CORONAL DIAGNOSTIC SPECTROMETER

SoHO

CDS SOFTWARE NOTE No. 21

Version 1.0 July 1, 1994

PROFILE FITTING TO CDS/SUMER DATA

N. Brynildsen

Institute of Theoretical Astrophysics
P.O.Box 1029 Blindern
N-0315 OSLO
NORWAY

INSTITUTE OF THEORETICAL ASTROPHYSICS
UNIVERSITY OF OSLO
1994

INTRODUCTION

This document describes the line fitting routines tested out for use on CDS spectra. In Oslo we have experience with Gaussian line fitting on High Resolution Telescope and Spectrograph (HRTS) data, both from rocket flights and from the Spacelab 2 Mission in 1985. CDS and HRTS spectra differs quite a lot when it come to spectral range and resolution and these differences influences the line fitting. Thus spectrograms from Solar EUV Rocket Telescope and Spectrograph - SERTS (supplied by Bill Thompson) were used to substitute for CDS spectra.

The following fitting routines have been applied:

- CURVEFIT from IDL'S USERLIB (see Bevington 1969 for method).
- AMOEBA from Numerical Recipes (IDL version by Bill Thompson).
- LSTSQR supplied by Bill Thompson.
- MOMENT, a routine that calculates the moments of a spectral line.

The fitted functions are Gaussian or combinations of Gaussian profiles. All routines depend on selecting a set of initial values for the parameters to be fitted.

In the following we describe the applicability of the fitting functions on both HRTS, CDS and SUMER data. This is done by a) fitting curves to observed spectrograms from HRTS and SERTS, and b) applying the fitting routines to simulated CDS and SUMER spectra, including a realistic model for the noise, which these data will contain. The latter procedure was used to check the uncertainty of our fitting results and the dependence on the choice of initial parameter values. Some simple non-fitting routines used to get line intensities are described in the appendix.

LINE FITTING ON HRTS SPECTRA

The line fitting of HRTS data has mainly been performed on the C IV line at 1548 Å emitted at 100 000 K in the solar transition region. This spectral line often shows a multi component line profile.

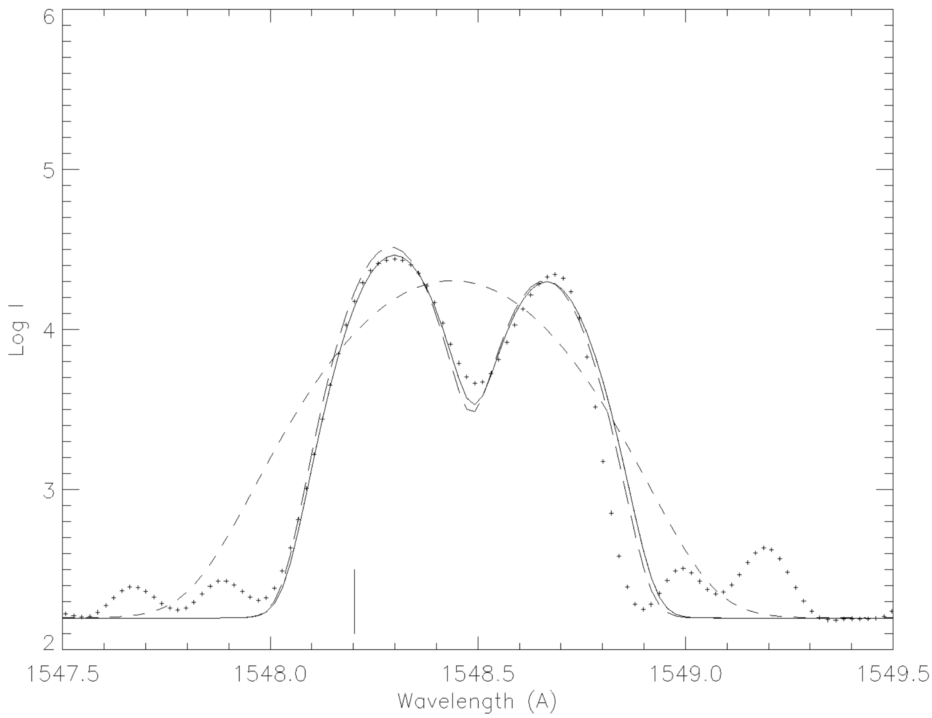

Figure 1. An example of a fitted C IV 1548 Å line with 2 line components from HRTS Spacelab 2 data. The laboratory wavelength (1548.202 Å) is marked in the plot. The observed spectrum is the dotted line, the solid line is fit with CURVEFIT and LSTSQR, long dashed lines is the AMOEBA fit and the short dashed line is the MOMENT fit. The parameters for the different fits are given in Table 1. The spectral window used for the fit is as shown in the figure i.e. $\lambda\lambda$=1547.5 - 1549.5 Å.

This spectral line has therefore been fitted with 1, 2 or 3 Gaussian line components. The IDL routine CURVEFIT has been used lately, earlier a FORTRAN program applying the same fitting procedures, from Bevington (1969), was used.

A typical HRTS C IV profile consists of 30-50 pixels. The pixel size of the HRTS data is 0.02 Å (i.e. 4 km s^{-1} at 1548 Å), and the Doppler width of the C IV line is typically 0.1 Å. These values differ strongly from the corresponding parameters in the CDS spectra.

Experience has shown that the fitting routine requires a shift between two line components of about one Doppler width in order to successfully fit them with two Gaussians. A line width of 0.1 Å at 1550 Å corresponds to a velocity of 20 km s^{-1}.

In Figure 1 and Table 1 a representative HRTS C IV profile is fitted with the four different fitting routines. We note that both CURVEFIT, LSTSQR and AMOEBA give reasonable fits. One would not expect a good fit for MOMENT to this line profile, which contains two distinct components.

TABLE 1

Parameter values for the various line fitting routines using HRTS data. A two component profile was fitted with the start parameters given in the first row. The fitted profiles are shown in Figure 1.

		1. line component		2. line component	
	Width	Log(I)	v (km s^{-1})	Log(I)	v (km s^{-1})
Start parameters	0.10	4.40	9.7	4.10	58.1
CURVEFIT	0.11	4.46	18.8	4.30	89.8
LSTSQR	0.11	4.46	18.8	4.30	89.9
AMOEBA	0.11	4.52	17.1	4.30	88.8
MOMENT	0.27	4.31	45.7		

LINE FITTING ON CDS SPECTRA

The pixel size of the NIS on CDS is 0.068 Å and 0.114 Å (about 60 km s^{-1}) in the second and first order, according to the CDS Blue Book. The typical Doppler line width of a line in this wavelength range is 0.05 Å. Using the results from the HRTS data, this means that two components must be separated by 0.05 Å to be fitted as separate line components. At a wavelength of 350 Å, 0.05 Å corresponds to a velocity of 45 km s^{-1}. The main problem, however, is the resolution. In the HRTS range a typical Doppler width is 5 times the resolution, but in the NIS range they are comparable.

The CDS test data, from SERTS, has 4 spectral lines in a 120 by 120 pixel raster; He II (303.78 Å), Fe XVI (335.40 Å), Fe XVI (360.76 Å) and Mg IX (368.06 Å). In none of these lines is it possible to distinguish multi component line profiles. The spectral lines, however, only have 3-5 pixels across the line profile. This means that spectral lines in the NIS range is most reasonably fitted with a single Gaussian line component. Results of the fits are shown in Figure 2 and Table 2.

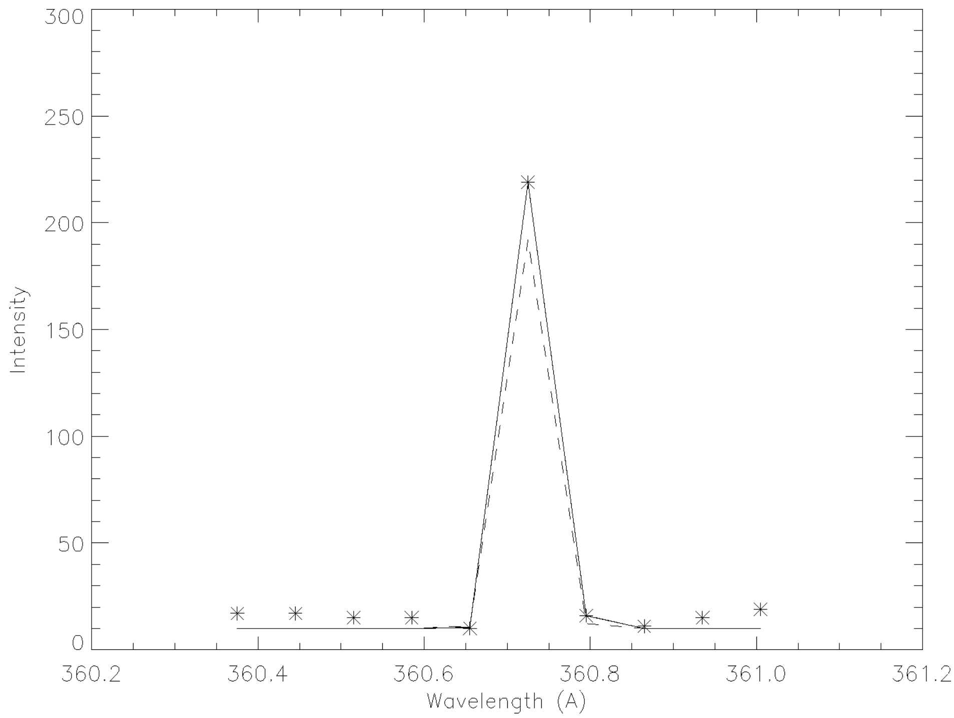

Figure 2. Example of a fitted Fe XVI line (360.76 Å) from CDS/SERTS test data, parameters for these fits are given in table 2. The line styles used are the same as in Figure 1, the only exception being that the observed data are marked with asterisks.

TABLE 2

Parameter values for the line fitting routines shown in Figure 2 from the CDS test data.

	Width	Int.	v (km s^{-1})
Start parameters	0.05	219	4.2
CURVEFIT	0.03	235	7.3
LSTSQR	0.03	240	8.1
AMOEBA	0.03	245	8.9
MOMENT	0.03	193	2.1

SYNTHETIC SPECTRA IN THE CDS RANGE

An IDL program was made to test the fitting of spectral line profiles with noise in the CDS wavelength range using *simulated* line profiles. Profiles with one or two line components were constructed. For the two component cases a range of relative component amplitude ratios and relative shifts were used. The total line intensities in counts/sec corresponds to typical values found in the CDS Blue Book. The exposure times selected were 1, 10 and 100 seconds. A 20 pixel wide window was used, and each line component had a fixed Doppler line width of 0.063 Å. The rest wavelength for the line was set to 554 Å. The width of a pixel was taken as 0.07 Å.

LSTSQR was used, but test runs with AMOEBA did not give significantly different results.

For a realistic simulation a background noise of 1 count/pixel/sec was added to the profiles. Furthermore photon noise was added. Figure 3 shows an example of a synthetic profile with and without noise.

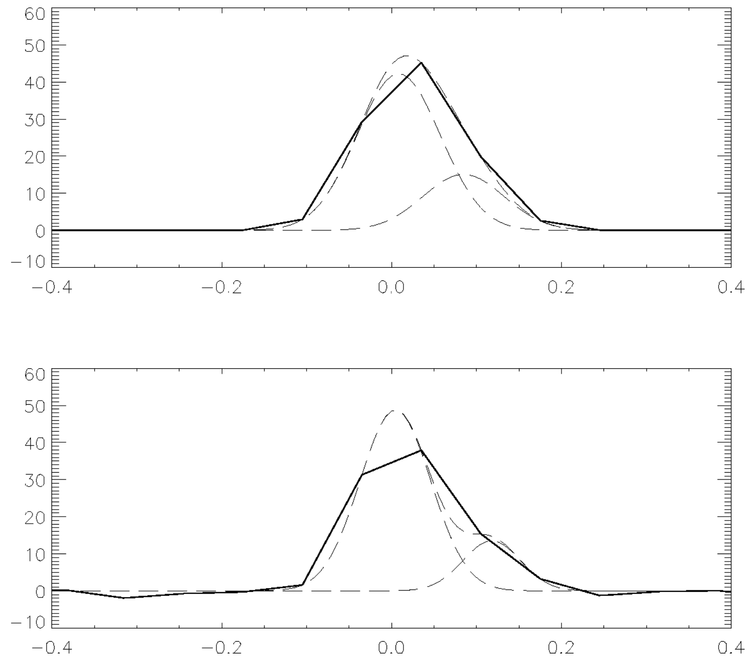

Figure 3. Example of a synthetic two component spectral line in the CDS range (solid line). The parameters in this figure corresponds to the first row in Table 3. The upper panel shows the profile with no noise and the lower shows the profile with noise added. The two fitted profiles using LSTSQR, and their sum, are also drawn with dashed lines.

In Figure 4 histograms for the distribution of fitted parameters are shown for one hundred profile fits with initial parameters as given in the first row for the first run in Table 3.

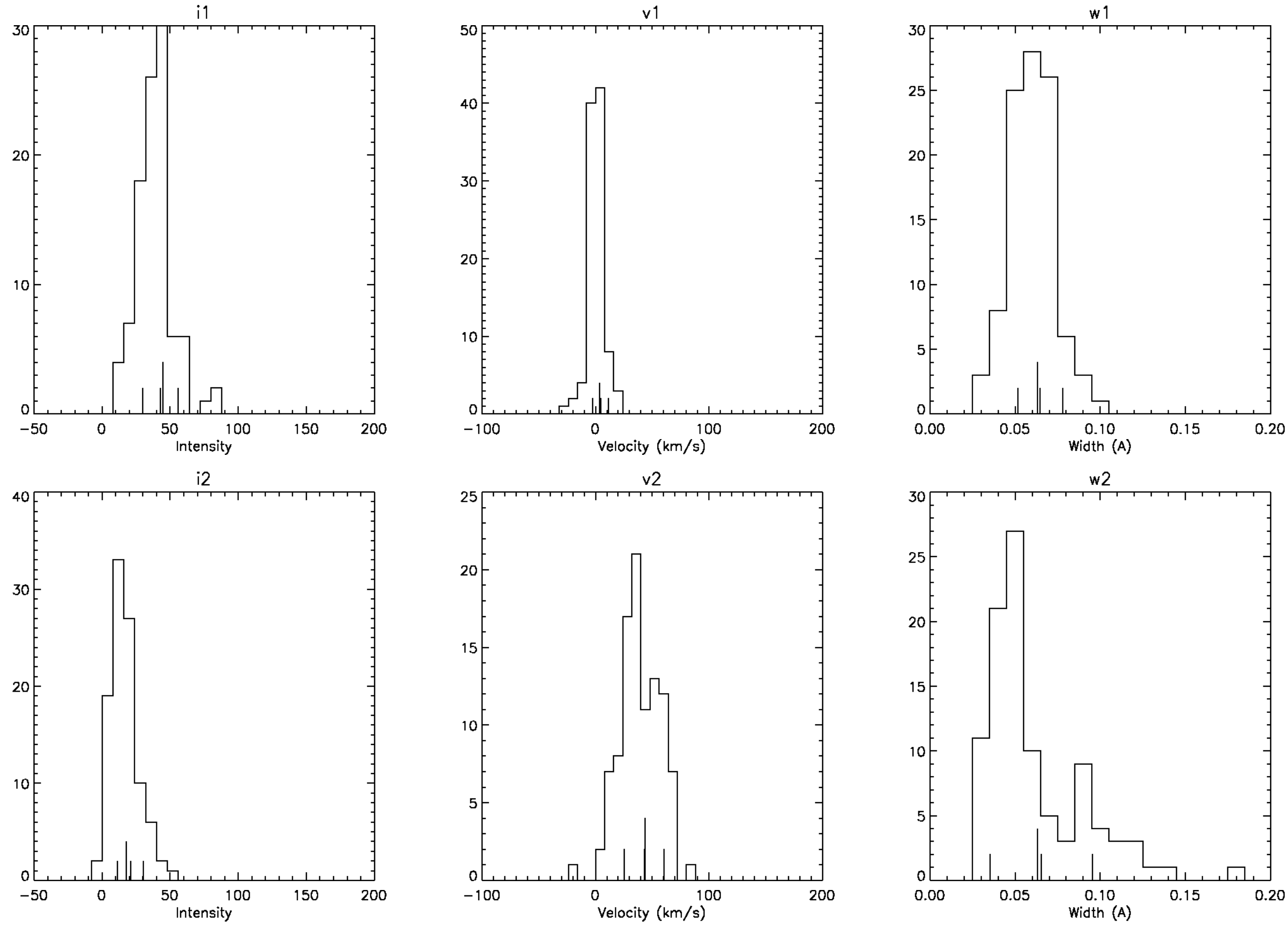

Figure 4. Histograms for the distribution of fitted parameters from simulating a set of one hundred line profiles fitted with parameter values as given in the first row for each run in Table 3. The parameters for a profile without noise and the mean and standard deviation for each parameter is also marked in the figure.

The resulting mean values (m) and standard deviations (s) were calculated for a set of one hundred fits for each selected set of line profile parameters for a given line profile configuration. Table 3, 4 and 5 give the results of some of these runs. They differ in the total counts and exposure times used. The parameters listed are the maximum intensities of the two line components (I_1 and I_2), the shifts of the components (v_1 and v_2) in km s^{-1} and the Doppler width of the components (w_1 and w_2) in Å.

TABLE 3

Results of synthetic line fitting for CDS. In the table each set of three rows give 1) the parameters of a two component line profile, 2) the mean value (m) and 3) the standard deviations (s) of each fitted parameter for one hundred fitted profiles with added noise. The difference from one set of parameters to the next, in the table, is that one parameter changes i.e. v_2 or I_2.

	I_1	v_1	w_1	I_2	v_2	w_2	*Total counts*	*Exp.time*	*Noise*
	44.8	3.8	0.063	17.9	43.6	0.063	100	1	1
m:	42.8	4.4	0.065	20.8	42.9	0.065			
s:	13.0	6.9	0.013	9.6	17.4	0.030			
	44.8	3.8	0.063	17.9	21.8	0.063	100	1	1
m:	43.4	1.6	0.066	21.7	27.0	0.051			
s:	9.1	3.9	0.013	6.7	9.5	0.016			
	44.8	3.8	0.063	17.9	10.9	0.063	100	1	1
m:	46.3	0.1	0.064	20.2	21.6	0.057			
s:	9.3	4.6	0.012	7.6	11.4	0.018			
	44.8	3.8	0.063	17.9	3.8	0.063	100	1	1
m:	47.2	3.7	0.063	18.7	2.3	0.062			
s:	5.7	4.6	0.007	6.4	18.0	0.023			
	31.3	3.8	0.063	31.3	43.6	0.063	100	1	1
m:	30.3	3.9	0.062	31.3	42.1	0.065			
s:	10.0	9.4	0.014	9.6	9.5	0.017			
	31.3	3.8	0.063	31.3	87.2	0.063	100	1	1
m:	30.8	2.7	0.065	29.4	87.2	0.068			
s:	7.5	5.8	0.016	5.6	7.6	0.013			
	31.3	3.8	0.063	31.3	3.8	0.063	100	1	1
m:	28.9	2.7	0.068	28.9	4.3	0.068			
s:	3.1	2.7	0.005	3.1	2.7	0.005			

TABLE 4

Results of synthetic line fitting for CDS. For further explanation see heading of Table 3.

	I_1	v_1	w_1	I_2	v_2	w_2	*Total counts*	*Exp.time*	*Noise*
	44.8	3.8	0.063	17.9	43.6	0.063	10	10	1
m:	50.8	5.0	0.057	27.2	49.2	0.051			
s:	18.2	9.7	0.022	14.5	24.3	0.054			
	44.8	3.8	0.063	17.9	21.8	0.063	10	10	1
m:	48.6	1.6	0.063	24.8	28.7	0.037			
s:	14.9	7.3	0.023	11.2	20.7	0.042			
	44.8	3.8	0.063	17.9	10.9	0.063	10	10	1
m:	49.4	-0.2	0.063	27.4	19.7	0.042			
s:	17.7	5.6	0.025	15.6	18.5	0.049			
	44.8	3.8	0.063	17.9	3.8	0.063	10	10	1
m:	56.0	3.4	0.053	25.0	3.2	0.049			
s:	14.3	7.4	0.015	12.3	22.7	0.037			
	31.3	3.8	0.063	31.3	43.6	0.063	10	10	1
m:	39.7	2.8	0.060	41.6	45.5	0.050			
s:	15.3	10.2	0.028	18.0	18.2	0.029			
	31.3	3.8	0.063	31.3	87.2	0.063	10	10	1
m:	36.6	1.0	0.063	37.4	88.7	0.061			
s:	16.7	11.6	0.042	14.0	12.8	0.031			
	31.3	3.8	0.063	31.3	3.8	0.063	10	10	1
m:	33.7	1.8	0.062	33.7	1.8	0.062			
s:	10.7	5.7	0.019	10.7	5.7	0.019			

TABLE 5

Results of synthetic line fitting for CDS. For further explanation see heading of Table 3.

	I_1	v_1	w_1	I_2	v_2	w_2	*Total counts*	*Exp.time*	*Noise*
	44.8	3.8	0.063	17.9	43.6	0.063	1	100	1
m:	123.	2.1	0.041	47.8	46.9	0.063			
s:	71.9	17.3	0.049	56.8	105.	0.115			
	44.8	3.8	0.063	17.9	21.8	0.063	1	100	1
m:	96.6	0.9	0.018	21.3	3.4	0.059			
s:	89.4	17.5	0.081	74.1	102.	0.140			
	44.8	3.8	0.063	17.9	10.9	0.063	1	100	1
m:	99.6	0.8	0.019	38.1	12.7	0.017			
s:	114.	18.8	0.050	68.6	43.8	0.096			
	44.8	3.8	0.063	17.9	3.8	0.063	1	100	1
m:	126.	-1.9	0.034	39.8	4.2	0.007			
s:	106.	16.9	0.055	90.0	22.7	0.099			
	31.3	3.8	0.063	31.3	43.6	0.063	1	100	1
m:	86.1	-8.4	0.044	95.4	43.8	0.029			
s:	85.5	16.3	0.081	88.2	51.9	0.076			
	31.3	3.8	0.063	31.3	87.2	0.063	1	100	1
m:	83.2	-0.2	0.025	88.9	87.1	0.028			
s:	212.	21.2	0.081	84.2	48.1	0.083			
	31.3	3.8	0.063	31.3	3.8	0.063	1	100	1
m:	77.7	2.6	0.030	78.4	3.0	0.030			
s:	60.2	20.5	0.052	60.4	20.5	0.052			

SYNTHETIC SPECTRA IN THE SUMER RANGE

The IDL program that produces synthetic line profiles with noise has also been used to simulate a spectral line at 1400 Å in the SUMER range, with a pixel width of 0.04 Å (from the SUMER Red Book). The noise characteristics used for CDS were adopted for SUMER as well. Line strengths were 2.5, 25 and 250 counts/pixel/sec. The results are similar to those found for CDS, despite the different spectral resolutions (see Table 6, 7 and 8 which contains results for different input values of total counts and exposure times). This somewhat unexpected result will be discussed in the conclusion.

In Figure 5 a typical example of a two component line profile is shown without and with added noise. The parameters are as shown in the first row of Table 6.

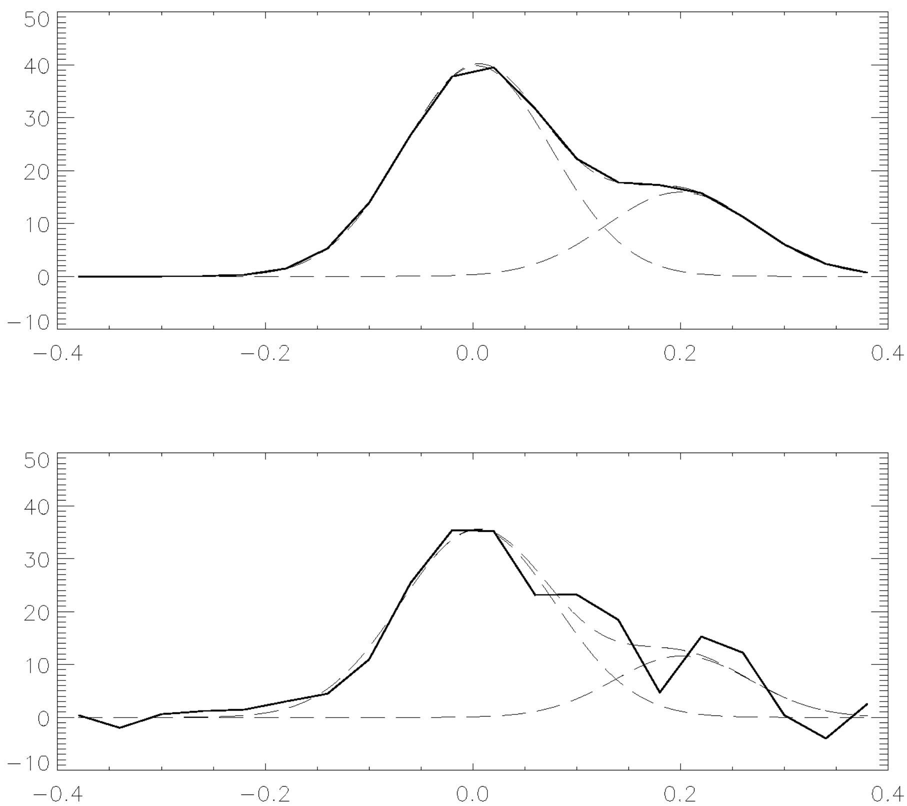

Figure 5. Example of a synthetic two component spectral line in the SUMER range (solid line). The parameters in this figure corresponds to the first row in Table 6. The upper panel shows the profile with no noise and the lower shows the profile with noise added. The two fitted profiles and their sum are also shown with dashed lines.

TABLE 6

Results of synthetic line fitting in SUMER range. For further explanation see heading of Table 3.

	I_1	v_1	w_1	I_2	v_2	w_2	*Total counts*	*Exp.time*	*Noise*
	40.3	0.9	0.100	16.1	42.9	0.100	250	1	1
m:	38.8	0.5	0.098	17.7	40.4	0.105			
s:	4.8	2.6	0.015	3.0	8.2	0.037			
	40.3	0.9	0.100	16.1	21.4	0.100	250	1	1
m:	37.6	1.1	0.090	25.0	17.7	0.080			
s:	11.9	5.0	0.025	13.0	9.4	0.038			
	40.3	0.9	0.100	16.1	10.7	0.100	250	1	1
m:	41.5	-0.2	0.087	27.2	12.4	0.077			
s:	11.2	3.2	0.022	13.0	6.5	0.034			
	40.3	0.9	0.100	16.1	0.9	0.100	250	1	1
m:	42.0	0.2	0.101	20.4	-0.6	0.078			
s:	7.2	2.2	0.012	8.6	4.1	0.035			
	28.2	0.9	0.100	28.2	42.9	0.100	250	1	1
m:	28.8	0.9	0.096	28.6	42.7	0.099			
s:	4.1	3.6	0.019	3.0	4.4	0.021			
	28.2	0.9	0.100	28.2	21.4	0.100	250	1	1
m:	28.9	2.2	0.084	31.4	20.4	0.083			
s:	13.7	4.8	0.034	12.6	7.9	0.035			
	28.2	0.9	0.100	28.2	0.9	0.100	250	1	1
m:	28.5	0.8	0.100	28.5	0.8	0.100			
s:	2.2	1.1	0.007	2.2	1.1	0.007			

TABLE 7

Results of synthetic line fitting in SUMER range. For further explanation see heading of Table 3.

	I_1	v_1	w_1	I_2	v_2	w_2	*Total counts*	*Exp.time*	*Noise*
	40.3	0.9	0.100	16.1	42.9	0.100	25	10	1
m:	39.8	1.0	0.099	24.8	42.0	0.082			
s:	8.5	3.4	0.036	8.3	13.8	0.075			
	40.3	0.9	0.100	16.1	21.4	0.100	25	10	1
m:	50.0.4	0.0	0.078	32.5	14.9	0.071			
s:	13.4	4.1	0.032	15.6	10.0	0.059			
	40.3	0.9	0.100	16.1	10.7	0.100	25	10	1
m:	45.0	0.2	0.078	32.5	14.8	0.071			
s:	13.4	4.1	0.032	15.6	10.0	0.059			
	40.3	0.9	0.100	16.1	0.9	0.100	25	10	1
m:	42.4	-0.0	0.086	31.1	1.0	0.066			
s:	10.7	3.3	0.027	17.6	4.2	0.048			
	28.2	0.9	0.100	28.2	42.9	0.100	25	10	1
m:	30.8	-1.1	0.090	30.4	41.7	0.096			
s:	8.1	3.4	0.045	8.8	8.2	0.041			
	28.2	0.9	0.100	28.2	21.4	0.100	25	10	1
m:	88.7	2.7	0.088	30.9	24.5	0.083			
s:	10.3	4.3	0.044	13.3	11.2	0.063			
	28.2	0.9	0.100	28.2	0.9	0.100	25	10	1
m:	28.2	0.7	0.101	28.2	0.7	0.101			
s:	3.7	2.3	0.015	3.7	2.3	0.015			

TABLE 8

Results of synthetic line fitting in SUMER range. For further explanation see heading of Table 3.

	I_1	v_1	w_1	I_2	v_2	w_2	*Total counts*	*Exp.time*	*Noise*
	40.3	0.9	0.100	16.1	42.9	0.100	2.5	100	1
m:	116.	-0.2	0.047	34.9	45.3	0.029			
s:	92.3	4.3	0.100	63.4	80.6	0.192			
	40.3	0.9	0.100	16.1	21.4	0.100	2.5	100	1
m:	128..4	0.7	0.027	50.0	21.4	0.042			
s:	90.6	4.0	0.084	69.1.6	43.7	0.142			
	40.3	0.9	0.100	16.1	10.7	0.100	2.5	100	1
m:	89.6	-0.6	0.050	38.7	24.8	0.010			
s:	64.2	5.2	0.076	61.0	35.3	0.091			
	40.3	0.9	0.100	16.1	0.9	0.100	2.5	100	1
m:	98.6	0.1	0.041	30.8	0.9	0.077			
s:	76.7	4.5	0.063	70.2	4.3	0.153			
	28.2	0.9	0.100	28.2	42.9	0.100	2.5	100	1
m:	85.6	1.0	0.047	74.0	38.1	0.065			
s:	101.	4.7	0.125	92.0	49.4	0.117			
	28.2	0.9	0.100	28.2	21.4	0.100	2.5	100	1
m:	87.6	-1.0	0.046	93.2	21.7	0.031			
s:	76.6	5.3	0.100	93.7	25.2	0.088			
	28.2	0.9	0.100	28.2	0.9	0.100	2.5	100	1
m:	63.7	0.7	0.063	63.7	0.7	0.063			
s:	55.1	4.6	0.075	55.1	4.6	0.075			

LINE FIT DEPENDENCE ON START PARAMETERS

In the synthetic spectra described above, the actual parameters of the profile without noise were used as a first guess or start parameters of the fit. We wanted to find out how dependent the final result was on the choice of start parameters.

Figure 6 show results for a two component fit with the following set of parameters; intensities at line maximum 44.8 and 17.9 counts/sec, line shifts 3.8 and 43.6 km s^{-1}, and both components had a Doppler width of 0.063 Å. All start parameters were multiplied with the same factor, where 1.0 corresponds to the values given above. The factor used went from 0.2 to 2.5. The background noise was 1 count/pixel/sec, and the fitting routine was LSTSQR. From Figure 6 it is seen that factors from 0.7 to 1.5 give results within a given 2 sigma range (see Figure 4) for all six parameters.

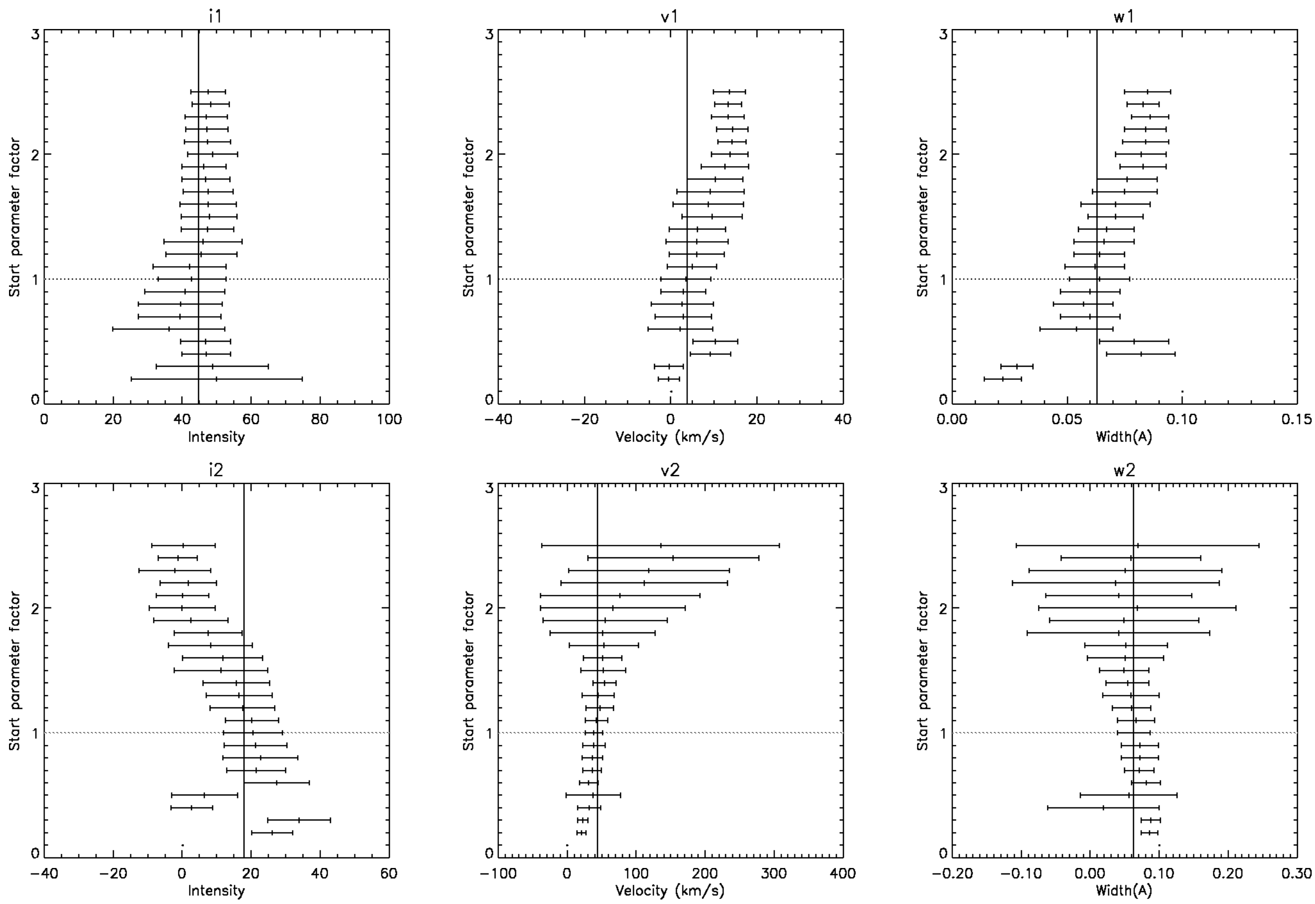

Figure 6. Dependence of spectral line fitting on the given start parameters.

CONCLUSION

Several test runs have given very similar results for CURVEFIT and LSTSQR, but LSTSQR seems to give a better fit in some difficult cases. AMOEBA has proved to be slightly more dependent on the start parameters than the other routines. MOMENT, for a one component fit, is dependent on the selected background level and on how many pixels across the spectral lines that are used in the fit.

From the SERTS data that has been available as CDS test data, it must be concluded that one component Gaussian line components will be used in most of the cases encountered. However the simulated profiles gave successful two component fits. Furthermore the results for CDS and SUMER gave similar results. With the spectral resolution given for CDS it would only be possible to fit with two components if the initial guess of parameters is good i.e. one must 'know' the position of the different components. The difficulty lies in how to select the first guess or start parameters to be used for the fit. When there is only one component in the line fit all goes pretty well, but with more components (2 or 3), it has been found that the result is dependent on the selected start values. From the synthetic spectra it was found that the start parameters must be within a range from 0.7 - 1.5 from the actual value to give a fit within 2 sigma. One way of overcoming this problem would be to interactively set the cursor at the position where one think there is a line component, but this would be an extensive task!

There is one common problem to all the fitting procedures, and that is how to set the background or 'continuum' level. For large amounts of data it will be necessary to automate the process of finding the best level to be subtracted before the fitting is done.

APPENDIX

This appendix describes additional routines added for the calculation of line intensities.

TOTAL INTENSITY FROM PIXEL VALUES

A very simple way to calculate the total line intensity of a given spectral line is just adding up all the pixel values across the line profile and multiply the sum with the spectral resolution. For some lines in the CDS spectral range this may actually be the most reasonable way of "fitting" the line profiles. A procedure that does this has been made.

TOTAL INTENSITY FROM I_0 AND FWHM

The total line intensity can be calculated using the measured intensity at line maximum and the full width at half maximum (FWHM). The total line intensity I_{tot} is given by the following equation, with I_0 as the maximum intensity in the line and $\Delta\lambda_D$ as the lines Doppler width:

$$I_{tot} = I_0 \int_{-\infty}^{\infty} e^{-\left(\frac{\lambda}{\Delta\lambda_D}\right)^2} d\lambda$$

Solving this equation, we get:

$$I_{tot} = \sqrt{\pi}\ I_0\ \Delta\lambda_D$$

The full width at half maximum (FWHM) can be expressed as:

$$FWHM = 2\sqrt{ln2}\ \Delta\lambda_D$$

Inserting this in the equation above then gives the following expression for the total line intensity:

$$I_{tot} = \frac{\sqrt{\pi}}{2\sqrt{ln2}}\ I_0\ FWHM$$

The maximum intensity in the line and the FWHM can easily be found in a given line profile. With this FWHM a new maximum intensity can, if the line is asymmetric, be found, and the total line intensity be calculated. This way of calculating the line intensities has been included in the line fitting software.

Any comments or suggestions are appreciated and can be directed to the author on the following address.

Address:

Nils Brynildsen
Institute of Theoretical Astrophysics
P.O.Box 1029, Blindern
N-0315 OSLO, Norway
Fax No.: +(47) 22 85 65 05

Phone: +(47) 22 85 50 08
Internet: nilsbr@astro.uio.no

REFERENCES

Bartoe, J.-D. F. and Brueckner, G. E., 1975, J. Opt. Soc. Am.,**65**, 13, (HRTS).

Bevington, P.R. 1969, Data Reduction and Error Analysis for the Physical Sciences (New York, McGraw-Hill).

Harrison, R. A., The CORONAL DIAGNOSTIC SPECTROMETER for SOHO: Scientific Report, ver. 5.0, May 1993, (The Blue Book).

Neupert, W. M., Epstein, G. L., Thomas, R. J. and Thompson, W. T., 1992, Solar Phys., **137**, 87, (SERTS).

Wilhelm, K., THE SUMER SPECTROMETER FOR SOHO, version 2.0, January 1992, (The Red Book).