

An information-theoretical methodology for profile parameter estimation of solar coronal spectra[★]

A. Nakagawa

Faculty of Science and Technology, Ryukoku University, Seta, Ôtu 520-2123, Japan
e-mail: nakagawa@rins.ryukoku.ac.jp

Received 24 November 2005 / Accepted 18 April 2006

ABSTRACT

Aims. An attempt is made to establish an information-theoretical methodology for profile parameter estimation of the spectral lines of the solar corona.

Methods. In this formalism, “description lengths” of a target line profile are defined for every possible profile model and are used to identify the optimal one among possible models. To properly reflect statistical properties of count numbers representing spectral profiles, a so-called α scale is introduced and incorporated into the definition of description lengths. Formulae of error propagation from measured count numbers to derived profile parameters are also given.

Results. The advantage of this methodology, which enables mutually consistent treatment of profiles of diverse properties in particular, has been used to examine how the change in spectroscopic characteristics is associated with the coronal structure.

Key words. Sun: corona – Sun: UV radiation

1. Introduction

Measurements of the solar corona by ultraviolet spectral lines provide us with information concerning coronal physical conditions that cannot be obtained with other types of remote observations (e.g., Pottasch 1963; Raymond & Doyle 1981; Kohl & Withbroe 1982; Withbroe et al. 1982; Noci et al. 1987; Brosius et al. 1996; Li et al. 1998; Del Zanna & Bromage 1999; Noci 2000; Parenti et al. 2000). Continuous monitoring of these coronal lines has become possible for the first time with the advent of the Solar and Heliospheric Observatory (SOHO) space probe. The UltraViolet Coronal Spectrometer (UVCS), the Solar Ultraviolet Measurement of Emitted Radiation (SUMER), and the Coronal Diagnostic Spectrometer (CDS) instruments on board, designed to observe coronal ultraviolet lines, have been providing detailed spectral profiles nearly a decade since its launch in 1995 (Harrison et al. 1995, 1997; Kohl et al. 1995, 1997; Wilhelm et al. 1995, 1997; Mason et al. 1997).

We are concerned with determining spectral profile parameters. We exemplify it for ultraviolet line profiles: in particular, those of the H I Ly- α line at 121.56 nm, and the O VI doublet at 103.19 nm and 103.76 nm, three major UVCS target lines in heliocentric altitudes 1.5–10 R_{\odot} . The techniques developed in this paper are also applicable to line profiles obtained elsewhere by other spectrometers, in particular, those by SUMER and CDS at lower altitudes. It may be possible to apply our methodology to spectral profiles of non-coronal origins as well.

In most cases, profiles obtained by UVCS measurements can be approximated by a single Gaussian function, or else a Gaussian accompanied by a secondary Gaussian and/or a background continuum, at least over a limited spectral interval of interest (e.g., Kohl et al. 1996, 1997, 1999a, 1999b; Peres et al. 1997; Dobrzycka et al. 1999b; Miralles et al. 2001a, 2002; see

also Hassler et al. 1997; Tu et al. 1998; Moran 2003). In these situations, observed profiles can be represented by a small number of parameters, such as line intensities or line widths. Also, the knowledge of these profile parameters is a prerequisite for a spectral diagnosis of the solar corona (e.g., Cranmer et al. 1999; Strachan et al. 2000; Zangrilli et al. 2002; Antonucci et al. 2004).

Derivation of spectral profile parameters should be straightforward; given a measured spectral profile and once having specified a fit interval and a model profile, one should find a best fit by the least-square method, for example (e.g., Kohl et al. 1997). Even for those with at least some experience, carrying out such a procedure is not a difficult task. Various kinds of practical difficulties, however, may arise when treating big volumes of actual profile data. Thus, some sort of a unifying criterion is required, especially, for a consistent treatment of the profiles of diverse properties. Obviously, using a residual sum of squares, χ^2 , is not adequate for this purpose, because merely by reducing the fit interval the corresponding reduction of χ^2 is attained immediately. One example requiring such a consistent treatment is given in spectral measurements along slices of the solar corona including both a polar coronal hole and an equatorial streamer, which will be treated below.

In this paper, an attempt will be made to establish a unified methodology for profile parameter estimation, with the aid of information theory. The specific problem here is how to select an adequate spectral model and/or an adequate fit interval, for a given spectral profile. This can be regarded as a “data-based model selection problem”. Here we employ the minimum description length (MDL) principle for this purpose (Rissanen 1989; Han & Kobayashi 2002). Under the MDL principle, for a given data set, one selects an optimal among several possible models as the one that gives the least description length of the data. Thus the central issue here is how to construct a description length suitable for the problem under consideration.

[★] Appendix A is only available in electronic form at <http://www.edpsciences.org>

In Sects. 2.1 and 2.2, we briefly introduce the idea of description lengths, using a histogram partition problem and a curve-fit problem. Both will be incorporated into the definition of the description length for profile parameter estimation (Sect. 2.4). In the case of profile parameter estimation, an intrinsic difficulty in defining a description length arises from a statistical property of count numbers constituting spectral profiles. In Sect. 2.3, for the consistent treatment of count number statistics, we introduce a scale transformation of count numbers, called the “ α scale”. The use of the α scale is effective, or sometimes indispensable, for the parameter estimate of actual profiles (Sect. 3.2). The description length defined in Sect. 2.4 is constructed using count numbers normalised with the α scale. The main portion of our proposal is thus given in Sects. 2.3 and 2.4.

The UVCS observations during the last solar minimum have revealed that some streamers exhibit a striking difference between HI Ly- α images and those in the OVI lines; along the location of the streamer axis in the HI Ly- α images, there is a local minimum of the OVI brightness (Noci et al. 1997; Raymond et al. 1997). The UVCS observation treated below is a typical example of this type of streamer, and it is classified as *Category I* by Noci et al. (1997). A derivation procedure of profile parameters using actual coronal line profiles observed by the SOHO/UVCS is then described in Sect. 3. The instantaneous FOV (field of view) of our observation crosses both the streamer axis and its neighbouring northern coronal hole. The advantage of an information-theoretical approach is best demonstrated in a mutually consistent treatment of profiles having such diverse properties (Sect. 3.3). This ability allows us to reveal the change in spectroscopic characteristics associated with the coronal structure. The relevant observations and their solar-physical consequences are treated in Sect. 4.

Although description lengths are suited to use in selecting the optimal fit model for a given profile, we cannot use them to determine the quality of fit parameters (Sect. 3.3). A better criterion for this purpose is given by estimating the uncertainties of the derived parameters. Uncertainties of the derived profile parameters come from errors, both statistical and observational, in count numbers representing spectral profiles. The formula for this error propagation is given separately in Appendix A.

2. Description lengths

In this section, we establish the definition of the description length used for profile parameter estimation.

2.1. MDL principle

To introduce the idea of a “description length” in the MDL principle, let us consider a *histogram partition problem*, given n integer numbers over a fixed domain, to find out the histogram that is most appropriate for expressing the distribution of these integers. Assume that, by partitioning the domain into k subintervals having an equal length, Δ , we have a histogram whose j th subinterval contains n_j integer numbers each ($j = 1, 2, \dots, k$). The description length of the given integers is then defined as

$$l(\Delta) = (k - 1) \log n + \log \binom{n}{n_1 n_2 \dots n_k} + n \log \Delta. \quad (1)$$

The terms on the right-hand side, respectively, can be interpreted as the description lengths of encoding the numbers, (n_1, n_2, \dots, n_k) , of encoding to which subinterval each integer

belongs (the parentheses in the equation denote the multinomial distribution), and of encoding the positions of these integers within respective subintervals. The description length, $l(\Delta)$, is defined for each domain partitioning (i.e., histogram). Under the MDL principle, we are to select, among possible histograms, the one that gives the minimum of these $l(\Delta)$'s, as the most suitable for expressing the data distribution.

For the case $n_j = O(n)$ and $k = O(1)$, we may rewrite (1), using the Stirling's approximation $\log(n!) \sim \log \sqrt{2\pi n} + n(\log n - 1)$, as

$$l(\Delta) = n \left\{ \sum_j \left[-\frac{n_j}{n} \log \frac{n_j}{n} \right] + \log \Delta \right\} + \frac{k-1}{2} \log n + O(1), \quad (2)$$

where $O(n)$ and $O(1)$ are used to represent truncation terms at least of the orders n^1 and n^0 , respectively. In the case where the domain cannot be specified beforehand (as opposed to the above assumption that the domain is preformed), $k - 1$ should be replaced with $k + 1$. It is known that, in the context of information theory, (2) is a generally valid formula for an *asymptotically universal optimal* description length; given a model of a data distribution, the second term represents a description length for specifying a model, while the first term denotes that of the data under the specified model.

The description length of the histogram partition problem above can be considered as that of a “model-less” model on data, meaning that it does not make any assumption on a data distribution.

2.2. Curve fitting

The second example concerns a *curve-fit problem*: given a set of n data points, $(\lambda_i, C_i)_i$, to find a best curve fit, $C = f(\lambda)$, having m parameters under the condition that the error distribution of C_i is known to be $p_i(C)$ (here we have in mind that symbols λ and C denote the wavelength and the count number, respectively). The description length of this curve fit problem is defined as

$$l(f) = \sum_i \left[-\log p_i(f(\lambda_i)) \right] + \frac{m}{2} \log n. \quad (3)$$

Again, the second term on the right-hand side can be interpreted as the description length for specifying a model (i.e., parameters of the curve) and the first term that of the data (i.e., C_i 's) under the specified model. If C_i is a continuous variable, the first term should be $-\log[p_i(f(\lambda_i)) \delta C_i]$, defined over an infinitesimal interval, δC_i .

2.3. Scale transformation

Spectral profiles observed by UVCS are expressed in terms of count numbers in each pixel of its XDL (cross delay line) detectors. Making a description length correctly reflect the statistical properties of these count numbers is essential for its proper definition. Relevant considerations will be given here.

Even in an ideal case without systematic instrumental errors, count numbers can differ for each measurement. In this ideal case, count numbers obey Poisson statistics, in which the uncertainty of a count number, C , is given by

$$1\sigma = \sqrt{C}.$$

By noting

$$\sqrt{C \pm \sqrt{C}} \sim \sqrt{C} \pm \frac{1}{2} + O\left(\frac{1}{\sqrt{C}}\right), \quad (4)$$

we can see that the degree of scatter of a count number scaled by a standard deviation is, under this “square-root scale”, normalised to 1/2, irrespective of its magnitude.

When a count number per pixel is not large, say less than 100, the scatter of the count number is due mainly to this Poisson uncertainty ($1/\sqrt{100} = 10\%$). When the count number amounts to several thousands, by contrast, the errors with instrumental origin, such as the pixel-to-pixel variation of detectors, could exceed that of Poisson statistics.

Giving a rigorous model of these instrument-specific error factors, however, is not a simple task. We therefore introduce a working model, instead. For a restricted purpose of profile parameter estimation, a simple model such as introduced here works well, as will be demonstrated in the next section.

One error model we employ here is that having an error distribution as

$$1\sigma = \sqrt{C + (\alpha C)^2}, \quad (5)$$

where α is a small number reflecting instrument quality in total. This is a model that is reduced to Poisson statistics for small count numbers and becomes the one with linear dependence ($\sim\alpha C$) for large count numbers. The scale transformation for this model is given by

$$t_\alpha(C) = \frac{1}{\alpha} \text{Arcsinh}[\alpha \sqrt{C}] = \frac{1}{\alpha} \log[\alpha \sqrt{C} + \sqrt{1 + \alpha^2 C}]. \quad (6)$$

In fact, under the error model (5), we may have

$$\begin{aligned} & \frac{1}{\alpha} \text{Arcsinh}[\alpha \sqrt{C} \pm \sqrt{C + (\alpha C)^2}] \\ &= \frac{1}{\alpha} \text{Arcsinh}[\alpha \sqrt{C}] \pm \frac{1}{2} + O\left(\alpha + \frac{1}{\sqrt{C}}\right), \end{aligned}$$

implying the normalisation of a standard deviation to 1/2. Hereafter we refer to the error model (5) as “ α distribution”, and the scale transformation (6) as “ α scale”. As expected, α scale (6) is reduced to square-root scale in the $\alpha \rightarrow 0$ limit; thereby, α scale can be considered as an extension of the square-root scale with a control parameter. In the examples of Sect. 3, we employ $\alpha = 0.15$ for the LYA channel and $\alpha = 0.06$ for the OVI channel, otherwise stated. Estimation shows that the pixel-to-pixel variations of the UVCS detectors are $\sigma = 10\%$ and 4% for the LYA and OVI channels, respectively (Cosmo et al. 1999; Gardner et al. 2002).

2.4. Description lengths

Having introduced necessary components, we are now in a position to give a definition of a description length for observed line profiles, for the purpose of estimating the profile parameters. Assume that we are given a spectral profile including a target spectral line. We tentatively fix a spectral interval for profile fit, and divide this interval into three subintervals; we refer below to the middle one as interval M , only on which an actual profile fit will be done, and the respective side ones as intervals L and R . The MDL principle is employed to determine the optimal profile model and the optimal subinterval M for the profile fit. Let n_M , n_L , and n_R denote the number of data points of respective subintervals; and we write $n_{\text{total}} = n_M + n_L + n_R$. The description length is defined for all the data points.

As in the examples given in Sects. 2.1 and 2.2, the description length of a profile, under the concept of the MDL principle, can be expressed as the sum of i) the description length of a profile model and ii) that of profile data under the specified model.

With regard to ii), it is natural to employ the description length of a curve fit (Sect. 2.2) for the data points on the interval M , and that of a histogram partition (Sect. 2.1) for those on the intervals L and R . The original expressions in the preceding subsections, however, need to be changed for such a composite model. In addition, alterations are also required for the mutually consistent treatment of the statistical properties of data points. This should be done for the sake of consistency, not only between data points within the respective subintervals, but also among those between different subintervals.

For the interval L , we use count numbers transformed by the α scale for doing histogram partition. Under this transformation, count numbers are normalised as statistical variables, having a fixed value of $1\sigma = 1/2$. Accordingly, in defining a description length ii), we use transformed count numbers in units of σ , and also supplement the description length with a term reflecting the situation that count numbers are statistical variables (note that, when defining (2) in Sect. 2.1, integers there are assumed to be exact numbers). The description length is thus written as

$$\begin{aligned} l_L &= n_L \left\{ \sum_j \left[-\frac{n_j}{n} \log \frac{n_j}{n} \right] + \log\left(\frac{\Delta}{\sigma}\right) + \log(\sqrt{2\pi}\sigma) \right\} \\ &= n_L \left\{ \sum_j \left[-\frac{n_j}{n} \log \frac{n_j}{n} \right] + \log(\sqrt{2\pi}\Delta) \right\}. \end{aligned} \quad (7)$$

Here $\log(\sqrt{2\pi}\sigma)$ is the supplementary term noted above. The definition of l_R on the interval R is given in the same manner.

For ii) of the interval M , we give the description length, again using transformed values $t_\alpha(f(\lambda_i))$, as

$$l_M = \sum_i \left\{ -\log \left[p_i(t_\alpha(f(\lambda_i))) \delta t_\alpha(C_i) \right] \right\},$$

where $f(\lambda)$ denotes a model profile fit with the data points on the interval M (the fit itself is to be done using count numbers expressed in the linear scale, rather than using count numbers transformed by the α scale). For convenience, we approximate the error distribution p_i as a Gaussian distribution satisfying $1\sigma = 1/2$:

$$\begin{aligned} p_i(t_\alpha(f(\lambda_i))) &\sim \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\{t_\alpha(C_i) - t_\alpha(f(\lambda_i))\}^2}{2\sigma^2}\right] \\ &= \sqrt{\frac{2}{\pi}} \exp\left[-2\{t_\alpha(C_i) - t_\alpha(f(\lambda_i))\}^2\right]. \end{aligned}$$

To make the definition consistent with l_R , we set $\delta t_\alpha(C_i) = 1$. We thus have

$$l_M = \sum_i \left[2\{t_\alpha(C_i) - t_\alpha(f(\lambda_i))\}^2 + \log \sqrt{\frac{\pi}{2}} \right]. \quad (8)$$

With regard to i), we follow the prescription for the *category classification problem*: hence,

$$l_{\text{model}} = \frac{m + (k_L + 1) + (k_R + 1)}{2} \log n_{\text{total}}, \quad (9)$$

where m denotes the number of fit parameters on the interval M , and k_L and k_R represent the numbers of partition of the intervals L and R , respectively.

The total description length l_{total} is given by the sum of (7), (8), and (9):

$$\begin{aligned} l_{\text{total}} &= l_M + l_L + l_R + l_{\text{model}} = \left(l_M + \frac{m}{2} \log n_{\text{total}} \right) \\ &\quad + \left(l_L + \frac{k_L + 1}{2} \log n_{\text{total}} \right) + \left(l_R + \frac{k_R + 1}{2} \log n_{\text{total}} \right). \end{aligned} \quad (10)$$

Table 1. Profile models and total description lengths.

Row#	Histogram model	Single Gaussian	Single Gaussian with const.	Double Gaussian	Double Gaussian with const.	Triple Gaussian	Triple Gaussian with const.
34	119.73	64.19 ^a	37.36	31.65 ^b	32.68	35.58	39.85
17	116.76	47.23	29.77 ^c	31.95	40.93	39.06	34.46

^a Fig. 1; ^b Fig. 2; ^c Fig. 3.

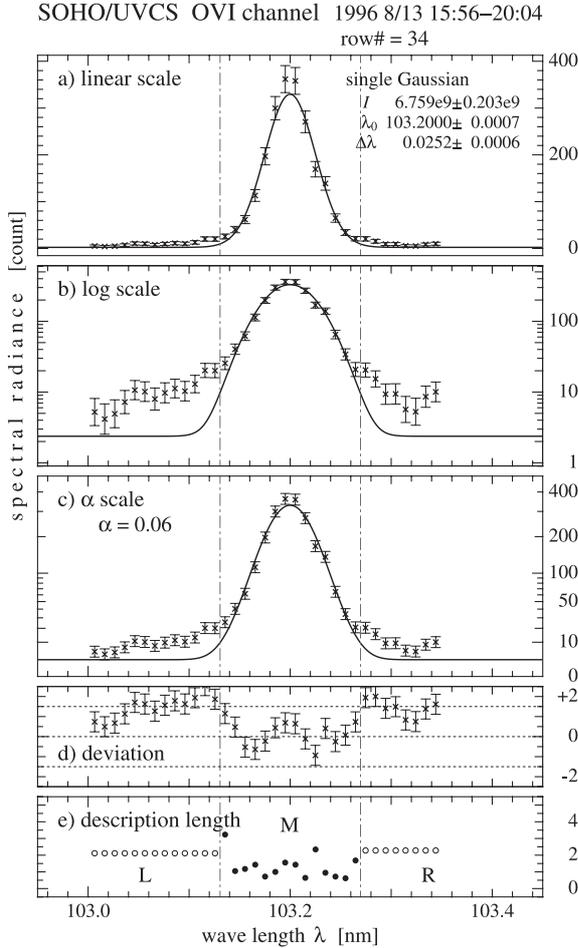


Fig. 1. O VI 1032 Å profile observed by the UVCS OVI channel and its fit to a single Gaussian. Panels **a)–c)** represent an identical observed profile with error bars equivalent to $\pm 1\sigma$ being attached in **a)** linear scale, **b)** logarithmic scale, and **c)** α scale ($\alpha = 0.06$). In each panel, the ordinate is graduated in count numbers. The curves therein represent the optimal fit to a single Gaussian, with assuming a detector dark count rate of 3×10^{-5} counts per second per pixel. Derived profile parameters are given within panel **a)**; I is in units of $\text{ph s}^{-1} \text{cm}^{-2} \text{sr}^{-1}$, and λ_0 and λ are in units of nm. Panels **d)** and **e)** show deviations, in an α scale, from the fit curve and the description lengths of individual data points, respectively.

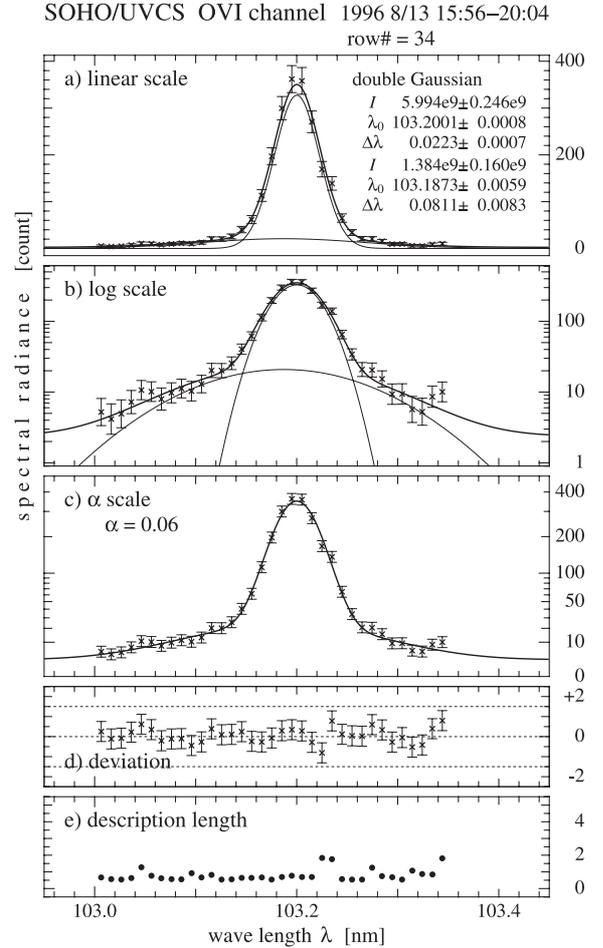


Fig. 2. Double Gaussian fit of the same profile as Fig. 1.

In the following, we adopt a single or multiple Gaussian with/without an additive constant as a fit function:

$$f(\lambda) = C_0 + C_A g_A(\lambda) + C_B g_B(\lambda) + \dots$$

$$g_A(\lambda) = \frac{1}{\sqrt{2\pi}} \frac{\delta}{\Delta\lambda_A} \exp\left[-\frac{(\lambda - \lambda_{0A})^2}{2(\Delta\lambda_A)^2}\right], \quad (12)$$

where profile parameters are C_0 , C_A , λ_{0A} , $\Delta\lambda_A$, etc.

On this extreme right-hand side, the terms concerning the subinterval L are limited to

$$l_L + \frac{k_L + 1}{2} \log n_{\text{total}}. \quad (11)$$

This implies that, in determining the optimal histogram partition on the subinterval L , we must minimise (11). Note here that in its last term n_{total} appears, rather than n_L .

3. Profile parameter determination

UVCS observed an equatorial streamer at the East limb on 13 August 1996. The instantaneous FOV of the UVCS entrance slit, being a strip in the direction parallel to the limb tangent ($37.8'$ and $41.3'$ long for the LYA and OVI channels, respectively), was located at a nominal heliocentric height $r = 1.5 R_\odot$, with a position angle 60° (measured counterclockwise from the

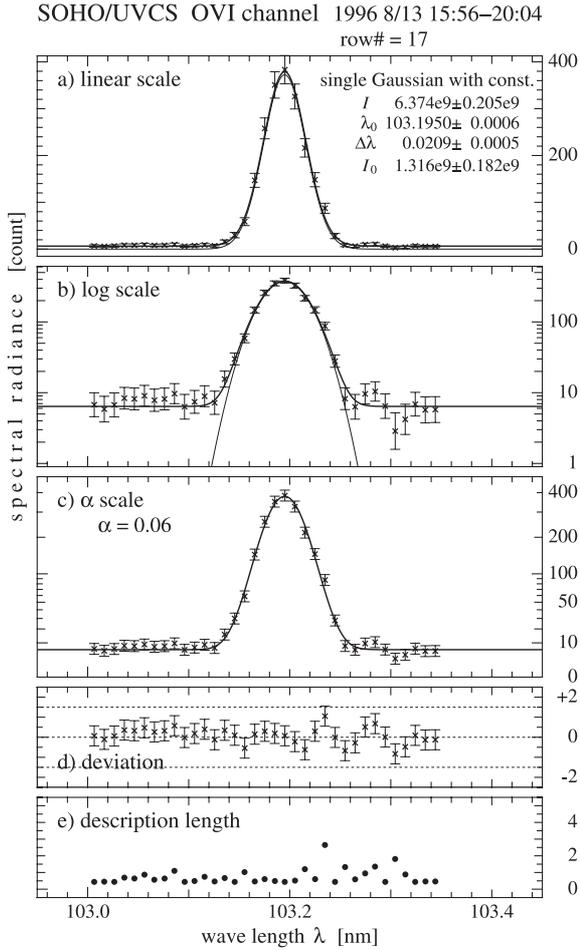


Fig. 3. O VI 1032 Å profile and its fit to a single Gaussian with an additive constant.

solar north). The slit widths were $50 \mu\text{m}$ ($14''$) for the LYA channel and $75 \mu\text{m}$ ($21''$) for the OVI channel. The spatial binning along the slit was $42''$ for both channels. The integration times were about 4 hours. The raw data was processed with the standard UVCS calibration procedure (Gardner et al. 1996), using the DAS33 software (dated June 2001), and was also corrected for the so-called device effect (Kohl et al. 1997; Nakagawa 2003). If the stray light contribution is significant, we simply dispense with the use of such highly “contaminated” profiles, rather than trying to subtract the expected levels from them (our estimates of stray light levels are $3 \times 10^8 \text{ ph s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1}$ for Figs. 1–3, 6, and $3 \times 10^{10} \text{ ph s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1}$ for Figs. 4, 5, 7 below; see Gardner et al. (1996), Kohl et al. (1999a) and Zangrilli et al. (1999). In Fig. 8, we do not draw contours where derived profile line intensities exceed stray light intensity estimates). We assume a detector random dark count rate of 3×10^{-5} counts per second per pixel (Kohl et al. 1999a).

3.1. Model selection

We exemplify here how profile model selection and profile parameter estimate are done using the MDL principle, by an O VI 1032 Å profile observed in an equatorial streamer employing the OVI channel (Fig. 1). The curves in panels a–c represent the optimal fit to a single Gaussian, with assuming a detector dark count rate as noted above. The actual curve fit calculation has been done only using data points belonging to the interval M

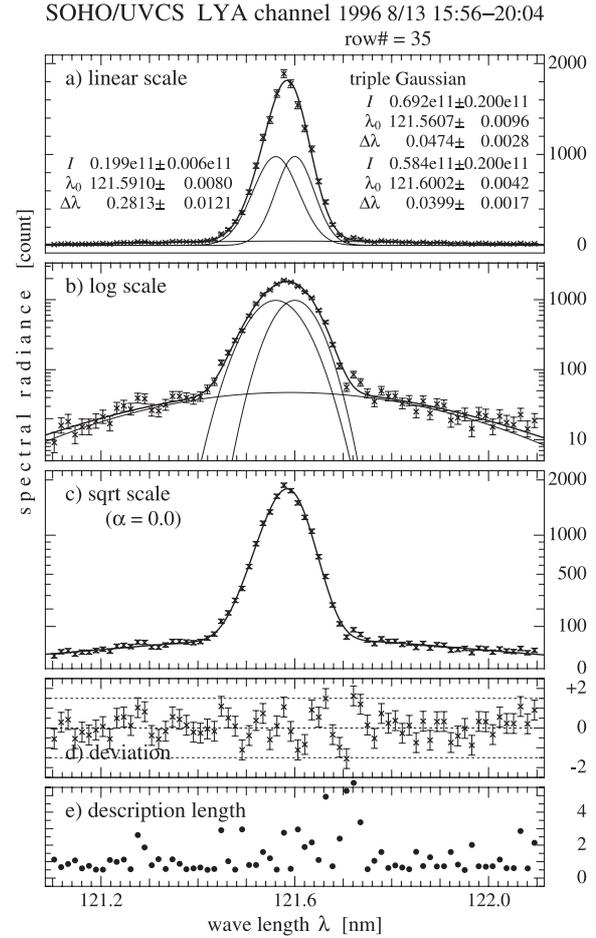


Fig. 4. HI Ly- α profile observed by the UVCS LYA channel and a triple Gaussian as an optimal model for the profile, under the assumption of Poisson statistics ($\alpha = 0$).

given in panel e. The interval M is the one that is selected by the MDL principle, as the most suitable for a profile fit to a single Gaussian.

Values of derived parameters and their $\pm 1\sigma$ error widths are given within panel a. Line widths should be corrected for an instrumental broadening with $\sqrt{(\Delta\lambda)^2 - (\Delta\lambda_{\text{corr}})^2}$ (Kohl et al. 1999a). (We adopt $\Delta\lambda_{\text{corr}} = 0.0108 \text{ nm}$ for Figs. 1–3, and 6, and $\Delta\lambda_{\text{corr}} = 0.0126 \text{ nm}$ for Figs. 4, 5, and 7. In Figs. 1–5 the uncorrected profiles are displayed, while in Figs. 6 and 7 corrected values are plotted.) Here, $\Delta\lambda$'s are those defined in (12), having relations with the “ $1/e$ half width” as $\Delta\lambda_{1/e} = \sqrt{2} \Delta\lambda$, and the $FWHM$ as $\Delta\lambda_{FWHM} = \sqrt{8 \log 2} \Delta\lambda$.

For a visual inspection on the suitability of a fit function for a profile, it is convenient to use a logarithmic scale (panel b), in which a Gaussian is represented as a parabola, or an α scale (panels c and d), in which error bars are of equal length. From these panels one can discern the discrepancy of the observed profile from the fit curve as a narrow core in the line centre and a broad wing on both sides of it, indicating the inadequacy of a single Gaussian as a line model for the given profile.

For the same profile, a fit to a double Gaussian is depicted in Fig. 2. Although it is a general rule that a more complicated model fits data better, its cost, scaled in terms of a description length, of describing such a complicated model becomes a disadvantage, from the point of view of the MDL principle. As shown in

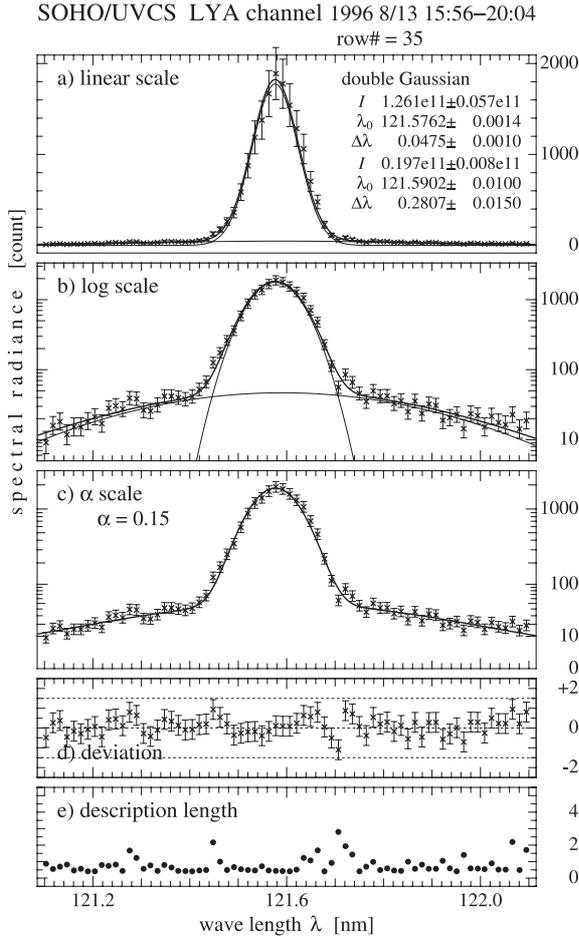


Fig. 5. Double Gaussian, as an optimal model for the same profile as Fig. 4, under the assumption of α distribution with $\alpha = 0.15$.

Table 1 (row# 34), the fit to a double Gaussian, giving the least description length, is optimal among the line models listed.

As another example, we give, in Fig. 3, a profile for which a single Gaussian is concluded to be optimal (row# 17 in Table 1). This O VI 1032 Å profile is observed in the core of the same streamer as in Fig. 2.

3.2. The α distribution

We now consider an example in which α distribution is effectively used for the parameter estimate. Figure 4 shows a HI Ly- α profile observed by the LYA channel. When we assume Poisson statistics ($\alpha = 0$) for this profile, a triple Gaussian is selected as an optimal model by the MDL principle (Fig. 4b).

When α distribution with $\alpha = 0.15$ is applied to the same profile as above, by contrast, a simpler, double Gaussian model is found to be optimal (Fig. 5). Larger lengths of error bars in Fig. 5a than in Fig. 4a reflect the contrasting assumptions on the values of α , which also make the profile appearances in the α scale dissimilar (Figs. 4c and 5c).

The employment of α distribution is useful, in particular, when count numbers are sufficiently large. In such cases, the scatter of count numbers caused by non-uniformity in detector pixel efficiencies can eventually exceed the level of pure Poisson uncertainty; for example, the count number 1000 corresponds to the Poisson uncertainty $1/\sqrt{1000} \sim 3.2\%$. In an extreme case, to cope with such false scatter, we may be led to select an

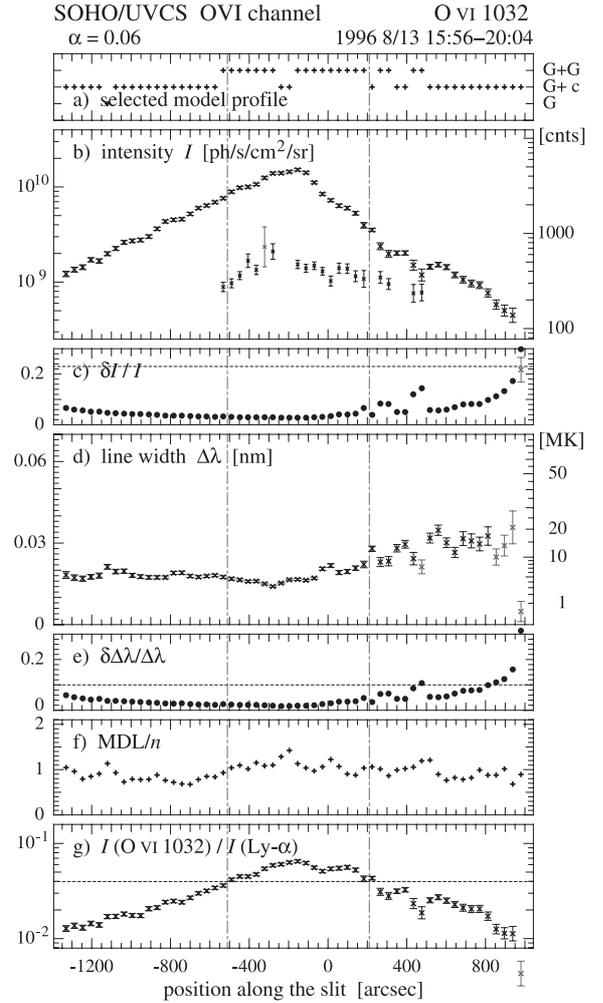


Fig. 6. Derived profile parameters and estimated relative errors for the O VI 1032 Å line, assuming α distribution with $\alpha = 0.06$. In panels b and d, the ordinates are also graduated in total count numbers and MK (megakelvin), respectively.

unnecessarily complicated model as optimal. This can be avoided by employing an α distribution with an appropriately chosen value of α . We will get a simpler optimal model by adopting larger α .

3.3. Reliability criteria

Along an instantaneous FOV of the UVCS entrance slit, several tens of profiles with identical measurement parameters are typically obtained simultaneously. For a resulting set of profiles, of which the magnitude of count numbers and/or the spectroscopic properties may vary appreciably, to keep mutual consistency of the parameter estimate and to make certain that no acceptable curve-fit is excluded unnecessarily or vice versa are quite important. In this subsection, we shall see how this is done in the current framework.

Results of the parameter estimate along the slice of the corona are given in Fig. 6 (O VI 1032 Å) and Fig. 7 (HI Ly- α). The FOVs corresponding to the entrance slit positions are drawn in Figs. 8a and 8b for the respective lines. For Figs. 6 and 7, as with Figs. 1–5, we do not use the count number average along the slit length, except for the count accumulation due to the on-board spatial binning.

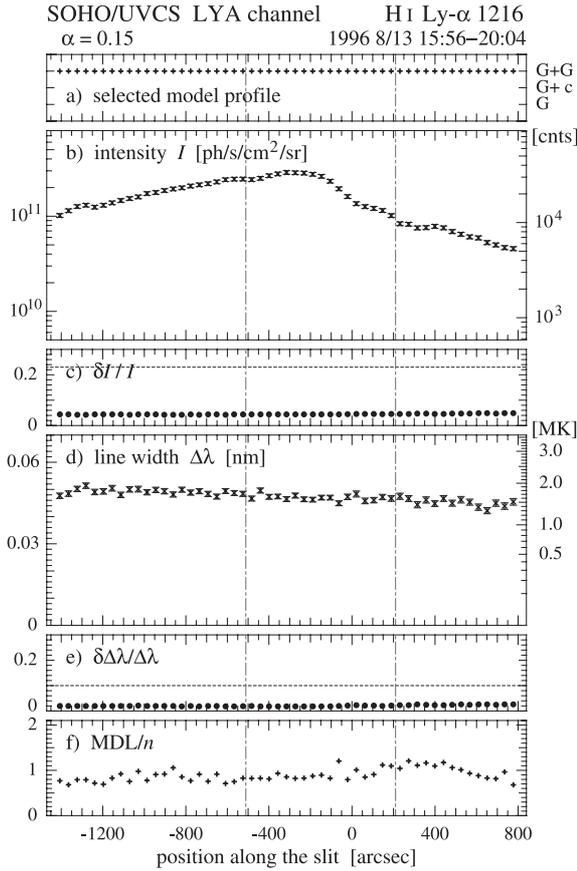


Fig. 7. Derived profile parameters and estimated relative errors for H I Ly- α line, assuming α distribution with $\alpha = 0.15$.

Panel a shows line models selected by the MDL principle. For the H I Ly- α case (Fig. 7a), double Gaussians (“G+G”) are selected exclusively. In contrast, the single Gaussian with constant (“G+c”) and the double Gaussian share the nomination for the O VI 1032 Å case (Fig. 6a). Figure 2 (Fig. 3) offers an example profile whose optimal fit is given by the “G+G” (“G+c”) model. For the O VI profiles in the region $>+500''$ of Fig. 6, count numbers are not enough to resolve multiple components, which yields the selection of a simpler, single Gaussian as optimal.

Estimated profile parameters are depicted in panels b and d. In Fig. 6b, line intensities are shown for both narrow and broad components, the former being more intense for all the profiles labelled as “G+G”; because the stray light contribution is not negligible for the H I Ly- α broad component, they are not depicted in Fig. 7b. The line widths of broad components (panel d) are not shown for both lines because of their lower reliability.

As an indicator of the reliability of parameter derivation, one can use the relative magnitude of 1σ error width. In panels c and e of Figs. 6 and 7 these values of 1σ are drawn for respective profile parameters. By tentatively setting the thresholds to 0.1 for the line width (panel e) and to $0.23 \sim 0.1 \times \log 10$ for the logarithmic line intensity (panel c), we draw respective parameters in panels b and d in grey when they exceed these thresholds. Note that, as seen from panel f, the description lengths per data point, by contrast, cannot be an indicator of the fit quality.

Drawn in Fig. 8 are contour maps of the line intensities (of narrow components) by synoptic observation on 14 August 1996. Contour lines are depicted in black when 1σ is below $0.1 \times \log 10$ in relative magnitude, in dark grey when below

$0.13 \times \log 10$, and in light grey when below $0.2 \times \log 10$. We can see from the figures overall consistency in excluding unacceptable curve-fits, with the value of 1σ . Obviously, small-scale “fluctuations” of contour curves, in particular, in the coronal-hole regions of higher altitudes, being a reflection of the accuracy of estimated parameters, should not be considered as an observational fact. By lowering the above thresholds, we will obtain contours of higher quality, although the portion of contours drawn in black has to be reduced accordingly.

4. Coronal structure and profile characteristics

In this section, we are concerned with the solar-physical implication of the derived profile parameters, by starting from the identification of the coronal structure. In Figs. 6 and 7, we define the boundary between a coronal hole and an equatorial streamer at the position $+210''$ in the abscissae, with the hole being its left-hand side. In Fig. 6d, a well-known broadening of the O VI lines is seen in the coronal hole region (Kohl et al. 1997, 1998, 1999b; Cranmer et al. 1999; Frazin et al. 1999; Zangrilli et al. 1999), while in the H I Ly- α case (Fig. 7d), the line width is a bit narrower in the coronal hole than in the streamer, at this altitude of observation (Suleiman et al. 1999; Zangrilli et al. 1999). It has been reported that the coronal hole/streamer boundary during activity minimum is clearly discerned (Miralles et al. 1999; Dobrzycka et al. 1999a; Spadaro et al. 2005).

Our streamer, classified as Category I, can be divided into regions of a streamer core and streamer legs, the former being a region of intensity reduction in the O VI image (Noci et al. 1997). Noting that the intensity ratio of O VI 1032 Å to H I Ly- α varies gradually in the streamer region (Fig. 6g), we define here a tentative position of the core/leg boundary at $-510''$ in Fig. 6. The reduction of the O VI doublet intensities in a streamer core is accounted for by a decrease in the O^{5+} abundance ratio there (Raymond et al. 1997; Ofman 2000; Marocchi et al. 2001; Vasquez & Raymond 2005), while the reduction at a coronal hole is attributed to a Doppler dimming of O VI lines (Li et al. 1998; Kohl et al. 1998; Cranmer et al. 1999; Zangrilli et al. 2002).

Given the above identification of the coronal structure, we can point out that spectroscopic characteristics of the O VI doublet change with the coronal structure (Fig. 6a). In the streamer leg (between $-510''$ and $+210''$ of Fig. 6), O VI profiles exhibit the presence of a broad component, typically shown in Fig. 2, while this component is below the observational limit ($\leq 10^9$ ph s $^{-1}$ cm $^{-2}$ sr $^{-1}$) in the streamer core (e.g., Fig. 3). In the coronal hole, broad components are identified only on the left-hand side of $+600''$ in Fig. 6. As stated above, in this particular measurement, statistics are not sufficient for resolving multiple components on the right-hand side (under the nomenclature given in Appendix A, the fits therein belong to the domain ii) of Fig. A.4). In the UVCS observations on 5–13 August 1996 (not shown), we can in fact identify a broad component ($\Delta\lambda \sim 0.08$ nm) having an intensity that is several times weaker than a narrow component ($\Delta\lambda \sim 0.02$ nm).

Multiple-component fits to the O VI doublet have been reported repeatedly, in particular, for the coronal hole profiles (Kohl et al. 1997, 1999a,b; Miralles et al. 2001a,b, 2002). For tenuous coronal hole lines, the three-dimensional configuration of the corona may easily lead to a contamination of the line-of-sight by non-hole origins (e.g., nearby streamers). In an equatorial hole observation, Miralles et al. (2001b) ascribe the broad component to the equatorial hole in the plane of the sky and the narrow component to foreground/background streamers. The

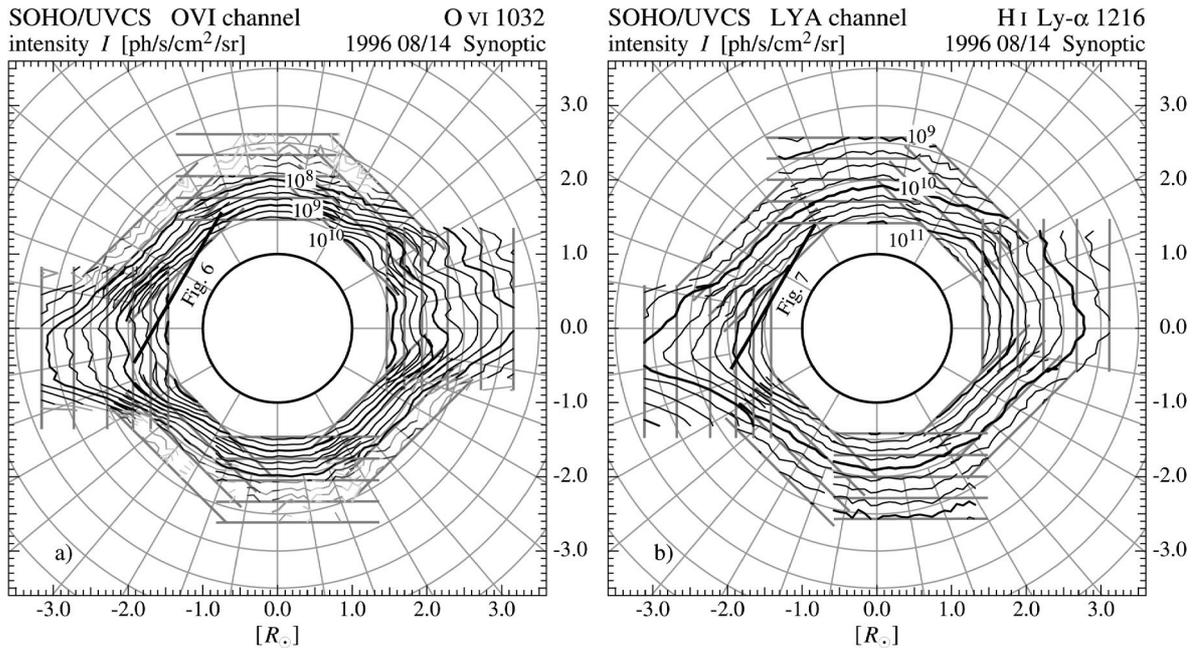


Fig. 8. Line intensity contours by the UVCS synoptic observation on 14 Aug. 1996: **a)** O VI 1032 Å and **b)** HI Ly- α . Grey bars represent slit positions. Also depicted are black bars corresponding to slit positions for Figs. 6 and 7.

same interpretation was also applied to a high-latitude hole observation at the time of the magnetic polarity reversal (Miralles et al. 2001a).

For O VI profiles observed within stable polar coronal holes during activity minimum, however, the account due to a contamination by a nearby streamer is not plausible. This situation is similar to that of Kohl et al. (1997, 1999a, 1999b). In such cases, both narrow and broad components must have a coronal hole origin. As can be recognised from Fig. 1 in Kohl et al. (1999a) and Figs. 1 and 2 in Kohl et al. (1999b), the O VI 1032/1037 ratio is larger between narrow components (>3) than between broad components (~ 2) and the fraction of the broad component intensity grows with altitude. At heliocentric height $2.0 R_{\odot}$, the entire narrow component can be interpreted as straylight (Kohl et al. 1997). Also, plumes and interplume regions, when dealt with separately, show different narrow/broad component intensity ratios; the broad component is a higher fraction of the total intensity within interplume regions than in plumes (Kohl et al. 1997).

The presence of the broad wing in coronal hole O VI profiles deserves attention in several respects. First, analyses have shown that the O VI line width is larger in interplume regions than in plumes (Antonucci et al. 1997; Noci et al. 1997; Giordano et al. 2000). This can merely reflect the difference in broad component fraction of the line intensity between both coronal regions; a higher broad-component fraction accounts for the reported larger line width there, when one disregards the presence of multiple components, as in past analyses.

Second, in pursuit of the source location of the fast wind, outflow velocity has been investigated for plumes and interplume regions (Giordano et al. 2000; Teriaca et al. 2003). One of the decisive parameters there is the O VI 1032/1037 line ratio. Important to note here is that, as stated above, narrow and broad component pairs yield rather different values for this ratio. Also, for observations at lower heights ($1.7 R_{\odot}$), Kohl et al. (1997) argued the possibility that dense polar plumes could be the source region of the narrow component and interplume regions that of the broad one. Lower values of the O VI line ratio, which would

result from collecting pure interplume contributions (i.e. broad component), indicate the occurrence of the Doppler dimming, implying a higher outflow velocity. This would reinforce the hypothesis that the fast wind emanates from interplume regions; the resulting outflow velocity estimate might become higher than those reported.

Third, the broad wing in coronal hole profiles was not incorporated appropriately in the existing empirical models (Kohl et al. 1999a; Cranmer et al. 1999; Zangrilli et al. 2002; Antonucci et al. 2004). Obviously, one major reason for this is that less luminous coronal hole emissions yield count numbers that are not high enough to resolve profiles into multiple components, unless integration time is sufficient (say, more than several hours). Also a large spectral binning (0.06 nm) used in the daily synoptic observations, which was employed as input to the models of Kohl et al. (1999a) and Cranmer et al. (1999), prohibits resolving multiple components. We need empirical determination of the altitudinal (and latitudinal) changes of coronal hole O VI profiles with sufficient resolution for resolving multiple components.

Fourth, it is interesting to note that *in-situ* solar wind measurements also revealed high energy tails or non-Maxwellian shapes of heavy ion velocity distributions, as well as of electron and proton ones (Marsch 1991; Feldman & Marsch 1997 and references therein). Although there is a preliminary study in this regard (von Steiger & Zurbuchen 2003), the relationship between optically observed coronal hole profiles and particle distribution functions measured *in-situ* needs further investigation.

Acknowledgements. I am grateful to Prof. G. Noci, and to Drs. G. Poletto, M. Romoli, L. Maccari and L. Zangrilli for their kind instructions and hospitality during my stay at Arcetri, Firenze, where part of this work was initiated. I thank the referee for his/her careful comments that helped to improve this paper.

References

- Antonucci, E., Giordano, S., Benna, C., et al. 1997, Signature of Open Magnetic Field Lines in the Extended Solar Corona and of Solar Wind Acceleration. in Proc. Fifth SOHO Workshop, The Corona and Solar Wind Near Minimum Activity, ed. A. Wilson, The Netherlands, ESA Publications Division, 175

- Antonucci, E., Dodero, M. A., Giordano, S., Krishnakumar, V., & Noci, G. 2004, *A&A*, 416, 749
- Brosius, J. W., Davila, J. M., Thomas, R. J., et al. 1996, *ApJS*, 106, 143
- Cosmo, M. L., Smith, P. L., Atkins, N., et al. 1999, *Proc. SPIE*, 3764, 161
- Cranmer, S. R., Kohl, J. L., Noci, G., et al. 1999, *ApJ*, 511, 481
- Del Zanna, G., & Bromage, B. J. I. 1999, *J. Geophys. Res.*, 104, 9753
- Dobrzycka, D., Cranmer, S. R., Panasyuk, A. V., Strachan, L., & Kohl, J. L. 1999a, *J. Geophys. Res.*, 104, 9791
- Dobrzycka, D., Strachan, L., Miralles, M. P., Kohl, J. L., & Gardner, L. 1999b, *Space Sci. Rev.*, 87, 177
- Feldman, W. C., & Marsch, E. 1997, in *Cosmic Winds and the Heliosphere*, ed. J. R. Jokipii, C. P. Sonett, & M. P. Giampapa (Tucson: U. Arizona Press), 617
- Frazin, R. A., Ciaravella, A., Dennis, E., et al. 1999, *Space Sci. Rev.*, 87, 189
- Gardner L. D., Kohl J. L., Daigneau P. S., et al. 1996, *Proc. SPIE*, 2831, 2
- Gardner, L. D., Smith, P. L., Kohl, J. L., et al. 2002, *ISSI SR-002*, 161
- Giordano, S., Antonucci, E., Noci, G., Romoli, M., & Kohl, J. L. 2000, *ApJ*, 531, L79
- Han, T. S., Kobayashi, K. 2002, *Mathematics of Information and Coding* (Providence, RI: Am. Math. Soc.)
- Harrison, R. A., Sawyer, E. C., Carter, M. K., et al. 1995, *Sol. Phys.*, 162, 233
- Harrison, R. A., Fludra, A., Pike, C. D., et al. 1997, *Sol. Phys.*, 175, 123
- Hassler, D. M., Wilhelm, K., Lemaire, P., & Schüle, U. 1997, *Sol. Phys.*, 175, 375
- Kohl, J. L., & Withbroe, G. L. 1982, *ApJ*, 256, 263
- Kohl, J. L., Esser, R., Gardner, L. D., et al. 1995, *Sol. Phys.*, 162, 313
- Kohl, J. L., Strachan, L., & Gardner, L. D. 1996, *ApJ*, 465, L141
- Kohl, J. L., Noci, G., Antonucci, E., et al. 1997, *Sol. Phys.*, 175, 613
- Kohl, J. L., Noci, G., Antonucci, E., et al. 1998, *ApJ*, 501, L127
- Kohl, J. L., Esser, R., Cranmer, S. R., et al. 1999a, *ApJ*, 510, L59
- Kohl, J. L., Fineschi, S., Esser, R., et al. 1999b, *Space Sci. Rev.*, 87, 233
- Li, X., Habbal, S. R., Kohl, J. L., & Noci, G. 1998, *ApJ*, 501, L133
- Marocchi, D., Antonucci E., & Giordano, S. 2001, *Ann. Geophys.*, 19, 135
- Marsch, E. 1991, in *Physics of the Inner Heliosphere 2: Particles, Waves and Turbulence*, ed. R. Schwenn, & E. Marsch (Berlin: Springer-Verlag), 45
- Mason, H. E., Young, P. R., Pike, C. D., et al. 1997, *Sol. Phys.*, 175, 143
- Miralles, M. P., Strachan, L., Gardner, L. D., et al. 1999, *Space Sci. Rev.*, 87, 277
- Miralles, M. P., Cranmer, S. R., Kohl, J. L. 2001a, *ApJ*, 560, L193
- Miralles, M. P., Cranmer, S. R., Panasyuk, A. V., et al. 2001b, *ApJ*, 549, L257
- Miralles, M. P., Cranmer, S. R., Kohl, J. L. 2002, *Cyclical Variations in the Plasma Properties of Coronal Holes*, in *Proc. SOHO 11 Symposium, From Solar Min to Max: Half a Solar Cycle with SOHO*, ed. A. Wilson, The Netherlands: ESA Publications Division, 351
- Moran, T. G. 2003, *ApJ*, 598, 657
- Nakagawa, A. 2003, *Opt. Rev.*, 10, 47
- Noci, G. 2000, in *Plasmas in the Universe*, ed. B. Coppi, A. Ferrari, & E. Sindoni (Amsterdam: IOS Press), 227
- Noci, G., Kohl, J. L., & Withbroe, G. L. 1987, *ApJ*, 315, 706
- Noci, G., Kohl, J. L., Antonucci, E., et al. 1997, *The Quiescent Corona and Slow Solar Wind*, in *Proc. Fifth SOHO Workshop, The Corona and Solar Wind Near Minimum Activity*, ed. A. Wilson, The Netherlands: ESA Publications Division, 75
- Ofman, L. 2000, *Geophys. Res. Lett.*, 27, 2885
- Parenti, S., Bromage, B. J. I., Poletto, G., et al. 2000, *A&A*, 363, 800
- Peres, G., Ciaravella, A., Betta, R., et al. 1997, *SOHO Observations of the North Polar Solar Wind*, in *Proc. Fifth SOHO Workshop, The Corona and Solar Wind Near Minimum Activity*, ed. A. Wilson, The Netherlands: ESA Publications Division, 587
- Pottasch, S. R. 1963, *ApJ*, 137, 945
- Raymond, J. C., & Doyle, J. G. 1981, *ApJ*, 247, 686
- Raymond, J. C., Kohl, J. L., Noci, G., et al. 1997, *Sol. Phys.*, 175, 645
- Rissanen, J. 1989, *Stochastic Complexity in Statistical Inquiry* (Singapore: World Scientific)
- Spadaro, D., Ventura, R., Cimino, G., et al. 2005, *A&A*, 429, 353
- von Steiger, R., & Zurbuchen, T. H. 2003 *Temperature Anisotropies of Heavy Solar Wind Ions from Ulysses-SWICS*, in *Solar Wind Ten*, ed. M. Velli, R. Bruno, & F. Malara, New York: AIP, 526
- Strachan, L., Panasyuk, A. V., Dobrzycka, D., et al. 2000, *J. Geophys. Res.*, 105, 2345
- Suleiman, R. M., Kohl, J. L., Panasyuk, A. V., et al. 1999, *Space Sci. Rev.*, 87, 327
- Teriaca, L., Poletto, G., Romoli, M., & Biesecker, D. A. 2003, *ApJ*, 588, 566
- Tu, C.-Y., Marsch, E., Wilhelm, K., & Curdt, W. 1998, *ApJ*, 503, 475
- Vásquez, A. M., & Raymond, J. C. 2005, *ApJ*, 619, 1132
- Wilhelm, K., Curdt, W., Marsch, E., et al. 1995, *Sol. Phys.*, 162, 189
- Wilhelm, K. K., Lemaire, P., Curdt, W., et al. 1997, *Sol. Phys.*, 175, 75
- Withbroe, G. L., Kohl, J. L., Weiser, H., & Munro, R. H. 1982, *Space Sci. Rev.*, 33, 17
- Zangrilli, L., Nicolosi, P., Poletto, et al. 1999, *A&A*, 342, 592
- Zangrilli, L., Poletto, G., Nicolosi, P., Noci, G., Romoli, M. 2002, *ApJ*, 574, 477

Online Material

Appendix A: Error propagation

In the main text, we treat a curve-fit problem to a Gaussian function with/without an additive constant, assuming α distribution for count number statistics. The purpose of this appendix is to provide analytical formulae of error propagation from count numbers to derived profile parameters, for the Gaussian-fit problem with α distribution.

It is not unusual that the actual setting of measurement parameters is a compromise among various observational restrictions. The theoretical formulae derived below can be used not only for obtaining error estimates of profile parameters, but also for finding an appropriate arrangement of measurement parameters prior to measurements.

A.1. Least-square method

We establish the notation here by introducing the least-square method: for a given set of measurements, $(\lambda_i, C_i)_i$, to determine parameters, p_j , of a fit curve, $f(\lambda)$, by minimising a square-sum

$$S = \sum_i \left[\frac{C_i - f(\lambda_i)}{\sigma_i} \right]^2,$$

where σ_i is the standard deviation of C_i . The condition for p_j is expressed as the so-called ‘‘normal equation’’. Using the matrices defined by

$$A_{ij} = \frac{1}{\sigma_i} \frac{\partial f(\lambda_i)}{\partial p_j}$$

$$y_i = \frac{C_i}{\sigma_i}$$

$$x_j = p_j$$

we write the normal equation as

$$\tilde{A}y = \tilde{A}Ax, \quad (\text{A.1})$$

where \tilde{A} denotes the transposed matrix of A . A convenient way to solve the normal equation is to use the LR decomposition of the matrix $A = LR$, where R is an upper-triangular, square matrix and L is a rectangular matrix whose columns form an orthonormal set. Therefore, the normal Eq. (A.1) can be immediately solved as

$$x = R^{-1} \tilde{L}y. \quad (\text{A.2})$$

Now, one can derive a formula representing the error propagation from the observed values y to the fit parameters x . We assume a condition that the variables, C_i 's, are mutually independent; this can be written in a matrix notation as $\langle (y - \hat{y})(y - \hat{y}) \rangle = I$, where $\hat{\cdot}$ represents a ‘‘true’’ value and $\langle \cdot \rangle$ denotes an average taken over measurements. When this condition holds and with the help of (A.2), one may obtain

$$\begin{aligned} \langle (x - \hat{x})(x - \hat{x}) \rangle &= R^{-1} \tilde{L} \langle (y - \hat{y})(y - \hat{y}) \rangle LR^{-1} \\ &= (\tilde{R}R)^{-1} = (\tilde{A}A)^{-1}. \end{aligned}$$

The error matrix $(\tilde{A}A)^{-1}$ is symmetric, but not necessarily diagnostic.

A.2. Error matrix

One of the difficulties in obtaining analytical formulae of $(\tilde{A}A)^{-1}$ arises from the treatment of the α distribution; we resolve this by employing the power series expansion.

As a fit function, $f(\lambda)$, adopted here is a Gaussian function having an additive constant:

$$\begin{aligned} f(\lambda) &= C_0 + Cg(\lambda) \\ g(\lambda) &= \frac{1}{\sqrt{2\pi}} \frac{\delta}{\Delta\lambda} \exp\left[-\frac{(\lambda - \lambda_0)^2}{2(\Delta\lambda)^2}\right], \end{aligned} \quad (\text{A.3})$$

where $\delta \equiv \lambda_i - \lambda_{i-1}$ denotes the wavelength interval. Thus, parameters to be determined p_0, p_1, p_2 , and p_3 are C_0, C, λ_0 , and $\Delta\lambda$, respectively. The reason to incorporate the effect of the additive constant, C_0 , is that it may considerably affect the magnitude of error propagation. Also, even for a multiple Gaussian case, the effect on the primary component could be closely approximated by replacing the remaining components with C_0 .

From this point we assume that λ_i 's are located so densely, as compared to $\Delta\lambda$, that a summation over the model Gaussian function can be approximated by integration, so that we may, for example, have

$$\sum_i g(\lambda_i) \sim \int dt \frac{1}{\sqrt{2\pi}\Delta\lambda} \exp\left[-\frac{(t - \lambda_0)^2}{2(\Delta\lambda)^2}\right] = 1.$$

This assumption can be regarded as fulfilled for observations properly designed for a target line spectrum. Under the same premise, we may also have

$$\begin{aligned} \sum_i \left(\frac{\lambda_i - \lambda}{\Delta\lambda}\right)^{2m} [g(\lambda_i)]^n &\sim \int dt \left(\frac{t - \lambda}{\Delta\lambda}\right)^{2m} [g(t)]^n \\ &= \left(\frac{1}{\sqrt{2\pi}} \frac{\delta}{\Delta\lambda}\right)^{n-1} \frac{(2m-1)!!}{n^{m+(1/2)}}. \end{aligned} \quad (\text{A.4})$$

Let us start from the evaluation of $\tilde{A}A$. For the fit function $f(\lambda)$ of (A.3), we have

$$\begin{aligned} \frac{\partial f}{\partial C_0} &= 1 \\ \frac{\partial f}{\partial C} &= g(\lambda) \\ \frac{\partial f}{\partial \lambda_0} &= \frac{C}{\lambda} \frac{\lambda - \lambda_0}{\Delta\lambda} g(\lambda) \\ \frac{\partial f}{\partial \Delta\lambda} &= \frac{C}{\lambda} \left[\left(\frac{\lambda - \lambda_0}{\Delta\lambda}\right)^2 - 1 \right] g(\lambda). \end{aligned}$$

Thus, for example, the 00-component of $\tilde{A}A$ satisfies

$$(\tilde{A}A)_{00} = \sum_i \frac{1}{\sigma_i^2} = \sum_i \frac{1}{C_i(1 + \alpha^2 C_i)}$$

(also recall the definition (5) of α distribution). Here the adoption of $C_i \sim C_0 + Cg(\lambda_i)$ leads to

$$(\tilde{A}A)_{00} = \sum_i \frac{1}{[C_0 + Cg(\lambda_i)] [1 + \alpha^2 C_0 + \alpha^2 Cg(\lambda_i)]}.$$

We have to notice here that the summation (or its equivalent in the integration form) in the preceding expression cannot be evaluated analytically, because the Gaussian, $g(\lambda)$, appears in the denominator. Instead, by using

$$\begin{aligned} \bar{u} &= \frac{\alpha^2 C}{1 + \alpha^2 C_0} \\ \bar{v} &= \frac{C}{C_0} \end{aligned}$$

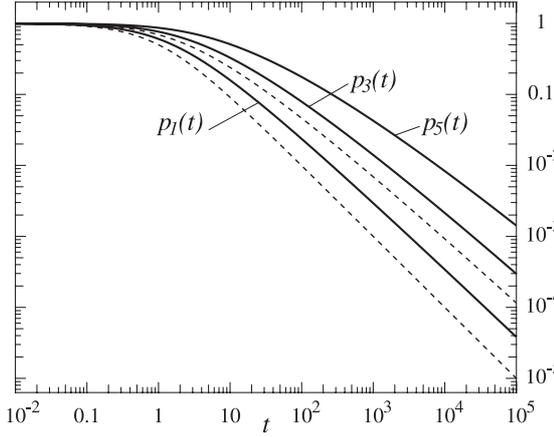


Fig. A.1. Profiles of functions $p_1(t)$, $p_3(t)$, and $p_5(t)$. Also plotted are $p_0(t) = 1/(1+t)$ and $p_2(t) = \log(1+t)/(-t)$.

we resort to the series expansion as

$$\begin{aligned} \frac{1}{[1 + \alpha^2 C_0 + \alpha^2 C g(\lambda_i)]} &= \frac{1}{(1 + \alpha^2 C_0) [1 + \bar{u} g(\lambda_i)]} \\ &\sim \frac{1}{(1 + \alpha^2 C_0)} [1 - \bar{u} g + (\bar{u} g)^2 - (\bar{u} g)^3 + \dots] \\ \frac{1}{[C_0 + C g(\lambda_i)]} &= \frac{1}{C_0 [1 + \bar{v} g(\lambda_i)]} \\ &\sim \frac{1}{C_0} [1 - \bar{v} g + (\bar{v} g)^2 - (\bar{v} g)^3 + \dots]. \end{aligned}$$

Therefore, by collecting terms of same order in g , we obtain

$$\begin{aligned} &\frac{1}{[1 + \alpha^2 C_0 + \alpha^2 C g(\lambda_i)] [C_0 + C g(\lambda_i)]} \\ &\sim \frac{1}{C_0 (1 + \alpha^2 C_0)} [1 - (\bar{u} + \bar{v}) g + (\bar{u}^2 + \bar{u}\bar{v} + \bar{v}^2) g^2 + \dots] \\ &= \frac{1}{C_0 (1 + \alpha^2 C_0)} \sum_{k=0}^{\infty} \frac{\bar{u}^{k+1} - \bar{v}^{k+1}}{\bar{u} - \bar{v}} [-g(\lambda_i)]^k \\ &= \frac{1}{C_0 (1 + \alpha^2 C_0)} - \frac{1}{C} \sum_{k=1}^{\infty} [\bar{u}^{k+1} - \bar{v}^{k+1}] [-g(\lambda_i)]^k. \end{aligned}$$

Since they are the summation over the powers of $g(\lambda_i)$, the above expressions can now be evaluated analytically with the aid of (A.4). In fact, by introducing a function

$$p_1(t) = 1 - \frac{t}{\sqrt{2}} + \frac{t^2}{\sqrt{3}} - \frac{t^3}{\sqrt{4}} + \dots$$

and variables

$$\begin{aligned} u &= \frac{1}{\sqrt{2\pi}} \frac{\delta}{\Delta\lambda} \bar{u} = \frac{1}{\sqrt{2\pi}} \frac{\delta}{\Delta\lambda} \frac{\alpha^2 C}{1 + \alpha^2 C_0} \\ v &= \frac{1}{\sqrt{2\pi}} \frac{\delta}{\Delta\lambda} \bar{v} = \frac{1}{\sqrt{2\pi}} \frac{\delta}{\Delta\lambda} \frac{C}{C_0}, \end{aligned} \quad (\text{A.5})$$

we may write the resultant expression simply as

$$(\tilde{A}A)_{00} = \frac{N}{C_0 (1 + \alpha^2 C_0)} + \frac{1}{C} [\bar{u}^2 p_1(u) - \bar{v}^2 p_1(v)],$$

where N denotes the number of data points.

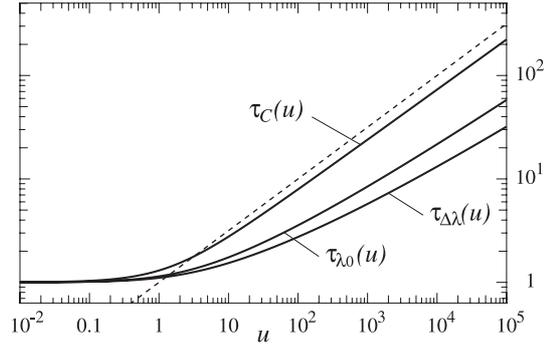


Fig. A.2. Profiles of the model functions $\tau(u)$'s representing relative errors of profile parameters C , λ_0 and $\Delta\lambda$.

In a similar way, by using functions

$$\begin{aligned} p_3(t) &= 1 - \frac{t}{\sqrt{2^3}} + \frac{t^2}{\sqrt{3^3}} - \frac{t^3}{\sqrt{4^3}} + \dots \\ p_5(t) &= 1 - \frac{t}{\sqrt{2^5}} + \frac{t^2}{\sqrt{3^5}} - \frac{t^3}{\sqrt{4^5}} + \dots, \end{aligned}$$

we may also have

$$\begin{aligned} (\tilde{A}A)_{11} &= \frac{1}{C} [p_1(u) - p_1(v)] \\ (\tilde{A}A)_{22} &= \frac{C}{(\Delta\lambda)^2} [p_3(u) - p_3(v)] \\ (\tilde{A}A)_{33} &= \frac{C}{(\Delta\lambda)^2} \{3[p_5(u) - p_5(v)] \\ &\quad - 2[p_3(u) - p_3(v)] + [p_1(u) - p_1(v)]\} \\ (\tilde{A}A)_{01} &= (\tilde{A}A)_{10} = -\frac{1}{C} [\bar{u} p_1(u) - \bar{v} p_1(v)] \\ (\tilde{A}A)_{03} &= (\tilde{A}A)_{30} = -\frac{1}{\Delta\lambda} \{[\bar{u} p_3(u) - \bar{v} p_3(v)] - [\bar{u} p_1(u) - \bar{v} p_1(v)]\} \\ (\tilde{A}A)_{13} &= (\tilde{A}A)_{31} = \frac{1}{\Delta\lambda} \{[p_3(u) - p_3(v)] - [p_1(u) - p_1(v)]\}, \end{aligned}$$

with all the other elements being zero.

The above procedure using power-series expansion can be justified because respective functions are analytical, having finite radius of convergence. Shown in Fig. A.1 are $p_1(t)$, $p_3(t)$, and $p_5(t)$, as well as $p_0(t) = 1/(1+t)$ and $p_2(t) = \log(1+t)/t$ (functions $p_j(t)$ defined above are related to the so-called polylogarithmic function as $p_j(t) = \text{Polylog}(j/2, -t)/(-t)$). These are monotonically decreasing for $t \geq 0$, converging to zero for $t \rightarrow \infty$.

A.3. Single Gaussian

Let us evaluate $(\tilde{A}A)^{-1}$ for a simplified version, namely for the case of single Gaussian: $C_0 = 0$ in (A.3). We thus need not consider the elements related to C_0 . In the $C_0 \rightarrow 0$ limit (or equivalently $v \rightarrow \infty$), $p_1(v) = p_3(v) = p_5(v) = 0$. Hence, we may have

$$\begin{aligned} (\tilde{A}A)_{11} &= \frac{1}{C} p_1(u) \\ (\tilde{A}A)_{22} &= \frac{C}{(\Delta\lambda)^2} p_3(u) \\ (\tilde{A}A)_{33} &= \frac{C}{(\Delta\lambda)^2} [3p_5(u) - 2p_3(u) + p_1(u)] \end{aligned}$$

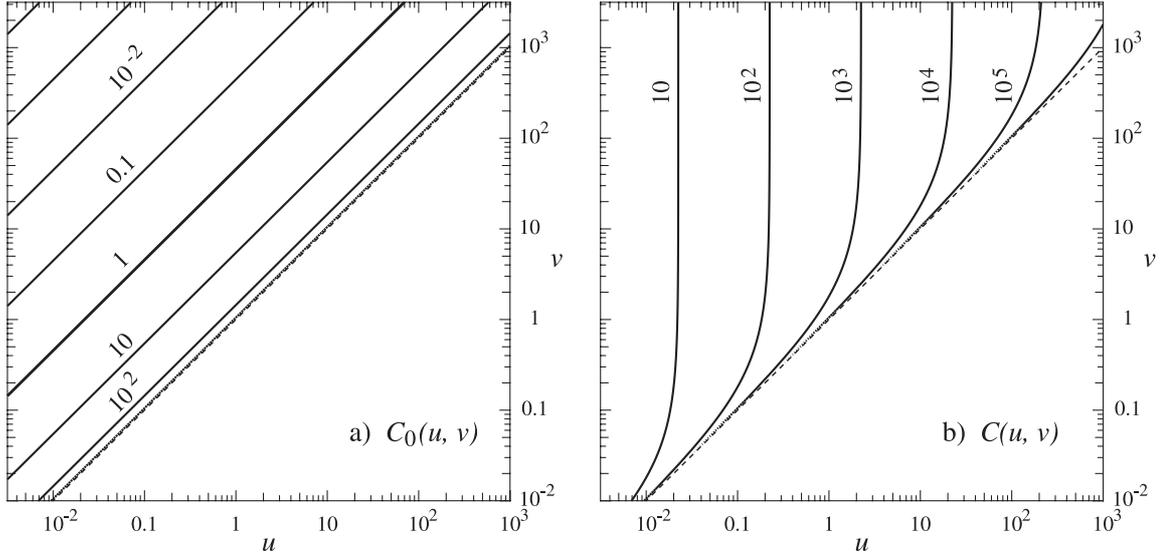


Fig. A.3. Contours of **a)** $C_0(u, v)$ and **b)** $C(u, v)$ for $\alpha = 0.15$ and $\Delta\lambda/\delta = 4$.

$$(\tilde{A}A)_{13} = (\tilde{A}A)_{31} = \frac{1}{\Delta\lambda} [p_3(u) - p_1(u)]$$

with

$$u = \frac{1}{\sqrt{2\pi}} \frac{\delta}{\Delta\lambda} \alpha^2 C. \quad (\text{A.6})$$

The inverse, $(\tilde{A}A)^{-1}$, gives the formulae for the error propagation of parameters, C , λ_0 , and $\Delta\lambda$. With the functions defined as

$$\tau_C(u) = \sqrt{\frac{3p_5(u) - 2p_3(u) + p_1(u)}{3p_5(u)p_1(u) - p_3(u)^2}}$$

$$\tau_{\lambda_0}(u) = \sqrt{\frac{1}{p_3(u)}}$$

$$\tau_{\Delta\lambda}(u) = \sqrt{\frac{2p_1(u)}{3p_5(u)p_1(u) - p_3(u)^2}}$$

we write the *relative errors* of these parameters as

$$\frac{\sqrt{[(\tilde{A}A)^{-1}]_{11}}}{C} = \frac{\tau_C(u)}{\sqrt{C}}$$

$$\frac{\sqrt{[(\tilde{A}A)^{-1}]_{22}}}{\Delta\lambda} = \frac{\tau_{\lambda_0}(u)}{\sqrt{C}}$$

$$\frac{\sqrt{[(\tilde{A}A)^{-1}]_{33}}}{\Delta\lambda} = \frac{\tau_{\Delta\lambda}(u)}{\sqrt{2C}}.$$

As expected, the dependence of relative errors on the total count number are inverse-proportional to its square ($\propto 1/\sqrt{C}$). The effect of the α distribution is expressed as functions, $\tau(u)$'s. Characteristic here is that this effect emerges as a function of $(\delta/\Delta\lambda)C$, through u of (A.6). For the case of Poisson statistics (that is, $\alpha = 0$, or $u = 0$), $\tau_C(u) = \tau_{\lambda_0}(u) = \tau_{\Delta\lambda}(u) = 1$. In contrast, for the case of $\alpha > 0$, these functions are larger than 1, as Fig. A.2 displays, indicating a departure from the Poisson statistics. However, it is only when $u \gtrsim 1$, or

$$C \gtrsim \sqrt{2\pi} \frac{\Delta\lambda}{\delta} \frac{1}{\alpha^2},$$

that the deviation from the value of the Poisson statistics becomes significant (C , λ_0 and $\Delta\lambda$ are in decreasing order in this deviation). Note here that the increase (or gradient) in $\tau(u)$ does not exceed that of $\sqrt{u} \propto \sqrt{C}$ (depicted as a broken line in Fig. A.2).

By utilising the above results, we may give a rough estimate on the error bound of parameters as

$$\max\left(\frac{1}{\sqrt{C}}, \sqrt{\frac{1}{\sqrt{2\pi}} \frac{\delta}{\Delta\lambda} \alpha}\right)$$

if the measurement error is described by the α distribution and a fit function is single Gaussian. The former one is effective for $u < 1$. The latter is a substitution for the value of $u > 1$ with that of $u = 1$; this is given by substituting $C = \sqrt{2\pi}(\Delta\lambda/\delta)(1/\alpha^2)$, obtained from $u = 1$, into $\tau(u)/\sqrt{C} \sim 1/\sqrt{C}$. For the UVCS measurement, the latter is of the same order as (more specifically, a fraction of) α , because in normal conditions the value of $\Delta\lambda/\delta$ is between 1 and 10. This means that to attain the relative error of less than 10%, the total count number, C , should exceed several hundreds.

A.4. Single Gaussian with constant

We here consider the case of fit functions having a non-vanishing constant term ($C_0 \neq 0$). The relation between (u, v) and (C_0, C) , given by (A.5), is examined first. As an example, by setting $\alpha = 0.15$ and $\Delta\lambda/\delta = 4$ (these correspond to a typical situation for the UVCS measurement by its LYA channel with the spectral binning being 1), we depict the contour maps of

$$C_0 = \frac{u}{v-u} \frac{1}{\alpha^2}$$

$$C = \sqrt{2\pi} \frac{\Delta\lambda}{\delta} \frac{uv}{(v-u)\alpha^2} = \sqrt{2\pi} \frac{\Delta\lambda}{\delta} v C_0$$

in Figs. A.3a and A.3b, respectively. Here, $u \geq v$ corresponds to the domain of invalid parameters, because the conditions $C_0 > 0$ and $C > 0$ necessarily yield $u < v$. As noticed from the figures, C_0 and C depend essentially on $v - u$ and u , respectively, apart from in the neighbourhood of $u = v$. Actual values of C_0 and C , however, are also determined by α and $\Delta\lambda/\delta$, along with u and v .

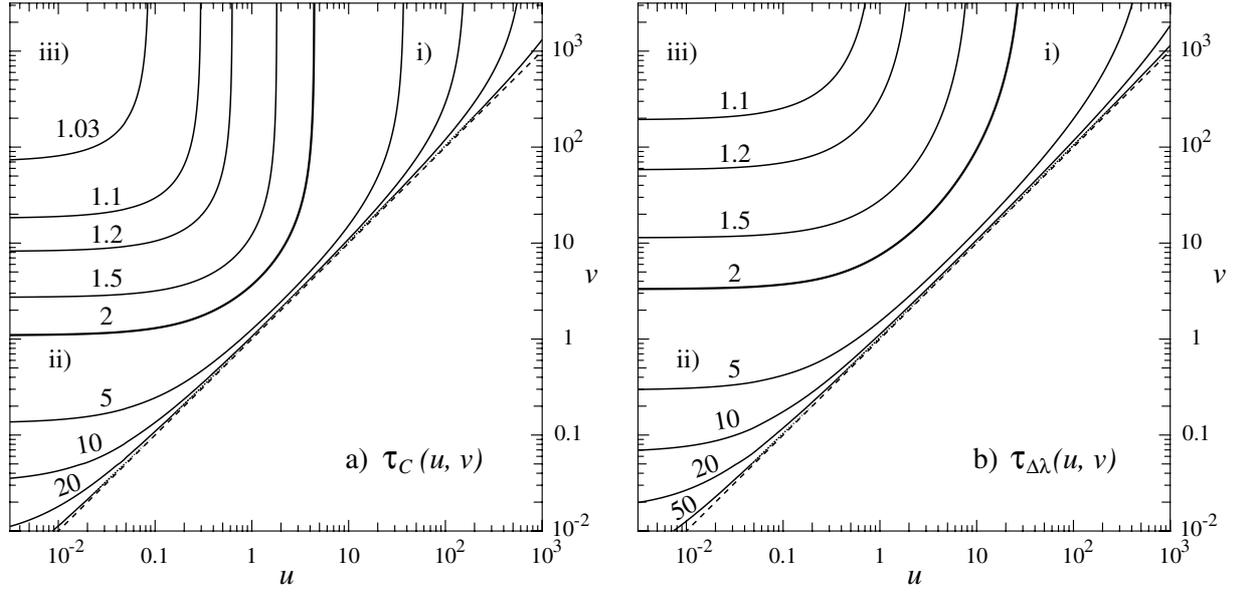


Fig. A.4. Contours of the model functions **a)** $\tau_C(u, v)$ and **b)** $\tau_{\Delta\lambda}(u, v)$ representing the relative errors of respective profile parameters, C and $\Delta\lambda$, for $N = 64$, $\alpha = 0.15$ and $\Delta\lambda/\delta = 4$.

Now, we return to the error matrix. Shown in Fig. A.4 are contour maps of

$$\tau_C(u, v) = \frac{\sqrt{[(\tilde{A}\tilde{A})^{-1}]_{11}}}{\Delta\lambda} \sqrt{C}$$

$$\tau_{\Delta\lambda}(u, v) = \frac{\sqrt{[(\tilde{A}\tilde{A})^{-1}]_{33}}}{\Delta\lambda} \sqrt{2C}$$

for the case of $N = 64$, $\alpha = 0.15$, and $\Delta\lambda/\delta = 4$. Again, $\tau(u, v)$ represents the relative error of profile parameters, with the modification of \sqrt{C} -dependence; τ is the multiplicative factor to the case of Poisson statistics.

We consider these contour maps by dividing the domain into three: i) $u \gtrsim 10$, ii) $v \lesssim 2$, and iii) the remaining domain (this division is explicable in the light of the fact that $\tilde{A}\tilde{A}$ is a function of $p_j(u) - p_j(v)$ and the functional forms of p_j). The domain iii) corresponds to the case approximated well by Poisson statistics, while the deviation from it is notable in the other two domains. To the case i) one can apply the results of the preceding subsection concerning the effects of the α distribution with assuming $C_0 = 0$. For example, the fit of Fig. 5 yields $(u, v) = (23.7, 58.5)$,

which belongs to this case i). For the case ii), by contrast, the effect of $C_0 \neq 0$ is significant. The value of $\Delta\lambda/\delta$ being typically within the range between 1 and 10 for the UVCS measurement, the condition for the non-zero C_0 to be ineffective in the parameter estimation is $C/C_0 > 5 \sim 50$ under the condition $v > 2$. If the ratio C/C_0 becomes lower than this value, the accuracy of Gaussian fit is reduced accordingly. In principle, in the domain ii), the accuracy of estimated profile parameters is not high owing to the low value of C there (see Fig. A.3b).

As examined above, with the use of u and v , instead of C_0 and C , one can consider i) the effect by the α distribution and ii) that of $C_0 \neq 0$ separately. Moreover, the amount of relative errors of profile parameters is, as a function of N , $\Delta\lambda/\delta$, α , u , and v , determined mostly by u and v ; its dependence on N , $\Delta\lambda/\delta$ and α is not strong (less than 10%). The exception is when N is not sufficient for expressing a Gaussian, namely, when the value of N is not higher than that of $\Delta\lambda/\delta$ (roughly speaking $N < 8\Delta\lambda/\delta$). In this case, the evaluation by (A.4) is inaccurate, and the formulae that succeed are thus not applicable.