MULTISPEC: A Code for the Extraction of Slitless Spectra in Crowded Fields

Jesús Maíz Apellániz
May 4, 2005

ABSTRACT

I present MULTISPEC, a code for the automatic extraction of multiple spectra from slitless exposures of dispersing elements such as objective prisms, grisms, or gratings. The code has been optimized for crowded fields composed of point sources such as resolved stellar clusters. MULTISPEC obtains the spectra by multiple profile fitting and is a spectral analog of a profile-fitting photometry package such as DAOphot or HSTphot. It requires a series of calibration files, of which the ones for the STIS objective prism have been calculated and are presented here as well.

Introduction

First-order spectroscopy using a small slit is an inefficient use of a bidimensional detector since only a small fraction of the pixels play a role in the detection of the signal. A long slit is an improvement if one needs to measure several sources in the same field but, unless all of your targets are aligned, is still quite inefficient since several exposures are needed to measure all the sources. Several alternatives (e.g multiple fibers, integral field units, micromirrors, or microshutters) are present in some current or planned telescopes to maximize the simultaneous collection of spectral data from multiple sources, but in others (e.g. HST) the only available alternative is the use of slitless exposures and a dispersing element, be it a grating, an objective prism, or a grism.

The use of slitless exposures invokes several possible complications. The first one, spectra overlapping due to crowding, is illustrated in Fig. 1, which shows slitless exposures of a single star, a sparse field, and a crowded one. In the first case (one source) it is straightforward to assign the origin of each detected count to its
source. In the second case (few tens of sources) we have the same situation for most stars in the field (though one of the bright sources in this specific example turns out to be a binary and there is a significant overlap between the two spectra). In the third case (few hundreds of sources) most sources overlap to some degree with neighboring ones and the assignment of an origin to a given detected count can be quite complicated.

A second problem is the treatment of extended objects, either sources or background. An extended source observed in slitless mode will be in general spread out in both the dispersion and the cross-dispersion directions. The latter is only a minor complication if there is no significant velocity structure in the extended emission; it can be solved by integrating over a larger window (if using aperture extraction) or convolving a model of the spatial profile with the instrumental cross-dispersion profiles (if doing profile fitting). The former is more serious, since it translates into a degrading of the spectral resolution. An extended background is an added problem for slitless exposures as compared to ones with a slit. We can distinguish two different origins for the background: astronomical objects, such as nebular emission from an HII region or a planetary nebula or the continuum from an unresolved stellar population; and foreground sources, such as geocoronal emission lines, atmospheric emission lines and continuum, zodiacal light, or earthshine. Foreground sources are usually rather uniform and they only complicate the extraction by degrading the S/N ratio for weak sources. Astronomical objects, on the other hand, can be a more serious problem, since they are likely to be spatially variable and to introduce confusion and larger uncertainties in the extraction.

A third problem, related to the first one, is the identification and positioning of the individual sources. If crowding is severe, as in the right panel of Fig. 1, it may be hard not only to separate different spectra but even to identify how many there are. But even if crowding is not severe, as in the middle panel of Fig. 1, the question remains as to how one measures the offset (or zero position) in the spectral direction.

In this Instrument Science Report I present MULTISPEC, a code that can be used to extract multiple spectra from slitless exposures. I start by addressing how each of the problems described above are attacked. Later, I describe how the different modules that make up the code are structured. I end up discussing the requirements to install and use MULTISPEC.

---

1We use the term background here in the general meaning of anything that is not the source, even though those photons may actually originate behind, at the same distance as, or in front of the source.
Techniques

The goal of extracting multiple spectra from a slitless exposure is, in several ways, similar to that of extracting the photometry of a collection of point sources in an image. In the latter case, when one is analyzing a sparse field the most straightforward method is aperture photometry with a small radius. However, for a crowded field the aperture photometry of a star can easily be contaminated by counts produced by neighbors, and a profile-fitting (sometimes called crowded-field) code such as DAOphot, Dophot, or HSTphot does a better job. For slitless spectroscopy, crowding is a larger problem than for photometry because a given source produces a significant number of counts over a larger number of pixels. Hence, given the same number of sources, chances of having an overlap are higher. Therefore, I decided to use profile fitting as the technique for MULTISPEC instead of aperture extraction, the spectroscopic equivalent of aperture photometry.

There are several issues which need to be addressed in fitting profiles to multiple spectra:

- A difference between a photometry- and a spectroscopy-profile-fitting code is that the former has no preferred direction in the detector and, for that reason, it uses 2-D profiles as its fundamental unit. Spectroscopic exposures, on the other hand, have a distinct asymmetry in that one direction shares the dual character of being an angular – or spatial – and a wavelength – or spectral – coordinate (that direction will be called here the dispersion direction). Furthermore, the position of the object in the perpendicular angular coordinate (the cross-dispersion direction) is fixed over its extent. For those reasons, it makes more sense to fit 1-D profiles in the cross-dispersion direction once we have determined the position of the object on the 2-D detector (i.e. its trace).

- Another difference between the two types of code is the need for auxiliary exposures. A photometry package can use one image to identify the sources and to extract the magnitudes of the objects. Doing so in a spectroscopy code is more difficult, given the possible extensive overlap among sources. For that reason, an image (or images) of the same field is a preferred alternative for source identification.

- The use of two different types of observations, an image (or images) for source identification and a slitless spectral exposure for the measurement of spectra, introduces an added complication: the need to accurately know the relative geometric distortion between the two in order to produce a one-to-one correspondence between the positions in one and the other throughout the field covered. These geometric distortions can be calculated by obtaining calibration spectral slitless exposures and images of a crowded field. The images are used to generate a simulated slitless exposure of the field without any distortion which can then be compared with the real exposure. Note that, due to the degeneracy between spatial and spectral coordinates in the dispersion direction, the precision of the geometric distortion solution in that direction is likely to be poorer than in the cross-dispersion direction unless the presence of absorption/and or emission lines is ubiquitous in the sources, their intensities and wavelengths are well known, and the dispersing element can resolve them.

As is also the case for crowded-field photometry, a detailed knowledge of the spatial (cross-dispersion) profile is required for a crowded-field spectroscopy code. Such a profile is likely to be wavelength dependent. If the spatial profile is not well known, systematic errors can be introduced, which would be small for bright stars and caused by an incorrect aperture correction but possibly large for dim stars, especially those close to bright ones.

- The ability of a profile-fitting code to accurately deconvolve the fluxes of a close pair is expected to depend heavily on the S/N ratio of the data (Porter et al. 2004). Therefore, we should expect different
uncertainties for stars which have the same number of counts at the same wavelength depending on the presence or absence of neighbors.

With those issues in mind, the problem can be established in the following way. For the sake of notation simplicity, we will assume that the dispersion direction is parallel or nearly-parallel to the $x$ axis (this is the case for the STIS modes). Then, the 1-D profile fitting can be done in the $y$ direction and we can work directly with column data. Let $C(i, j)$ be the number of counts at the pixel with coordinates $(i, j)$ after the background has been subtracted. We are ultimately interested in measuring the spectral fluxes $F_k(\lambda), (k = 1, N)$ of the $N$ stars present in the exposure but our immediate goal is to find the contribution $C_k(i, j)$ of star $k$ to pixel $(i, j)$ subject to the restriction:

$$C(i, j) = \sum_{k=1}^{N} C_k(i, j).$$  \hfill (1)

$C_k(i, j)$ is the result of the discretization in the vertical (cross-dispersion) direction of:

$$F_k(\lambda_k(i)) \cdot t_{\exp} \cdot P_{\lambda}(y - y_k(i)) / s(\lambda_k(i)),$$  \hfill (2)

where $\lambda_k(i)$ is the relationship between wavelength and the horizontal pixel index $i$ for star $k$ (which depends on the dispersion relation, the position of the star in the reference image, and the geometric distortion\(^2\), $P_{\lambda}(y - y_k(i))$ is the normalized cross-dispersion profile centered at the continuous vertical coordinate $y = y_k(i)$, $y_k(i)$ is the position of the center of the profile of the star $k$ at the pixel with horizontal index $i$, $t_{\exp}$ is the exposure time, and $s(\lambda)$ is the sensitivity function in units of $(\text{erg} \cdot \text{s}^{-1} \cdot \text{cm}^{-2} \cdot \text{Å}^{-1})/c(\text{cts} \cdot \text{s}^{-1} \cdot \text{pixel}^{-1})$. $P_{\lambda}(y)$ is a function that has to be determined from high S/N observations that allow for a determination of the cross-dispersion profile at the sub-pixel level. Finally, $y_k(i)$ is calculated from [a] the position of the star in the image, [b] the relative geometric distortion between the image and the spectral exposure (plus any possible offsets and/or rotations), and [c] the trace produced by the spectrum of a star centered on the detector (described by a function $y(x)$), which we will assume does not vary for stars located off-center (apart from the relative displacement in $x$ and $y$ derived from [a] and [b]).

If we now select the column defined by $i$, we will find that the traces of $N'$ stars ($N' \leq N$) cross it (for some dispersing elements, e.g. the STIS objective prism, the trace of an individual star does not span all the columns in the detector). Then, for that column we have $N_j$ equations (where $N_j$ is the number of pixels in a column) defined by Eq. 1 and $N'$ unknowns (the $F_k(\lambda_k(i))$ for the $N'$ stars). If we impose the reasonable condition that $N' < N_j$ (certainly, some crowding may be expected but it would be hard to measure more than one spectrum per pixel!), we find that the system is overdetermined. We can solve for such a system minimizing the $\chi^2$-like function:

$$g_i = \sum_{j=1}^{N_j} \left( C(i, j) - \sum_{k=1}^{N'} C_k(i, j) \right)^2 = \sum_{j=1}^{N_j} \left( C(i, j) - \sum_{k=1}^{N'} \frac{F_k(\lambda_k(i)) \cdot t_{\exp} \cdot \text{discretize}_j[P_{\lambda}(y - y_k(i))]}{s(\lambda_k(i))} \right)^2$$  \hfill (3)

\(^2\)Throughout this discussion I will assume for simplicity that $F_k(\lambda)$ is a quantity already discretized in wavelength and I will refer to discretization only in the spatial $y$ coordinate.
for the $N^\prime$ unknowns $F_k(\lambda_k(i))$. In Eq. 3 we discretize the cross-dispersion profile by integrating the sub-pixel precision $P_x(y - y_k(i))$ over the extent of a vertical pixel and, in that way, we take into account the different centering possibilities (e.g. pixel-centered or edge-centered vertical profiles).

Minimizing a $\chi^2$-like function which is as highly-non-linear and has a potentially large number of parameters as the one in Eq. 3 is easy to program but dangerous to apply blindly. The main reason is that accurate initial guesses for the parameters are required: otherwise, the algorithm can easily end up in a local minimum which is not the correct one. Furthermore, the existence of noise or incorrect background subtraction or knowledge of the cross-dispersion profile can also lead to unphysical (i.e. negative) results for $F_k(\lambda_k(i))$. These issues can be avoided by:

- Providing as good an initial guess for the parameters as possible. In MULTISPEC, I combine multicolor photometry of the sources with an a priori knowledge of their nature to estimate $F_k(\lambda_k(i))$ beforehand by spectral energy distribution (SED) fitting to the photometry.

- Using a code to solve Eq. 3 capable of placing constraints on the parameters in order to detect unphysical or unrealistic cases and to minimize the consequences on the calculated fluxes of neighboring sources.

Another issue is the treatment of the background, which needs to be subtracted before Eq. 3 is solved. The algorithm I use to calculate the background in a potentially crowded field involves the following steps:

1. Identification of the region of the detector where sources provide the dominant contribution to the total number of counts in a given pixel. That zone will be called the source region and the complementary zone the background region.

2. An optional estimation of the residual effects of the neighboring sources on the background region and a subtraction of that estimate from the original data pixel by pixel.

3. Calculation of a background map in the background region, possibly using spatial smoothing.

4. Calculation of the background in the source region by interpolating from the background calculated in the previous step.

5. An optional iteration to improve the definition of regions in the first step and the contamination from sources in the second one.

If the background is complex, the implementation of the above algorithm is far from trivial and it is likely to require fine tuning of the region selection and smoothing parameters in order to be accurate. Furthermore, if the background is rapidly varying and bright close to some sources, a precise solution may not be possible from the data (this is a larger problem for slitless spectroscopy than for photometry) and we may have to settle for an approximate one. In the next section I will discuss the options available in MULTISPEC for fine-tuning the background calculation.

Even after all the above measures have been taken, it is possible that some sources are nearly exactly aligned in the dispersion direction so that it will be impossible to unambiguously deconvolve their individual fluxes using that exposure. For that reason, a code should be able to identify those cases and warn the user of the problem. Also, it would be convenient to have a post-processing tool that combines the data from exposures with different orientations in order to use alternative data when such a circumstance arises.
MULTISPEC modules and their use

MULTISPEC is a profile-fitting crowded-field slitless-spectroscopy code that has been written in IDL and which consists of five different modules: PREPDATA, GENGUESS, CALCBACK, MAINFIT, and WRITEOUTPUT. Each module can be called separately or, alternatively, a wrapper named MULTISPEC itself can be used to call each of them sequentially. Figure 2 shows the basic flow diagram for the five modules that make up MULTISPEC and briefly describes the tasks performed in each one of them. Figure 3 shows a more detailed flow diagram that includes the auxiliary subroutines and the three IDL packages required (all of which are freely available from the cited web pages). Figure 4 shows the public (non-hidden) keywords available for the wrapper and each of the five modules.

MULTISPEC was originally written for the STIS objective prism and in the current version that is the only available mode. However, with the proper calibration files, other observing modes can be easily included. In this section we discuss the general characteristics of the software and in the next one we will concentrate on the STIS objective prism as well as on how new modes can be incorporated.

The current version of MULTISPEC can handle dithered data which can be combined beforehand into a single field or analyzed individually at each dither position, since it allows for the treatment of a number of image/slitless spectral exposure pairs. The user needs to provide the FITS files for the slitless spectral exposure and the image (in the latter case a dummy can be used if it does not exist) as well as a photometry list for all the stars he/she is interested in. The photometry can be extracted from the images using e.g. DAOPHOT or DOPHOT and two filters have to be provided in order to estimate the SED for each star (see the PREPDATA subsection for details).

The wrapper MULTISPEC

We start describing the wrapper MULTISPEC, which can be used to sequentially execute all the modules. Besides the four specific wrapper keywords listed in Figure 4, any of the module keywords can be used here so that they are passed along. Also, the first two wrapper keywords (DIROUT and RESTORE) are universal, meaning that they should also be used with the individual modules. A typical MULTISPEC execution could then be, e.g.:

MULTISPEC, DIROUT='ngc604', INPUT='ngc604_a.dat', EXPOS=1, $  
FIXPROF=3000., PROC=2.0, CRBACK=15.0, NCELLXBACK=32, NCELLYBACK=8

Such a command is equivalent to consecutively executing the following modules one by one:

PREPDATA, DIROUT='ngc604', INPUT='ngc604_a.dat', EXPOS=1, $  
FIXPROF=3000., RESTORE=restore
GENGUESS, DIROUT='ngc604', RESTORE=restore
CALCBACK, CRBACK=15.0, NCELLXBACK=32, NCELLYBACK=8, DIROUT='ngc604', $  
RESTORE=restore, PASS=1
MAINFIT, DIROUT='ngc604', RESTORE=restore, PASS=1
CALCBACK, CRBACK=15.0, NCELLXBACK=32, NCELLYBACK=8, DIROUT='ngc604', $  
RESTORE=restore, /SKIPWRITE, PASS=2
MAINFIT, DIROUT='ngc604', RESTORE=restore, PASS=2
WRITEOUTPUT, DIROUT='ngc604', RESTORE=restore, PASS=2

Here is a description of each of the specific wrapper keywords:
Figure 2: Basic flow diagram and description of the tasks executed by the MULTISPEC modules.
Figure 3: Detailed MULTISPEC flow diagram.

In addition, MULTISPEC requires that the following libraries be installed:

**IDL astronomy library**
http://idlastro.gsfc.nasa.gov

**jmaplot**
http://www.stsci.edu/~jmaiz

**Markwardt IDL library**
http://cow.physics.wisc.edu/~craigm/idl/idl.html
Figure 4: Keywords available for the MULTISPEC wrapper and for the individual modules.

DIROUT: Output directory.
RESTORE: File name for saving and restoring.
PROC: Number of iterations.
SKIP: Number of steps to skip.

INPUT: Name of main input file name.
EXPOS: Exposure to process.
MLIM: Limiting magnitude.
FIXPROP: Red wavelength limit for fixed cross-dispersion profiles.

NPOINTSMIN: Minimum number of points required to apply individual recentering.

CRBACK: Critical value for background.
ALTBACK: Flag to define background areas from the data instead of the model.
NCELLXBACK: Number of background cells in x.
NCELLYBACK: Number of background cells in y.
PASS: Iteration number.

PASS: Iteration number.

NOSL:EXPW: Flag to skip the last slitless exposure writing.
PASS: Last iteration number.
• **DIROUT** (default: ‘out’): Directory where output files are written.

• **RESTORE** (default: dataset name): MULTISPEC uses a series of IDL .sav files to transfer data between its different modules and RESTORE is used as the root for the name of the file. By default, it is set by PREPDATA (the first module) to be the HST dataset name of the spectrum (e.g. ‘o6bz05010’). Note that if the wrapper is not used (see the module-by-module example above), then the keyword must be used in order to let each module know which data to process. In such a case, the RESTORE keyword can be used as an output for PREPDATA and as an input for the rest of the modules.

• **PROC** (default: 1.5): This keyword determines the number of iterations for the calculation of the background and the profile fitting. The available options are:
  - 1: Both the background calculation and the profile fitting are performed once.
  - 1.5: An additional background calculation is performed after profiles are fitted (but no second profile fitting). Note that this option implies a last call to CALCBACK after WRITEOUTPUT which is not represented in Figures 2 to 4 for simplicity.
  - 2: A second profile fitting is performed after the background is calculated for the second time.

See the PASS keyword in CALCBACK, MAINFIT, and WRITEOUTPUT if you prefer to use module-by-module execution.

• **SKIP** (default: 0): Number of modules to skip before starting the execution. This option is useful to restart the execution from a certain point, possibly changing some keywords. More specifically, the starting points for the possible values are:
  - 0: PREPDATA.
  - 1: GENGUESS.
  - 2: CALCBACK (first pass).
  - 3: MAINFIT (first pass).
  - 4: WRITEOUTPUT if PROC = 1 or 1.5, CALCBACK (second pass) if PROC = 2.
  - 5: CALCBACK (second pass) if PROC = 1.5, MAINFIT (second pass) if PROC = 2.
  - 6: WRITEOUTPUT if PROC = 2.

**PREPDATA**

PREPDATA executes the preliminary steps required in MULTISPEC. It reads the different text and FITS files, selects the corresponding reference values from the information in the headers, calculates the best SED model consistent with the available photometry, obtains the transformation between image and spectral spatial coordinates and, optionally, recenters the data based on a cross-correlation between the real slitless spectral exposure and an artificially generated one based on the calculated SEDs. The current version of PREPDATA is prepared to work only with STIS objective-prism data but it can be easily adapted to other configurations. The following keywords are available with PREPDATA:

• **INPUT** (default: ‘input.dat’): Name of the main text input file name. A sample file is shown here:

  ngc4214/phot.dat  
  Text file with the stellar photometry.
visit3  FITS file directory.
stisnuv.f25cn182  First filter name.
stisnuv.f25cn270  Second filter name.
o6bz02isq o6bz02010  First image/spectrum exposure pair.
o6bz02iwq o6bz02020  Second (optional) exposure pair. More lines may be added.

The stellar photometry file should have one row for each star with the following seven columns:

- x coordinate.
- y coordinate.
- Right ascension in degrees (optional, it can be substituted by a dummy value).
- Declination in degrees (optional, it can be substituted by a dummy value).
- Magnitude in the first filter.
- Magnitude in the second filter.
- An integer numerical identifier for the object.

Currently, three choices are available for filters (stisnuv.f25cn182, stisnuv.f25cn270, and stisnuv.f25qtz) and only \( \log Z/Z_\odot = 0.5 \) high-gravity Kurucz models with 30 000 K and varying extinction using the Cardelli et al. (1989) law with \( R_V = 3.1 \) can be used for the initial SED estimates. In the future, I plan to integrate MULTISPEC with CHORIZOS (Maíz-Apellániz 2004), thus providing a wide range of SED families to be used.

- **EXPOS** (default: 1): Number of the image/spectrum exposure pair to process.
- **MLIM** (default: 80.0): Maximum magnitude limit to include in the extraction.
- **FIXPROF** (default: not used): Critical wavelength for the use of a constant cross-dispersion spatial profile (see the section on the STIS objective prism for an explanation).

PREPDATA requires two user-provided text files, the main one specified by INPUT and the photometry list which is named in the first row in the main one, and at least two user-provided FITS files, one image\(^3\) and one slitless spectral exposure. It also requires a number of calibration files that are described in the section on the STIS objective prism.

PREPDATA reads the file ‘geomdistopy\_final.output3’ in the geomdistop directory to check whether the user wants to use a specific recentering. That file contains three columns, the first one with dataset names and the last two with pixel displacements in \( x \) and \( y \). If the dataset being analyzed is included in that list (the user may edit it), then the displacements are used for recentering. If the dataset being analyzed is not in the list, then PREPDATA performs a cross-correlation between the real slitless spectral exposure and a simulated one derived from the photometry (see also the next subsection) to determine the displacements.

**GENGUESS**

GENGUESS uses the SED derived for each star from the photometry provided in the previous section to generate the following initial estimates: (1) the number of counts as a function of wavelength for each

\(^3\)Note, however, that all the image information is derived from the photometry list, not from the FITS file itself. The reason for including the FITS file name is simply for tracking down the header information and writing it into the headers of the final FITS tables.
star, \( C'_k(i) = \sum_j C_k(i, j) \); (2) the corresponding slitless spectral exposure obtained by placing each star at its corresponding location and adding the result for all the stars convolved by the corresponding cross-dispersion spatial profile, \( C(i, j) \); and (3) the additional parameters \( y_k(i) \) and \( \text{discretize}_j[P_\lambda(y - y_k(i))] \) for the multiple-profile fitting at each column. The initial estimates in (1) and (2) will be used to calculate the background and as initial guesses for the multiple-profile-fitting algorithm in the next two modules. From now on we will call the result of (2) the \textit{data estimate} and the slitless spectral exposure the \textit{real data}.

GENGUESS also calculates small corrections to the positions of the individual stars in the slitless spectral exposure \( (y_k(i)) \) if the geometric distortion is not known to a high degree of precision. One keyword is available with GENGUESS:

- **NPOINTSMIN** (default: 6): Minimum number of points to have the small position correction calculated for an individual star. In order to calculate the correction, GENGUESS first collapses the real data into 10-pixel horizontal bins and calculates for each star how many of those bins have a high S/N and are located far away from neighbors, making them good candidates for recentering (i.e., the default value implies a strong S/N over \( 60 = 6 \times 10 \) pixels without nearby stars). The user can manually decrease this keyword if the number of stars in the field is low.

GENGUESS produces as output the positions of all the individual spectra. The output is in the three forms of: [a] a human-readable text file, [b] a PS file with the positions drawn on top of the real data (see Fig. 5), [c] a file that can be read as a region list by SAOimage DS9 in order to overplot the positions on e.g., the FITS file of the real data. In the graphical output, red is used for stars where the small position correction has been applied individually and green for those where an average of the nearby stars has been used instead.

**CALCBACK**

CALCBACK calculates the background of the slitless spectral exposure. The following keywords are available:

- **CRBACK** (default: automatic): Critical value (in counts) used to define the background region. If not set, it is calculated from the image statistics.
- **ALTBACK** (default: not used): By default, the background region is defined from the data estimate (generated by GENGUESS in the first iteration and by MAINFIT in the second one). If this flag is set (by adding /ALTBACK to the command line), then the real data is used to define the background region.
- **NCELLXBACK** (default: 16): Number of background cells in the \( x \) direction. The number must be of the form \( 2^n \), with \( n \) an integer between 0 and 8.
- **NCELLYBACK** (default: 16): Number of background cells in the \( y \) direction. The number must be of the form \( 2^n \), with \( n \) an integer between 0 and 8.
- **PASS** (default: 1): Iteration number. The allowed values are 1, 1.5 (for the additional background calculation described for the PROC keyword in the wrapper), and 2.

CALCBACK uses as its first background estimate the subtraction of the data estimate from the real data. It then calculates a (broad) second background estimate by a weighted mean of the first estimate in each of the cells (as defined by NCELLXBACK and NCELLYBACK), and robustly smoothing the result. It then defines
Figure 5: Sample PS file output from GENGUESS. Note that the bright star in the lower left quadrant is a binary, so it appears as a double green rectangle (two stars, no individual recentering) instead of a single red one (one star, recentering). The stars without a rectangle have not been selected for fitting because they are fainter than MLIM. The STIS objective-prism field shown is the same one as in the middle panel of Fig. 1.
Figure 6: Different 2-D data used or generated by MULTISPEC for the crowded field case shown in Fig. 1. [upper left] Spectral exposure. [upper center] Associated exposure in imaging mode. [upper right] Modeled data. [lower left] Partial residual (exposure minus modeled data). [lower center] Final background. [lower right] Full residual (exposure minus modeled data and background). All panels use the same color scale and count limits. Note that for the STIS NUV-MAMA the plate scale (in arcseconds/pixel) for an image is 83.4% that of an objective-prism spectral exposure (compare the above left and above center panels).

the background and source regions, either from the estimate or from the real data, by (a) using CRBACK to provide a first estimate of the source region, (b) expanding it to include the neighboring pixels, and (c) defining the background region as the complementary of the source region. A third background estimate is calculated by using (a) the first estimate in the background region and (b) the second estimate in the source region. The third estimate is then robustly smoothed once more to produce the final background.

As previously mentioned, the background calculation in a slitless spectral exposure is far from trivial. It is expected that the user will have to try different values of the keywords until he/she is satisfied with the results. In order to simplify the process, every time CALCBACK is executed it produces three FITS files: [1] the data estimate, which is updated by MAINFIT after the first iteration with the newly calculated values (Fig. 6, upper right panel); [2] the first background estimate (e.g. the residual generated by the subtraction of the data estimate from the real data, Fig. 6, lower left panel); and [3] the final background (Fig. 6, lower center panel).
MAINFIT

MAINFIT is the core of MULTISPEC, since it is where the spectrum of each star is extracted. This is done by scanning the detector column by column and simultaneously fitting the contribution of each one of the stars present in that column, as described in the previous section. The fitting algorithm uses as initial guess the values calculated in GENGUESS and places two constraints on each of the individual values of $F_k(\lambda_k(i))$: a minimum of 0 and a maximum of three times the initial guess. The first one is included to avoid unphysical negative fluxes (e.g. when a dim spectrum is hidden in the wings of a very bright one or immersed in a bright background). The second one is included to avoid “runaway” fits when two spectra of very different intensity have separations of a small fraction of a pixel. Also, a short and a long wavelength limit are imposed on the fit for every star by calculating the values where the estimate of $C_k(i) < 10$ counts (this is equivalent to the imposition of a minimum S/N for the detection of a star by crowded-field photometry codes).

MAINFIT has one keyword available:

- **PASS** (default: 1): Iteration number. The allowed values are 1 and 2.

MAINFIT updates the values generated by GENGUESS in case a second iteration is used and prepares the data for their use by WRITEOUTPUT.

A sample fit for a section of a column is shown in Fig. 7.

WRITEOUTPUT

WRITEOUTPUT organizes and writes the final data. Unless the NOSLEXPW keyword is used (or PROC = 1.5 is set from the wrapper), it generates two FITS files with the final versions of the data estimate.
(as generated by the last iteration of MAINFIT) and the residual; e.g. the updated versions of the first two FITS files generated by CALCBACK. It also produces a FITS table for each of the stars with the following columns:

- **lambda**: Wavelength in Å.
- **flux**: Spectral flux, $F_k(\lambda)$, in erg s$^{-1}$ cm$^{-2}$ Å$^{-1}$.
- **sflux**: Uncertainty in the spectral flux in erg s$^{-1}$ cm$^{-2}$ Å$^{-1}$.
- **cts**: Counts, $C_k^*(i)$.
- **scts**: Uncertainty in counts.
- **cts0**: Initial guess used for the counts.
- **res11**: Residual counts integrated over 11 pixels centered around $y$.
- **res61**: Residual counts integrated over 61 pixels centered around $y$.
- **fact**: Sensitivity function, $s(\lambda)$, in (erg s$^{-1}$ cm$^{-2}$ Å$^{-1}$)/(cts s$^{-1}$ pixel$^{-1}$).
- **back**: Background value in counts.
- **flag**: Flag with the possible values of 0 (good fit), 1 (fit at edge of possible range, see previous subsection), and 2 (not fitted).
- **distany**: $y$ distance (in pixels) to the nearest fitted spectrum.
- **distm1**: $y$ distance (in pixels) to the nearest fitted spectrum with at least $1/10 \times C_k^*(i)$.
- **distp0**: $y$ distance (in pixels) to the nearest fitted spectrum with at least $1 \times C_k^*(i)$.
- **distp1**: $y$ distance (in pixels) to the nearest fitted spectrum with at least $10 \times C_k^*(i)$.
- **distp2**: $y$ distance (in pixels) to the nearest fitted spectrum with at least $100 \times C_k^*(i)$.
- **chitest**: Reduced $\chi^2$ for the seven pixels in the vertical direction centered around $y$.
- **x**: $x$ coordinate in IRAF style (center of pixel at lower left corner has $x = 1$) for this wavelength.
- **y**: $y$ coordinate in IRAF style (center of pixel at lower left corner has $y = 1$) for the center of the cross-dispersion spatial profile.

WRITEOUTPUT also populates the header of the FITS table with information about the datasets used, the star (coordinates, magnitudes, identification), and the MULTISPEC run.

The following keywords are available in WRITEOUTPUT:

- **NOSLEXPW** (default: not used): Flag to deactivate the writing of the last versions of the data estimate and residual.
- **PASS** (default: 1): Last iteration number used in MAINFIT. The allowed values are 1 and 2.
Spectra post-processing

The FITS table that WRITEOUTPUT generates for each star is the end product for a single slitless spectral exposure. However, the user may want to improve the result by combining different exposures with the same orientation to increase the S/N ratio or from different orientations in order to eliminate superpositions between stars. Indeed, the reason why WRITEOUTPUT includes columns such as back, flag, distany, distm1, distp0, distp1, distp2, and chitest is that the user may use them as criteria for such combinations. The choice of criterion (or combination thereof) to combine spectra has only been partially explored with the available data at the current time: more work is planned for the future. It is also possible to use the res11 and res61 columns to check for possible residuals due to e.g. an incorrect assumption of the cross-dispersion spatial profile.

Using MULTISPEC

As previously mentioned, MULTISPEC has been written with the STIS objective prism in mind. Indeed, it has already been used by Maíz-Apellániz and Bohlin (2005) to calibrate that STIS mode using spectral exposures of the standard star HS 2027+6051. An example of those exposures is shown in the left panel of Fig. 1. An issue discovered by Maíz-Apellániz and Bohlin (2005) for the STIS objective prism for \( \lambda > 3000 \) Å is the existence of contamination by the flux at shorter wavelengths. As explained there, a solution that minimizes the problem is to use a constant cross-dispersion profile for \( \lambda > 3000 \) Å and adding the fit residual over 11 pixels. MULTISPEC does this by using FIXPROF=3000.0 in PREPDATA and recovering the residual from the res11 in the table output. Of course, when doing the latter one has to check for the existence of nearby stars.

MULTISPEC requires a number of calibration files, some of which are standard STIS reference files and some of which have to be specifically created for the code. Of the first type there are five:

- APDSTAB: Aperture description table.
- DISPTAB: Dispersion coefficients table.
- PCTAB: Photometric correction table.
- PHOTTAB: Photometric conversion table.
- SPTRCTAB: 1-D spectrum trace table.

The selection of each type of file is determined from the corresponding keyword in the FITS header of the slitless spectral exposure. See http://www.stsci.edu/hst/stis/calibration/reference_files for more information on STIS reference files.

MULTISPEC also uses four specific calibration files that must be created. All of them are in human-readable text format and are as follows:

- geomdistopyxy_output: A file with four columns corresponding to the geometric distortion coefficients \((c_x, c_y, d_x, d_y)\) for the transformation between the image and the slitless spectral exposure. The coefficients are defined in Maíz-Apellániz and Ubeda (2004). Here, a fourth order polynomial is used and the coefficients are listed in the order: \((0,1), (1,1), (0,2), (1,2), (2,2), (0,3), (1,3), (2,3), (3,3), (0,4), (1,4), (2,4), (3,4), (4,4)\). The normalization is such that \(c_{x,1,1} \approx c_{y,0,1} \approx d_{x,1,1} \approx d_{y,0,1} \approx 1.0\) and \(c_{x,0,1} \approx c_{y,1,1} \approx d_{x,0,1} \approx d_{y,1,1} \approx 0.0\).
• geomdistopxy_final.output3: The pixel displacement list as defined in the PREPDATA subsection.

• prof.dat: A file with eight columns that defines the parameters for the cross-dispersion profile. The sub-pixel functional form used is a triple Gaussian defined by:

\[
P_{\lambda}(y) = \sum_{i=1}^{3} \frac{I_i}{\sigma_i \sqrt{\pi}} e^{-\left(y/\sigma_i\right)^2}.
\]

(4)

Since the cross-dispersion profile is discretized into pixels, MULTISPEC integrates \(P_{\lambda}\) over the extent of each pixel. \(P_{\lambda}\) must be normalized to 1, i.e. \(\sum_j \text{discretize}_j[P_{\lambda}(y)] = \int P_{\lambda}(y)dy = 1\), implying that \(I_1 + I_2 + I_3 = 1\). The eight columns in the file are: pixel number (starting at 0), wavelength, \(I_1\), \(I_2\), \(I_3\), \(\sigma_1\), \(\sigma_2\), \(\sigma_3\).

• syntnuvcols.dat: A three-column table with the synthetic colors in the Vega magnitude system as a function of reddening. The three columns are: [1] \(E(4405 - 5495)\), the monochromatic equivalent to \(E(B-V)\), see e.g. Maíz-Apellániz 2004; [2] the synthetic stisnuv.f25cn182 - stisnuv.f25cn270 color; and [3] the synthetic stisnuv.f25cn182 - stisnuv.f25qtz color.

MULTISPEC has been written in IDL and at the present time a beta distribution is available from http://www.stsci.edu/~jmaiz. The distribution requires the following three packages: the IDL astronomy library (available at http://idlastro.gsfc.nasa.gov), the Marwardt IDL library (available at http://cow.physics.wisc.edu/~craigm/idl/idl.html), and jmaplot (available from the same web site as MULTISPEC). Also, to avoid possible errors and conflicts between routine names, it is recommended that the beta version be run from the same directory where it is installed and that the current directory be the first entry in the IDL path. MULTISPEC has been tested only under IDL 6.0.

I would like to thank Linda Dressel and Paul Goudfrooij for their comments reviewing this ISR.

Bibliography