ANALYSIS OF FABRY-PEROT SPECTRA OF LIDAR BACKSCATTER ECHOES

M. Conde

ABSTRACT

A numerical procedure is presented for analysis of backscatter spectra recorded by an atmospheric Doppler lidar experiment. The recorded spectra are modelled by an analytic function of five parameters. The procedure solves for the set of model parameters that yields the best least-squares fit between the model and the data. The model is formulated to allow for Doppler shifted and broadened scattering from both aerosols and from multiple molecular species, for a continuum background, and for the effects of instrumental broadening. All five model parameters are fitted in the signal domain. Three parameters are fitted by a direct analytic solution and two by numerical searching. If required, one or more parameters may be held fixed at externally-supplied values, while the remainder are fitted accordingly. The method's performance was examined by applying it to numerically simulated test spectra, for several example lidar configurations. Some implications of these results for the design of actual lidar experiments are discussed.

1 INTRODUCTION

Lidar systems operate by illuminating the atmosphere with short duration and highly directional pulses of laser photons. Measuring properties of the backscattered photons as a function of time allows measurement of atmospheric properties as a function of distance from the lidar. A high resolution spectrometer in the receiver system can measure the wavelength spectra of the backscattered photons. The atmospheric kinetic temperature and line-of-sight wind as a function of range from the lidar can then be inferred from these spectra. The spectrometer of choice for this application is a single or dual etalon Fabry-Perot spectrometer (FPS).

Analysis of a recorded spectrum proceeds by fitting an analytically-described model spectrum to it. The model is formulated as a function of several parameters. These parameters correspond to the desired geophysical quantities (wind and temperature) plus, usually, several other geophysically useful quantities that will be estimated as a by-product.

Designing an effective analysis scheme is a three-part problem. First, the functional form of the model spectrum must be specified. Second, we need a scalar expression defining the ‘goodness-of-fit’ between the recorded spectrum and a particular instance of the model. Third, we need a scheme for determining the set of parameters that yields the best possible fit between the model and the recorded spectra, using this goodness estimator. For a single spectrum, the set of best-fit parameters are the experimental estimators of the desired geophysical quantities.

2 THE MODEL SPECTRUM

Experimentally recorded spectra are, inevitably, instrumentally broadened by the finite spectral bandwidths of both the transmitting laser and the receiving Fabry-Perot spectrometer. However, there is no need for the analysis to represent these two broadening functions independently; a single

function can correctly account for the combined effects of both. Indeed, in a typical lidar experiment, this combined effect is what is actually measured when directly viewing the transmit laser. We refer to this as the 'instrument function' and denote it mathematically with the symbol I.

It is assumed that actual lidar returns from the atmosphere will arise due to both Rayleigh scattering from molecules and Mie scattering from aerosols. (Inelastic processes, such as Raman and Brillouin scattering, will be ignored.) Because aerosol particle masses are large compared to the atmosphere's mean molecular mass, the spectral width of aerosol backscatter is much smaller than the spectral width of the Rayleigh component of the return signal. The Fabry-Perot instrument function should be chosen to be comparable to the Rayleigh Doppler width. Thus, the aerosol backscatter spectral width would appear negligible to the Fabry-Perot system, for all reasonable values of atmospheric temperature. It is assumed that aerosols drift with the neutral wind, so that the peak wavelength of the aerosol return spectrum will be the same as the peak wavelength of the molecular Rayleigh return spectrum. Finally, we assume that there will be some background continuum present in the recorded signal. These various components of the received signal must be incorporated into the model spectrum.

A scanning FPS does not record continuous spectra. Rather, a set of discrete samples of the return spectrum are obtained by stepping the FPS passband in wavelength, and measuring the received intensity at each step. (Because the instrument scans its passband in wavelength over time, care must be taken to avoid possible spectral distortions due to time variation of the return spectrum. Possible distortions due to this effect are suppressed by co-adding a large number of short duration spectral scans during each integration interval.) If each spectral scan contains N discrete samples, then at the end of the scan we will have a set of values, $\{Y_n\}$ where n varies from 0 to N–1. The set $\{Y_n\}$ is a discrete approximation to the spectrum of backscattered laser photons, modified by the spectral response of the instrument.

From the discussion above, we will choose to model $\{Y_n\}$ with a set of samples, $\{s_n\}$ of an analytic function (of wavelength) of the form

$$S = I*B \tag{1}$$

where

I = instrument function,

B = backscatter spectrum and

* denotes the convolution operator.

Rayleigh backscatter will occur from all atomic and molecular species in the atmosphere. The mixing ratios of these species (and therefore their relative backscatter intensities) are well known throughout the height range accessible to the lidar. Relative backscatter intensities due to the molecular species can thus appear as constants within the model. The return spectrum due to each molecular species will be modelled as a Gaussian function of wavelength. As each species has a unique mass, the Doppler width of the backscatter spectrum due to each species will be different. However, all species are assumed be at the same temperature – so only one free parameter is needed in the model to describe these different widths. All scattering species, including the aerosol, can be assumed to travel with the wind at the same velocity. Thus, again, we only need one model parameter to describe the Doppler shift.

For the purpose of describing the backscatter spectrum, the aerosol could potentially be treated as just another atmospheric species, albeit one of very high mass. However, we treat it separately from the molecular species, for two reasons. First, the aerosol abundance is variable, yielding a relative backscatter intensity that cannot be described effectively by an *a-priori* constant. Second, the bandwidth of the aerosol return spectrum would (for a reasonably optimised experiment) be narrow compared to our wavelength sampling interval. Thus, we cannot adequately represent the aerosol return spectrum using the actual experiment's wavelength sampling interval. Rather, we must use a model that only ever requires evaluation of the convolution of the aerosol spectrum with the instrument function.

From these arguments, we can rewrite equation 1 as the instrument function convolved with a sum of three separately treated contributions: a continuum background, aerosol Mie scattering, and Rayleigh scattering summed over all contributing molecular species. We write this as

$$s(\lambda) = \sum_{j=0}^{2} a_j S_j(\lambda), \tag{2}$$

where the a_j are the backscatter intensities of each of the 3 contributing terms and

$S_0(\lambda)$ = 1 represents the continuum component,

$S_1(\lambda)$ = $I(\lambda) * \delta(\lambda - \lambda_0)$; ($\delta$ denotes the Dirac delta function)

= $I(\lambda - \lambda_0)$, the aerosol term,

$S_2(\lambda)$ = $I(\lambda) * \sum_{k=0}^{K-1} qk \exp - \left(\frac{\lambda - \lambda_0}{w_k(T)} \right)^2$, the total Rayleigh contribution arising from K different molecular species,

$w_k(T)$ = is the width in wavelength units of the backscatter from the kth species for a temperature of T,

λ_0 = the Doppler shifted peak wavelength of the return signal.

We refer to the three functions $S_0(\lambda)$, $S_1(\lambda)$ and $S_2(\lambda)$ as the 'basis functions' for our model. In the expression for S_2 we have used q_k to represent the *relative* contribution to the Rayleigh backscatter from the kth molecular species. Values for q_k are supplied to the analysis as pre-determined constants, with the property that

$$\sum_{k=0}^{K-1} q_k = 1. \tag{3}$$

Generally, the model contains five free parameters that we need to fit, i.e. a_0, a_1, a_2, λ_0 and T. However, potentially, values for some of these parameters could be supplied as constants estimated by other means. For example, above the aerosol layer, a_0, a_1 and T can be estimated directly from the height profile of the spectrally-integrated return signal. Thus, the analysis program is formulated to allow any combination of parameters to remain fixed at their externally-supplied starting values.

3 DEFINING THE BEST FIT

We use the conventional χ^2 parameter to describe the degree to which the model spectrum departs from the recorded spectrum, i.e.

$$\chi^2 = \sum_{n=0}^{N-1} \frac{(y_n - s_n)^2}{\sigma_n^2}, \tag{4}$$

where σ_n^2 is the variance of the nth element of the observed spectrum, y_n. Providing that our model accurately describes the functional form of the recorded spectrum, then the best possible choice of model parameters yields the minimum χ^2, with an expected value of N minus the number of free parameters in the model. If we denote the set of model parameters as $\{p_j\}$ then the best choice for $\{p_j\}$ occurs when

$$\frac{\partial \chi^2}{\partial p_j} = 0, \tag{5}$$

for all j. This condition provides an analytic expression that could potentially be used to calculate the optimum $\{p_j\}$. However, using the model function described by equation 2, it is not possible to solve equation 5 for all $\{p_j\}$.

4 STRATEGIES FOR FINDING OPTIMUM VALUES FOR THE MODEL PARAMETERS

In designing an analysis scheme, it is important to distinguish between the *criterion* used to define the optimum model parameters and the *procedure* used to find parameters that satisfy this criterion. So far we have described the criterion, but not the procedure.

If this is not feasible to solve equation 5, χ^2 can be evaluated numerically for various trial (p_j).and the results used to drive a search procedure to locate the χ^2 minimum. The latter approach has usually been adopted by most Fabry-Perot spectral analysis algorithms, because it is usually not possible to solve the peak position and width parameters analytically.

The simplest search strategy, conventionally described as a 'grid search', proceeds by minimising χ^2 with regard to each model parameter in turn, while all other parameters are held fixed. Optimisation over the complete set of parameters is repeated cyclically until satisfactory convergence is obtained. This algorithm is very stable, although it can converge slowly if χ^2 variations with respect to two or more of the parameters are not independent. However, the grid search will eventually be successful even in such cases, indicating that repeated independent optimisation of each parameter is viable as a generalised strategy.

Another consequence of this is that is does not matter *how* χ^2 is minimised for a given parameter, nor whether the same technique is used for each one. This conclusion leads to the hybrid approach adopted for the lidar analysis problem – a grid search in which, at each iteration, some of the parameters are computed analytically and some are obtained by searching.

5 ANALYTICALLY CALCULATING THE INTENSITIES

We can solve equation 5 analytically for the three intensities appearing in our model, i.e. a_0, a_1, and a_2. Several algorithms exist for this. The method presented below (known as 'directly solving the

normal equations') is the most straightforward to describe, and is presented to illustrate that an analytic solution does indeed exist. However, this method can be unstable if the basis $\{s_n\}$ are not clearly distinguished by the data. (By this we mean that a region exists around the global χ^2 minimum within which the values of two or more coefficients of the basis functions can be changed in a complementary fashion, so the resulting χ^2 perturbation is not statistically significant.) In the case of our model function for lidar returns, this situation could arise if the aerosol and molecular return spectra are not distinguishable within the noise on the recorded spectrum. The problem manifests itself during numerical inversion of the curvature matrix – which can become near singular. A better (but more complex) algorithm in such cases is singular value decomposition.

5.1 Direct solution of the normal equations

Substituting equation 2 into equation 4 gives

$$\chi^2 = \sum_{n=0}^{N-1} \frac{\left(y_n - \sum_{j=0}^{2} a_j (S_j)_n \right)^2}{\sigma_n^2}, \tag{6}$$

where the notation $(S_j)_n$ refers to the jth basis function evaluated at the nth discrete wavelength. Equation 5 thus becomes

$$\frac{\partial \chi^2}{\partial a_i} = -2N - \sum_{n=0}^{N-1} \frac{(S_i)_n \left(Y_n - \sum_{j=0}^{2} a_j (S_j)_n \right)}{\sigma_n^2} = 0. \tag{7}$$

Equation 7 describes three simultaneous linear equations in three unknowns a_0, a_1, and a_2. Explicitly,

$$\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_0)_n Y_n = a_0 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_0)_n (S_0)_n + a_1 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_1)_n (S_0)_n + a_2 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_2)_n (S_0)_n$$

$$\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_1)_n Y_n = a_0 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_0)_n (S_1)_n + a_1 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_1)_n (Sl)_n + a_2 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_2)_n (S_l)_n$$

$$\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_2)_n Y_n = a_0 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_0)_n (S_2)_n + a_1 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_1)_n (S_2)_n + a_2 \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_2)_n (S_2)_n$$

which can be expressed as

$$\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_i)_n Y_n = \sum_{j=0}^{2} \left\{ a_j \sum_{n=0}^{N-1} \left[\frac{1}{\sigma_n^2} (S_i)_n (S_j)_n \right] \right\} \tag{8}$$

for i = 0, 1, 2. Equation 8 can be expressed even more compactly in matrix form as

$$\mathbf{b} = \mathbf{A}\mathbf{a}, \tag{9}$$

where **b** is a row vector given by

$$\mathbf{b} = \left(\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_0)_n Y_n, \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_1)_n Y_n, \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_2)_n Y_n \right) \tag{10}$$

a is a column vector given by

$$\mathbf{a} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix},$$

and **A** is a 3 x 3 square matrix defined by

$$A_{ij} = \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} (S_i)_n (S_j)_n. \tag{11}$$

A is conventionally referred to as the 'curvature matrix'.

Multiplying both sides of equation 9 by $\mathbf{A}^I$ yields an analytic expression for the model's intensity coefficients, **a**,

$$\mathbf{a} = \mathbf{b}\mathbf{A}^I \tag{12}$$

Although this process may appear complex, it is numerically very simple to evaluate **a**. The analysis program merely constructs the row vector **b** and the curvature matrix **A** using the summations defined by expressions 10 and 11 respectively. A standard matrix inversion routine is then called to evaluate $\mathbf{A}^I$ the result is inserted into the very simple calculation of **a** defined by equation 12.

There are a couple of points to note. First, these calculations must be done in the *signal* domain (we have not derived expressions for optimising equation 6 analytically with respect to **a** in the transform domain). Second, values for the parameters T and λ_0 are needed to evaluate the functions S_1 and S_2. This means that initial estimates for T and λ_0 must be generated before the first attempt to solve for **a**. On each subsequent iteration, the latest estimates of T and λ_0 are used. Third, as mentioned above, there may be times when one or more of the a_0, a_1 and a_2 will be determined outside of this analysis, and simply supplied as constants. Thus, a practical program will need to implement these possibilities by allowing the matrix expressions to operate for any of the cases of 3, 2 or 1 unknown(s).

5.2 Singular value decomposition

The singular value decomposition (SVD) algorithm has been described by Press et al. (1986); repetition of the description is beyond the scope of this work. SVD also provides an analytic solution for the set of coefficients $\{a_j\}$ but does so in a fashion that avoids directly inverting the curvature matrix **A** which, potentially, could be near-singular for noisy spectra. Note that SVD is *not* applicable for finding coefficients that appear non-linearly within our model function. Thus, SVD cannot be used to evaluate T and λ_0.

The lidar analysis was implemented using the programming language 'IDL', which includes a pre-written SVD subroutine. This implementation can return not only estimates of the coefficients $\{a_j\}$

but also estimates of their variances. (Variances of the coefficients are also available from the normal equations, although discussion of this has been omitted above.) Note, however, that in this application there is an additional contribution to the variances of parameters $\{a_j\}$ due to the variances of T and λ_0. Analytic solutions for the linear parameters offer no way to allow for this, so it is not appropriate to use variances returned by them. Rather, we must separately compute the parameter variances simultaneously for all parameters, as is described later.

SVD is known to be significantly slower than solving the normal equations. However, tests on numerically synthesised data have shown that a PC using a 233 MHz Pentium MMX processor can run the complete analysis described here on one 128-channel spectrum in an average of around 0.45 seconds. This is considered easily fast enough, even for near-real-time data processing.

6 ESTIMATING TEMPERATURE AND DOPPLER SHIFT

6.1 The Levenberg-Marquardt search

Because we cannot solve equation 5 analytically for T and λ_0, these parameters must be determined by searching. There are many possible strategies for doing this. One possibility would be to use a grid search in the Fourier transform domain, as is done by many existing routines used to analyse Fabry-Perot spectra of the airglow and aurora. This method is attractive because the number of Fourier coefficients needed to describe the data and the model function is typically a factor of ten smaller than the number of observation points in the signal domain. However, this advantage is partially lost because the grid search is very inefficient (it makes no use of the gradient of the fitting function).

The approach described below operates in the signal domain and attempts to use as much analytic information as possible. Although we cannot *solve* equation 5 analytically, we can compute analytic expressions for the partial derivatives of χ^2 with respect to T and λ_0. These expressions will be used to implement the Levenberg-Marquardt (LM) algorithm to find simultaneous solutions to

$$\frac{\partial\chi^2}{\partial T} = 0 \quad \text{and} \quad \frac{\partial\chi^2}{\partial\lambda_0} = 0.$$

The LM method is particularly elegant in that it implements a smoothly-varying combination of two search strategies. When the initial guess for the model parameters is a long way from optimum, the LM search uses an inefficient but numerically stable steepest-descent search. As the parameter estimates approach their solution, the search strategy changes smoothly, placing progressively more weight upon an analytic solution of a quadratic-expansion of χ^2. The two strategies differ in the method used to generate each successive new estimate for the model parameters.

The following description of the LM procedure closely follows that presented by Press et al. (1986). Let us denote the set of J parameters that we wish to optimise as $\{p_j\}$ to distinguish them from the set $\{a_j\}$ that we solved above using direct analytic methods. We can also regard the set $\{p_j\}$ as a vector, in which case we denote it simply as $\mathbf{p}$. Search procedures operate by generating a sequence of estimates for $\mathbf{p}$. The steepest descent method generates a new estimate for $\mathbf{p}$ by perturbing the current one according to

$$\mathbf{p}_{m+1} = \mathbf{p}_m - \gamma\nabla\chi^2(\mathbf{p}_m), \qquad (13)$$

where γ is some coefficient chosen to be small enough that a single iteration does not overshoot the χ^2 minimum. Conversely, the quadratic expansion of χ^2 method generates its sequence of estimates using a relation of the form

$$\mathbf{p}_{m+1} = \mathbf{p}_m + \mathbf{D}^{I}[-\nabla\chi^2(\mathbf{p}_m)], \tag{14}$$

where $\mathbf{D}$ is a JxJ matrix describing the second cross-partial derivatives of χ^2 with respect to the parameters $\mathbf{p}$.

Consider our expression for χ^2, equation 4. Using the chain rule, the gradient of χ^2 with respect to $\mathbf{p}$ can be written in component form as

$$\frac{\partial\chi^2}{\partial p_j} = -2\sum_{n=0}^{N-1}\frac{(Y_n - s_n)}{\sigma_n^2}\frac{\partial s_n}{\partial p_j}, \tag{15}$$

from which we define

$$\beta_j = \sum_{n=0}^{N-1}\frac{(Y_n - s_n)}{\sigma_n^2}\frac{\partial s_n}{\partial p_j} = -\frac{1}{2}\frac{\partial\chi^2}{\partial p_j}. \tag{16}$$

Taking a second partial derivative gives

$$\frac{\partial^2\chi^2}{\partial p_j \partial p_l} = 2\sum_{n=0}^{N-1}\frac{1}{\sigma_n^2}\left[\frac{\partial s_n}{\partial p_j}\frac{\partial s_n}{\partial p_l} - (Y_n - s_n)\frac{\partial^2 s_n}{\partial p_j \partial p_l}\right]. \tag{17}$$

Press et al. (1986) explain that it is conventional to drop the second-order partial derivatives from the above expression – i.e. the right-hand term of the square brackets inside the summation. This is because near the solution ($Y_n - s_n$) should be small if our model is well chosen. Further, retaining the higher derivatives is often numerically destabilising in practice. From this we define

$$\alpha_{jl} = \sum_{n=0}^{N-1}\frac{1}{\sigma_n^2}\frac{\partial s_n}{\partial p_j}\frac{\partial s_n}{\partial p_l} \approx \frac{1}{2}\frac{\partial^2\chi^2}{\partial p_j \partial p_l} \tag{18}$$

Now both of equations 13 and 14 can be expressed in a general form as

$$\mathbf{p}_{m+1} = \mathbf{p}_m - \boldsymbol{\delta}, \tag{19}$$

where δ is the (vector) increment added to $\mathbf{p}_m$ in order to generate $\mathbf{p}_{m+1}$. Thus, for the steepest descent method we have

$$\delta_l = \gamma\beta_l. \tag{20}$$

It can be shown that $\boldsymbol{\alpha} = {}^1/_2\mathbf{D}$ in equation 14, so that for the quadratic expansion method

$$\sum_{j=1}^{J}\alpha_{lj}\delta_l = \beta_l. \tag{21}$$

The Levenberg-Marquardt search combines these expressions by taking two further steps. First, based on dimensional arguments, the γ is chosen to be a vector whose lth element is given by

$$\gamma_l = \frac{1}{\lambda\alpha_{ll}},$$

so that

$$\delta_l = \frac{1}{\lambda \alpha_{ll}} \beta_l. \tag{22}$$

In the above equations, λ is a non-dimensional parameter whose value we can freely vary depending on the search behaviour that we desire. The second step is to note that both equations 22 and 21 can be written as a single expression

$$\sum_{j=1}^{J} \alpha'_{lj} \ \delta_l = \beta_l, \tag{23}$$

where

$$\begin{aligned} \alpha'_{ll} &= \alpha_{ll}(1+\lambda) \\ \alpha'_{lj} &= \alpha_{lj} \quad \text{for}(j \neq l). \end{aligned} \tag{24}$$

For very large values of λ, α' is dominated by its diagonal elements so that equation 23 approaches 20. Conversely, as λ approaches zero, equation 23 approaches equation 21. Thus, simply by varying λ, we can smoothly vary the strategy for generating $\mathbf{p}_{m+1}$ between a steepest descent method and a quadratic expansion of χ^2 method. The former is most useful a long way from the χ^2 minimum, whereas the latter works best close to the minimum.

Press et al. (1986) describe the procedure for actually implementing an LM search as follows:

1. Generate an initial estimate of the parameters, $\mathbf{p}$, and evaluate $\chi^2(\mathbf{p})$.
2. Pick a starting value for λ, with 0.001 being recommended.
3. Evaluate the vector $\boldsymbol{\beta}$ and the matrix $\boldsymbol{\alpha}'$.
4. Solve equation 23 for $\boldsymbol{\delta}$ by inverting $\boldsymbol{\alpha}'$.
5. If $\chi^2(\mathbf{p} - \boldsymbol{\delta}) \geq \chi^2(\mathbf{p})$ increase λ by a factor of 10 and go back to (3) .
6. If $\chi^2(\mathbf{p} - \boldsymbol{\delta}) < \chi^2(\mathbf{p})$ decrease λ by a factor of 10 and replace $\mathbf{p}$ with $\mathbf{p} - \boldsymbol{\delta}$, and go back to (3) if the stopping criterion has not been satisfied.
7. Terminate the iteration when differences between successive χ^2 estimates are too small to be statistically meaningful (i.e. <<1).
8. Set $\lambda = 0$ and evaluate the covariance matrix of standard errors, $\mathbf{C} = \boldsymbol{\alpha}^{-1}$.

6.2 Application to the lidar analysis

The above description has described the Levenberg-Marquardt search in general terms. For the lidar problem, there are only two parameters that we need to find by searching, i.e. T and λ_0 Thus, β is a two-element vector and $\boldsymbol{\alpha}'$ is a 2x2 matrix. Because these each have a small number of elements, solving equation 23 is not computationally expensive.

Application of the LM method to any specific problem, including that of the lidar analysis, requires deriving expressions for

$$\frac{\partial \chi^2}{\partial p_j}$$

so that β and α' can be evaluated. Now equations 16 and 18 indicate that we do not need to evaluate

$$\frac{\partial \chi^2}{\partial p_j}$$

directly; we can adequately approximate these terms provided we know the derivatives of the basis functions,

$$\frac{\partial s_n}{\partial p_j}.$$

These latter expressions are more simply obtained, as is done below.

6.3 The partial derivative with respect to Doppler shift

Expanding equation 2 to explicitly show all its terms yields

$$s_n = a_0 + a_1\left[I(\lambda - \lambda_0)\right]_n + a_2\left[I * \sum_{k=0}^{K-1} q_k \exp - \left(\frac{\lambda - \lambda_0}{w_k(T)}\right)^2\right]_n,$$

from which we obtain the result that

$$\frac{\partial s_n}{\partial \chi_0} = a_1\left[\frac{\partial}{\partial \lambda_0} I(\lambda - \lambda_0)\right]_n + a_2\left[I * \sum_{k=0}^{K-1} \frac{2(\lambda - \lambda_0)}{w_k^2(T)} q_k \exp - \left(\frac{\lambda - \lambda_0}{w_k(T)}\right)^2\right]_n. \qquad (25)$$

Notice that the derivative is explicitly evaluated in the second (molecular scattering) term of equation 25 above, but not in the first (aerosol scattering) term. For the second term, we have evaluated the derivative using the property that

$$\frac{\partial}{\partial u}\left[I(\lambda) * f(\lambda, u)\right] = I(\lambda) * \left[\frac{\partial}{\partial u} f(\lambda, u)\right]. \qquad (26)$$

Now because $I(\lambda - \lambda_0) = I(\lambda) * \delta(\lambda - \lambda_0)$, we can indeed rewrite

$$\frac{\partial}{\partial \lambda_0} I(\lambda - \lambda_0)$$

in the same form as the left-hand side of equation 26 above. However, because we can neither represent (much less differentiate) a discretely-sampled delta function, we cannot use the right-hand side of equation 26 to evaluate the aerosol derivative in practice. Further, because $I(\lambda)$ is an experimentally determined function (i.e. it is not analytically described), we can never express

$$\frac{\partial}{\partial \lambda_0} I(\lambda - \lambda_0)$$

analytically – we must compute it numerically. The computer program actually does this by applying the Fourier derivative theorem to the discrete Fourier transform $\{I_n\}$. ($\{I_n\}$ is the discrete representation of $I(\lambda)$).

6.4 The partial derivative with respect to temperature

As before, we expand equation 1 to show all its terms. However, this time we must explicitly include the expression that relates $1/e$ Doppler half width of the molecular backscatter (in measurement units) to kinetic temperature, i.e.

$$w(T) = \xi\lambda_0\sqrt{\frac{2kT}{Mc^2}},$$

where ξ the coefficient needed to convert the wavelength units (of λ_0) to 'channels' in our discretely sampled spectra. Our expression thus becomes

$$s_n = a_0 + a_1\left[I(\lambda-\lambda_0)\right]_n + a_2\left[I * \sum_{k=0}^{K-1} q_k \exp-\left(\frac{\lambda-\lambda_0}{\xi\lambda_0\sqrt{\frac{2kT}{M_k c^2}}}\right)^2\right]_n.$$

Differentiating with respect to T gives

$$\frac{\partial s_n}{\partial T} = a_2\left[I * \sum_{k=0}^{K-1} \frac{M_k c^2}{2kT^2}\frac{(\lambda-\lambda_0)^2}{\xi^2\lambda_0^2} q_k \exp-\left(\frac{\lambda-\lambda_0}{\xi\lambda_0\sqrt{\frac{2kT}{M_k c^2}}}\right)^2\right]_n. \qquad (27)$$

Equation 27 completes the information needed to calculate the vector β and the α' used in the LM search.

7 THE COMPLETE LIDAR ANALYSIS SCHEME

We now consider how the methods above can be combined to form a complete lidar analysis scheme. We have constructed the analysis in two parts: one to solve analytically for parameters appearing linearly in our model s_n, and one to solve for the non-linear parameters. We perform these two parts iteractively. The procedure is:

1. Generate initial guesses for T and λ_0.

2. Generate new estimates for a_0, a_1 and a_3 using the SVD method and the T and λ_0 estimates.

3. Generate new estimates for T and λ_0 using an LM search and the current a_0, a_1 and a_3 estimates.

4. Evaluate χ^2 using the updated parameters.

5. Go back to (2) until successive χ^2 estimates are not distinguishable with statistical significance.

One slight complication is that the LM method itself is an iterative procedure. Thus, the analysis could either leave the linear parameters constant throughout each *complete* LM search or, alternatively, it could re-calculate the linear parameters between every step in the LM search. In practice, the former approach is faster.

8 ESTIMATING THE PARAMETER VARIANCES

Both the SVD and the LM portions of the complete solution can return estimates of the variances of the parameters. However, while it is simple to independently compute *values* of a subset of the model parameters (given values for each of the remaining parameters), it is not simple to independently compute the *variances* of the subset (even given values and variances of the remaining parameters). Fortunately, standard non-linear least-squares fitting techniques (e.g. Bevington, 1969; Press et al. 1986) provide a generalised method for computing the variances of all the model parameters simultaneously, as follows.

Let us denote the complete set of L variable parameters in our model as $\{v_l\}$ where l varies from 0 to L–1. (Using the notation of the previous sections, $\{v\}$ would be formed from the union of the sets $\{a\}$ and $\{p\}$ and L = 5 if we fit all parameters.) We construct a complete L x L curvature matrix, **c**, encompassing all parameters using

$$c_{jk} = \sum_{n=0}^{N-1} \frac{1}{\sigma_n^2} \frac{\partial s_n}{\partial v_j} \frac{\partial s_n}{\partial v_k}. \qquad (28)$$

The parameter variances are then obtained directly from the diagonal elements $\mathbf{c}^{\mathbf{I}}$. That is, if we write $\varepsilon = \mathbf{c}^{\mathbf{I}}$, then

$$\sigma^2(v_l) = \varepsilon_{ll} \qquad (29)$$

9 PERFORMANCE EVALUATION

9.1 Simulating spectra

To examine the performance of the method, we apply it to simulated Fabry-Perot spectra, generated numerically from precisely known values of the five model parameters. By applying the method to large sets of such spectra, we can examine the speed and stability of convergence, as well as the statistical distributions of the fitted parameters. In particular, we need to know if the means of the parameter estimates match the values used to generate the spectra. Further, we will compute the variances of the distributions of parameter estimates over the whole set of simulated spectra, and see if these are consistent with the parameter variances returned by the fit routine after fitting an individual spectrum. The basic characteristics used to simulate the experiment were:

Laser wavelength	532 nm
Etalon gap	15.8 mm
FPS instrumental finesse	19.7
Number of spectral channels	128

Doppler shift by one channel	~40 ms^{-1}
Atmospheric composition	75% N_2, 25% O_2
Doppler width	250 Kelvins
Doppler shift	20 channels
Molecular peak height	50 counts
Aerosol peak height	15 counts
Spectral background	45 counts/channel

In simulating the instrument profile, an attempt was made to include several of the 'less desirable' characteristics that frequently occur in real instruments. Thus, the simulated instrument profile was generated by summing 3 Gaussian functions, and then convolving these with an asymmetric (right-angle triangle) shaped function. For the Gaussians, the dominant one was chosen to be very narrow (corresponding to an equivalent N_2 temperature of 8 Kelvins). The remaining two Gaussians heights were only 30% and 10% that of the main one, but their spectral widths were 5 and 10 times greater, respectively. The full-width at half height (FWHH) of the triangle function was set to be twice that of the main Gaussian's FWHH. The instrument function was deliberately displaced slightly off-center of the spectrum; its centroid occurs at 40% of the scan range. A noise-free version of this instrument function was used to generate the simulated sky return spectra. However, the version supplied to the fitting routine had a small noise term added. This was to simulate the real experiment, in which the instrument function can only be obtained by measurement – a process that inevitably introduces noise. Figure 1 illustrates the simulated instrument profile supplied to the program, including measurement noise.

9.2 Some example fits

We now present three examples of the results of fitting to individual spectra. Figure 2 presents an example of a simulated sky return spectrum, with the fitted function superimposed, after 0, 3, and 12 iterations of the procedure. To test the fitting routine strenuously, the initial estimates of peak position and temperature were allowed to vary randomly over a much wider range than would occur in practice. In Figure 2 the initial peak position estimate was clearly to high. Nevertheless, convergence was good, as is illustrated by Figure 3, which depicts the history of the estimates for each parameter after each iteration. Indeed, tests have shown that the procedure can converge from initial guesses very much worse than shown here. The only (minor) difficulty occurs if the initial temperature is absurdly low and the initial position is several peak widths away from its correct location. In this case there is little overlap between peaks in the data and the model, yielding a poor indication of which way the position and temperature parameters need adjusting. However, the program can test for this condition, and can respond effectively.

The program actually terminated after a total of 21 iterations, when the reduced χ^2 value was 1.20. (21 iterations is more than usual; the average number of iterations at termination is only 13.) Included among the diagnostic information returned by the program is the sky spectrum signal/noise ratio. This is computed from the power spectrum of the sky profile, and is defined as the power spectrum's fundamental component divided by the average value of it's high-frequency, noise-

dominated components. The signal/noise ratio was 748 for the sky spectrum shown in Figure 2a. The final parameter estimates and their corresponding uncertainties were:

	Value	Uncertainty
Temperature	243.6 Kelvins	26.3 Kelvins
Doppler shift	19.9 channels	0.24 channels, or 9.6 ms^{-1}
Molecular	48.9 counts	2.9 counts
Aerosol	16.3 counts	4.1 counts
Background	46.4 counts/channel	0.58 counts/channel

The sequence of χ^2 values after each iteration is shown in Figure 4.

Figure 5 shows an example of the fit results for a spectrum with a low signal/ratio, i.e. only 91. The program terminated after a total of 13 iterations, when the reduced χ^2 value was 0.84. The final parameter estimates and their corresponding uncertainties were:

	Value	Uncertainty
Temperature	272.0 Kelvins	90.0 Kelvins
Doppler shift	19.7 channels	0.70 channels, or 28 ms^{-1}
Molecular	48.3 counts	8.6 counts
Aerosol	21.4 counts	12.4 counts
Background	43.7 counts/channel	1.8 counts/channel

Figure 6 shows an example of the fit results for a spectrum with a high signal/ratio, i.e. 2 194. The program terminated after a total of 22 iterations, when the reduced χ^2 value was 0.91. The final parameter estimates and their corresponding uncertainties were:

	Value	Uncertainty
Temperature	257.3 Kelvins	16.4 Kelvins
Doppler shift	19.98 channels	0.14 channels, or 5.6 ms^{-1}
Molecular	48.3 counts	1.7 counts
Aerosol	16.3 counts	2.4 counts
Background	45.5 counts/channel	0.34 counts/channel

9.3 Distributions of the fit results

Having seen how the method behaves in a few selected examples, we now consider the statistical distributions of fit results for a large (1500 element) set of simulated spectra. The amplitude of the

noise term added to each spectrum was varied randomly, so that the set of spectra spanned a range of signal/noise ratios. Figure 7 presents scatter diagrams of the values returned for each of the five model parameters, plotted against signal/noise ratio. It is apparent that the parameter estimates cluster about their correct values for all five parameters and at all noise levels. As expected, the parameter estimates become more widely scattered (i.e., the parameter variances increase) at higher noise levels. While generating the results in Figure 7, the method took an average 13 iterations and 0.45 seconds of computer time per spectrum. Thus, it took around 12 minutes to process all 1 500 spectra.

Figure 8 presents scatter diagrams of the parameter standard deviations (i.e. $\sqrt{\sigma^2}$ returned by the fit, plotted against signal/noise ratio. To test if the variances calculated in equation 29 appear reasonable, we compared these values to the distributions of parameter estimates within 7 subsets the entire 1 500 spectra. That is, we divided the spectra in 7 subsets, spanning 7 contiguous intervals of signal/noise ratio. Within each of these subsets we computed the variances of the parameter estimates. The square root of these values represents the actual standard deviations that were obtained within that subset; in Figure 8 these values are superimposed as heavy black diamonds onto the scatter plots of individual standard deviation estimates. The variances (and hence standard deviations) returned from fitting an individual spectrum generally appear to be consistent with actual variances over a large set of trials. There is perhaps a suggestion that the uncertainties in Doppler shift and aerosol scattering intensity may be slightly underestimated for signal/noise ratios below ~100. However, this is an inexact test of the routine, and the agreements are close. Further, spectra with signal/noise ratios this low are unlikely to be of much use anyway. Overall, refinement of the error estimation method does not seem justified for now.

9.4 The effects of fitting fewer parameters

As mentioned earlier, the analysis program has been written to allow one or more of the model parameters to be held 'frozen' at some externally-supplied initial value. This facility was included as it is anticipated that several of the parameters could potentially be estimated independently by other means, for a typical lidar experiment. It is instructive to see how inclusion of such *a-priori* knowledge can improve the fit results for the remaining parameters. Figure 9 shows the parameter standard deviations for a second set of 1 500 spectra, statistically equivalent to those used to generate Figure 8. However, in this case, the fits were constrained to only search for temperature, Doppler shift, and molecular backscatter intensity. The aerosol and background intensities were held fixed, at their correct values.

The uncertainties in temperature and molecular scattering intensity were dramatically reduced by applying *a-priori* knowledge of the background and aerosol intensities. A discernible, but less dramatic, improvement was also seen for the Doppler shift parameter. The computational load was also reduced. On average, the routine converged after only nine iterations and took 0.32 seconds per spectrum.

9.5 The effects reduced instrumental finesse

The simulations presented so far have used an instrument function finesse ~20, which is toward the upper limit of what is likely to be achievable for a practical FPS. We now consider the results of a simulation using a broader instrument function, with a finesse of ~9, as depicted in Figure 10. This instrument function was generated using the same procedure as previously, except that the main Gaussian's equivalent N_2 temperature was specified to be 24 Kelvins (rather than 8 Kelvins). The

widths of the other contributing functions were scaled relative to that of the main Gaussian, so all component functions were similarly broadened, yielding a final instrument function width roughly twice that used previously. (Note, however, that at 14.3 channels, the FWHM of this function is still significantly less that the FWHM of 22.9 channels for scattering from N_2 at 250 Kelvins.)

Figure 11 shows the parameter standard deviations for a another set of 1 500 spectra, in this case generated using the broadened instrument function depicted in Figure 10. As can be seen, the parameters worst effected were molecular and aerosol scattering intensities, followed temperature. Some degradations also occurred for the background intensity and Doppler shift, but these were less significant. It would seem that a high finesse instrument is most critical for distinguishing between aerosol and molecular scattering, and for measuring temperature. If wind is the only quantity of interest, there is less incentive to achieve high finesse in the Fabry-Perot system.

10 CONCLUSIONS

A new method has been presented for analysing Doppler spectra from an atmospheric lidar experiment. The method allows for the effects of both aerosol and multiple-species molecular scattering, a continuum background, and instrumental broadening of the recorded spectra. The analysis proceeds by least-squares fitting a five-parameter model function to the recorded spectra. Values for three of these parameters are computed analytically, and two by numerical searching. One or more parameters can be held fixed at externally-supplied values, while the remainder are computed accordingly. Applying the method to simulated spectra indicated that it converges stably over a wide range of signal/ratios, and even given very poor initial guesses for the two search parameters. Parameter estimates from sets of 1 500 trial spectra indicated that the means of the parameter estimates closely matched the values used to generate the model spectra. Further, the variances of the sets of results were, in most cases, consistent with the variances returned by individual fits. There was perhaps a suggestion that the variances in Doppler shift and aerosol scattering intensity were slightly underestimated for signal/noise ratios ~100. However, the discrepancy did not appear large enough to warrant further refinement of the error estimation for now.

The simulations indicated that significant improvement in the accuracy of estimation of some parameters can be obtained by minimising the number of parameters solved simultaneously, and by making the instrument function as narrow as possible.

Finally, simulations like these have proved useful not just for testing the analysis procedure. They have also yielded useful guidance for the design of the lidar experiment itself. For example, it is apparent that there is less incentive to achieve a narrow instrument function if Doppler shift is the only quantity needed with high precision.

11 ACKNOWLEDGMENTS

This work was supported by the Australian Antarctic Division. The author visited the Antarctic Division's Atmospheric and Space Physics section between July and September 1998, at the kind invitation of Dr. Ray Morris.

REFERENCES

Press, W.H., Flannery, B.P., Teukolsky, S. A. *and* Vetterling, W.T. (1986). Numerical recipes: the art of scientific computing. Cambridge University Press.

Bevington, P.R.(1969). Data reduction and error analysis for the physical sciences. McGraw-Hill Book Company, New York.

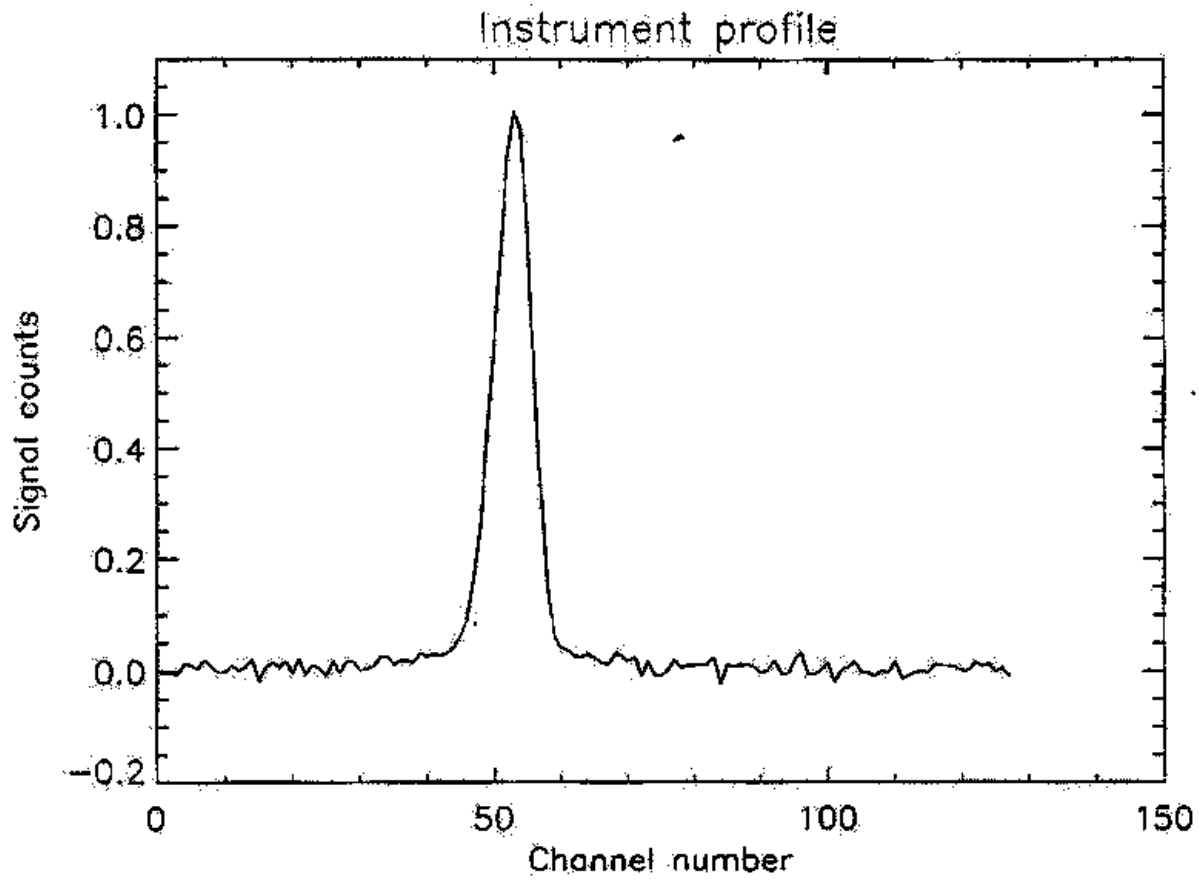

Figure 1: The simulated instrument profile used to test the analysis procedure. This is the version actually supplied to the analysis routine, including the noise term.

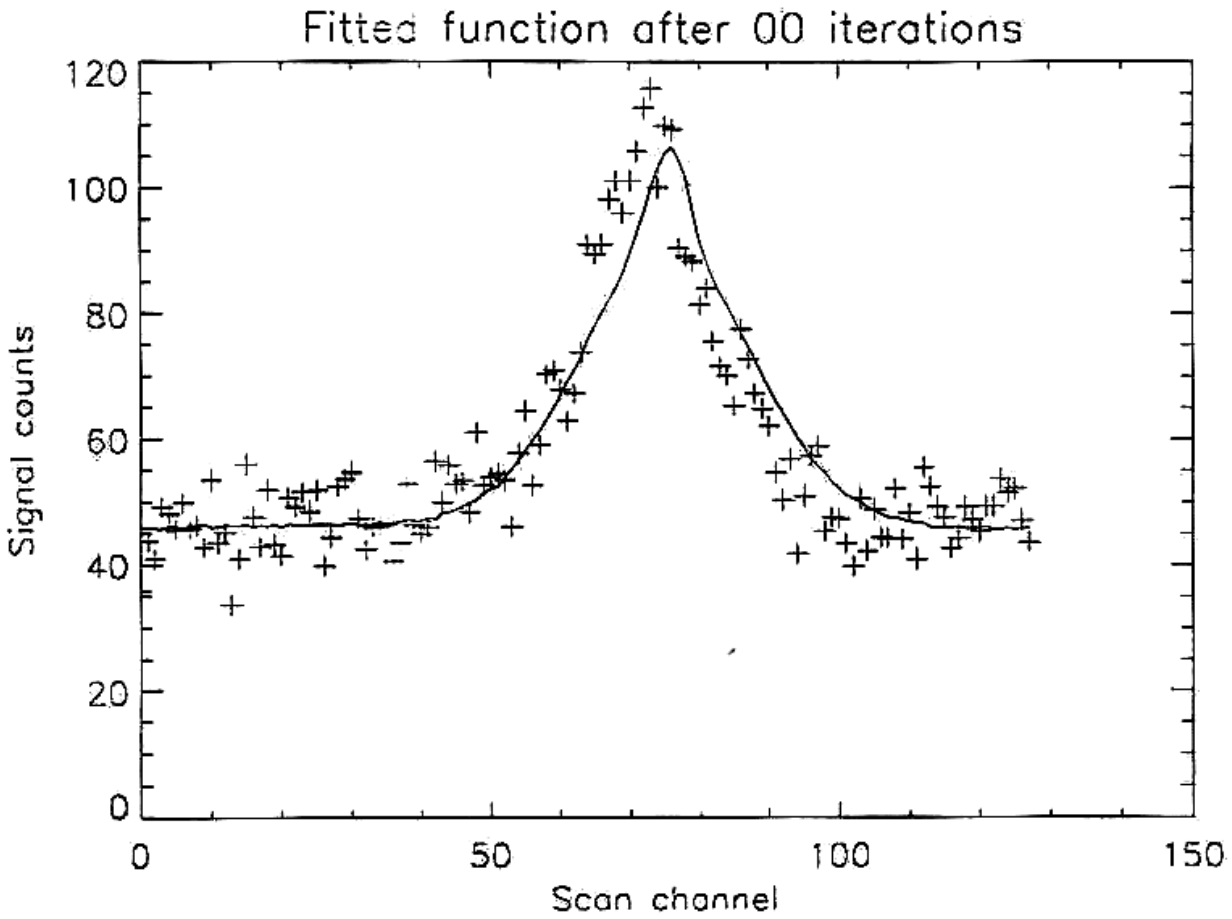
Fitted function after 00 iterations
Signal counts
Scan channel
0
20
40
60
80
100
120
0
50
100
150

Figure 2 a:

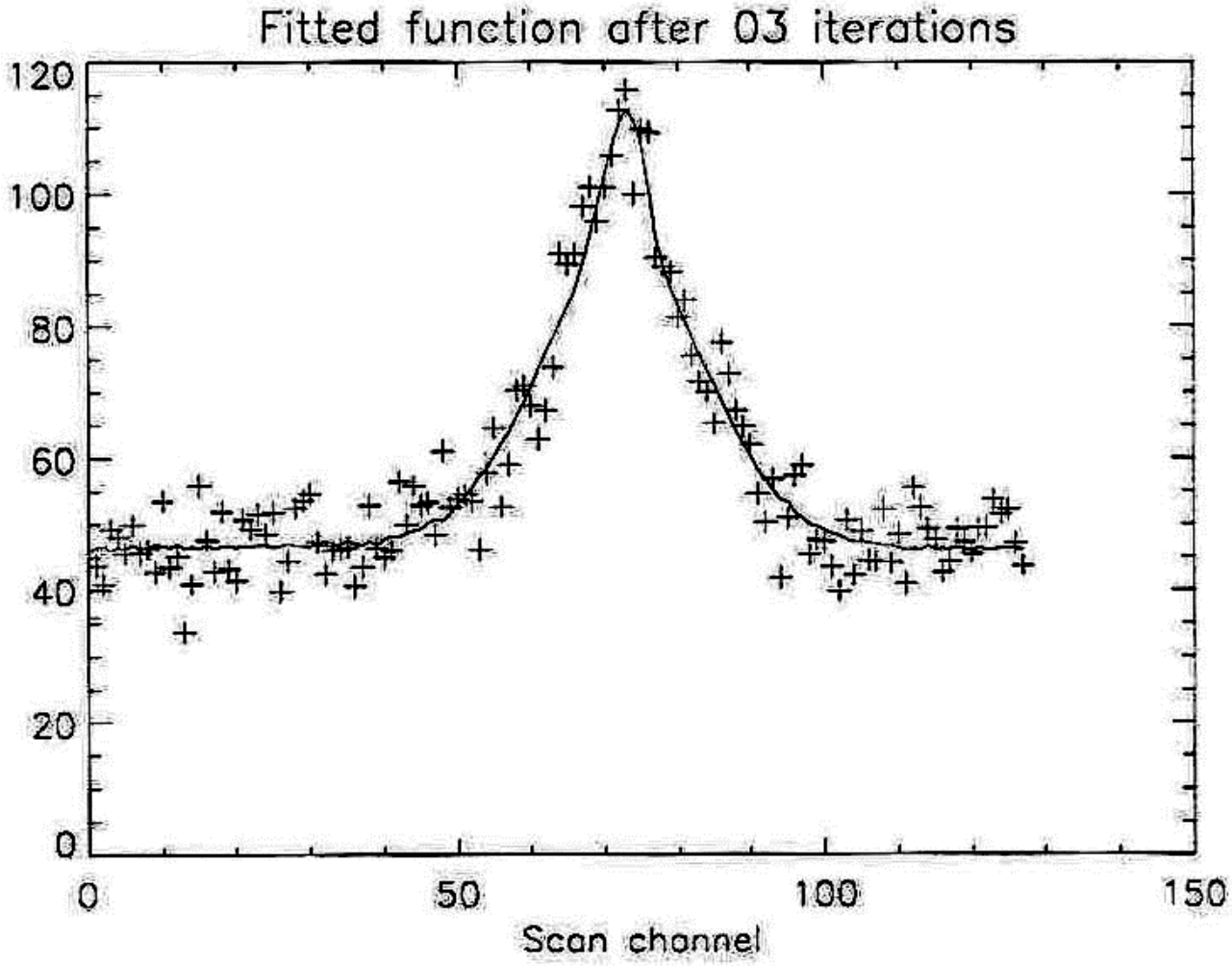
Fitted function after 03 iterations
120
100
80
60
40
20
0
0
50
100
150
Scan channel

Figure 2 b:

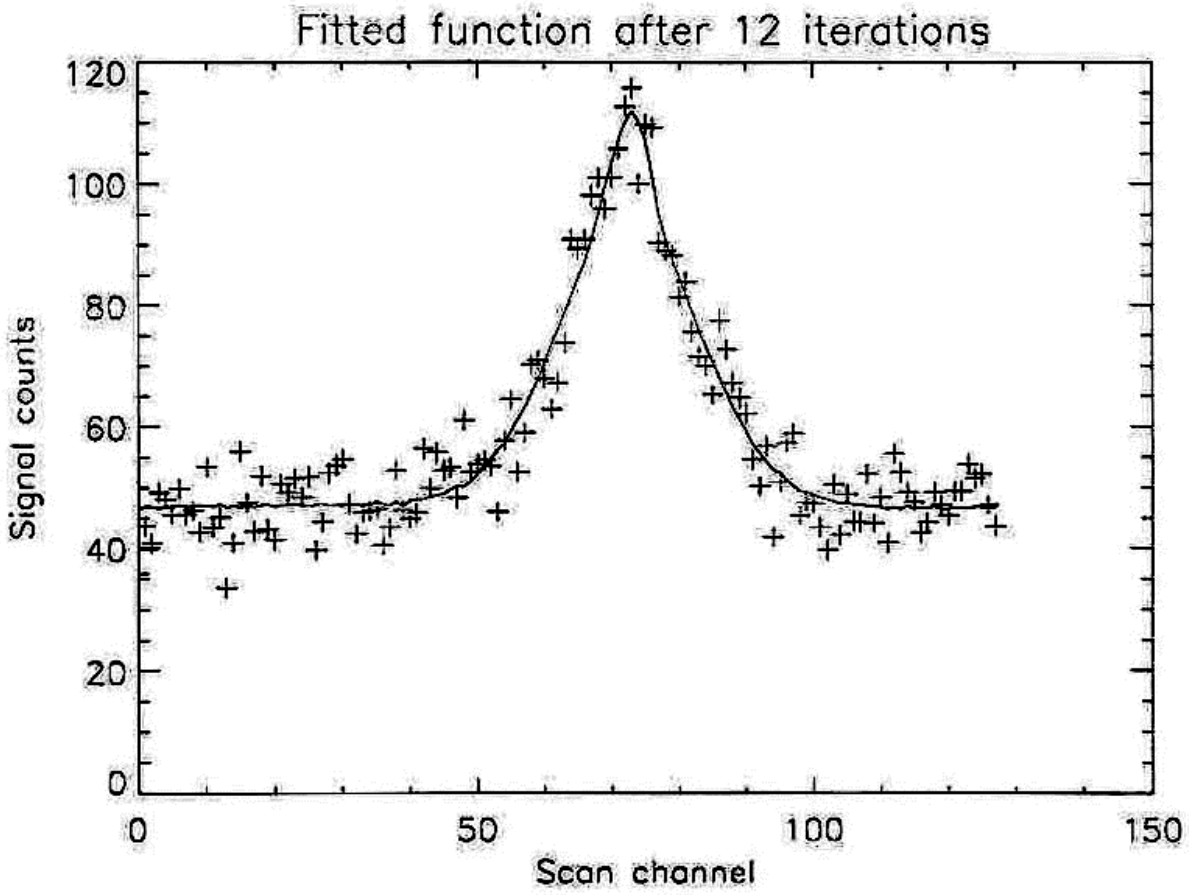

Figure 2 c: These three panels show a simulated Fabry-Perot spectrum (including noise) and the fitted function after 0, 3, and 12 iterations respectively. The function at zero iterations uses intensities (a_0, a_1, and a_2) calculated from the initial guesses for temperature and Doppler shift. The change to the fitted function between 3 and 12 iterations was minor, but discernable.

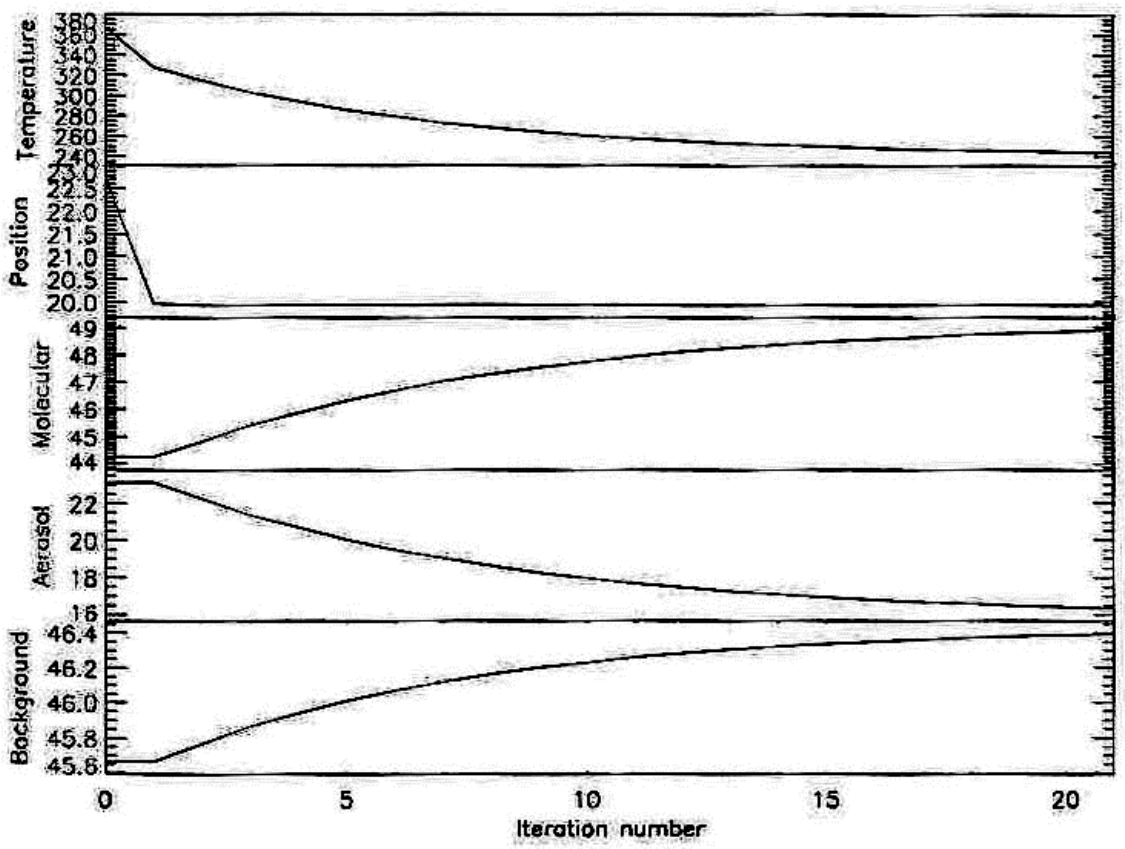

Figure 3: These 5 panels show the sequence of estimates for each of the five model parameters, as functions of the analysis iteration number. Each parameter converged stably. Note the Y-axis label "Position" for the second panel down is actually synonymous with "Doppler shift".

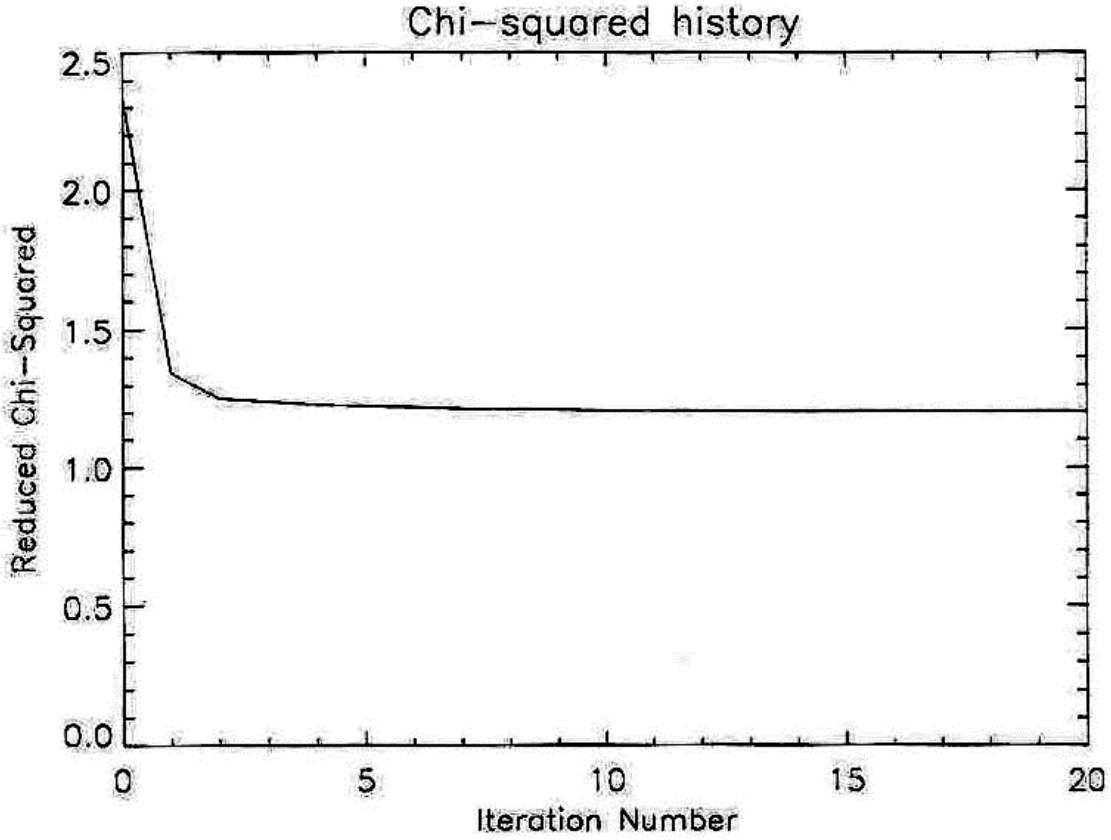

Figure 4: The sequence of reduced χ^2 values as a function of iteration number, for the fit to the sky spectrum shown in figure 2.

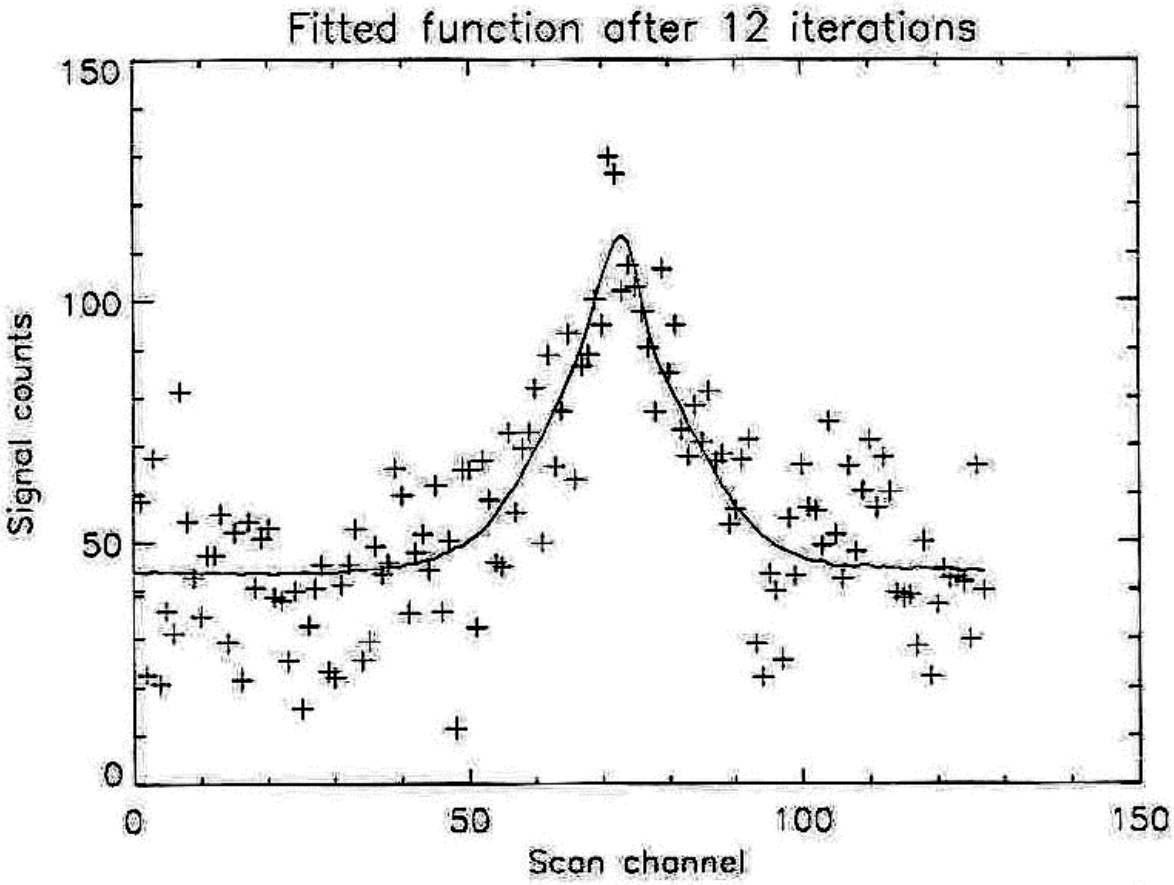

Figure 5: A low signal/noise ratio spectrum (i.e. sig/noise=91), and the corresponding fitted function.

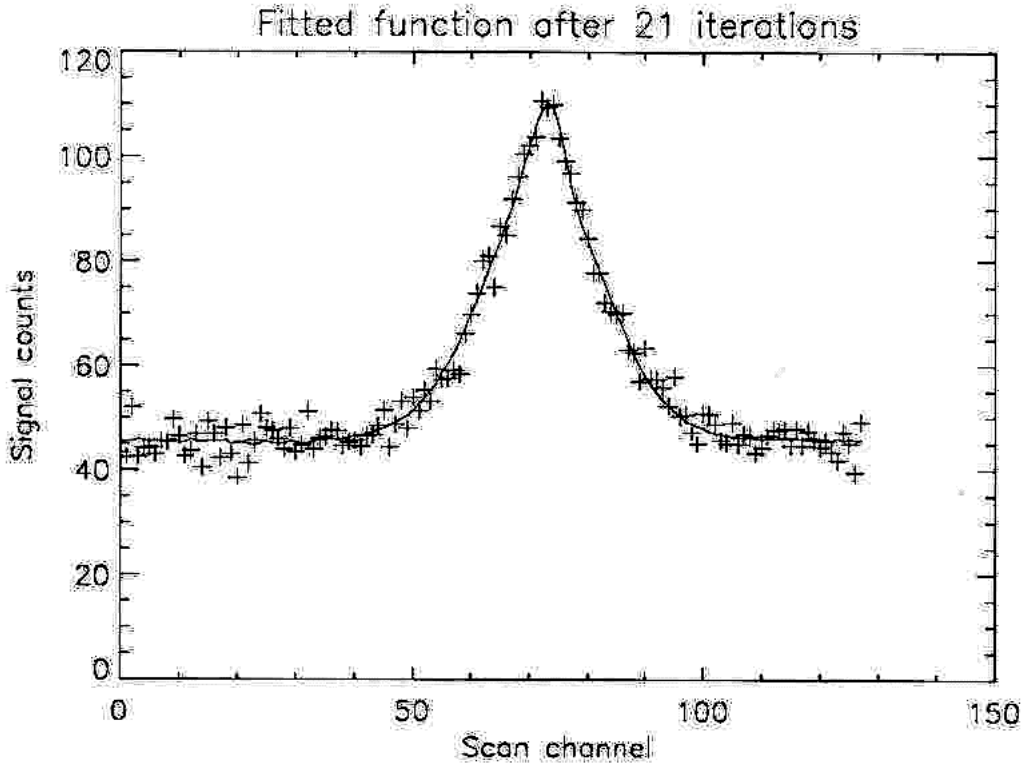

Figure 6: A high signal/noise ratio spectrum (i.e. sig/noise=2194), and the corresponding fitted function.

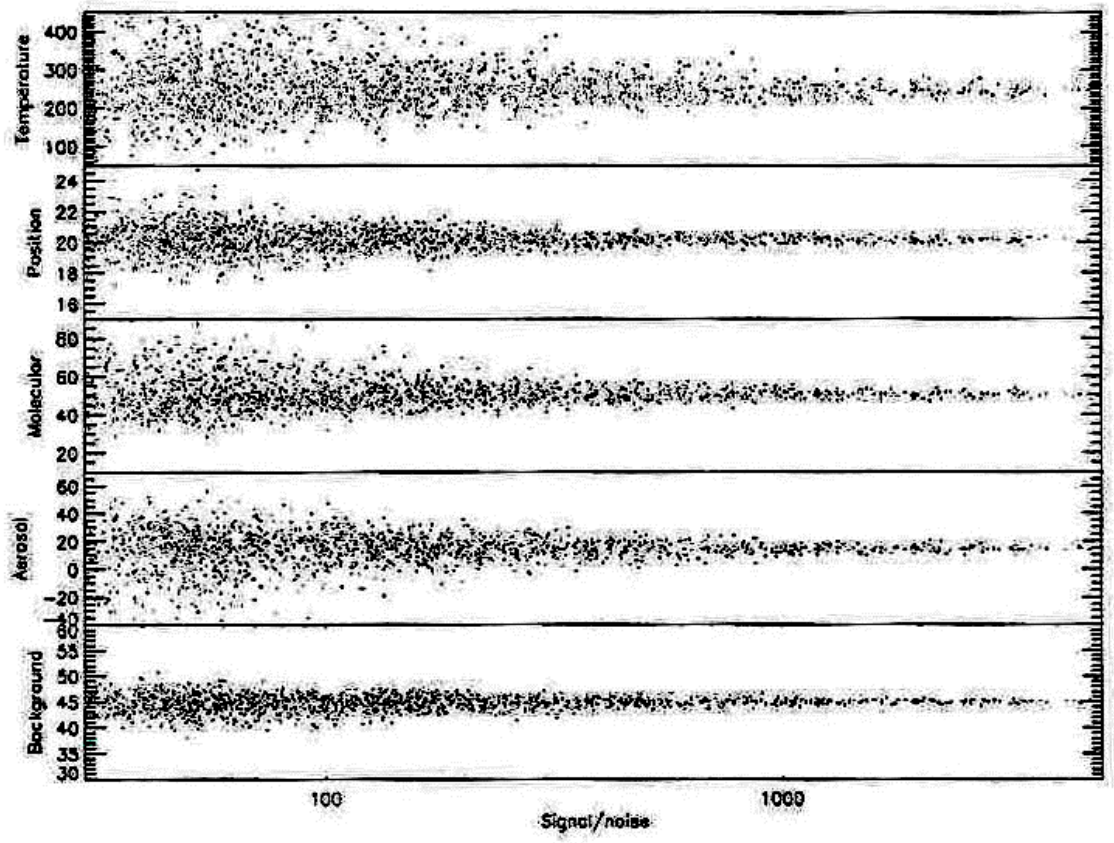

Figure 7: Scatter plots of parameter estimates versus signal/noise ratio, for a set of 1500 numerically generated test spectra.

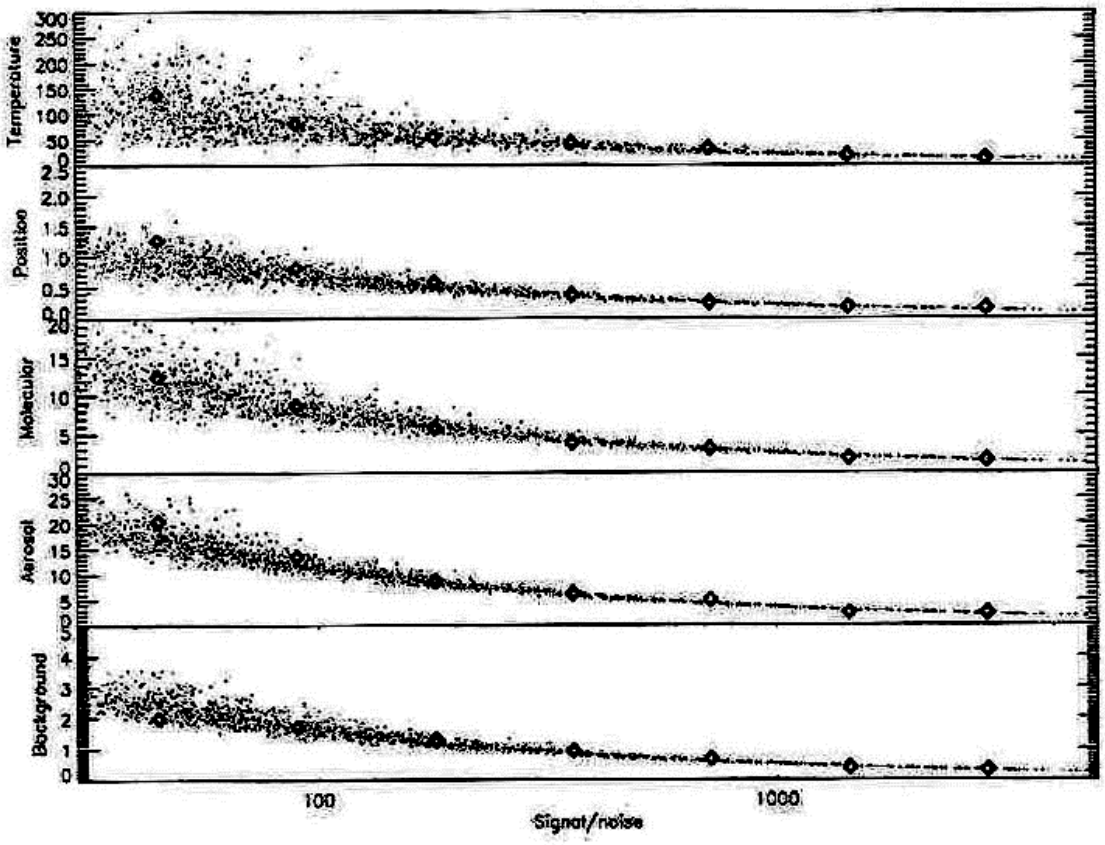

Figure 8: Scatter plots of parameter standard deviations returned by the analysis program, plotted against spectral signal/noise ratio (small points). Superimposed as heavy diamonds are the actual standard deviations of the parameter estimates obtained in 7 subsets of the spectra, spanning 7 contiguous intervals of signal/noise ratio.

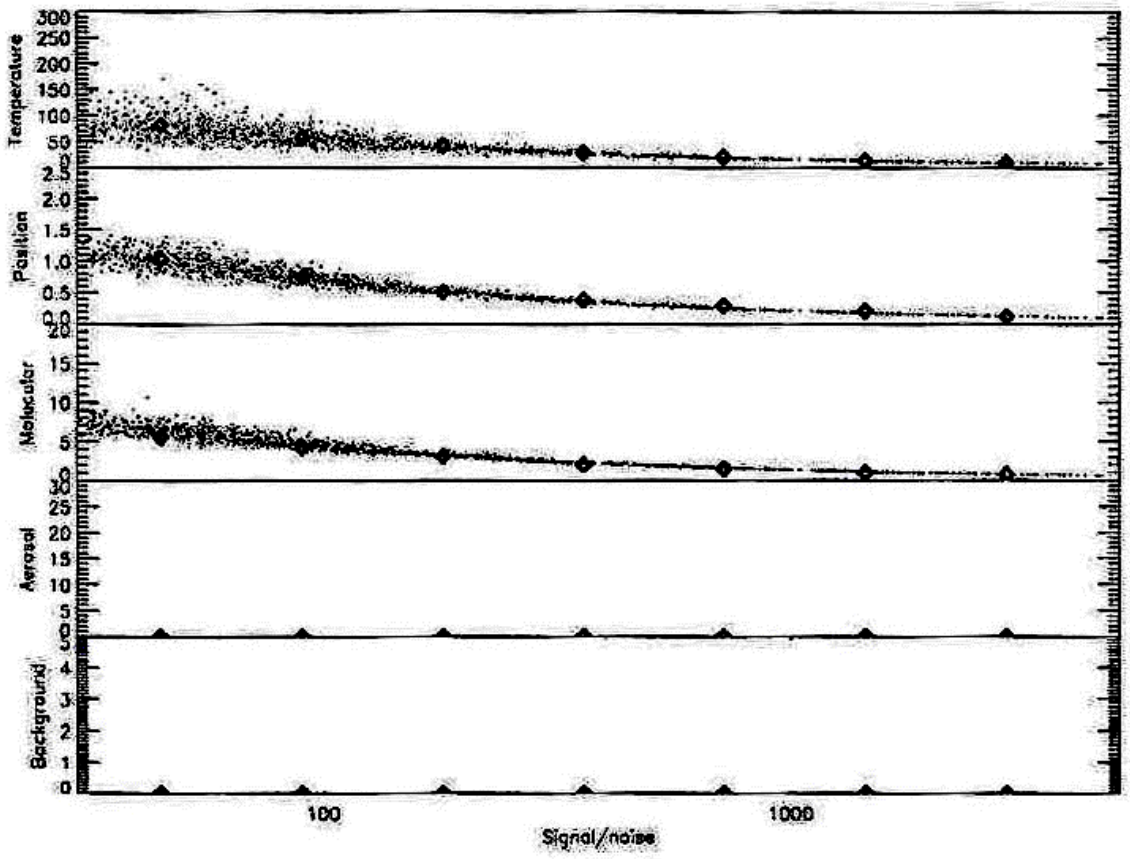

Figure 9: As before, this figure shows scatter plots of parameter standard deviations, both estimated for individual spectra and calculated from the distributions. In this case, only the top three parameters were fitted. The bottom two were held fixed at their correct values throughout the fit.

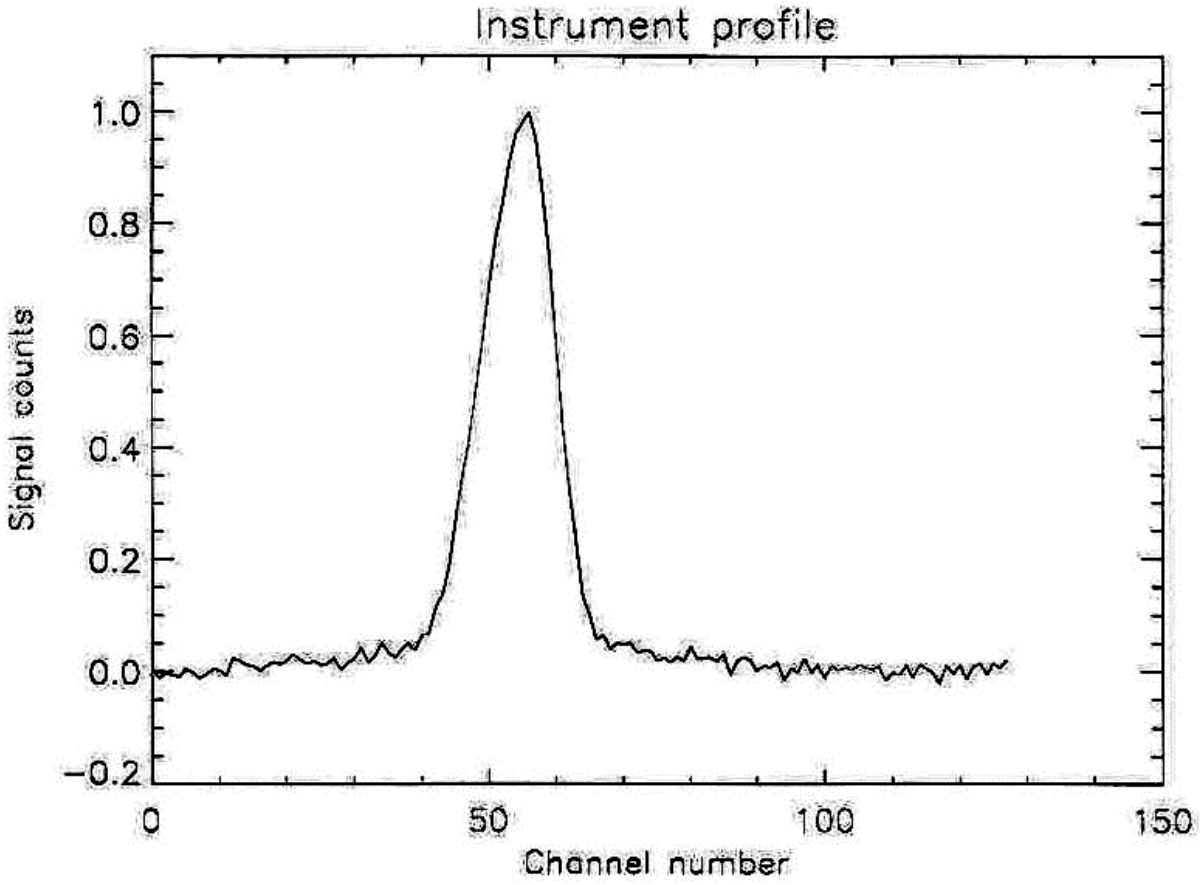

Figure 10: This figure depicts an instrument profile approximately twice the width of that used previously. Simulations using this profile were run to examine how reduced instrument finesse would degrade the recovery of model parameters.

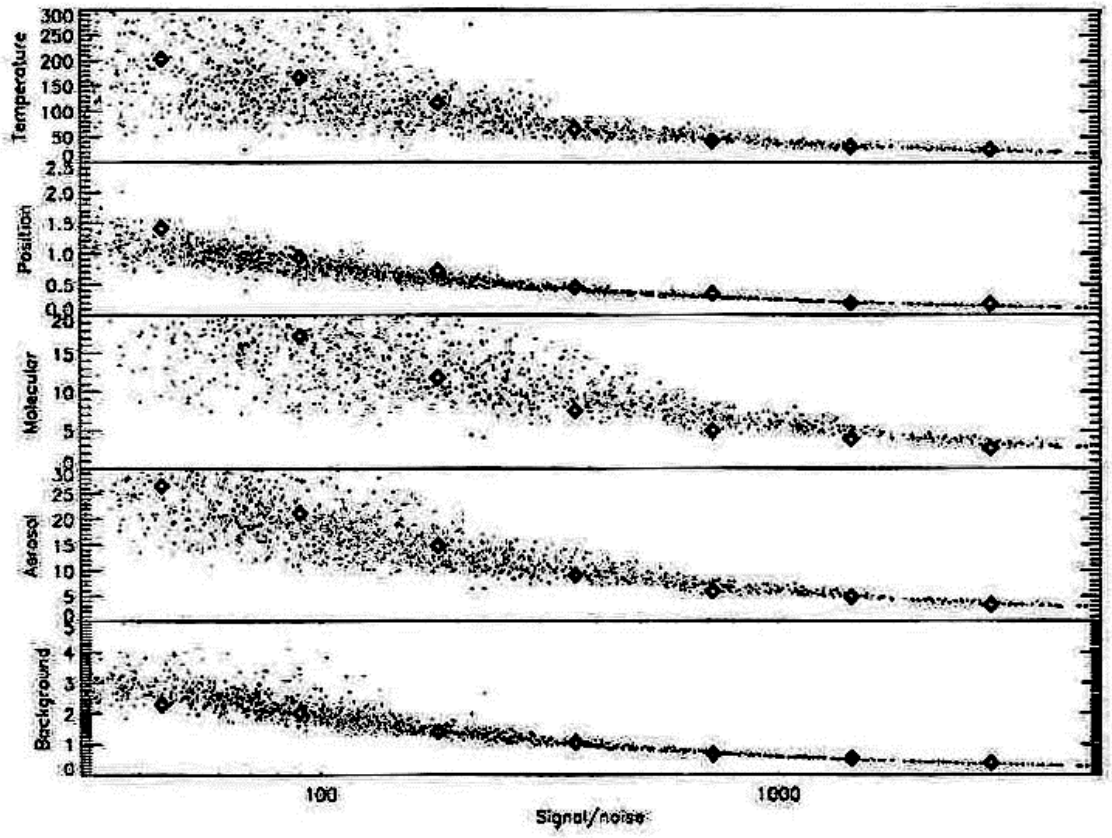

Figure 11: This figure shows scatter plots of parameter standard deviations from analysis of spectra generated using the broadened instrument function. Molecular and aerosol scattering intensities are the parameters most degraded by the reduced finesse, followed by temperature.