The Spectral Correlation Function – A New Tool for Analyzing Spectral Line Maps

Erik W. Rosolowsky

July 21, 1998

Contents

Abstract

At the current time,the study of the Interstellar Medium (ISM) can be approached from two different angles. The first is by analyzing observational results from radio astronomy, thereby deducing the structure of the ISM. The second method is to model the conditions in the ISM using Magneto-Hydrodynamic (MHD) simulations where the initial conditions and time evolution illuminate the properties of the ISM. The Spectral Correlation Function (SCF) is a new data analysis algorithm that measures how the properties of spectra vary from position to position in a spectral-line map. For each spectrum in a position-position-velocity data cube, the SCF measures the differences in shape and size between that spectrum and a specified subset of its neighbors. This algorithm is intended for use on both simulated and observational data cubes. We have shown that a histogram of the SCF for a map is a good descriptor of the spatial-velocity distribution of material. In this thesis, two observational and two theoretical data sets derived from MHD simulations are analyzed. By studying the SCF distributions for these cubes and the same data in randomized positions, the data indicate that the distributions for all data sets in their original configurations are similar, and these are much different from the data sets with their constituent spectra in randomized positions. The character of these similarities and differences can be used to analyze the data sets. The ultimate aim of the SCF project is to use the SCF both on its own, as a tool for simplifying the analysis of spectra maps, and as a method of evaluating how well simulations resemble observations.

Chapter 1

Historical Background and the State of the Art

This section of the paper is meant to give the reader a general introduction to the type of radio astronomical observations which inspired the development of the Spectral Correlation Function (SCF). The SCF is descended from the fusion of two paths for studying the Interstellar Medium (ISM). One path is the observational technique of generating spectral line maps by taking spectra from several positions in an astrophysical object and analyzing the data contained in these maps. The other path is the theoretical technique of using Magneto-Hydrodynamic (MHD) simulations to model the properties of the ISM. The SCF provides an interface between these two techniques. It is intended to be a tool for comparing the products of these two paths, as well as a method by which spectral line maps can be analyzed in their own right.

1.1 The Interstellar Medium

Outside of objects in the solar system, the most prominent astronomical objects are the stars in our galaxy. Apart from a few prominent nebulae, early astronomers thought that the galaxy was defined strictly in terms of these luminous bodies. It was only near the beginning of the twentieth century that astronomers began to suspect the presence of more material in the galaxy aside from what was visible. Only a few observers prior to this time noticed or paid much heed to the anomalous 'holes' in an otherwise uniform background of stars. In 1889, E. E. Barnard compiled an atlas of these dark patches and recognized them as clouds of obscuring material in the galaxy. The presence of such obscuring matter led other observers to posit the existence of dark, obscuring material between the stars that was not as dense or opaque as Barnard's dark clouds. Observations established that the light from distant stars was diminished by this material and corrections to distance estimates were established, drastically expanding the size-scale of the universe. Questions regarding the nature of this material remained to be addressed.

Since that time, a century of study has established a large body of knowledge about the ISM. In understanding the ISM, astronomers have come to split the ISM into its two principal components: gas and dust. The former is further subdivided into five basic cat-

egories, characterized by the names of the regions in which the gas is found (Dyson and Williams, 1980, 172 ff):

Name	Temperature (K)	Typical Form	Density (m^{-3})
Coronal Gas	5×10^5	Ionized	$< 10^4$
Diffuse Nebula (HII)	8000	Ionized and Atomic	$\geq 10^8$
Inter-Cloud	6000	Ionized and Atomic	3×10^5
Diffuse Atomic Clouds (HI)	70	Atomic	3×10^7
Cool Molecular Cloud	20	Molecular	$10^9 - 10^{10}$

Table 1.1: Classification of the Gas Components of the ISM

Astronomers study each of these regions individually, for different techniques are better suited for the study of certain regions than others. For example, the low energy environments of cool molecular clouds are particularly well suited for study using the techniques of radio astronomy because such low energy interactions produce low energy photons. Hence, the radio data sets analyzed using the Spectral Correlation Function will usually focus on these cool objects.

The dust component of the ISM exists primarily in the cooler regions and consists large atomic and molecular agglomerations called grains. These grains form from the material blown off cool stars by stellar winds. The cooler regions preserve the grains from destruction because the thermal motion of the ambient gas lacks sufficient energy to destroy the grains in collisions. These regions, characterized by high densities of molecules and dust grains, are referred to as molecular clouds. The material in this thesis deals primarily with the observations and modeling of these objects.

1.2 Spectral Lines and Spectral Line Maps

Spectroscopy has always been a tool of immense use in physics; it is even more so in astronomy. Spectroscopic observations of the ISM were first taken in the visible wavelengths, examining the prominent nebulae in the sky. Some nebulae had similar spectral features to stars near them and were categorized as reflection nebulae. Others exhibited spectral features associated with the recombination of hydrogen, indicating the presence of a large body of ionized gas. Observations continued in this fashion until 1951 when Ewen and Purcell observed interstellar neutral hydrogen in the ubiquitous 21 cm spin-coupling transition of the ground state. From this point, the door to observing the ISM in wavelengths outside the visible was opened. In 1963, the absorption caused by OH at $\lambda = 18$ cm was discovered in interstellar clouds. Soon thereafter, emission from several different molecules was discovered. The standard source of radiation at long wavelengths came from the rotational transitions of heavy molecules. The study of spectra from these molecular sources swiftly grew into a new field of astronomy.

Eventually, observational technology progressed to the point that spectra from individual portions of molecular clouds could be analyzed separately. Thus, a cloud could be

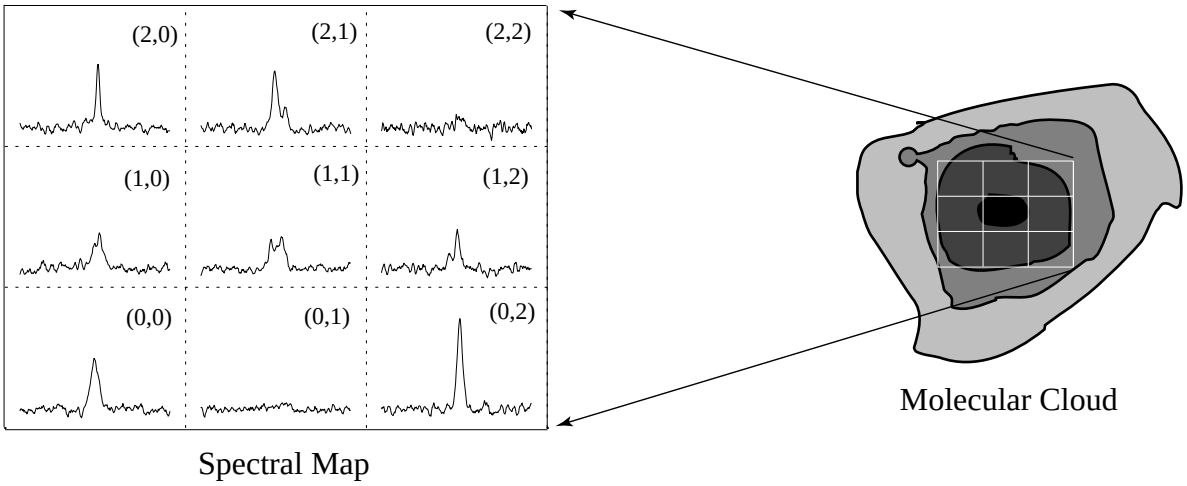


Figure 1.1: A sample spectral map.

'mapped' by taking a series of spectra over its extent, creating a grid with a spectrum at every intersection. An example of such a map appears in Figure 1.1.

Such maps have an inherently three dimensional coordinate system, with two axes representing position and the third representing the frequency of the spectrum. At every point in this three dimensional cube, there is an antenna temperature assigned. Antenna temperature is a conventional measurement in radio astronomy corresponding to the power received from incoming radiation. The use of the word 'temperature' is carried over from the first days of radio astronomy and is the same amount of power that a perfect resistor would emit in the given frequency range if the resistor were substituted in place of the antenna. In the spectra measured by radio telescopes, antenna temperature is the substitute for the intensity that one would normally find in a spectrum. When a spectral line is identified with a given absorption or emission feature, the difference between the rest frequency of the transition and the observed frequency is converted via the Doppler effect into a velocity. Thus, the cubes are often measured in position-position-velocity space.

1.3 Data Analysis Techniques

The analysis of a spectral-line data cube can be very time consuming, especially when the size of the maps grows beyond a few hundred pixels. As a consequence, several analytical tools have been developed to speed analysis. All of these tools involve condensing a large portion of the map to fewer data by assigning a characteristic number for a spectrum to represent that spectrum. Examples include assigning the peak antenna temperature or the mean velocity shift from a spectrum to that spectrum's point in the map, thereby reducing the three dimensional cube to a two dimensional map. These maps can be analyzed quickly and the salient features can be selected for more precise analysis. Such reductions can be perilous, however, for there is the risk of omitting too much data.

In addition to these rather simple methods, much effort has been invested in develop-

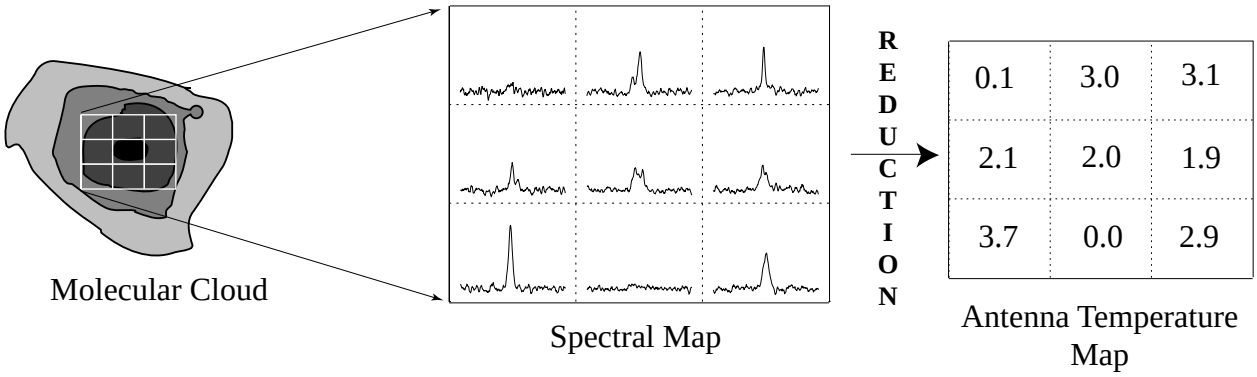


Figure 1.2: Example of creating an antenna temperature map from a spectral map by replacing each spectrum with the corresponding peak antenna temperature. The reduced maps are usually then converted to greyscale maps.

ing other analysis techniques for these immense data sets. These techniques can be divided into three basic categories:

- More rigorous analysis of the two dimensional maps
- Examinations of the three dimensional structure in the cubes
- Examining the relations between objects at different size scales in the cubes.

The following represents a brief summary of the methods used to date. It is by no means comprehensive; the only purpose is to illustrate what approaches are being tried and how the SCF differs from them. Briefly, the SCF can be summarized as a measure of how a given spectrum differs from its neighbors. These differences are further broken down into differences in shape, antenna temperature scale and velocity offset. By examining where groups of spectra are closely related and where they are not, the SCF provides new information about data cubes made from the ISM.

1.3.1 Two-Dimensional Analyses

These techniques of analysis carry the examination of two-dimensional maps to their most mathematically precise extent. As in simple analyses, the spectral maps have their constituent spectra reduced from a full spectrum to a characteristic number for that spectrum. A cartoon of this reduction appears in Figure 1.2. The most common characteristics are: peak temperature, integrated temperature, line width and line offset in velocity space. All of these characteristics are illustrated in the diagram found in Figure 1.3

Using these maps, the analyses are performed. The techniques of wavelet analysis have been brought to bear on these maps (Gill and Henriksen, 1990). These techniques center around convolving the "wavelet" function over the map at different scales, highlighting structures of a comparable shape. This analysis shares much with the concepts in Fourier analysis, except the particular analysis can select the wavelet shape.

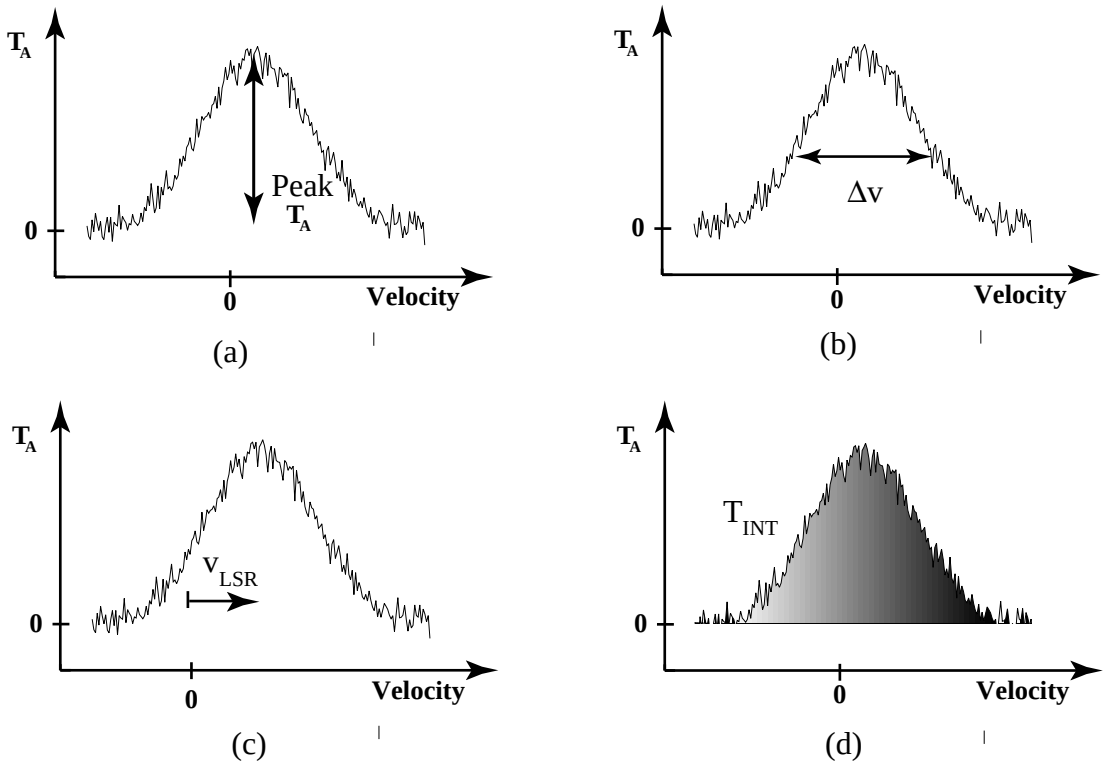


Figure 1.3: Diagram illustrating the definitions of the basic spectral line parameters. (a) Peak antenna temperature. (b) Line width (Δv). (c) Line offset (v_{LSR}). (d) Integrated temperature (T_{int}).

In contrast with the wavelet techniques, Houlahan and Scalo (1992) propose the use of structure tree statistics to characterize the properties of the ISM. This method focuses specifically on the hierarchical nature of the ISM. A map of the data with a single contour is plotted and the structure of the individual contour at a given value is examined. Specifically, the number of closed circles made by the contour is counted. As the value of the contour is raised (lowered) it splits (merges) into what are dubbed its children (parents). By examining the number of children/parents a certain contour has at a certain level, the hierarchical nature of the cloud can be characterized.

Another method of analysis is proposed by Wiseman and Adams (1994) who consider all of the clouds to be a set in a topological space. Structure is imposed upon this space by defining a pseudometric between the elements based upon their given properties. The metric functions used convert differences in the volume and mass into distances in this space. By examining the distribution of distances between these elements, a statistical analysis of the cloud is thereby performed.

All of these methods rely upon the contraction to a two-dimensional map and the subsequent sacrifice of data. The results presented by each of these methods illuminate a different aspect of each of the maps. Still, the SCF extends beyond them by considering the entirety of a spectrum and its relation to its neighbors when reducing the cubes to two dimensional maps.

1.3.2 Three Dimensional Structure Analyses

Not all techniques sacrifice the third dimension of the data. The methods discussed in the following section consider the spectral maps as a cube of data in three dimensional space. From these cubes, surfaces of constant antenna temperature can be considered and their structure examined.

The prime example of realizing the map as a three dimensional structure is found in the work of Williams *et al.* (1994). In this research, the CLUMPFIND routine is developed to examine data cubes. This routine searches for the peaks in the three dimensional array and then appropriately sorts the neighboring points into 'clumps' corresponding to the 'parent' peaks. The algorithm's primary task is to sort which neighbors belong to which peaks if the choice is not obvious (i.e. those measurements at saddlepoints). Once all the clumps are accounted for, statistical analyses on the resulting structures are performed, examining mass, peak temperature, velocity and size of the clump.

Studies of the fractal nature of the ISM offer illuminations into its structure by using these three dimensional analyses. An example of fractal analysis is presented in Elmegreen and Falgarone (1996). The fractal approach relies upon the observation that there are similarities between the two and three-dimensional structure of the ISM when it is viewed on specific size scales. The nature of these scale-similarities is characterized by the fractal dimension of the data. Different formation models for the ISM predict different fractal dimensions, thus models of the ISM can be evaluated by comparing the fractal dimensions produced by simulations with those observed in the ISM.

The SCF differs from both of these methods in that it considers the elements of a map as spectra rather than considering the whole map as a three dimensional array. By making this distinction, the SCF is able to locate similarities between structures at different rest velocities which would not appear in three dimensional analysis.

1.3.3 Scale Investigations

The recognition of similarities at different size scales motivated fractal analyses. Observations also indicate that there are relations between some observed quantities and the size scale over which these quantities are measured. The fiducial example is given by Larson (1981) who noticed that there exists a power law relationship between the width of a given line and the scale over which the width is measured (that is $\Delta v \propto \ell^a$). This observation probes the nature of turbulent motion in the ISM where the specific kind of turbulence is related to the exponent a. Such analysis is continued in other papers and the scales on which such relationships break down can be used to understand the behavior of turbulence in the ISM (Goodman et al., 1997).

In addition to testing correlations among size-scale and other simple observables, astronomers have also worked to examine more complex relationships. In these studies, the similarities between basic properties of two spectra within a map are compared in terms of the distance in pixels between those spectra. An example of this kind of analysis is the use of the autocorrelation functions (Dickman and Kleiner, 1985). In this analysis, a certain characteristic of the spectra (for example, the velocity offset) is compared between all the spectra a set distance apart. The average value of the difference between these characteristic

values is averaged over the map and the result characterizes that parameter on the given size scale. The size scale is then changed and the behavior of the function is then analyzed. Other types of functions are also used but their primary purpose is the development of relations between observed spectral quantities and size scales.

A final method of interest is that of Principal Component Analysis (PCA) established by Heyer and Schloerb (1997). The key to PCA is to realize that the various line parameters in a map are related. For example, the amount of material giving rise to a line may influence both the peak antenna temperature and the width of the line. In realizing this, the data can be reduced to a few independent variables called the principal components. These components represent all significant variation in the data. The analysis of how well these principal components model the data as well as studying the maps of these components will produce more information about the clouds.

The principal difference between these methods and the SCF is that they focus upon the relationships between objects at set scales often larger than those considered in the analysis of neighbors that the SCF produces. The SCF also focuses on different qualities than do the above studies and most importantly, it preserves the three dimensional nature of the data that these studies reduce in other fashions. In some fashions, the SCF is most akin to autocorrelation functions; however, the notion of scale is much more malleable with the SCF than it is with these other analyses.

1.4 The Origins of the Spectral Correlation Function

The above methods of analysis present several manners in which the data cubes can be examined. In all cases, the methods involve the reduction of the data to a manageable and easily interpreted set. These reductions choose to include or exclude specific aspects of the data in their condensation; and naturally, much of the character of the data is lost or obscured. The SCF offers another approach to analyzing data cubes, preserving some aspects of the data and eliminating others. In particular, the SCF focuses on the similarities among constituent spectra in a data cube in shape, size, and velocity offset.

Certain reductions of data sets eliminate the uniqueness of the data set that generates the reduced data set. In other words, many different data sets will produce the same reduced data set. A particularly relevant example would be the reduced data set created by generating a histogram of the peak antenna temperatures for the constituent spectra in the cube. Naturally, a cube with the same spectra but in randomized positions will produce the same histogram. This type of uniqueness loss is the fundamental problem that the SCF was formulated to help solve. Specifically, the SCF set out to probe the uniqueness of MHD simulations of the ISM and the spectral maps that these cubes produce. The goal of any simulation of the ISM is to mimic the properties of an observed portion of the ISM as accurately as possible. Because it is impossible to establish the exact initial conditions that led to a particular object observed in a spectral map, these simulations must mimic the more general qualities of observational data. A proposed goal for the simulation projects is to develop a model that reproduces the distributions of spectral characteristics (e.g. antenna temperature, line widths, velocity offsets etc.) that are found in observations of the ISM. Falgarone *et al.* (1994) have proposed using the statistical moments of these distributions

to evaluate the degree to which simulations approximate the observed qualities of the ISM. In their paper, they find that the observed maps moments matched their simulations with a high degree of accuracy. To date, this is the only published work specifically evaluating hydrodynamic simulations by comparing them with real spectral maps.

As mentioned previously, the goal of reproducing characteristic distributions suffers from a lack of uniqueness in the data set from which the distributions are generated. Hence, the placement of spectra in the data cube can be random and the same distributions will be generated. Naturally, in the observed ISM, it is impossible to formulate a cube which will have the randomized positions of a possible theoretical cube. More specifically, the spectra in an observed map of the ISM are assumed to share some characteristics with their neighbors in pixel space because the physical regions from which the spectra are formed should share some physical similarities with their neighbors in real space. The hypothesis is that the shared characteristics of the neighboring spectra can be quantified through the development of an appropriate correlation function. These observations were prompted by the Falgarone *et al.* work on comparing spectra. In particular, the spectra in their final simulations seem to have radically dissimilar spectra next to each other, in contrast with the observed data with which they compared their work. It is this work which has prompted the development of the SCF to refine the evaluation criteria for MHD models of the ISM.

Chapter 2

The SCF Algorithm

The focus of the SCF project is to develop a function that represents the correlation between two spectra and generate a computer algorithm that efficiently calculates this correlation. The function is a type of two-dimensional autocorrelation function with additional parameters included to investigate the nature of the correlations over an entire spectrum. This task is subdivided into three distinct parts, each of which is described in the following sections. First, a meaningful correlation function must be defined. Next, such a function must be developed into a working computer algorithm. Finally, the presence of noise in the data must be addressed in order that a meaningful comparison between spectral maps can be made.

The basis for this plan can be found in Goodman (1997).

2.1 Mathematical Development of the SCF

2.1.1 The Deviation Function

To measure the correlation between two spectra, each spectrum is viewed as a function in velocity space: $T_A(v)$. A new function, called the deviation function, is defined based on these two functions:

$$[\delta_{1,0}(v)]^2 = [T^*_{A,1}(v) - T^*_{A,0}(v)]^2 \tag{2.1}$$

The deviation function is in units of antenna temperature squared $[K^2]$ and is identically zero for identical spectra. In order to develop a better sense for what the deviation function means, it is reduced to a single number by integrating in velocity space, resulting in a single number with units of $K^2 \cdot km/s$. This number is referred to as the deviation scalar or simply the deviation.

$$D(T_1, T_0) \equiv \int [\delta_{1,0}(v)]^2 dv = \int [sT^*_{A,1}(v - \ell) - T^*_{A,0}(v)]^2 dv \tag{2.2}$$

The free parameters s and ℓ have been introduced so that the deviation can be minimized between the two spectrum, allowing for spectra with similarities in shape and differences in amplitude or velocity offset to be recognized as having some correlation. The

introduction of these parameters is what distinguishes the SCF method from the other analysis routines discussed above. These parameters recognize more kinds of similarities as significant than do the three dimensional analyses or the correlation function methods.

The deviation is also examined with $s = 1$ and/or $\ell = 0$ to understand the differences between the two spectra. The values of these parameters offer hints to the relations between the two spectra. For example, if the value of ℓ is close to 0 for a certain correlation, the two spectra should have similar velocity distributions. If the value of s is close to 1 for the correlation, the two spectra should have similar temperatures over the velocity range.

The deviation function is non-negative by its definition so a deviation of zero is its lower bound resulting from the integral of the zero function. To determine the maximum value of the deviation, the functional definition 2.2 is expanded:

$$D(T_1, T_0) = s^2 \int [T^*_{A,1}(v - \ell)]^2 dv - 2s \int T^*_{A,1}(v - \ell) T^*_{A,0}(v) dv + \int [T^*_{A,0}(v)]^2 dv \qquad (2.3)$$

In the above equation, the cross term represents the correlation between the two functions. If the two spectra are radically dissimilar, this cross term will go to zero. For example, consider two lines radically separated in velocity space so that the near-zero wings of one function are multiplied by the peak of the other and vice versa. In the limit of poor correlation, the integral of the cross term will be zero and the maximum value of the deviation is simply the exterior terms:

$$D(T_1, T_0)_{max} = s^2 \int [T^*_{A,1}(v - \ell)]^2 dv + \int [T^*_{A,0}(v)]^2 dv \qquad (2.4)$$

This understanding of the range of the deviation function becomes essential in normalizing the results of the functions.

2.1.2 Normalization

In order to aid in interpreting the results of the deviation calculation, the value of the calculation can be normalized to the unit interval. The desired normalization is to have a value of 1 indicating perfect correlation and the value of 0 indicating the minimum correlation in the following fashion, thereby defining the Spectral Correlation Function $S(T_1, T_0)$:

$$S(T_1, T_0) \equiv 1 - \sqrt{\frac{D(T_1, T_0)}{s^2 \int T_1^2(v) dv + \int T_0^2(v) dv}} \qquad (2.5)$$

The normalization value in the denominator is chosen because it represents the maximum value of the deviation in the absence of absorption (c.f. Equation 2.4). Absorption in observed spectra only occurs in a few cases and the result is that the antenna temperature of a spectrum becomes negative. Thus, the cross term in Equation 2.3 will be positive if there is a large contribution due to absorption in one of the spectra. In this case, the final normalized deviation will be negative. This is because $\int [\delta(v)]^2 dv > s^2 \int T_1^2(v) dv + \int T_0^2(v) dv$. These effects will result in $S < 0$ and, because the scale factor s is derived using similar integrals, its value will also drop below zero.

2.2 Efficient Optimization of the Deviation

The SCF algorithm will necessitate the calculation of large numbers of deviation functions; hence these calculations will need to be done efficiently and accurately. Minimization of a function in parameter space is the subject of much work in the field of numerical analysis and there exist many routines to use in such minimization problems. Unfortunately, the deviation, viewed as a function of s and ℓ is more difficult to analyze than most functions to which these methods are applied. The main reasons for this difficulty lie in the parameter ℓ. These difficulties can be grouped into three major problems:

1. The value of $D(T_1, T_0)$ can only be evaluated for values of ℓ such that $T_A^*(v-\ell)$ is defined by the original spectrum. In other words if $v' = v - \ell$, then v' must be a velocity for which an antenna temperature is measured in the original spectrum.

2. Several minimization routines require the value of $\frac{\partial}{\partial \ell} D(T_1, T_0)$. Any accurate partial derivative of $D(T_1, T_0)$ with respect to ℓ requires evaluating $T_A^*(v-\ell)$ at several points. For noisy functions, the number of points evaluated must be quite high in order for the partial derivative to be of any use to a numerical optimization routine.

3. The function $T_A^*(v)$ has many local extrema because of noise and thus any minimization routines, in seeking a minimum with respect to ℓ, will find a local minimum. In order to insure this is the absolute minimum, most routines check to see if a randomly selected initial point will cause the algorithm to return to the already located minimum. If not, the values of the function at the two points are compared and the lower one is selected for the same displacement treatment. By numerous iterations of the program, this checking procedure will eventually discover the global minimum. However, for typical spectra, this would require a large number of iterations, far more than most programs are intended to use.

To develop an understanding of what will be required from a minimization routine, Figure 2.1 displays the deviation in $s-\ell$ parameter space. The two large humps in this figure represent the increase in deviation that occurs when both peaks in the spectra are present in the calculation window, but are not aligned so that the cancelation can occur. The drop off that occurs past these two peaks is because the spectra are not of infinite span in velocity space and the peaks migrate off the portions of the spectra which are being examined. By considering lines of constant ℓ, one important feature is noticed: the variations in s result in a parabolic variation of the deviation. This type of relationship is capitalized upon in the development of the minimization routine because the minimum, for a given value of ℓ, is easily calculated. By viewing the function as a paraboloid in s, a simple derivative will produce the appropriate minimizing value of s.

$$D(T_1, T_0) = s^2 \int [T_{A,1}^*(v-\ell)]^2 dv - 2s \int T_{A,1}^*(v-\ell) T_{A,0}^*(v) dv + \int [T_{A,0}^*(v)]^2 dv$$

$$\Rightarrow s_{min} = \frac{\int T_{A,1}^*(v-\ell) T_{A,0}^*(v) dv}{\int [T_{A,1}^*(v-\ell)]^2 dv} \tag{2.6}$$

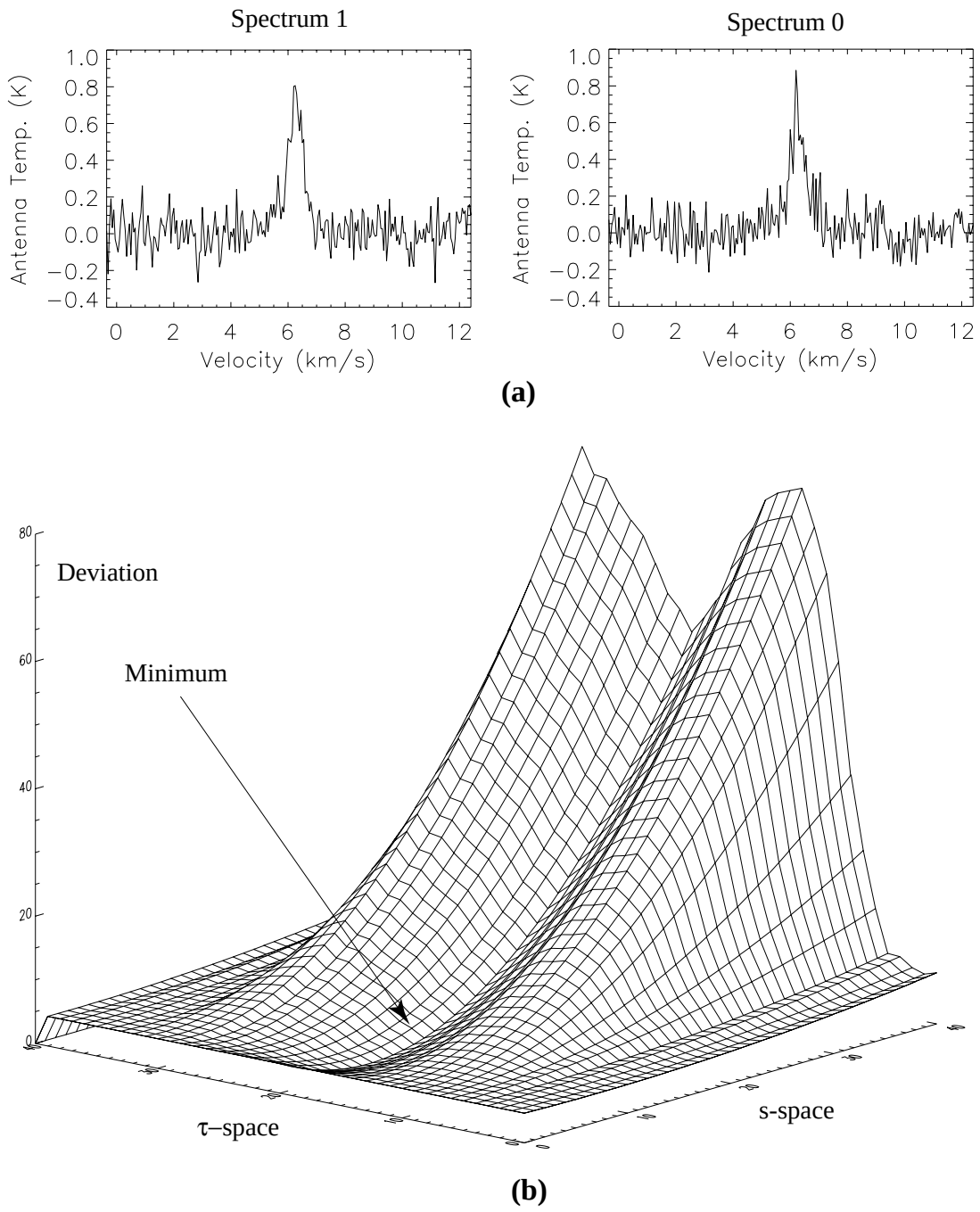


Figure 2.1: $D(T_1, T_0)$ viewed in parameter space. (a) The input spectra. (b) The parameter space.

Because there is only one local minimum, s_{min}, it is possible to minimize the function $D(T_1, T_0)$ with respect to s first and then with respect to ℓ. The presence of a single minimum in s-space is fortunate because it allows for the minimization of the deviation to be done serially. The process is to find the minimum value of the deviation for any given value of

ℓ by adjusting s and then selecting the minimum of all these values. The resulting value is guaranteed to be the minimum of the deviation because it is the minimum of the complete set of local minima. Computing the minimization in the opposite order does not necessarily guarantee that the set of local minima is complete and thus the global minimum might not lie among these. In addition, the fewer times the deviation is calculated, the better; for calculations with changing ℓ are very time-consuming.

2.2.1 Numerical Integration

The evaluation of the deviation, $D(T_1, T_0)$, requires the computation of the integrals in equation 2.3. In addition, minimizing s requires the computation of the same integrals. There are many numerical integration routines available on different computer platforms. However, efficiency is a priority in designing the algorithm and the simplest numerical method is used. The spectral data are sets of paired numbers $(v_i, T^*_{A,i})$ which can be plotted as $T^*_A(v_i)$ to yield the usual spectrum. Most spectral data are in the special case where the velocity abscissae are evenly spaced: $v_{i+1} - v_i = \ constant \ \equiv \delta v$ for all i. In order to integrate the spectrum over the velocity range, the values of the function are summed and the resulting total is multiplied by δv, corresponding to the rectangular estimate of the area under a curve.

$$\int f(v)dv \simeq \left(\sum_i f(v_i)\right) \cdot \delta v \tag{2.7}$$

This approximation is excellent, so long as the sampling of the data, δv, is smaller than the scale over which the function $f(v)$ varies appreciably. For Gaussians, an integration accurate to 1 part in 10^6 requires but 2 samples per half width. Unfortunately, most spectral data are noisy and thus not slowly varying on a scale larger than the sampling. The integral under the curve, therefore, will be significantly different from the total derived using the summation approximation. Each of the spectra involved in the function $f(v)$ has an associated value σ representing the noise in the spectrum. This value is the root-mean-squared value of the noise of the signal where it should ideally be zero and can be used as the error in the antenna temperature values. Thus, the error in the integration routine can be approximated by calculating the errors in quadrature. In all cases, the integrals are of the form $g = \int f(v)dv = \int T_i(v)T_j(v)dv$. Thus, the inherent error due to noise in the calculation is:

$$\partial f(v) = \sqrt{T_j^2(v)\sigma_i^2 + T_i^2(v)\sigma_j^2} \Longrightarrow \tag{2.8}$$

$$\partial g = \int \partial f(v)dv \simeq \left(\sum_i \partial f(v_i)\right) \cdot \delta v \tag{2.9}$$

In the above equation, a "∂" preceding a function indicates that it is the error in that function. By means of example, the error incurred by noise in the spectra is about 1 part in 4 whereas the error incurred by the summation integration is less than 1 part in 25 for a spectrum with signal to noise value of 3.18 and a channel width of 0.05 km/s. Thus, the error due to the approximation can be neglected in light of the error due to noise in the spectrum. The problems with correlation because of noise in the spectra will be addressed in the next section.

2.3 The Effect of Noise on the SCF

The presence of noise in the spectra interferes with the correlations between two spectra in a significant fashion. The reason for this is that noise changes the shape of the two spectra, and while the actual signal may be well correlated, the addition of noise will prevent the spectra from having similar shapes. There is therefore a bias in the SCF values favoring pairs of spectra with a high ratio of signal to noise.

In order to better estimate the effects of this signal to noise bias, a numerical simulation was performed. In this simulation, two perfect Gaussian spectra were compared using the spectral correlation function. The spectra had a FWHM value of 1.7 km/s and a uniform height. To each of these spectra, normally distributed noise was added in a fashion that set the signal to noise at a specific value. This was repeated 100 times to create a large body of spectra with the same signal to noise. 100 pairs of these noisy spectra were processed with the SCF, and the correlation functions were plotted as a function of signal to noise. Because each of these functions should have a value of 1 in the case of no noise, the value that the simulation yields should be the factor by which the correlation function is in error for a given signal to noise value.

In order to estimate the effects of the parameters s and ℓ on the behavior of the data, the spectra were compared using four different correlation functions with various combinations of the parameters. These functions are summarized in Table 2.1.

Function Name	s	ℓ	Spectral Property Examined
S_{ij}	Float	Float	Compares shapes of spectra
S_{ij}^{ℓ}	1	Float	Emphasizes similarity in shape and amplitude
S_{ij}^{s}	Float	0	Highlights similarity in velocity offset and shape
S_{ij}^{0}	1	0	Measures similarity in all properties

Table 2.1: Summary of the Correlation Functions Used in SCF analysis.

The behavior of these four correlation functions is plotted as a function of signal to noise in Figure 2.2. At low signal to noise values, the correlation functions which do not use a calculated value of s tend toward higher values than do those with the value given in Equation 2.6. This aberration is the result of the errors incurred by the numerical integration routines discussed previously. The noise in the spectra artificially reduces s factor in the calculation. This is because the squared spectrum in the denominator of Equation 2.6 is the square of a spectrum, allowing part of the noise to reinforce itself. The product of the noise from two different spectra has no such reinforcement. A more detailed discussion of the arguments behind this reasoning appears in Section 2.3.1.

2.3.1 The Seljak Correction

The signal to noise bias discussed in Section 2.3 must be circumvented in order that meaningful values of the SCF be calculated. In order to correct for these difficulties, Uros Seljak of the Harvard-Smithsonian Center for Astrophysics has proposed the following correction.

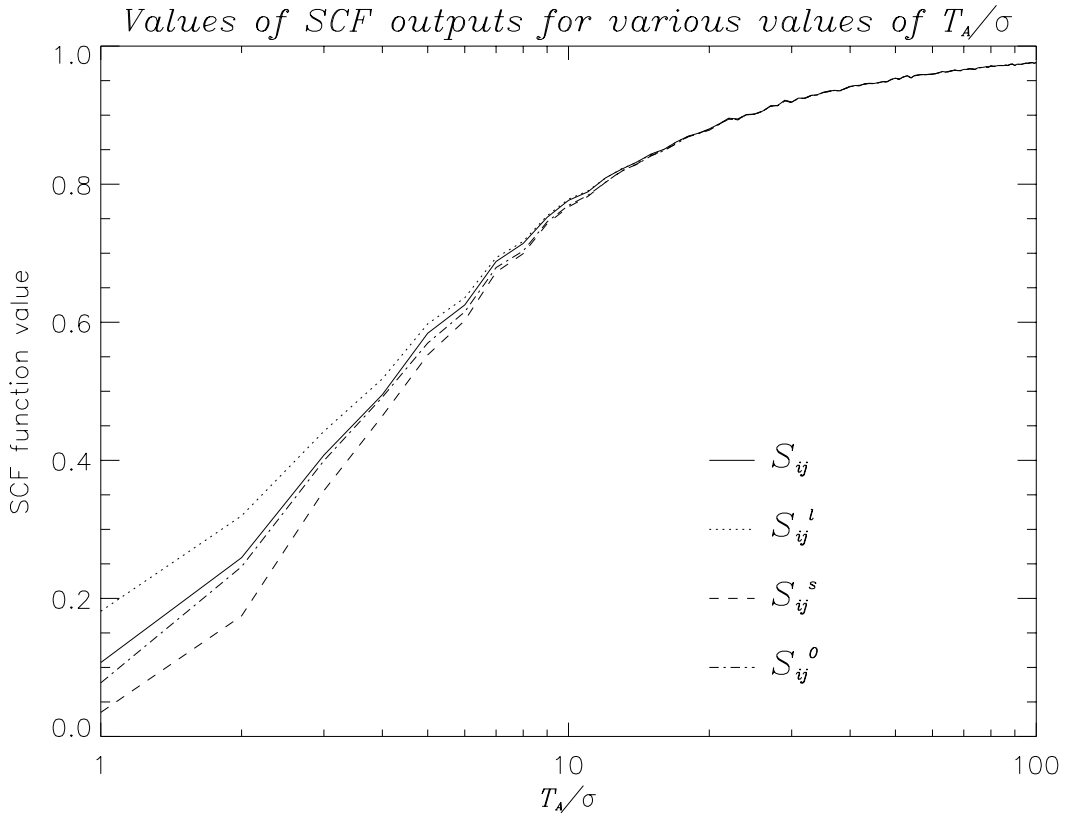


Figure 2.2: The behavior of the four correlation functions under the influence of different values of signal to noise.

Every spectrum can be considered as the sum of two different functions, a signal function and a noise function:

$$T_A(v) = S(v) + \mu(v) \tag{2.10}$$

Hence the deviation function can be expressed in terms of these two individual functions:

$$D(T_1(v), T_0(v)) = \int s^2\ T_1^2(v) - 2s\ T_1(v)T_0(v) + T_0^2(v)dv$$

$$= \int \left(s^2(S_1^2 + 2\mu_1 S_1 + \mu_1^2) - 2s(S_1 S_0 + \mu_1 S_0 + \mu_0 S_1 + \mu_1\mu_0) + S_0^2 + 2S_0\mu_0 + \mu_0^2 \right) dv \tag{2.11}$$

The principal insight in the Seljak correction is recognizing the individual terms of the integral in terms of their average values. For example, we can use the definition of the average value of μ_1^2 to evaluate the terms in the above integral:

$$\left\langle \mu_1^2 \right\rangle \equiv \frac{\int \mu_1^2(v)dv}{N_{pixels} \cdot \delta v_0} \Longrightarrow \int \mu_1^2(v)dv = \left\langle \mu_1^2 \right\rangle N_{pixels} \cdot \delta v_0 \tag{2.12}$$

In the above equation, δv_0 is the spacing between channels in the spectrum and $(N_{pixels} \cdot \delta v_0)$ is the range over which the noise is averaged.

The character of the noise in the spectra is assumed to be normally distributed noise with a mean of zero and a measured root-mean-squared deviation of σ. As a result of these features and the assumption that the noise and signal are statistically uncorrelated, the product $\langle \mu_i S_j \rangle$ or $\langle \mu_1 \mu_0 \rangle$ can be broken up into terms involving $\langle \mu_i \rangle$ multiplied by another term. According to our assumptions about the character of the noise, these terms are zero. Thus, equation 2.11 can hypothetically (see below) be reduced to:

$$D(T_1(v), T_0(v)) = \int s^2(S_1^2 + \mu_1^2) - 2sS_1S_0 + S_0^2 + \mu_0^2 \; dv \tag{2.13}$$

This reduction is accomplished by setting all terms that are linear in μ_i equal to zero using the argument given above. As given in equation 2.12, the values of $\int \mu_i^2 dv$ can also be calculated using the measured root-mean-squared value of the noise. Thus, reducing the calculation to only correlations between signal gives the corrected deviation:

$$D'(T_1(v), T_0(v)) = D(T_1(v), T_0(v)) - s^2 N_{pixels} \delta v_0 \sigma_1^2 - N_{pixels} \delta v_0 \sigma_0^2 \tag{2.14}$$

While this correction holds in the statistical limit, neglecting the terms that are linear in μ_i can be dangerous. Examining the rms deviation of these terms shows that the amount by which they deviate for a given signal, is larger than the correction applied because of the noise alone. Treating the signals as Gaussians the rms deviation of one of these terms is given by:

$$\Delta \left(\int \mu_i S_j dv \right) \simeq \sqrt{N_{pixels} \delta v_0 \sigma_i^2 A_j^2 \eta_j \sqrt{\frac{\pi}{2}}} \tag{2.15}$$

Here, η_j is the width of the Gaussian. The magnitude of these deviations are comparable to the calculated values for $\langle \mu_i^2 \rangle$. An assumption that these are identically zero is not necessarily a good one. Moreover, when the SCF algorithm is processed with the correction of equation 2.14 in mind, the results imply that the contribution given by terms like those in equation 2.15 are not normally distributed about zero. The precise cause of such effects are unknown, but it seems to imply that the signal and noise are not as separable as they might seem.

2.3.2 Signal Degradation

With the ingenuity of the Seljak correction frustrated by unknown complications, the bias for high signal to noise data must be corrected in some other fashion. The resulting correction is far from standard in the study of spectral maps. Essentially, if the spectra cannot have all bias due to signal to noise eliminated, the next best thing is to insure that all spectra have the same amount of bias. This is accomplished by reducing the signal to noise ratio for a spectrum to be a given level. This reduction is accomplished by adding noise to the spectra in a map to reduce them all to a given threshold level. Spectra with signal to noise ratios lower than the threshold value are rejected from the SCF analysis.

In order to gauge whether the results from such a degradation are meaningful, a simulation was performed. Instead of using artificial data, a set of observational data was used.

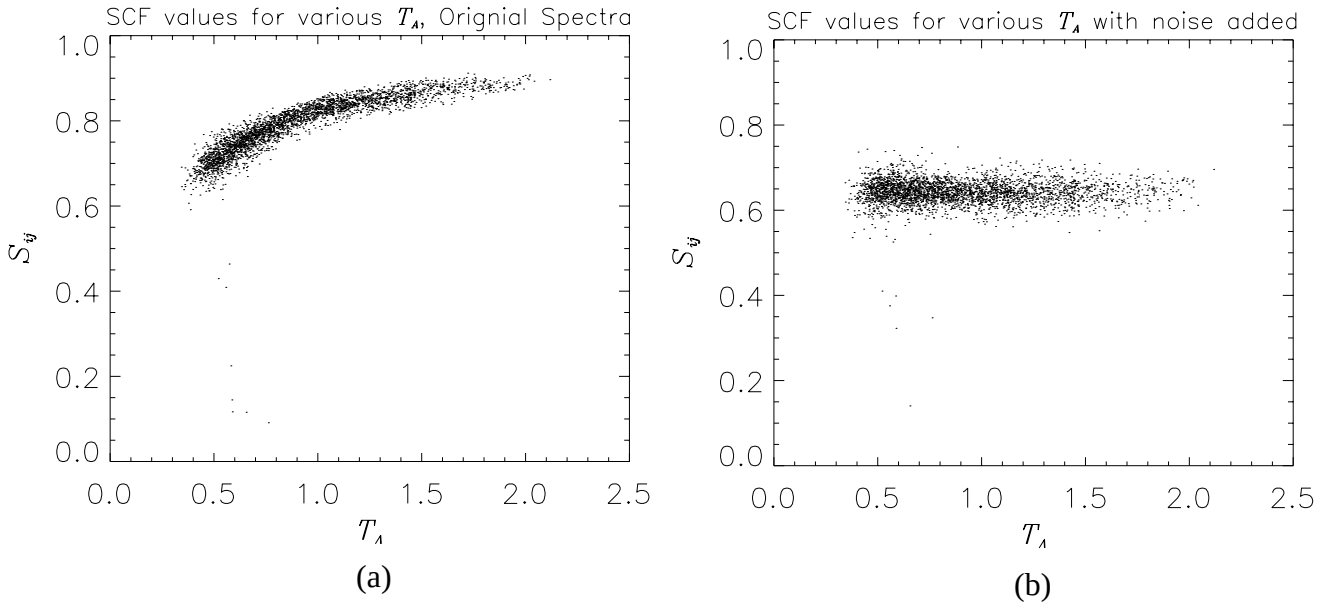


Figure 2.3: SCF as a function of antenna temperature for (a) the original data and (b) noise added to the spectra to create a uniform T/σ ratio of 5.

The degradation was performed 10 times on the same map (different sets of noise added) and the correlation function values were evaluated between the spectra. The threshold value for the signal to noise ratio was 5. The desired result was that the deviation for these values of the correlation functions would be close to zero indicating that similar results were arrived at using different sets of random noise. The results support this hope and lend credence to the method. Statistical moments of the resulting data are shown in Table 2.2. These results indicate that there is approximately a 0.2% error in using this method to eliminate the signal to noise bias.

Function	Mean	Deviation	Skewness	Kurtosis
S_{ij}	0.56	0.0021	-0.41	-0.84
S_{ij}^{ℓ}	0.57	0.0020	-0.25	-1.1
S_{ij}^{s}	0.53	0.0030	-0.46	-0.59
S_{ij}^{0}	0.54	0.0027	-0.33	-0.83

Table 2.2: Statistical moments for SCF outputs with varying noise but same T/σ

The only remaining question to ask is whether such a treatment does eliminate the bias that is observed towards correlation in higher signal to noise regions of a map. In order to explore this question, a data cube was processed with the SCF algorithm and the values of the SCF were calculated for the original data and with the noise added. Then, the results were plotted against their value of antenna temperature to determine whether any bias remained. The results appear in Figure 2.3.

As Figure 2.3 indicates, the process of adding signal to noise evens out an bias toward higher values of the antenna temperature. The mean values of the SCF are lowered a bit by this process, for some amount of correlation is lost in the process. The vertical spread in the points is not drastically altered in this process and thus the signal degradation performs the desired compensation.

Chapter 3

Making Maps with the SCF

In Chapter 2, the background for establishing a single function defining the correlation between two spectra was developed. The entire purpose of the SCF, however, is to analyze whole maps of spectra and thus the functions must be generalized in an appropriate way to create correlation maps of the data sets. In the following discussion, a method for doing this is established and then some examples of the correlation maps are explored in order to glean understanding about our results.

The SCF algorithm was written in the IDL software package to take advantage of the built-in functions, the multi-dimensional array processing, and graphics utilities. It would be easy to port the code to another language for speed in processing though the output would have to be analyzed using some graphics package, like PGPLOT or MONGO.

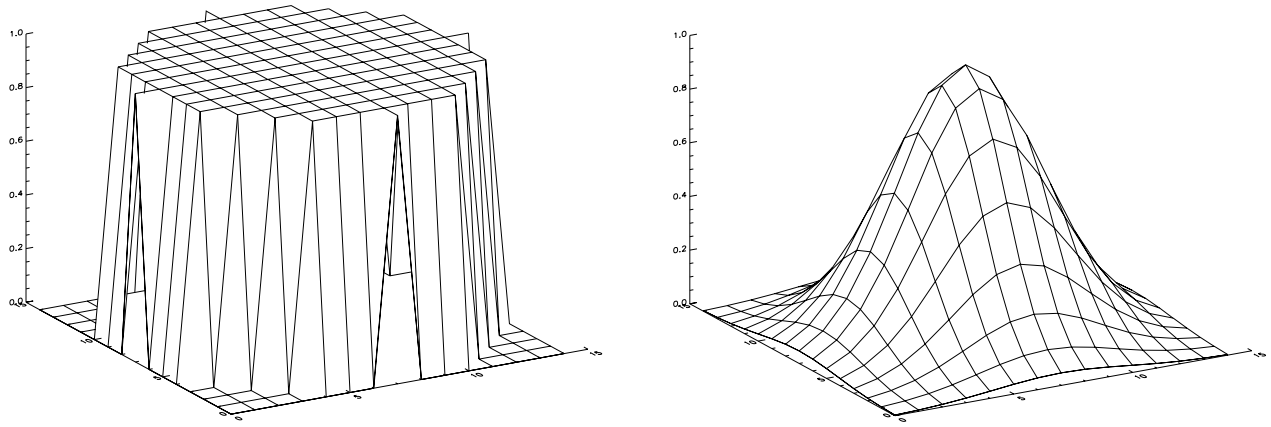
3.1 Creating Maps of the SCF

In order to generate the appropriate maps, a specific algorithm is developed.

1. Begin by considering a data cube with size $X \times Y$ and spectra at each (x_i, y_j) position in the cube. For purposes of the calculation, it is assumed that the spectra are taken on an evenly spaced grid. In the discussion to follow, (x_i, y_j) represent the real right ascension and declination coordinates of the cube and the coordinates (i, j) are the pixel coordinates in the data cube.

2. For each of these spectra, a Gaussian is fitted to the data, allowing for the estimation of certain line parameters. The parameters derived from the fit are peak antenna temperature (T_A^*), the width of the Gaussian in velocity space (Δv), the center of the Gaussian (v_{LSR}), the integrated area under the Gaussian (T_{int}) (generated with the summation integration discussed previously), and the rms noise of the spectrum (σ_{rms}). Note that most of these parameters are used only for the purpose of calculating the noise in the spectrum. The sole exception is that the width (Δv) is used to determine the relevant velocity ranges around a spectrum's peak that the SCF will consider.

3. Spectra with a signal to noise ratio below a selected threshold are rejected and all correlation calculations are performed without these rejected data.



(a) (b)

Figure 3.1: Two commonly used weight functions. (a) A flat, circular mask, used to eliminate any bias for the square directions in which the cube was sampled. (b) A Gaussian mask, also circularly symmetric, but also weights spectra closer to the base spectrum.

4. A test spectrum at (x_i, y_j) is selected to measure its correlation with nearby spectra.

5. The SCF requires two input parameters denoted by the variables r and q. r is the resolution of the SCF and represents the number of spectra for which the correlation with the test spectrum is measured. A resolution of r means to consider an $r \times r$ box centered on the test spectrum and to compare the test spectrum with every spectrum in the box. If desired, the results of the correlation calculations can be weighted by the distance away from the test spectrum. These weight functions emphasize regions in the data cube to eliminate certain biases. Examples of weight functions appear in Figure 3.1.

   The parameter q is a measure of what velocity range is compared with the other spectra in the $r \times r$ box. All calculations discussed in the following section are to be performed only on the velocity range

   $$v_{LSR}(i,j) - q\Delta v(i,j) \leq v \leq v_{LSR}(i,j) + q\Delta v(i,j) \tag{3.1}$$

   The Gaussian fit parameters in equation 3.1 are taken from the fit to the test spectrum at (x_i, y_j). All the integrations performed in the calculation of the deviation function and the SCF itself are taken over this velocity range as opposed to the full range of velocities within the spectral bandpass.

6. A comparison spectrum at (x_a, y_b) is selected with the provisions that $a \in [i - (r - 1)/2, i + (r - 1)/2]$, $b \in [j - (r - 1)/2, j + (r - 1)/2], a \neq i$, and $b \neq j$.

7. Letting the spectrum at (x_a, y_b) equal $T_1(v)$ and the spectrum at (x_i, y_j) equal $T_0(v)$, the deviation function, $D(T_1, T_0)$, is minimized by adjusting s and ℓ. The deviation function is then normalized according to prescription in equation 2.5.

8. The above step is repeated for each spectrum in the $r \times r$ box. The resulting values for $S(T_{a,b}, T_{i,j}), s_{a,b,i,j}$, and $\ell_{a,b,i,j}$ are averaged over all non-rejected values of a and b in the box using the selected weighting function. The resulting means are assigned to the position at (x_i, y_j) as $S_{ij}, s_{ij}, \ell_{ij}$.

9. The deviation function is then minimized between the test spectrum and all comparison spectra holding the value of s at 1. This represents the correlation between the areas of the spectra because differences in area can not be compensated by the scaling parameter s. The averaged values of the SCF and the lag, ℓ, over all applicable a and b are referred to as S^l_{ij} and ℓ^l_{ij} respectively.

10. The deviation function is then minimized holding $\ell = 0$, in an attempt to explore velocity offsets alone. The values of the SCF and $s_{a,b,i,j}$ in this case are referred to as S^s_{ij} and s^s_{ij} respectively.

11. The last correlation calculation to be performed over the box centered at (x_i, y_j) is to calculate the SCF with $\ell = 0$ and $s = 1$, the straight squared difference between each pair of spectra. The resulting averaged value is referred to as S^0_{ij}.

12. Finally, the above correlation calculations are performed for each spectrum in the cube serving as the base spectrum. This procedure yields values of the parameters $T^*_A, \Delta v, v_{LSR}, T_{int}, \sigma_{rms}, S, \ell, s, S^l, \ell^l, S^s, s^s$, and S^0 for every point (x_i, y_j). Maps of each of these parameters can be generated and in their analysis, the results of the SCF can be analyzed.

3.2 Interpreting Maps of the SCF outputs

The bulk of this section is to provide some insights into how the maps generated by the SCF can be interpreted. The results are different from any other functional output generated by other analysis routines; and in applying the analysis to simple situations, more insight can be gleaned about actual data. The understanding from these maps is preliminary and is used as a semi-qualitative descriptor of the behavior of the data.

To develop well-understood situations, the algorithm was applied to small cubes of completely artificially generated spectra. It is important to note that these do not represent the data from MHD simulations in any way. Rather, they are simple constructions of data cubes which are intended to display the behavior of the SCF with only one aspect of the data cube changing in a regulated fashion.

The data cubes generated had 484 spectra arranged in a 22×22 array. These spectra were Gaussians and had various amplitudes, widths, and/or velocity offsets. Several different configurations were analyzed including random parameters, sudden jumps, linear gradients, and constant variations in the parameters. A few general statements about the outputs from the SCF can be made.

The correlation function S represents the similarity in shape between the base spectrum and the surrounding spectral lines. In a test where the spectra were all Gaussians with

the same width, the value of S was 1 for the entire map, even for randomly distributed offsets and amplitudes. Since these all have the same shape, the correlation function accurately demonstrated the similarities of these lines.

The lag parameters ℓ and ℓ^l both represent the velocity by which the neighboring spectra are offset from the test spectra. For linearly changing values of v_{LSR}, this number is zero. This result is due to the fact that for any given test spectrum, the comparison spectra with higher velocities are balanced by those with lower velocities, so the mean shift is zero. Any deviations from this can be attributed to the discrete spacing of the grid points and the inability for the spectra to be shifted in order that a perfect match be obtained.

The lags are interesting because for smooth Gaussians, a velocity gradient of order m can be shown to have a corresponding lag field of order $m-2$, implying a relationship similar to that of taking the second derivative. Ordinarily, recognizing the relationship would be a boon to calculations, but the relationship only holds true when the peak velocity is the only thing differing between two spectra. When their shapes and amplitudes differ, the correlation might be maximized with a lag that is not the one predicted by the second derivative of the velocity field.

The scaling factors s and s^s represent the similarities between the antenna temperatures of the spectra. When the lag is turned on, the value s represents the similarity between the integrated areas of the spectra, with a value of 1 indicating equality. When the lag is off, the value, s^s, represents the mean, for all velocities, of the ratio between the antenna temperatures from the two spectra at given velocities. A value greater than one indicates that the base spectrum is consistently higher than comparison spectrum and a value of less than one indicates the opposite case. It is important to remember that the value of s is adjusted so that the differences in antenna temperatures are minimized over the whole spectrum, and the result does not correspond to the ratio of amplitudes at a given velocity.

The remaining correlation functions contain important information in what is held constant rather than what is allowed to vary. This subtle point stems from the realization that in allowing a parameter to vary any differences due to that parameter are eliminated by the SCF analysis. For example, S^l, the correlation with adjustable lag but not scaling of the spectra is helpful for comparing the similarities of line profiles. A high value will indicate that the spectra are similarly shaped and scaled though they may not necessarily be from regions with the same bulk velocity. On the other hand, the correlation function with zero lag but adjustable scaling, S^s, measures the similarity in velocity distribution at two positions without regard to the amplitude changes.

Finally the correlation function with the lag off, S^0, indicates simply the similarity of the two spectra in size, shape and velocity distribution. Strong correlations in this function are tied to features that are large compared to the resolution of the SCF; thus changing the resolution will highlight different structures. This result is because a large value of S^0 requires all the spectra in the resolution box to be very similar in all aspects.

In this section of the paper, several examples of artificially generated data are interpreted. The relevant maps are displayed and others are summarized in the text. The best approach to interpreting these data may be simply to leap in and look at some examples. In all these examples, the weight function is a uniform $r \times r$ box.

3.2.1 Velocity Jump

The first example contains the results of the SCF algorithm when the spectra in two regions of a map differ from each other by a sudden jump in velocity. In this case, the jump in v_{LSR} is from 1 km/s for low values of y to -1 km/s for high values of y. Physically, this jump corresponds to the spectra in the bottom half of the graph having a red-shift and the spectra in the top half of the graph being blue-shifted. These are the spectral maps of two bodies of gas: one coming towards the observer and one retreating from the observer. The results are shown in Figure 3.2.

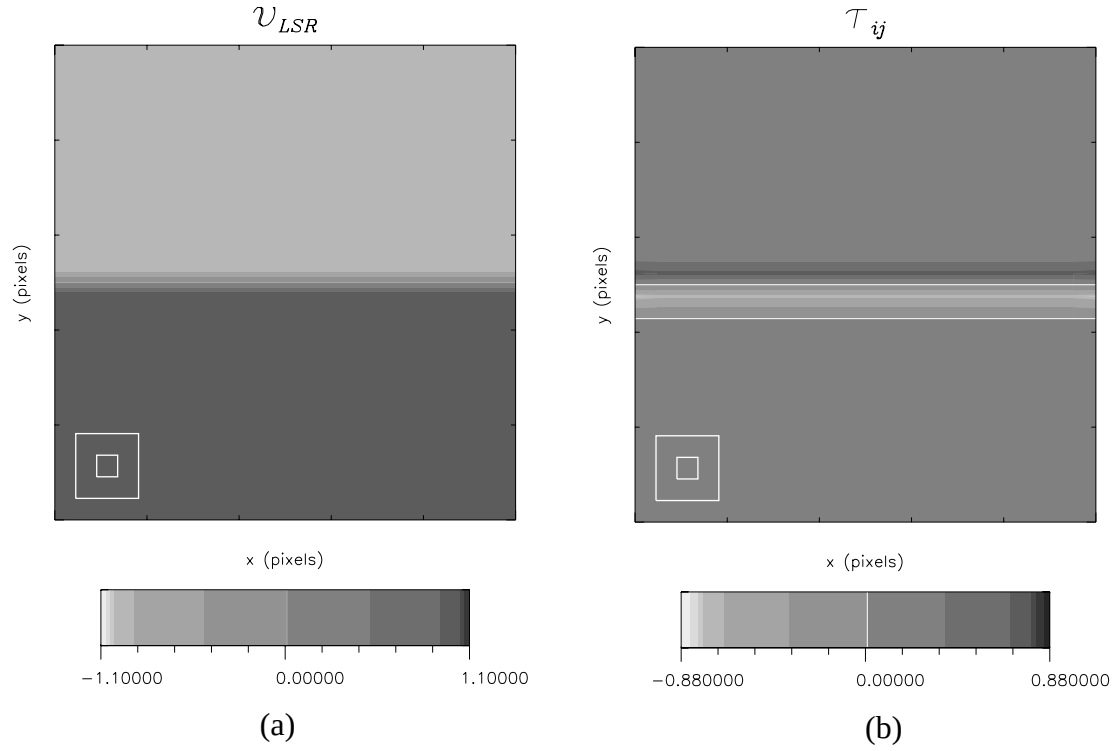


Figure 3.2: Correlation function outputs for a velocity jump. (a) The input velocity field. (b) The resulting field of lags

The figure on the left is a grey-scale representation of the fits to the v_{LSR} of the generated spectra. It depicts the sudden jump in v_{LSR} halfway up the data cube. The corresponding representation of the lag is displayed on the right. In all maps, the color bar extends for 10% beyond the actual values, which is why the extrema of the bar are at -1.1 and 1.1 respectively. The greyscale is selected to display a dynamic range of the data and color bars are provided for aid in interpreting the results.

As expected, the map displays a jump when it crosses the boundary between red-shifted spectra and the blue-shifted spectra. The drop in the value of ℓ before the jump is because the spectra on the border that are red-shifted are being correlated with spectra in

the blue-shifted region that have a negative v_{LSR} relative to the spectrum with which they are being correlated. There is a corresponding jump in the blue-shifted section where the spectra are correlated with spectra that have a greater value of v_{LSR}.

In all cases, the plots have a pair of nested boxes in the lower left-hand corner. These boxes represent the size of the pixel (inner box) and the size of the SCF $r \times r$ box where the correlations are conducted (outer box). In addition to these visual aids, an extra contour has been added at the zero level so that the examiner can distinguish between those regions with positive ℓ or v_{LSR} and those with negative values.

All other plots of the map yield what is expected. The SCF function is equal to 1 over the entire plot in cases when both scaling and lag (S_{ij}) and just lag (S^l_{ij}) are turned on. In the case where there is no lag turned on (S^s_{ij} and S^0_{ij}), the maps appear as shown in Figure 3.3.

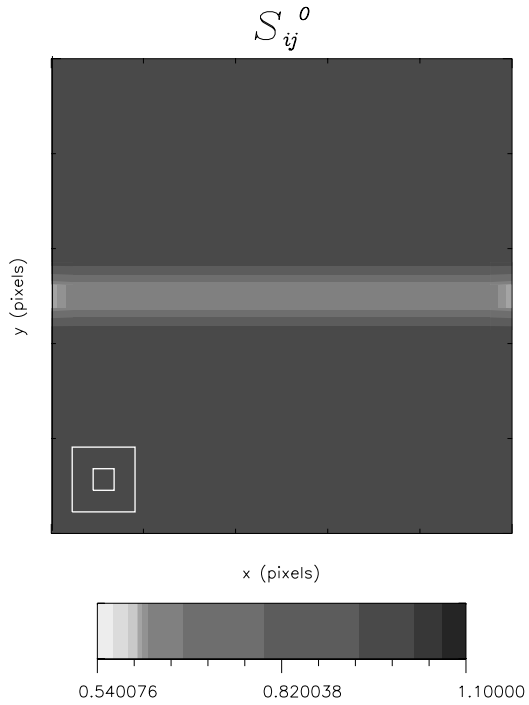


Figure 3.3: Map of S^0 in the case of a jump in v_{LSR}

The band in the center adequately represents the fact that the spectra in this region are not well correlated with their neighbors. The small boxes on the edge of even poorer correlation are the result of an edge effect. This effect results from the fact that pixels in the center of the map have eight spectra averaged together to yield a result while pixels on the edge only have five. It is because of effects like these that the r pixels on the border of the map are discarded in the actual analysis.

Because the artificial spectra were all of the same height, the scaling factor, s is uniformly 1. The parameters from the fitting indicate that the routine fits the Gaussians

well.

3.2.2 Amplitude Jump

The amplitude jump is almost exactly the same as the velocity jump, except that the jump is in the height of the Gaussian in the model spectra. The jump is from high amplitude at low y to low amplitude at high y, representing two different regions of gas, one with a higher peak antenna temperature than the other (See Figure 3.4).

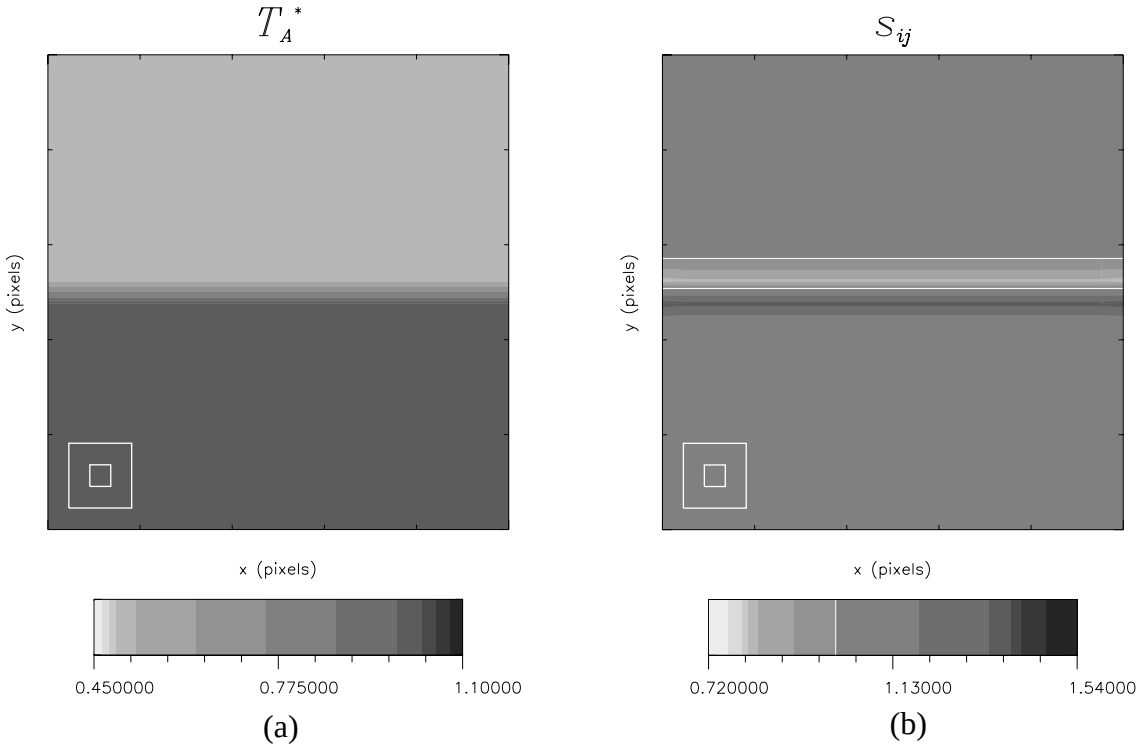


Figure 3.4: Maps of artificial data with a jump in T_A^*. (a) The amplitudes of the spectra. (b) The scaling factor s.

Again, there is an increase before and a dip after the change in the amplitude. The rise before the change may be a bit counter-intuitive; however, the way that s is defined is the amount by which the neighboring spectra are scaled up. Thus, the low spectra being scaled up to have the same form as the high spectra is represented by the jump and the high spectra being scaled down is represented by the dip. A border contour is plotted here as well, this time dividing regions with $s < 1$ from those with $s > 1$. Those with $s < 1$ indicate that the neighboring spectra had to be scaled down, so the spectrum here has, on average, a smaller integrated area than its neighbors.

The other plots generated by the SCF algorithm are similar to those generated in the case of the velocity jump. The SCF plots without scaling turned on are identical to those in the velocity jump without lag turned on (see Figure 3.3 and comments pertaining thereto).

Similarly, those plots of the SCF with scaling turned on are uniformly equal to one (with an error of 1 part in 10^7 due to fluctuations in the floating point calculations used on the computer).

3.2.3 Width Jump

While differences in lag and scale can be compensated for by the parameters ℓ and s respectively, differences in shape can not be so corrected; thus the SCF can be used to measure the differences in shapes between lines. The result is that the SCF defines correlated spectra to be those with similar shape. A change in width is tantamount to a change in shape.

In this set of data, the jump is from large widths for low y to small widths for large y. The relevant plots for the case of changing widths are found in Figure 3.5.

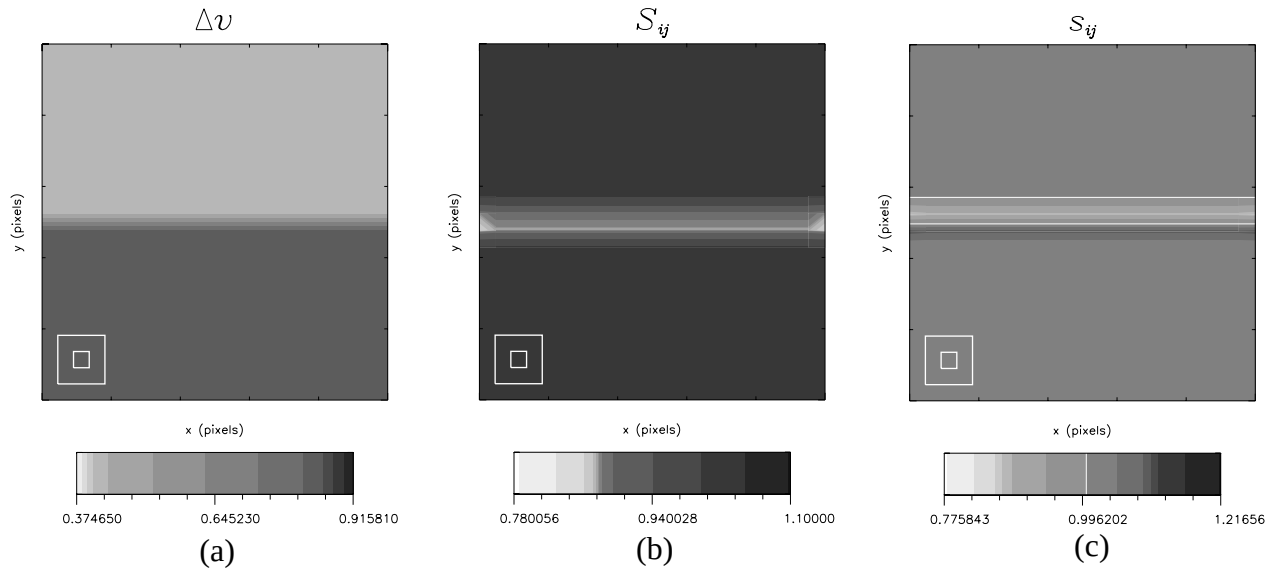


Figure 3.5: Relevant plots for a jump in Gaussian width. (a) The widths of the spectra. (b) The SCF values for the map. (c) The parameter s compensating for the differences in width.

The figure on the left represents the fit to the Gaussians, depicting the jump from large width to small width as y increases. The second plot, representing the SCF, shows how the width changing affects the data because the shapes are different along the interface between wide and narrow lines. There is less correlation between the spectra at that point. Finally, the last plot represents the value of s changing, which stands to reason because the SCF algorithm uses a changing scale factor to compensate for the differing widths.

3.2.4 Resolution

One of the input parameters that can be changed is the resolution, r, in order to regulate the size of the box over which the SCF parameters are evaluated. Again, the case of a changing v_{LSR} is used, but in this case there are two small clumps of gas with strong red-shifts against

background at rest relative to the observer. Like resolution in optics, the higher the value of r that is used, the more difficult it is to distinguish between two distinct features and to deduce their physical extent. The plots in Figure 3.6 illustrate this effect.

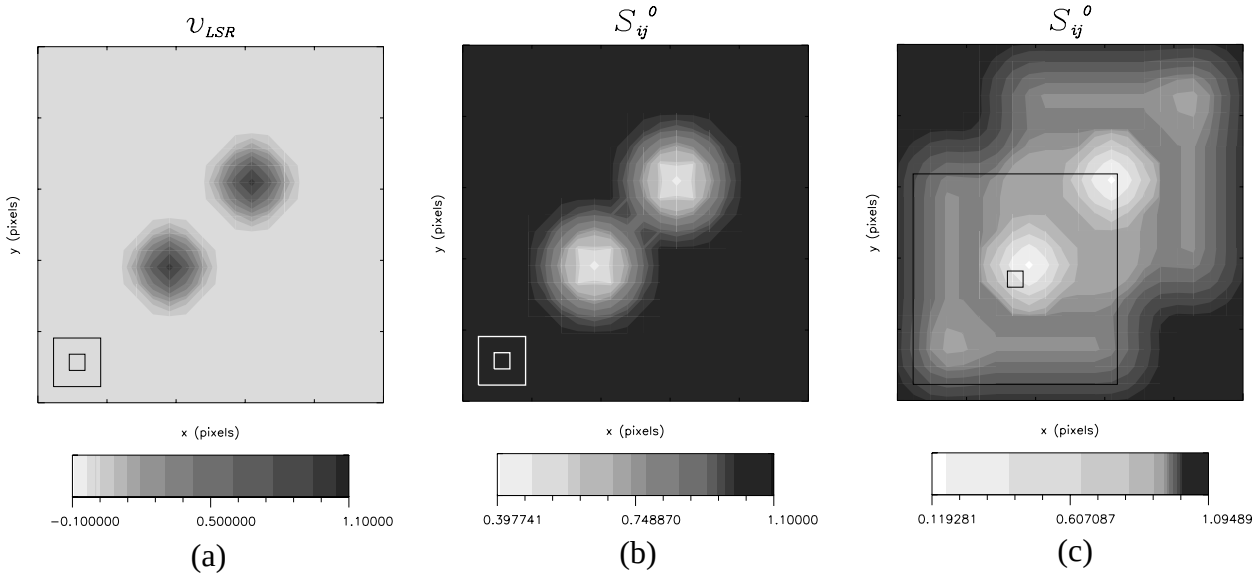


Figure 3.6: Relevant plots to display changing resolution. (a) The two-peaked velocity field. (b) The correlation function S_{ij}^0 at resolution 3. (c) The correlation function at resolution 13.

For the figure with r=13, it is impossible to determine whether the gas distribution has a large region of spectra that are not perfectly correlated or a smaller region that is pulling the averages down. For this reason, the smaller resolution can be used to pick out these details.

Sometimes it can be advantageous to use a larger value of r on purpose. The correlations between gas clumps at small scales can be detected with small r values, but these features tend to mask the larger scale correlations present in a data set. To illustrate this, a data set has been generated with small-scale and large-scale structure. The relevant plots appear in Figure 3.7.

The velocity field is one with a small-scale sinusoidal modulation, on order of a pixel size, and a large jump at the center from blue-shifted velocity to red-shifted velocity. The jump is not readily distinguishable in the correlation measure at low resolution. There is a feature there; however, it cannot be discerned from the smaller scale correlations. On the other hand, if the resolution is increased, the large-scale correlation of the gas on either side of the jump becomes more apparent. The increased resolution serves to highlight this feature. The jump will become more dominant as the resolution increases, washing out the small-scale modulation in the infinite limit.

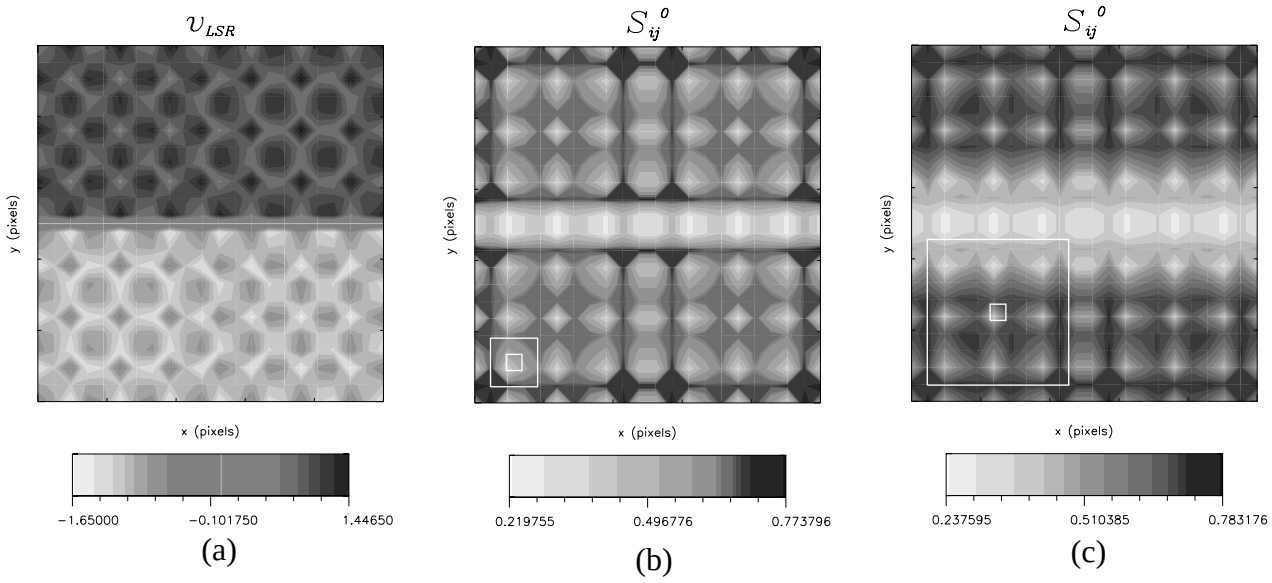


Figure 3.7: Plots displaying the ability for different resolutions to select for features of different size scales. (a) The velocity field used (b) S^0 with $r = 3$. (c) S^0 with $r = 9$.

3.2.5 Bringing it all together

The above differences are rather easy to understand individually; however, when all three basic effects are present in the same cloud, the interpretation is more difficult. In order to better understand such amalgamations of effects, the compensating factors s and ℓ can be turned on and off at will.

First, the advantages of normalization immediately become apparent. We can examine a clump in a cloud and compare its values for the various flavors of the SCF. With both scaling and lag turned on, the value of S represents the best correlation one can get between two spectra. This value serves as a basis for comparison. Then, we can turn off lag ($\ell \equiv 0$) and scaling ($s \equiv 1$) respectively and compare the resulting S_{ij}^s and S_{ij}^l. If one of these is significantly higher than the other, we can assert that the gas has similar velocity distributions ($S^s > S^l$) or that it is emitting similar amounts of radiation, but not moving coherently ($S^l > S^s$). If the SCF with both lag and scaling turned off is comparable to those values with it on, we can assert that the gas within one resolution box of the central spectrum has either a roughly uniform or a completely chaotic nature, corresponding to high and love values of the SCF respectively.

The next fashion in which the results of the SCF can be compared is along the lines of the actual value of the SCF. The function, after all, is a measure of the correlation between spectra and with the normalization factor included, different parts of the map can be compared.

Additionally, maps can be compared with each other. Regions with correlation close to one are similar in shape to the spectra within the resolution box. Similarly, the calculated values of the other SCF variations, S^s and S^l, indicate similarities in intensity and velocity

distribution over the same areas.

In addition to mapping the SCF and its subsidiary variations, the parameters s and ℓ that maximize the correlation for each point on the map are also calculated. These maps can be generated to give further information about the clouds.

The easiest aspect of these maps to consider is the magnitude of the lag parameter ℓ. The parameter takes on large magnitude values when calculated along the border between two differently moving regions of gas. It is similar to the velocity gradient of the cloud, but by altering the resolution of the SCF, the structure of a velocity shift can be interpreted. For example, the resolution at which the jump appears the most prominent is the size scale of the physical jump.

With these realizations about the SCF and its behavior, the analysis of 'real' data can begin.

Chapter 4

First Results of SCF Analysis

The data analyzed by the SCF thus far can be divided into the two categories of "observation and "simulation." The goal of the analysis is to discover any distinguishing characteristics between these two classes of data and to explain the differences among the individual members of these classes.

In these analyses, the same series of SCF routines was applied to four different data cubes, two observational and two theoretical. For these cubes, the data were all degraded to a signal to noise ratio of 5 (and spectra with a lower value were rejected). All data within 3 FWHMs of the peak in a given spectra were considered. From here, the SCF processed each cube using a weight function that gave full weight to all the data within two pixels of the base spectrum and ignored all spectra outside of this range. The correlation functions were all plotted; however, a more quantitative comparison was needed. It is for this reason that the use of statistical moments was implemented. This approach is the simplest method by which the data can be analyzed statistically. Once the SCF data are better understood, more sophisticated analyses can be applied. The first four moments of the SCF histograms were examined.

In order to examine the correlations more precisely, the spectral maps had their constituent spectra randomized to new positions. The purpose behind this shuffle was to investigate how necessary the original positions were to result in the derived value of the correlation. In the cubes, the data should exhibit a sudden drop in correlation because their neighbors are no longer the similar spectra with which the base spectra were well correlated. Hence, the change in the moments should also indicate the degree of the original correlation.

The results of the analyses indicate how well the SCF works and suggest some insights into the nature of the data cubes that is different from previous analytical routines.

4.1 The Data Sets

Four data sets were analyzed from various sources. The salient features of each cube are discussed below.

- **Heiles Cloud 2** - Observers at FCRAO observed Heiles Cloud 2 in 1996 in the $C^{18}O$ line at 109.782168 GHz. The resulting data cube consisted of 4800 spectra arranged in a grid of 50×96 pixels on the sky. The grid covered $58'$ in right ascension and

$40'$ in declination, centered on the cloud. The cloud is 140 pc distant and is located at $\alpha(2000) = 4^h30^m$ and $\delta(1950) = 25.8°$. These data correspond to a physical area of 2.3×1.6 pc and a spatial resolution of 0.02 pc. The spectra themselves have 256 channels of velocity running from -0.35 km/s to 12.45 km/s, with 0.05 $km/s/channel$. The peak emission from the cloud is at about $+6$ km/s.

- **L1512** - The data from this cloud is the among the first released data sets of the IRAM key-project (K003-92). The field is observed in the rotational transition of $^{12}\mathrm{CO}(2-1)$ at 230.538001 GHz. The dimensions of the cube were 40×80 pixels, spanning a region of $5' \times 10'$ on the sky. The distance to L1512 is 350 pc and the cloud is located at $\alpha(1950) = 5^h02^m$ and $\delta(1950) = 33.0°$. The physical size of the region, therefore, is 0.5×1.0 pc and the spatial resolution of the map will be 0.12 pc. Each spectra consisted of 864 velocity channels with a spacing of 0.025 $km/s/channel$, running from 7.1 km/s to 9.1 km/s. The peak of the emission is at about 8 km/s.

- **Theoretical Cube I** - This cube was generated by the preliminary MHD simulations of C. Gammie, E. Ostriker and J. Stone. The cube represents an under-resolved simulation and the data are entirely preliminary. The creators of this simulation have asked that their data be interpreted in a cautious and reasonable fashion.

  The cube itself consists of a 32×32 grid of spectra. Each of these spectra have 256 channels and a velocity channel width of 0.054 km/s. The particular transition simulated is the $^{13}\mathrm{CO}$ line at 110.20137 GHz. The angular size of the cube is irrelevant to these simulations, and the cube in question represents a spatial area of 2 pc square, implying the spatial resolution of the map is 0.06 pc.

- **Theoretical Cube II** - This cube was developed by Falgarone *et al.* from a hydrodynamic simulation and this simulation was evaluated specifically by comparing the statistical moments of the line parameters with those from a dense molecular cloud. The full details of this comparison have been published (Falgarone et al., 1994) and their results indicate that the simulated spectra and the point to point variations of the spectra are similar to those of molecular clouds.

  The spectra in this cube are laid out in a grid of 16×16 pixels. There are 512 channels in each spectrum and the channel width is 0.13 km/s. These velocities were derived from a spectra line with a rest frequency of 230 GHz. The physical size of the simulation is not given but the nature of the comparison suggests that the spatial resolution is should be comparable to that of the other cubes considered. The spectra were generated by assuming the gas to be optically thin allowing for each parcel of gas to emit at a given wavelength, altered only by the Doppler shift. By summing these contributions along a line of sight, the resulting spectra can be calculated.

4.2 Basic Results

4.2.1 Heiles Cloud 2

The Heiles Cloud 2 data set was the first data set that the SCF project attempted to reduce and the resulting data have evolved much over the course of the SCF's development. The primary results are summarized in Figures 4.1 and 4.2.

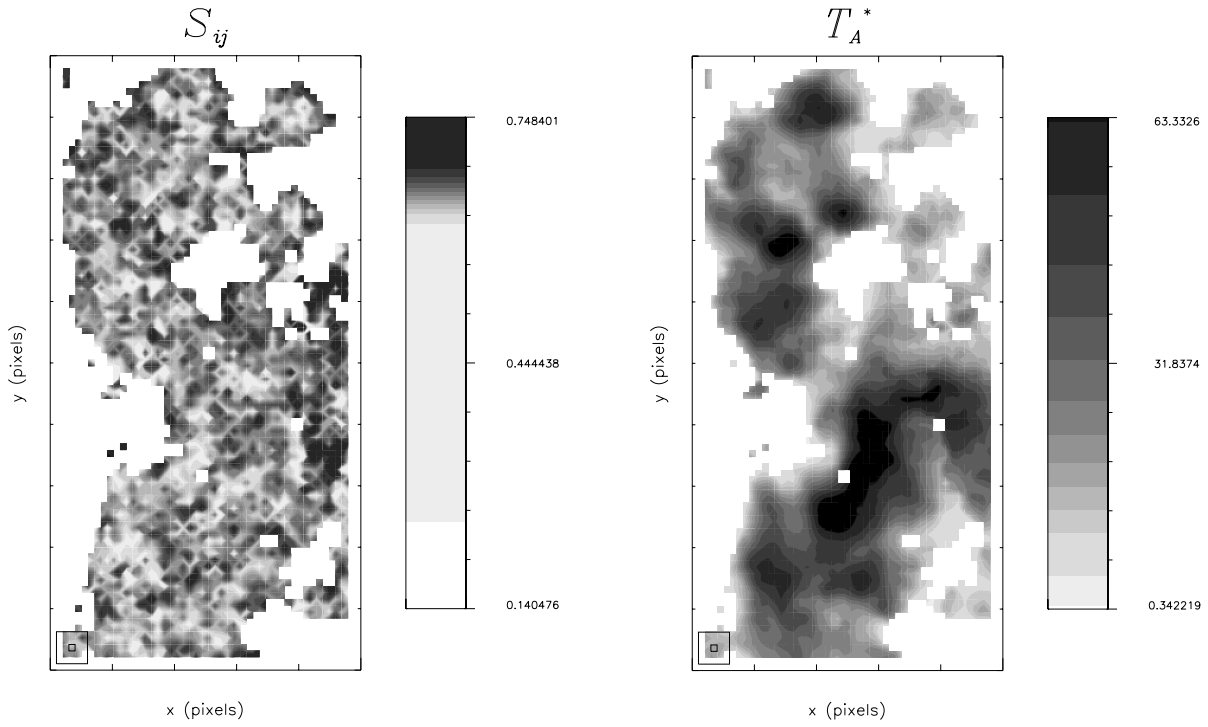


Figure 4.1: SCF results for the processing of Heiles Cloud 2. (a) S_{ij} map displaying regions of correlated spectral shapes. (b) T_A^* map showing peak antenna temperatures as a function of position.

In these diagrams, blank pixels are the locations of rejected spectra and the 2 pixel wide border around the map is rejected because of the edge effects discussed in Section 3.2.1. The maps are in a grey-scale designed to display a dynamic range of the data. As a result, different colors correspond to different values in each plot. In order to aid in interpreting these maps, color bars are placed with each map. As mentioned previously, the scales run from 10% above the highest value to 10% below the lowest value.

In examining Figures 4.1 and 4.2 a few things should be highlighted. First, the granular nature of the S_{ij} map is a result of the discrete nature of the maps. The effect could be eliminated by a larger resolution box; however, a larger box would sample too large a region of the cloud to be physically meaningful. Despite the granular nature of the data, one can see that the small knots of correlated spectra are not related in any way to a single

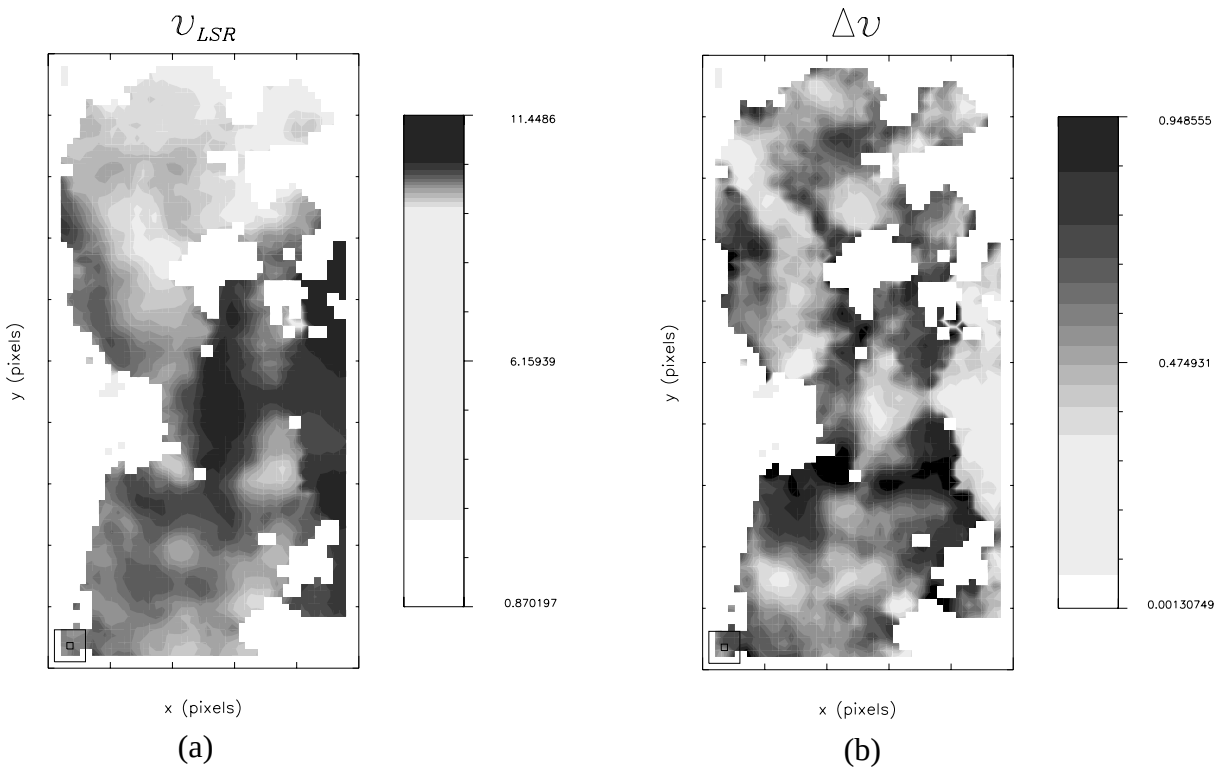


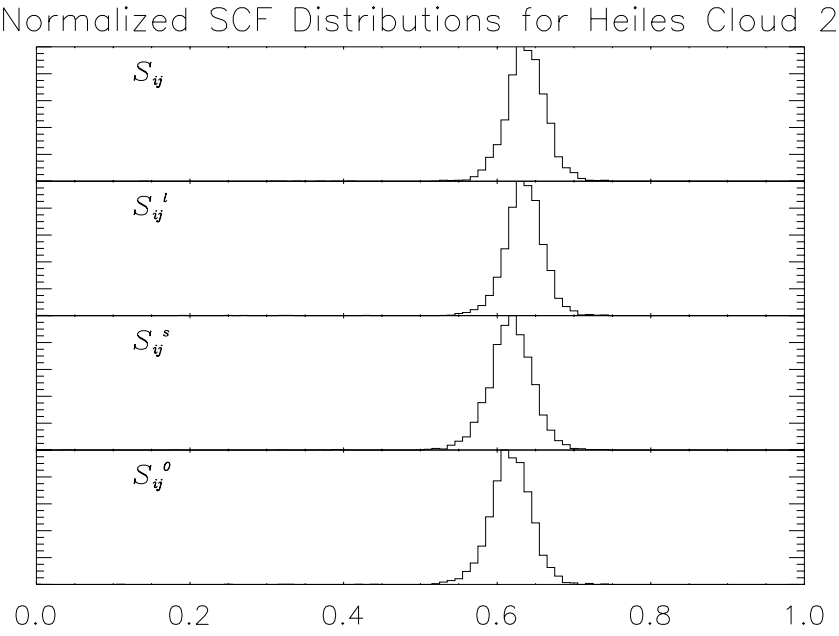
Figure 4.2: More SCF results for Heiles Cloud 2. (a) Map of the velocity offsets v_{LSR}. (b) Map of the spectral widths Δv.

map of line parameters. The lack of similarity to the T_A^* map indicates that a bias toward strong signal to noise has been reduced if not eliminated through the signal degradation routine. This statement is further supported by the analysis performed in Section 2.3.2 with this data set.

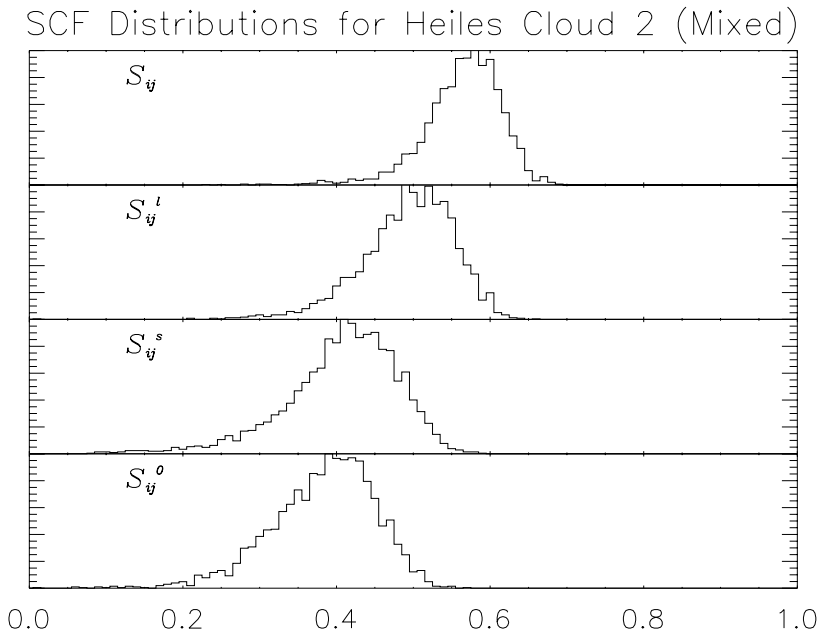
The meaning of the regions of correlation is, as yet, unclear. What can be said is that the SCF algorithm locates these small regions and finds them to be similar to their neighbors. No statement about the physical conditions in these regions has been made. Especially prominent regions can be examined by human observers, performing the usual analysis on a small subset of the map rather than the gargantuan data set.

The map results are difficult to interpret at this stage and statistical methods must be used to more fully understand these data. The two tools used are to generate histograms of the SCF and lag values over the map and to calculate the moments of the SCF histograms and the lag data. Figure 4.3 and Table 4.1 display each of these. In both the figure and the table, the original data are shown along with the data from the same SCF analysis performed on a map with the same spectra in random positions, a process referred to as mixing.

Perusal of Table 4.1 indicates that, on average, the unmixed data cube has a higher mean correlation and a larger skewness than the mixed data representing the fact that the spectra are better correlated and that the majority of the values are close to the mean with a tail off to smaller values. For the mixed data, the larger deviation and lower skewness



(a)



(b)

Figure 4.3: Histograms of the correlation functions generated by SCF analysis for Heiles Cloud 2. The histograms are normalized to the unit interval to emphasize the changing nature of the distributions. (a) Histograms of the data in their original positions. (b) Histograms with the data in randomized positions.

	Real Data				Mixed Data			
Parameter	Mean	Deviation	Skewness	Kurtosis	Mean	Deviation	Skewness	Kurtosis
S_{ij}	0.64	0.030	-3.0	39.	0.57	0.051	-1.2	3.5
ℓ_{ij}	0.00014	0.025	0.15	1.9	-0.004	0.22	0.081	-0.40
S^l_{ij}	0.63	0.033	-6.6	110	0.50	0.061	-0.88	2.1
ℓ^l_{ij}	4.1×10^{-5}	0.026	0.14	1.9	-0.0057	0.22	0.071	-0.41
S^s_{ij}	0.62	0.037	-5.3	81.	0.41	0.074	-0.95	1.5
S^0_{ij}	0.62	0.036	-6.7	110.	0.39	0.070	-0.87	2.2

Table 4.1: Statistical moments for SCF outputs for normal and mixed data for Heiles Cloud 2.

represents a wider distribution with a less drastic edge above the mean. The negative value of skewness adequately represents the presence of a tail of values towards lower correlations. The mean of the S_{ij} function for the mixed data is close to that for the unmixed map, indicating that the functions are roughly of similar shape over the entire map. Finally, it can be noted that the deviation of the distributions of v_{LSR} is 0.41, which is similar to the deviation of the values of ℓ for the mixed data. The deviations of ℓ from the unmixed map were much lower. This observation suggests that the unmixed map contains regions that exhibit similar velocity distributions, and these distributions are specific to areas of the cloud.

4.2.2 L1521

A similar analysis to that of Heiles Cloud 2 can be performed on the L1521 data. The observations are roughly, similar differing by a factor of two in spatial resolution. The data of primary use are the histograms and the statistical moments. As a result, only the maps of S_{ij} and T^*_A are displayed in Figure 4.4.

The maps of L1521 indicate that the regions of correlation are not the same as the regions of highest temperature. The regions of correlation stretch outside the areas of highest temperature, and the map appears to be more uniform than the temperature map. Again, viewing the histograms and statistical moments of the data yield insight into the nature of correlation in the observations of L1521. The statistical data appear in Figure 4.5 and Table 4.2.

In examining the statistical moments for L1521, some interesting features come to light. First, the difference in the value of all the correlation functions between the randomized and the original positions is smaller than it is for Heiles Cloud 2. The higher values of the SCF functions suggest that the spectra are more similar locally and the smaller shifts indicate a more global similarity. The histograms of the data when the spectra are in the mixed positions also display the long tail off toward lower values of the SCF that the Heiles Cloud 2 data display. The tails in the L1521 do not appear to develop as strongly in the S_{ij} and the S^ℓ_{ij} data for the mixed positions as they do for Heiles Cloud 2 data possibly suggesting that the spectra contributing to the tails have differences primarily in velocity. The differences between S^ℓ_{ij} and S^s_{ij} in the original values and their corresponding values for randomized

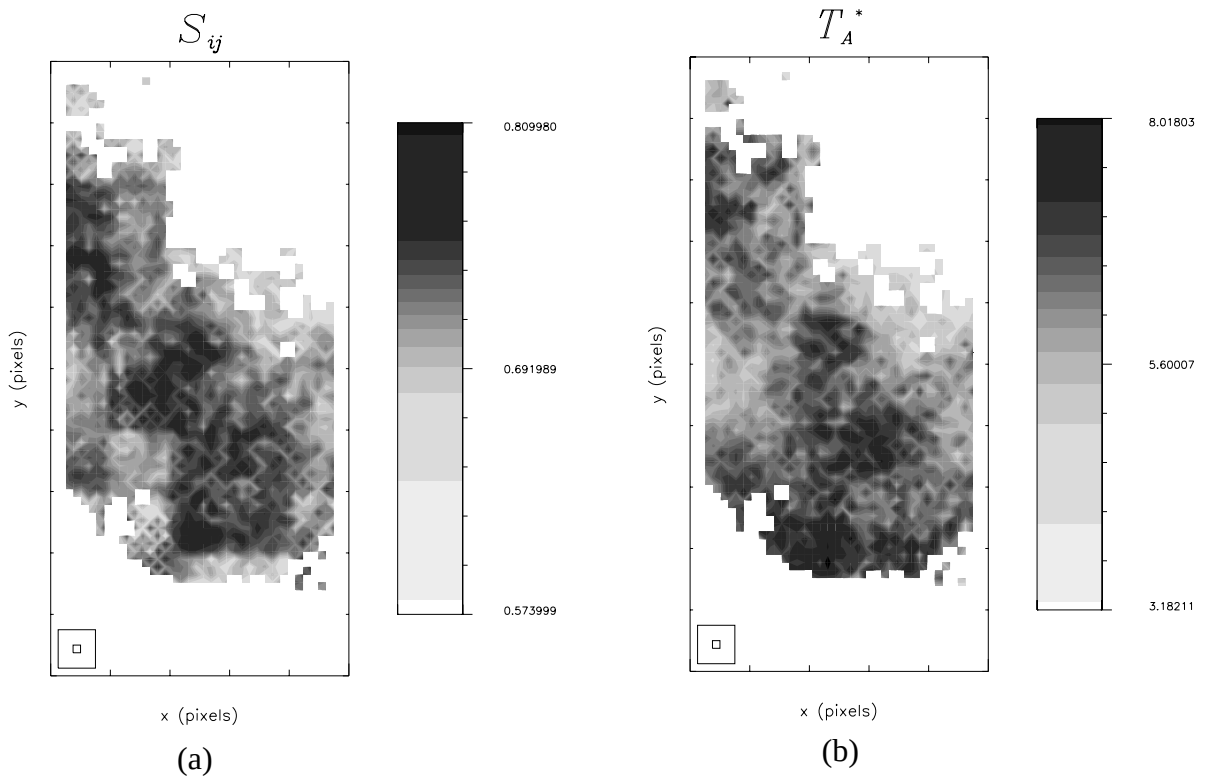


Figure 4.4: SCF results for L1521. (a) Map of the correlation function S_{ij} with lags and scaling on. (b) Map of the antenna temperature T_A^*.

	Real Data				Mixed Data			
Parameter	Mean	Deviation	Skewness	Kurtosis	Mean	Deviation	Skewness	Kurtosis
S_{ij}	0.72	0.040	-0.38	-0.24	0.65	0.036	-0.3	-0.18
ℓ_{ij}	0.00028	0.042	-0.23	0.93	-8.1×10^{-5}	0.12	0.52	0.21
S_{ij}^l	0.72	0.038	-0.23	0.93	0.64	0.041	-0.74	1.40
ℓ_{ij}^l	2.6×10^5	0.042	-0.22	0.93	1.2×10^{-5}	0.12	0.52	-0.22
S_{ij}^s	0.69	0.046	-0.42	0.20	0.57	0.069	-1.0	1.7
S_{ij}^0	0.69	0.043	-0.36	0.055	0.57	0.064	-0.95	1.3

Table 4.2: Statistical moments for SCF outputs for normal and mixed data for L1521.

positions indicate that the spectra are not significantly more correlated in velocity space or in amplitudes. These observations indicate that the spectra are more similar over the whole map than are the spectra in the Heiles Cloud 2 map. Examination of the map of T_A^* seems to support this hypothesis, for this map does not display several distinct regions like the Heiles Cloud 2 data indicate.

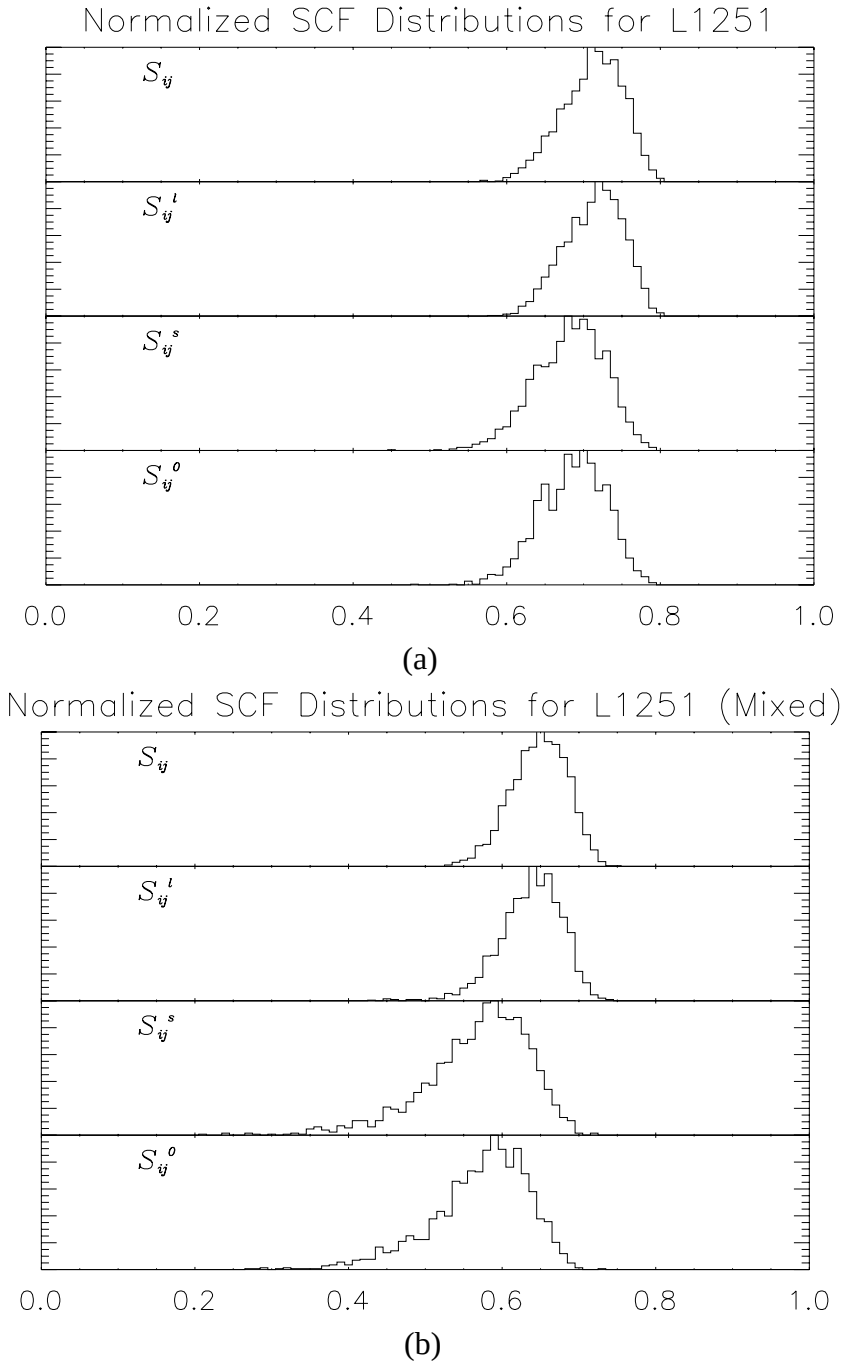


Figure 4.5: Histograms of the correlation functions generated by SCF analysis for L1521. (a) Histograms of the data in their original positions. (b) Histograms with the data in randomized positions.

4.2.3 Theoretical Cube I

The data set produced by Gammie, Ostriker, and Stone is the first theoretical data set examined by the SCF. Again, the authors do not wish these data to be interpreted as the

final results of their work, and the resulting cubes are only used as a probe of the effectiveness of the SCF. Maps of the correlation function compared with the antenna temperature appear in Figure 4.6. The standard statistical information appears as histograms in Figure 4.7 and as moments in Table 4.3.

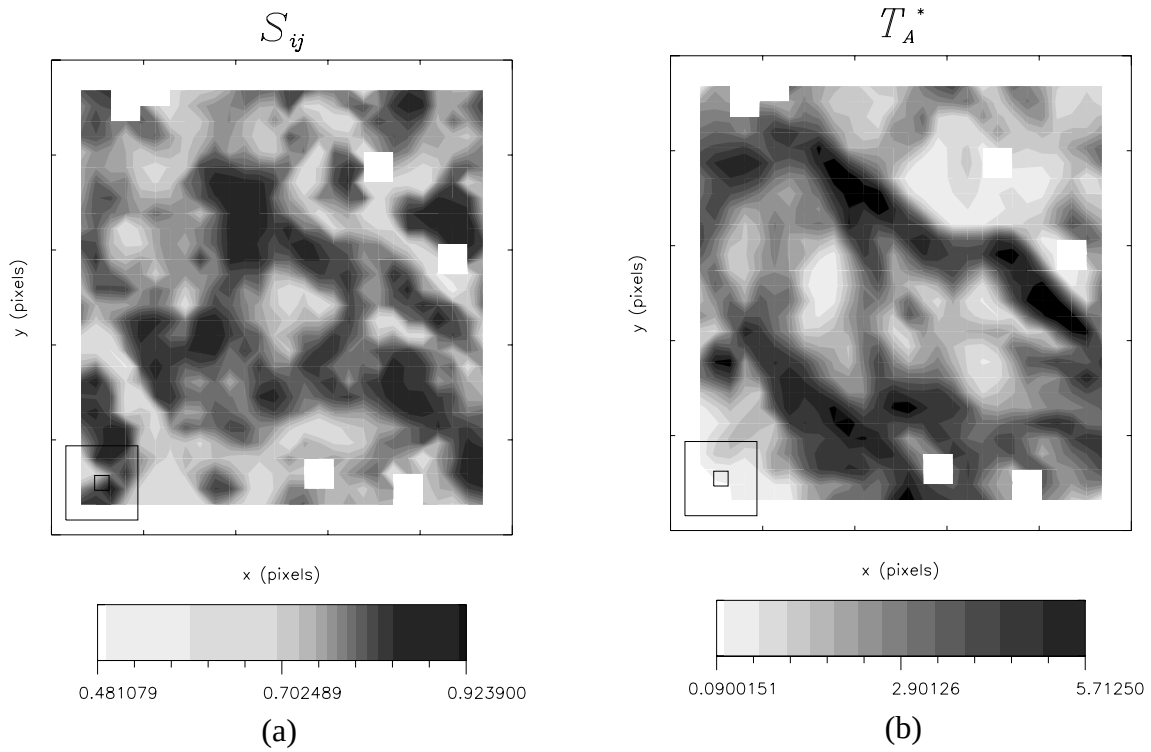


Figure 4.6: SCF results for Theoretical Cube I. (a) Map of the correlation function S_{ij} with lags and scaling on. (b) Map of the antenna temperature T_A^*.

The maps of the data indicate that the regions of high correlation are not necessarily related to the regions of highest peak emission. The regions of high peak emission are stretched out into structures that look similar to molecular clouds. The simulation has roughly the same spatial resolution as the maps of Heiles Cloud 2 and L1521 and the data indicate that the region involved is comparable in physical size to that of Heiles Cloud 2.

In observing the statistical information, some factors become readily apparent. First, $S_{il}^{s} > S_{ij}^{\ell}$ in the mixed map, which indicates, as a whole, the spectra are fairly well correlated in velocity space. The sudden drop when comparing these values to that of S_{ij} indicates that there are many dissimilar spectra from all parts of the the map. The moments of the SCF data behave similarly to those of the observed data, with the corresponding drops when the compensating factors s and ℓ are turned off as well as a decline when the positions of the spectra are mixed. The histograms of the data appear to be more spread out through the range of possible SCF values, an observation quantified by the higher deviation values for the data. The spreading out of the data is most pronounced when the spectra are examined in random positions. Finally, the distribution of the S_{ij}^{0} function for mixed spectra spans

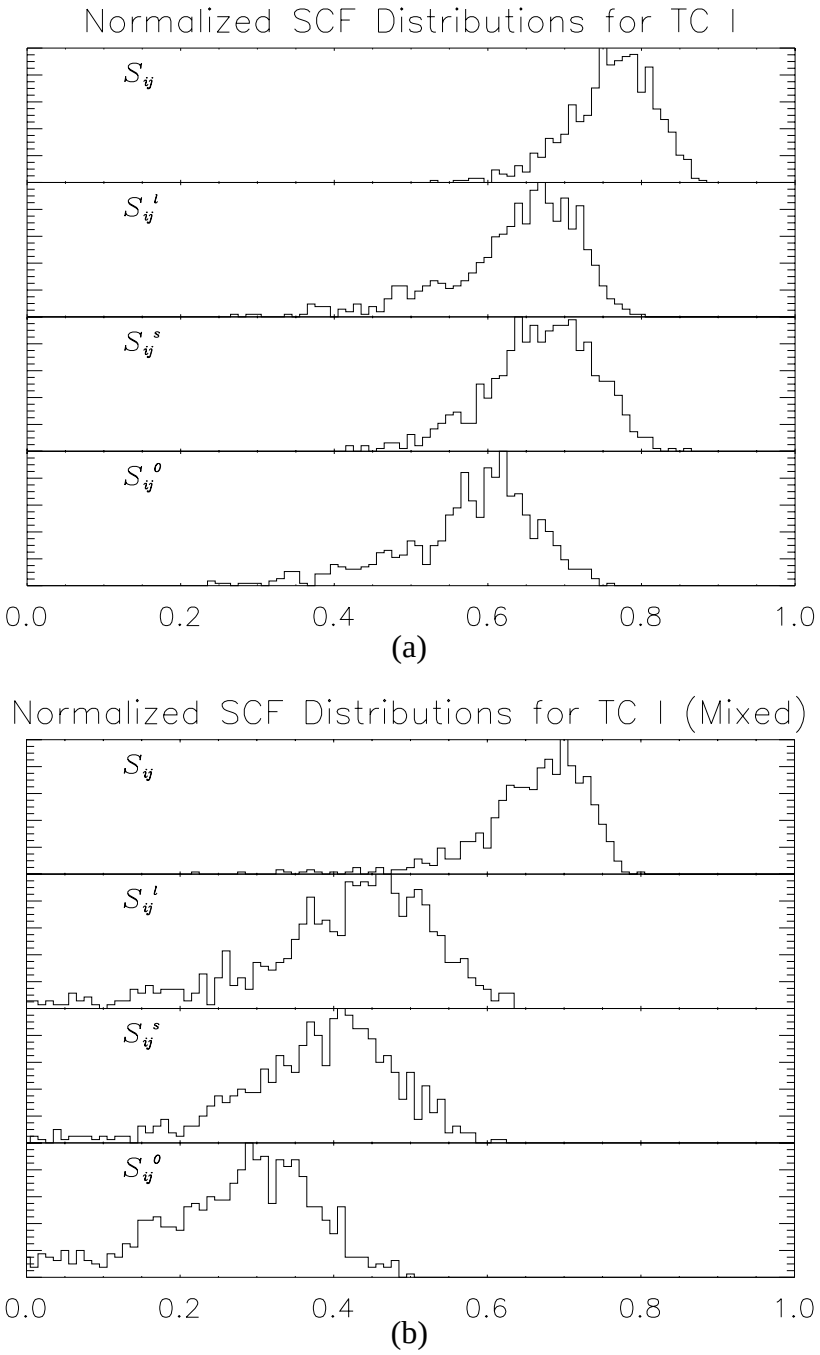


Figure 4.7: Histograms of the correlation functions generated by SCF analysis for Theoretical Cube I. (a) Histograms of the data in their original positions. (b) Histograms with the data in randomized positions.

a much lower range of values than do either of the data sets from real clouds, with several points approaching near zero correlation. This feature suggests that the spectra from this data cube have many different shapes, peak velocities, and heights, possibly spanning a

larger range than do those of the observed clouds.

	Real Data				Mixed Data			
Parameter	Mean	Deviation	Skewness	Kurtosis	Mean	Deviation	Skewness	Kurtosis
S_{ij}	0.76	0.056	-0.66	0.55	0.66	0.073	-1.6	4.8
ℓ_{ij}	6.1×10^{-6}	0.0017	-0.012	1.7	0.00036	0.0059	0.0028	-0.15
S^l_{ij}	0.63	0.086	-1.1	1.3	0.42	0.11	-0.98	0.96
ℓ^l_{ij}	5.3×10^{-5}	0.0019	0.15	1.2	0.00034	0.0060	0.0068	-0.17
S^s_{ij}	0.67	0.070	-0.51	0.29	0.38	0.10	-0.78	0.88
S^0_{ij}	0.58	0.090	-0.93	0.89	0.28	0.10	-0.55	-0.096

Table 4.3: Statistical moments for SCF outputs for normal and mixed data for Theoretical Cube I.

4.2.4 Theoretical Cube II

The cube produced by D.H. Porter and analyzed by E. Falgarone and her collaborators was the final data set to be analyzed using the SCF analysis techniques. As mentioned previously, this data set originally inspired the creation of the SCF to quantify the similarity between neighboring spectra in a data cube. The antenna temperature map of this cube as well as the SCF map appear in Figure 4.8.

The standard statistical analysis appears in Figure 4.9 and Table 4.4.

	Real Data				Mixed Data			
Parameter	Mean	Deviation	Skewness	Kurtosis	Mean	Deviation	Skewness	Kurtosis
S_{ij}	0.60	0.02	-0.71	0.48	0.59	0.030	-1.0	1.1
ℓ_{ij}	0.026	0.58	-0.080	-0.19	-0.0059	1.3	-0.13	-0.66
S^l_{ij}	0.59	0.026	-0.82	0.77	0.58	0.033	-1.0	0.84
ℓ^l_{ij}	0.027	0.58	-0.072	-0.20	-0.0056	1.3	-0.13	-0.67
S^s_{ij}	0.52	0.032	-0.45	0.32	0.45	0.06	-0.69	-0.02
S^0_{ij}	0.52	0.030	-0.30	0.046	0.45	0.052	-0.55	-0.17

Table 4.4: Statistical moments for SCF outputs for normal and mixed data for Theoretical Cube II.

The first thing that the maps of the SCF and the antenna temperature indicate is that this particular cube lacks distinct structures extending more than a few pixels across. Because the size scale is unclear, one cannot say whether this aspect of the map is because of the nature of the simulation or the size scale which is being examined. Nonetheless, widely dissimilar antenna temperatures appear next to each other in the map. This observation is borne out by the table of moments for the histograms. When the positions of the spectra are randomized, the mean hardly changes at all. The very small change in all correlation functions suggests that the original positions of the spectra were unimportant in composing the original map. The correlation values themselves are lower than any of the other cubes

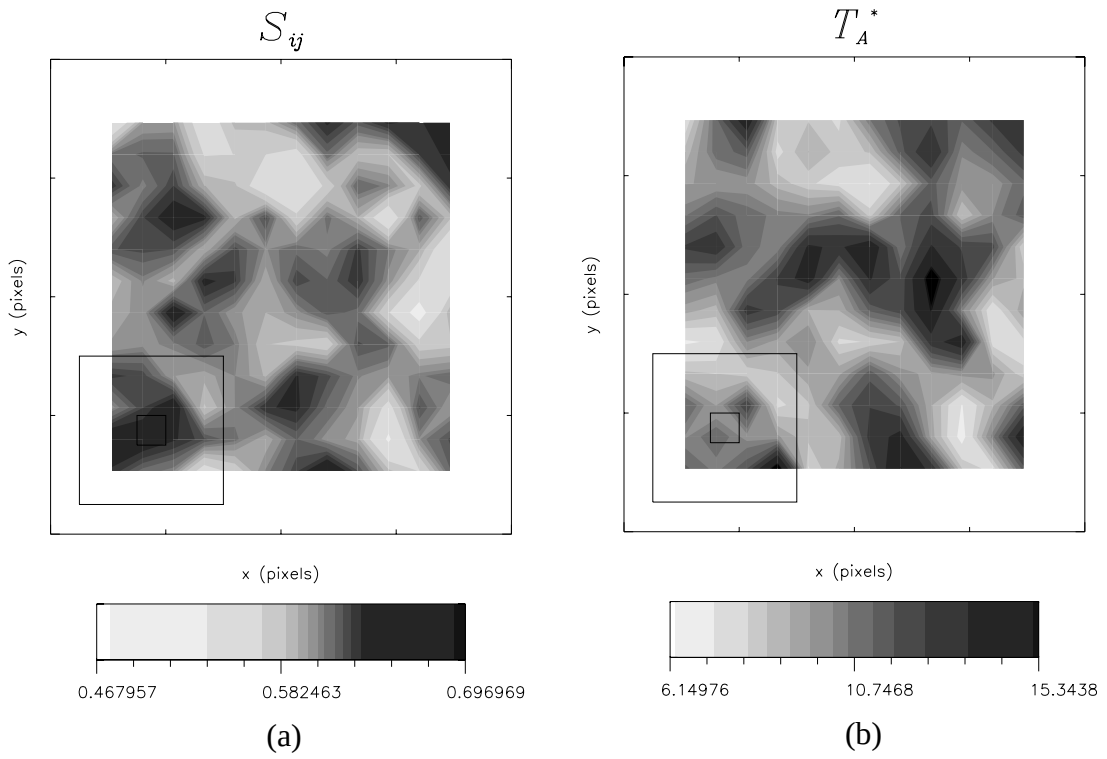


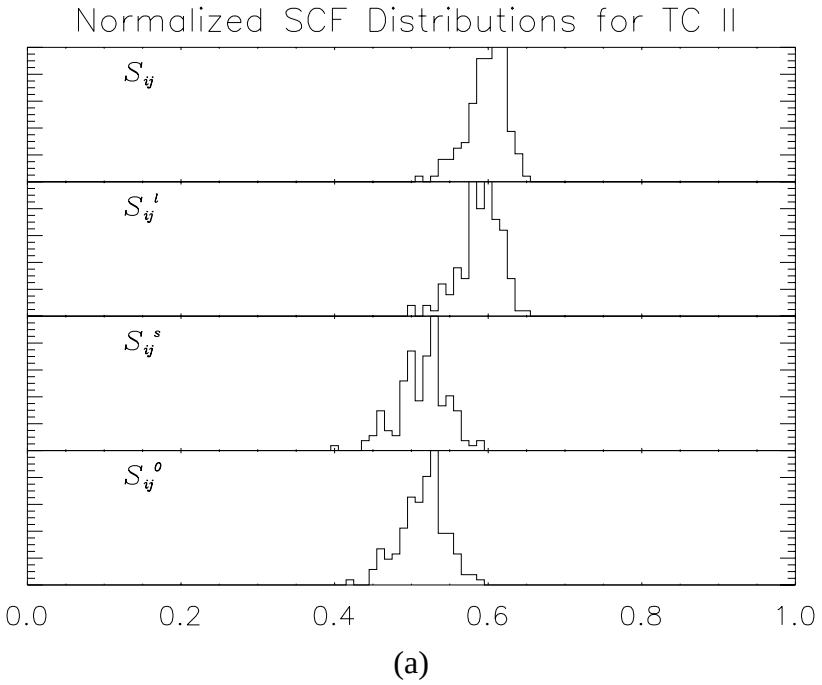
Figure 4.8: SCF results for the processing of Theoretical Cube II. (a) S_{ij} map displaying regions of correlated spectral shapes. (b) T_A^* map showing peak antenna temperatures as a function of position.

analyzed here, quantifying the observation that dissimilar spectra appear near each other in the map.

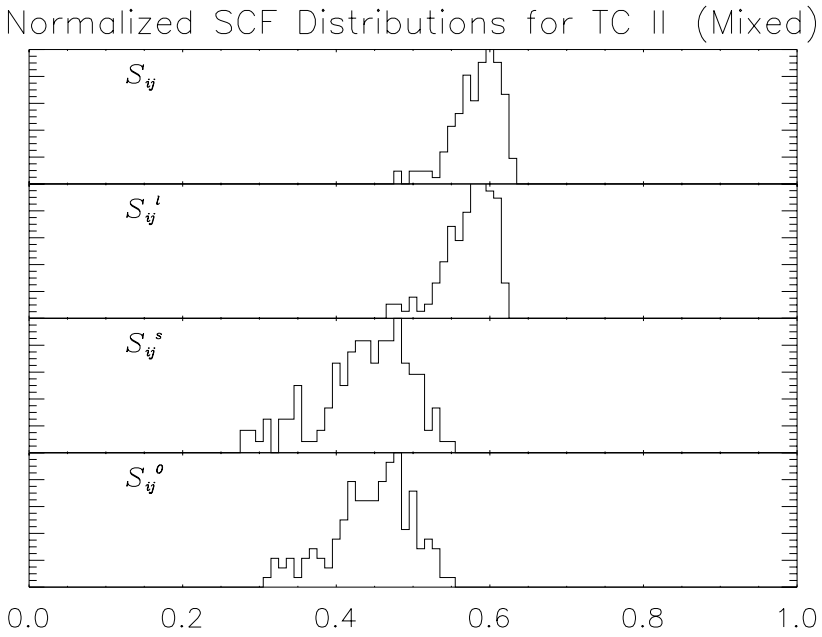
The mean value for S_{ij}^{ℓ} is larger than that of S_{ij}^{s} indicating that the spectra are more similar in amplitude than in velocity distribution. The small changes in the functions as these compensating parameters are turned on and off indicate that the similarities shared by the spectra are in both velocity, shape and antenna temperature. The values of the SCF seem to affirm the qualitative observation that the spectra in the cube do not share the similarity to their neighbors that is observed in real molecular clouds.

4.3 Conclusions

The preliminary work performed with the SCF algorithm indicates that it is performing its most basic task very well. It can distinguish between different data cubes and the interpretations of these differences offer some insights into the general nature of the data cubes examined. The most basic testimony to the success of the SCF function comes from the drops in values of the various correlation functions when examined with the spectra in randomized positions. Table 4.5 displays the various maps and the differences in the mean value of the correlation functions with real and random positions.



(a)



(b)

Figure 4.9: Histograms of the correlation functions generated by SCF analysis for Theoretical Cube II. (a) Histograms of the data in their original positions. (b) Histograms with the data in randomized positions.

The data in this table indicate that the SCF can distinguish between different types of theoretical data. Theoretical Cube I behaves much more like the Heiles Cloud 2 data than

	Heiles Cloud 2			L1521		
Function	Original	Mixed	Change	Original	Mixed	Change
S_{ij}	0.64	0.57	0.07	0.72	0.65	0.07
S_{ij}^{ℓ}	0.63	0.50	0.14	0.72	0.64	0.08
S_{ij}^{s}	0.62	0.41	0.21	0.69	0.57	0.12
S_{ij}^{0}	0.62	0.39	0.23	0.69	0.57	0.12
	TC I			TC II		
Function	Original	Mixed	Change	Original	Mixed	Change
S_{ij}	0.76	0.66	0.10	0.60	0.59	0.01
S_{ij}^{ℓ}	0.63	0.42	0.21	0.59	0.58	0.01
S_{ij}^{s}	0.67	0.38	0.29	0.51	0.45	0.06
S_{ij}^{0}	0.58	0.28	0.30	0.51	0.45	0.06

Table 4.5: Table of the mean values of the correlation functions for all data sets examined. The table compares the means for the original data sets and then for the data sets with randomized positions.

like the L1521 data in both the differences between real and randomized data and the change in the value of the SCF as the various compensatory factors are turned off. Theoretical Cube II behaves more like the L1521 data, but still exhibits some curious behavior which seems to be related to the qualitative observations of dissimilarity in the cube.

Without a larger body of knowledge about the spectral correlation function, it is impossible to say whether these theoretical models accurately represent the ISM. From these preliminary data, it is possible to say, however, that TC I behaves more similarly to the ISM than does TC II under the typical SCF analysis. Future investigation will indicate whether the differences are due to a size scale difference between the two simulations.

The SCF accurately represents the differences between neighboring spectra and recognizes their similarities. The results of the function using the regulated data sets indicates that the SCF behaves as would be expected. Further analysis of observed and theoretical data sets indicates that the SCF provides meaningful data about the spectral maps. From this point in the research, the problem of noise in the spectral maps is the primary issue which must be addressed. Once this problem is resolved, the interpretation of regions in the SCF maps can be commenced in earnest. Many different maps need to be processed both from radio observations and from MHD simulations so that the similarities and differences among maps can be understood in the context of existing analysis. The first part take of the SCF program has been a great success and the work will continue in the future.

4.4 Acknowledgments

I would like to thank several parties for their contributions to this paper. Primarily, I wish to thank Alyssa Goodman and Dave Wilner for their unending support and direction on this project in acting as advisors. I owe many thanks to Jonathan Williams for his help and orientation with IDL and insight with the project. Additionally, I would like to thank Marc Heyer of FCRAO for the use of the Heiles Cloud 2 data set before its official publication.

Edith Falgarone provided access to the L1512 data through the IRAM key project and I am very grateful for this addition to the analysis. Without the help of Charles Gammie, the chance to process the data from their MHD simulations would not be available. Finally, thanks are in order to Derek Lis who facilitated access to the simulation data from Falgarone *et al.* (1994).

Bibliography

Dickman, R. L. and Kleiner, S. C.: 1985, *Astrophysical Journal* **295**, 479

Dyson, J. and Williams, D.: 1980, *The Physics of the Interstellar Medium*, Manchester University Press, Manchester, England

Elmegreen, B. G. and Falgarone, E.: 1996, *Astrophysical Journal* **471**, 816

Falgarone, E., Lis, D., Phillips, T., Pouquet, A., Porter, D., and Woodward, P.: 1994, *Astrophysical Journal* **436**, 728

Fuller, G. and Meyers, P.: 1992, *Astrophysical Journal* **384**, 523

Gammie, C. and Ostriker, E.: 1986, *Astrophysical Journal* **466**, 814

Gill, A. G. and Henriksen, R. N.: 1990, *Astrophysical Journal* **365**, L27

Goodman, A., Barranco, J., Wilner, D., and Heyer, M.: 1997, *Astrophysical Journal* In Press

Goodman, A. A.: 1997, `http://cfa-www.harvard.edu/~agoodman/scf/scf.pdf`

Heyer, M. H. and Schloerb, F. P.: 1997, *Astrophysical Journal* **475**, 173

Houlahan, P. and Scalo, J.: 1992, *Astrophysical Journal* **393**, 172

Larson, R. B.: 1981, *MNRAS* **194**, 809

Mac Low, M.-M., Klessen, R. S., Burkert, A., and Smith, M. D.: 1998, *Phys. Rev. Lett.* In Press

Myers, P.: 1987, in D. Hollenbeck and H. Thronson Jr. (eds.), *Interstellar Processes*, pp 71 – 86, Dordrecht-Reidel, Boston

Passot, T., Vázquez-Semadeni, E., and Pouquet, A.: 1995, *Astrophysical Journal* **455**, 536

Pound, M. and Goodman, A. A.: 1997, *Astrophysical Journal* **482**, 344

Vázquez-Semadeni, E., Passot, T., and Pouquet, A.: 1996, *Astrophysical Journal* **473**, 881

Williams, J., de Gues, E., and Blitz, L.: 1994, *Astrophysical Journal* **428**, 693

Wiseman, J. and Adams, F. C.: 1994, *Astrophysical Journal* **435**, 708