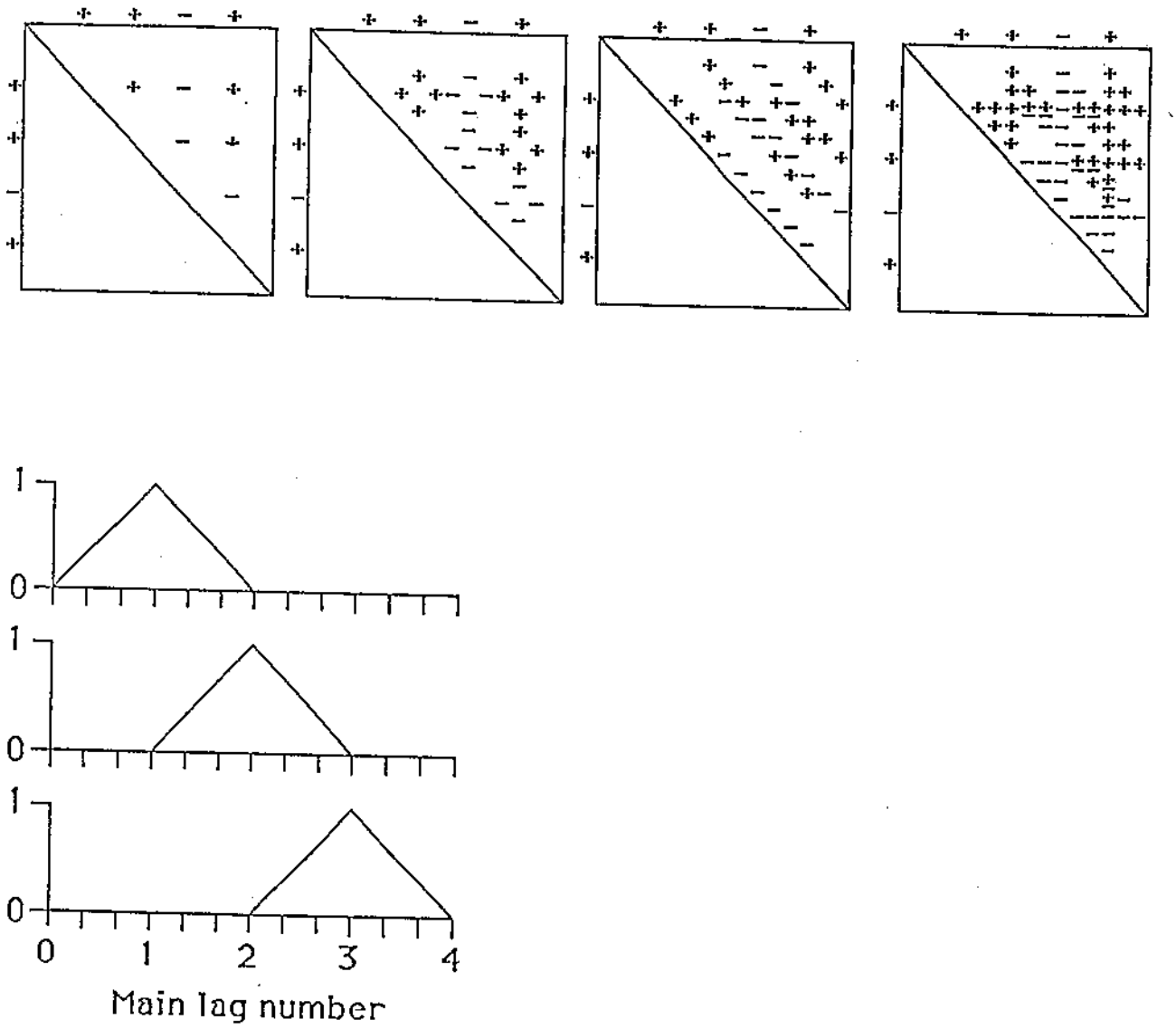


Fig. 23. The pulse ACF correction scheme for the 4–bit alternating code.

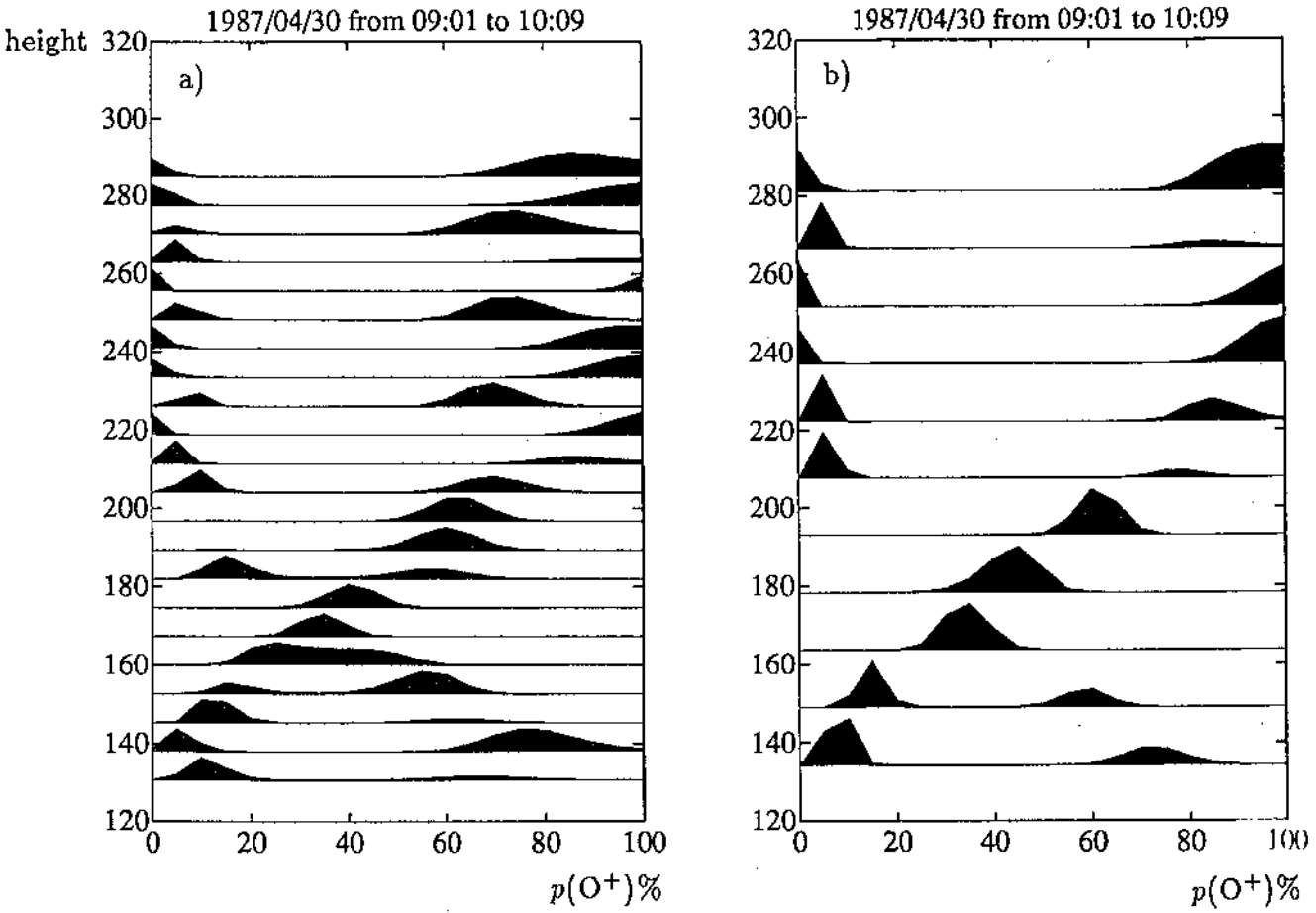


Fig. 24. The composition a posteriori distributions for the alternating code experiment on 30 April 1987 integrated over 3 dumps and a) 5 range gates b) 10 range gates.

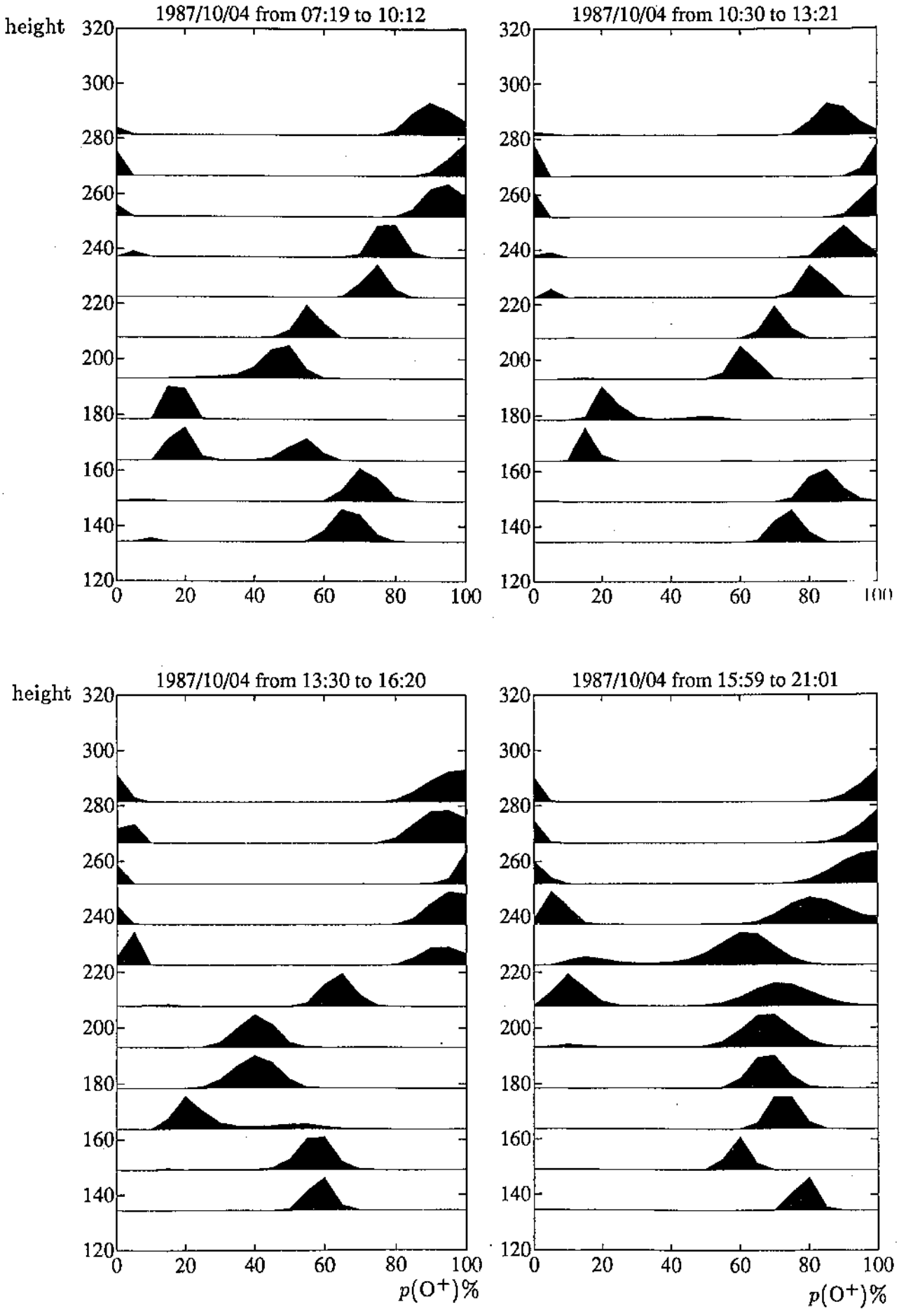


Fig. 25. The composition a posteriori distributions for the alternating code experiment on 4 October 1987 integrated over 6 dumps and 10 range gates.

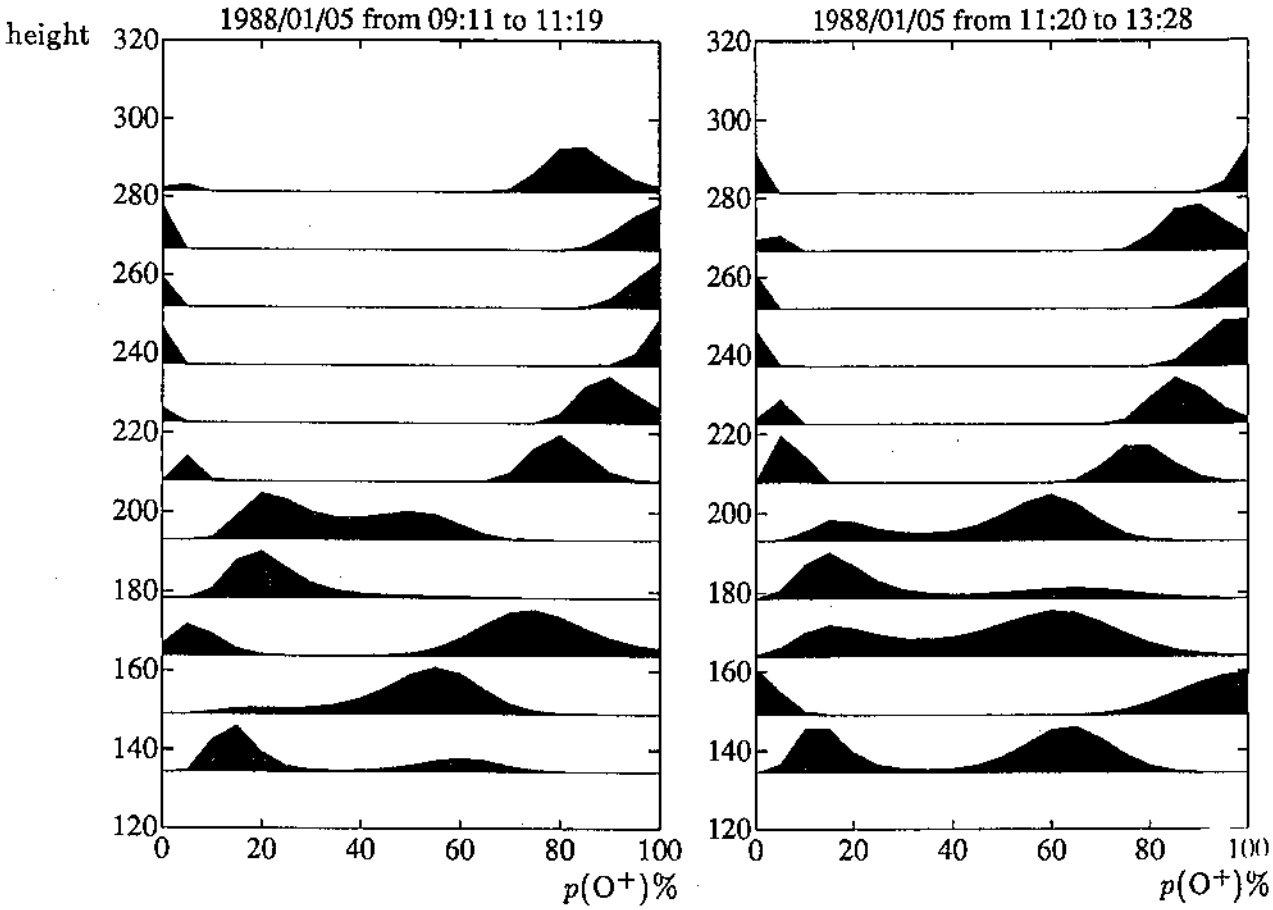


Fig. 26. The composition a posteriori distributions for the alternating code experiment on 5 January 1988 integrated over 6 dumps and 10 range gates.