WFIRST-AFTA TECHNICAL REPORT
Doc #: WFIRST-STScI-TR1502
Rev: B
Date: June 17, 2015

Planning for the Analysis of WFIRST-AFTA Grism Data: A Review of Current and Future Slitless Spectroscopy Software and Data Analysis Approaches

John W. MacKenty, Gabriel Brammer, Nor Pirzkal, Russell Ryan, Norman Grogin, and Stefano Casertano

Abstract

This report considers some of the data processing issues presented by WFIRST-AFTA slitless spectroscopic observations. The nature and limitations of this type of data is discussed and past experiences briefly considered. The basic steps required for the extraction and calibration of spectra from grism datasets are outlined and a detailed case study drawn from our experience with HST WFC3/IR data is presented. Existing and future efforts to extend and improve the handling of HST grism data are discussed with an eye towards identifying issues and approaches relevant to WFIRST-AFTA.

1. Introduction

This Technical Report examines data analysis approaches and software systems for calibrating and extracting information from slitless spectroscopic observations (also known as grism observations). Slitless spectroscopy analysis is a complex issue highly critical to the scientific success of WFIRST-AFTA as this mission proposes to dramatically extend prior experience with data of this type. The planned WFIRST-AFTA mission will include a grism to determine galaxy redshifts over a solid angle of several thousands of square degrees. These data will provide improved constraints on the properties of dark energy from measurements Baryon acoustic oscillations and growth of large-scale structure (see Spergel et al. 2015). These data will also support many other archival uses (e.g. galaxy evolution studies) plus the grism may be a key tool for Guest Observer (GO) observations (expected to be 25% of the observing time).

The analysis of the combined WFIRST-AFTA data will be very challenging, because their size exceeds by orders of magnitude the largest grisms surveys that have executed to date, and because the requirements on automation and control of systematics will vastly exceed current standards. To set the stage for the planning of the analysis of this dataset, we provide here a review of the current state-of-the-art in grism analysis software. We describe the challenges, strengths, and weaknesses of existing approaches and tools, with a particular focus on how these translate to WFIRST-AFTA. We also describe the concepts underlying several new

analysis strategies that are currently under investigation, and which may prove useful for WFIRST-AFTA data analysis.

This review starts from the perspective of those of us involved with the implementation of slitless capabilities for *HST*. We review the primary features, history, and issues connected with slitless spectroscopy in Section 2. In Section 3, the primary data analysis steps and associated calibrations are listed. Section 4 examines in detail the WFC3/IR Grism case and various practical difficulties are considered. Case studies of several of the current state of the art data analysis tools are presented in Section 5. Section 6 considers potential future approaches for extracting more information content from slitless spectroscopic observations. Finally, Section 7 raises some issues relevant to scaling to the very large surveys envisioned for WFIRST-AFTA.

2. Slitless Spectroscopy

A classical spectrometer projects the dispersed image of a slit onto a detector. The line spread function is determined by the width of the slit (assuming it is narrower than the object itself) and the properties of the dispersing element. If the slit is omitted, then the line spread function is determined by the morphology of the light source (and the properties of the dispersing element). A slitless spectrometer provides a spectrum of each point within the field of view. Of course, the spectra of these points are heavily overlapped along the dispersion direction. However, for the limiting case with a sufficiently small number of sources that overlap is unimportant and where the background is very much less than the source brightness, the full multiplex advantage is obtained.

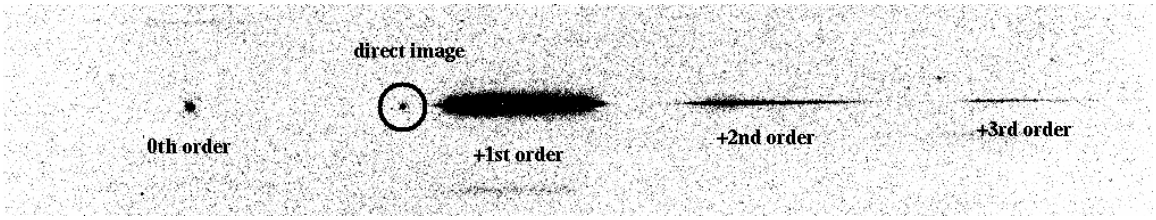

Figure 1: WFC3 Grism spectrum of a point source. The location of the source in a direct image is shown superimposed within a circle. The -1 spectral order is not shown in this figure but is comparable in brightness to the +2 order.

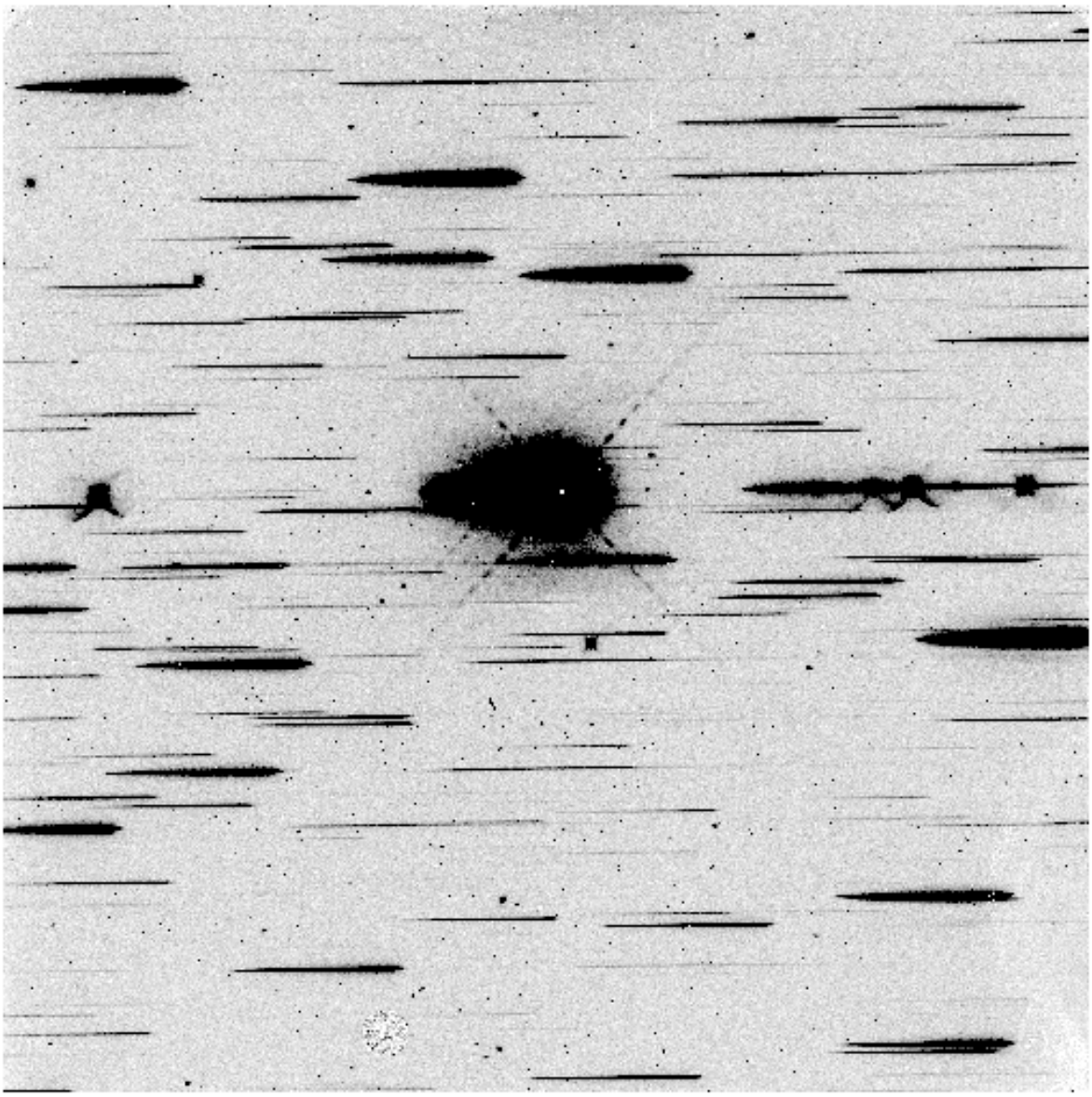

Figure 2: WFC3 G102 spectrum of the field around Planetary Nebula HB12. The first order spectrum of the PN is the large structure near the center of the frame, with diffraction spikes coming from bright emission lines of the compact central source. Emission lines are also seen in the additional spectral orders offset to the right. Numerous fainter stars are seen across the field (mostly in zero and first order).

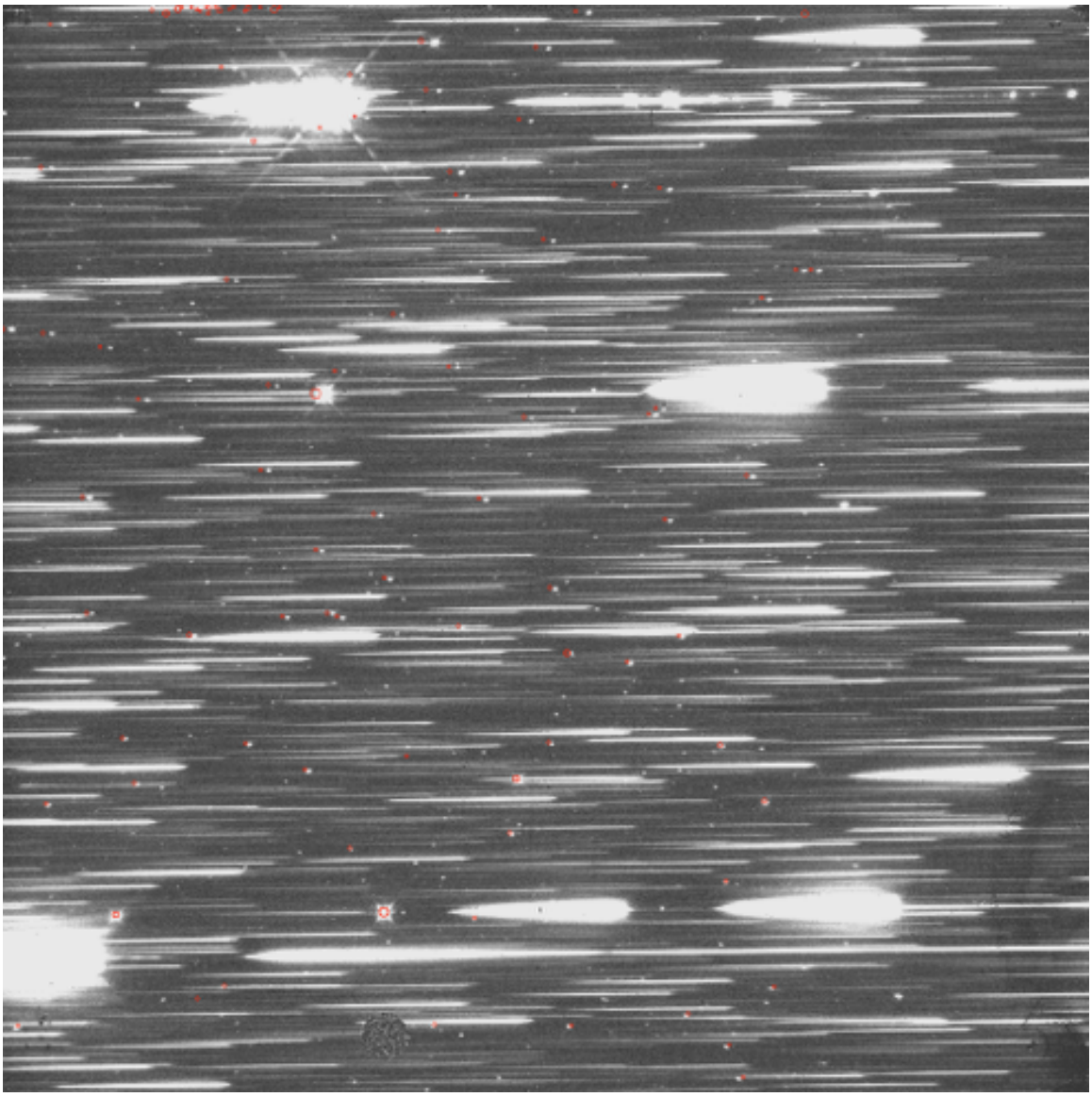

Figure 3: Deep WFC3 G141 grism made from stacking multiple orbits of data (here around the PN Vy 2-2) illustrates the complexity of overlapping spectra as scenes become confusion limited. The red circles denote zeroth order spectra.

2.1. Advantages and Drawbacks to Slitless Spectroscopy

When observing point or compact sources, slitless spectroscopy is attractive for several reasons: (1) no light is lost due to truncation of the source by the entrance slit, (2) considerable multiplex advantage is possible, (3) a priori knowledge of the locations of the sources of interest is not required, and (4) a relatively simple optical system is required. These advantages are offset by two major limitations: (1) the dispersed spectra may overlap, and (2) each wavelength along the spectrum has superimposed upon it the background of the entire spectral bandpass. For extended sources, an additional disadvantage is the convolution of the source morphology

along the dispersion direction with the line spread function resulting in decreased spectral resolution relative to a point source.

2.2. Practical Implementation Aspects

Slitless spectroscopy is a relatively simple augmentation to a camera, assuming that the camera provides access to a reasonable pupil (e.g. coincident with a filter mechanism). The typical solution in recent years is the inclusion of a “grism” in place of a filter. The grism consists of a prism wedge with a transmission grating replicated on one side. The purpose of the prism is to direct the first order spectra onto the approximate location of an undispersed image –thereby making optimal use of the field of view of the instrument. The dispersion is accomplished via the replication grating that is typically blazed to provide the most efficient possible first order (thereby minimizing the impact of the other orders). Note that the use of the prism implies that the zeroth order will be at least slightly dispersed. The inclusion of band limiting thin film filter(s) on one side of the prism is also often an aspect of the grism’s design.

An alternate approach successfully used for numerous ground-based surveys was the incorporation of an objective prism. Typically used with Schmidt-type telescope, this prism could cover all or part of the entrance aperture of the telescope allowing a planned balance between the direct image of each source and its single dispersed spectra.

In both designs, several practical considerations arise. The primary one is contrast. Since the background at each pixel of the detector is effectively the imaging background for a filter equal to the bandpass of the grism (ignoring possible vignetting at the edges), increasing the dispersion linearly decreases the contrast between the spectral continuum and the background. This tends to strongly favor low dispersion spectrometers—which also make more optimal use of the detector area in a camera type configuration. A significant exception to this behavior is the observation of compact sources with strong emission lines where increasing dispersion does not decrease contrast as long as the emission lines remain unresolved. Fundamentally, one requires sources relatively brighter compared to the background that would necessarily be the case for a spectrometer with input apertures (e.g. slits or fibers).

The second practical consideration is the handling of overlapping spectra. This may be due to either complex source morphologies and/or the surface density of sources in the field of view. Some degree of mitigation is possible by rotating the position angle of the dispersion. The special case of observations of transient phenomena is relatively simple (e.g. supernovae observations in distant galaxies with *HST*) and is usually solved via differencing data obtained at two epochs.

2.3. History

In the early days of photographic based astronomical observations, slitless spectroscopy was an important tool. Photographic plates provided large areas but relatively low quantum efficiency. Wide field astrographs and (later) Schmidt telescopes made excellent use of photographic plates. The inclusion of objective prisms converted these telescopes to highly

efficient spectrographs leading to key advances. The most famous result being the Harvard spectroscopic surveys started by Edward Pickering and analyzed by Annie Jump Cannon in the early 1900s. These data led to the fundamental classifications of stars based upon temperature and surface gravity forming the foundation of the modern understanding of stellar evolution. These surveys provided spectra of hundreds of thousands of stars down to 8th magnitude. As such, these surveys were little affected by sky brightness and required only moderate spectral dispersion.

More recent surveys have pushed to fainter sources where sky brightness becomes a factor. While these surveys (e.g. the Byurakan surveys for UV excess objects and UK Schmidt objective prism for AGNs) made important contributions, with the development of CCD detectors and highly multiplexed multi-fiber or slit-plate spectrometers, the use of slitless techniques from the ground has become less frequent. While CCD-based slitless surveys successfully identified large samples of emission-line galaxies over large areas of the sky (e.g., "KISS"; Salzer et al. 2000), it is worth noting that multi-fiber or slit-plate spectrometers provide much simpler datasets for automated analysis.

The GALEX mission provided a slitless spectroscopic mode covering the FUV (1340–1786 Å, *R*=200) and NUV (1770–2830 Å, *R*=90) channels with a grism that could be rotated on its axis to vary the position angle of the dispersed spectra within the 1.25º instrumental field of view (Morrissey et al. 2007). Rotating the grism allowed the identification of the optimal dispersion position angle for a given source that minimized contamination by spectra from other nearby sources. The GALEX pipeline processing provided two-dimensional image strips aligned with the dispersion axis and covering the brightest spectral orders yielding a result similar to the *HST* example shown in Figure 1 (Morrissey et al. 2007). In the deepest GALEX spectroscopic observations that covered the *Chandra* Deep-Field South (CDFS), some 342 exposures at independent orientations were obtained with a total integration time of nearly 100 hours (353 ks). Barger et al. (2012) used the independent orientations to construct a NUV spatial+spectral data cube analogous to an IFU covering 32', from which they identified 28 Lyman-α emitters at $0.67 < z < 1.16$. From the full suite of deep spectroscopic campaigns with GALEX, Cowie et al. (2010) identified 260 Lyman-α emission-line selected sources out to z=1.25 with a combination of an automated search for Gaussian features and visual inspection of spectra and line candidates.

2.4. *HST* Capabilities

The Hubble Space Telescope (*HST*) has implemented a slitless capability in most of its science instruments. First generation instruments WF/PC-1 and FOC (TBR) provided UV slitless modes that were very little used (due to spherical aberration the WF/PC-1 capability was used exactly once to support a calibration activity). The nature of the fix for spherical aberration implemented in WFPC2 precluded a slitless mode. However, both STIS and NICMOS provided significant slitless modes leading to the first large observing programs which used NICMOS to study the Hα emission in galaxies at z~1 (McCarthy P. et al. 1999). This also led to the development of the first supported data processing systems (see §5.1). The ACS provided both a prism in the SBC and a visible/red grism in the WFC that was extensively used for galaxy evolution studies

(see §5.3) and for the classification of higher redshift SN 1a (e.g. Rodney et al. 2012). Building upon the successes of NICMOS and ACS, the WFC3 camera included infrared grisms that have been extensively used for programs ranging from exo-planet transits observations to studies of the most distant galaxies detected to date. Since SM4, WFC3 grism observations (including parallels) have accounted for ~10% of all *HST* observing time. This has motivated major software efforts at STScI and elsewhere to provide supported tools and robust calibrations. This experience informs a major part of this Report.

The advantages of *HST* of lower sky levels (especially in the infrared) and higher angular resolution often make space-based slitless techniques more attractive than ground instruments. For certain applications at low spectral resolution the 2.4 meter *HST* outperforms the largest ground based telescopes in the near infrared even when faced with the limitations of slitless observations.

Table 1: Current HST Slitless Spectroscopic Modes (post-SM4)

Mode	Wavelength (nm)	Resolution (λ/dλ)	Sensitivity A0 star 1h 5σ (Vega Mags)
ACS/SBC PR110L	**115 – 180**	**79**	**20.9**
ACS/SBC PR130L	**125 – 180**	**96**	**21.5**
ACS/WFC G800L	**500 – 1050**	**100**	**24.4**
STIS/NUV PRISM	**115 – 360**	**2500 – 10**	**~21**
WFC3/UVIS	**190 – 450**	**70**	**21.9**
WFC3/IR G102	**800 – 1150**	**210**	**22.0**
WFC3/IR G141	**1075 – 1700**	**130**	**21.3**

Table 2: JWST Slitless Spectroscopic Modes

Mode	Wavelength (nm)	Resolution (λ/dλ)	Sensitivity 3h 5σ (AB Mags)
NIRCAM	**2400 – 5000**	**2000**	**~24**
NRISS	**1000 – 2500**	**150**	**~24**
MIRI	**5000 – 14000**	**100**	**23**

2.1. Future Mission Capabilities (JWST, Euclid, and WFIRST-AFTA)

In most respects, the JWST slitless spectroscopic capabilities are similar to *HST*'s (although NIRISS provides a specialized higher dispersion mode specifically aimed at exoplanet transit observations and NIRCAM's grisms provide only relatively high spectral resolution). Consequently it is reasonable to expect that the JWST data processing, calibration, and analysis requirements will be similar to our *HST* experience and will require only incremental advances in software tools. Thus we may view the creation of second-generation *HST* tools as positive investments in JWST.

However, the two planned large-scale infrared survey missions Euclid and WFIRST-AFTA pose interesting data processing, calibration, and analysis challenges in several areas. First, the scale of these projects requires the application of fully automated methods whereas prior experience involved considerable manual processing, exception handling, and quality assessment. Second, the primary scientific objectives of these missions require substantially improved control and knowledge of systematic effects. However, both missions employ similar instrumental approaches and offer similar capabilities to the current *HST* WFC3/IR grisms albeit at higher spectral resolution.

3. Generic Requirements for Software, Key Definitions, and Basic Calibrations

While the instrumental implementation of a slitless capability is relatively straightforward, the analysis of the resulting data has proven to be a major challenge. In this section we walk through the steps required to extract and calibrate spectra from slitless observations.

3.1. Key Data Items

3.1.1. Direct Image: An image of the field deep enough to detect all sources whose spectra are of interest and those bright enough to contaminate sources of interest via overlap. Often it is desirable to image a larger field to identify sources whose higher order spectra (e.g., +2 or -1) will appear in the spectroscopic image.

3.1.2. Spectroscopic Image: The actual dispersed image. Frequently multiple exposures with small positional offsets ("dithers") are obtained to improve spatial sampling and to avoid detector defects. Exposures obtained at multiple field orientations can help to identify and mitigate contamination of overlapping spectrum.

3.1.3. Source List: A list of all sources in the direct image. Contains at least the position of each source. Typically a size parameter is included in order to define which pixels are of interest for that source. Some analysis approaches benefit greatly from including multi-wavelength photometry to provide priors for the analysis of extracted spectra.

3.1.4. Spectral Trace: A function that defines the where in the grism exposure light from a position in the direct image (e.g., a source or a single pixel) at a specific wavelength will be found.

3.1.5. Extracted Spectrum: The resultant spectrum calibrated in wavelength and flux units. May be provided as either a one-dimensional spectrum or a two-dimensional image.

3.2. Data Processing Steps

3.2.1. Image Correction: Each image is first corrected for the usual instrumental effects including dark current and linearity (except that the flat field is not corrected for the spectroscopic image in a single step but is rather included later in the analysis).

3.2.2. Background Removal: As slit-less spectroscopic observations disperse every point in the image into a spectrum, the background at each point is the sum over the entire spectroscopic passband. The diffuse backgrounds arise from zodiacal light plus other unwanted sources (e.g. earth's illuminated limb and atmosphere). Frequently, this background has considerable spatial structure, especially at the edges, thus the subtraction of this background is required. If the spatial distribution of the background is constant, all that is required is scaling its amplitude and subtraction.

3.2.3. Source Identification: The direct image provides a source list providing the position and (optionally) shape and size of each source of interest. In the spectroscopic image, each source is associated uniquely with a spectral trace via a prior calibration.

3.2.4. Trace Extraction: The set of pixels in the spectroscopic image associated with each source is extracted into a two-dimensional spectrum.

3.2.5. Wavelength Calibration: A mapping between the two-dimensional spectrum and wavelength zero point and dispersion is obtained via a prior calibration. The zero point is obtained either from the direct image (assuming that the offset between the direct and spectroscopic images is known) or from features in the spectrum (e.g. emission lines or the ends of the grism transmission). The dispersion curve varies slowly within the field of view of the grism while the zero point is unique to each pixel (or source).

3.2.6. Flat Field Calibration: Each pixel in the two-dimensional spectrum is a (nearly) monochromatic illumination of the detector. The flat field is a combination of optical throughput variations and detector pixel response. The detector response is dependent upon both position and wavelength. A calibration correcting for the pixel response as a function of wavelength provides a flat field data cube for correcting the two-dimensional spectrum.

3.2.7. Geometric Distortion: The trace extraction, flat fielding, and wavelength calibration steps are very simple if the direct and spectroscopic images are not geometrically distorted. In practice, variations in both the projected pixel size and the geometric mapping complicate this process and constitute a major reason for the complexity of software systems. Logically, some of the geometric rectification steps need be applied before and some after the source identification while some geometric rectifications need to be applied during the trace extraction steps with consequences for the various calibrations.

3.2.8. Spectral Extraction: The two-dimensional spectrum of the source is collapsed into a one-dimensional spectrum. This may be weighted by knowledge of the source light distribution.

3.2.9. Flux Calibration: Finally, the relationship between detector units (e.g. counts) to astrophysical flux units is used to convert to flux units.

3.3. Required Calibration Reference Files

3.3.1. Detector (dark, bad pixels, linearity, direct image flat): These are identical to the standard imaging calibration files.

3.3.2. Spectroscopic background: This may be obtained from stacking large numbers of spectroscopic images with sources masked out. Since the structure of this background depends upon the spectrum of the zodiacal background and the stray light properties of the telescope+instrument system, it will be difficult to obtain such a calibration pre-launch.

3.3.3. Geometric distortion: The geometric distortions in the instrument may be modeled prior to launch or (preferred) measured from stepped observations of astrometric standard fields.

3.3.4. Direct image-to-trace mapping: This is a smoothly varying function of position within the field of view and may be obtained from observations of star fields.

3.3.5. Dispersion solution / wavelength calibration: This requires observations of compact sources with strong spectral features (e.g. planetary nebulae or M dwarfs)

3.3.6. Spectroscopic flat field data cube: This calibration is most easily obtained as part of a pre-launch calibration using a monochrometer.

3.3.7. Flux conversion: This requires observations of spectrophotometric standard stars.

4. Case Study: WFC3/IR Grism Reduction

4.1. Initial Frame Processing

Slitless grism exposures from the *HST* instruments are initially processed in much the same way as the normal imaging counterparts with the same instruments. This processing includes removal of the bias and dark-current, linearity corrections, populating the data-quality pixel flags and propagation of pixel uncertainties following the instrumental noise model (e.g., detector read noise and photon shot noise from the observed source + background counts). Differing from the typical processing of imaging filter exposures, the flat-field correction (flux calibration) of grism exposures is handled separately (§**Error! Reference source not found.**).

WFC3/IR exposures are taken with multiple non-destructive reads of the detector saved throughout the exposure. For example, a half-orbit 1400 s exposure might be composed of 14 separate non-destructive reads obtained at 100 s intervals (in addition to the "zeroth" read following the array reset). As part of the initial processing, the calwf3 pipeline performs an "up-the-ramp" analysis of the WFC3/IR pixel count to identify incident cosmic rays that are flagged as jumps in the otherwise (assumed) constant source + background count rates. In practice, however, WFC3/IR exposures frequently suffer time-variable background count rates *within* an exposure (see §4.2) and the magnitude of the variation is frequently large enough to confuse the hard-coded thresholds of the CR rejection algorithm. For such exposures, the pipeline CR rejection algorithm must be turned off and the CRs must be identified by other means, e.g., by comparing multiple dithered exposures of the same field. Even without adversely affecting the CR-rejection, the time-variable backgrounds can compromise the flux calibration and subsequent subtraction of the sky background (§4.2) and therefore the standard up-the-ramp averaging must be disabled for affected exposures.

4.2. Background removal

The backgrounds seen in *HST* grism exposures are dominated by scattered earthshine at optical wavelengths (ACS/WFC G800L) and zodiacal continuum (WFC3 ISR 2014-11) and an atmospheric He line[1] (1.083 μm, see WFC3 ISR 2014-03) in the near infrared (WFC3 G102 and G141). Each pixel in the grism exposure sees dispersed background flux from many wavelengths. In a sense the background is the converse of a dispersed object spectrum: rather than light from a single position on the sky being dispersed to many grism image pixels, a given grism image pixel sees light at different wavelengths coming from different positions on the sky. In the case of ACS/WFC G800L, the two-dimensional structure of the background shows large-scale gradients resulting from vignetting of the detector field of view. In the case of the IR grisms, the dispersed spectra cover a significant fraction of the detector width (~200 pixels = 20% of the field of view), and the overlapping vignetted spectral orders create a complicated 2D

[1] The Helium line can be strong at *HST*'s orbit at 550 km but is expected to be weak or nonexistent at the potential geosync and L2 WFIRST orbits (should be confirmed).

structure across the detector with high spatial frequency features/edges that are not easily modeled analytically.

Empirical master sky images have been created for the ACS/WFC G800L and WFC3/IR grisms to facilitate background subtraction (WFC3 ISR 2011-01) of grism exposures. These images are created by masking object spectra, normalizing to a common scale and averaging many archival science exposures. The master sky images can then be fit to and subtracted from individual exposures, removing the spatial structure resulting from the vignetted spectral orders. In the case of the WFC3/IR grisms, the two main components of the background—the zodiacal continuum and He emission line—have dramatically different spectral shapes that result in distinct two-dimensional structures in the images so they must be considered separately for optimal sky subtraction in the IR. We are preparing master sky images for the WFC3/IR grisms separated into these dominant sky components, which is determined for *each non-destructive read* of thousands of archival exposures (Brammer, ISR in prep). The background of any given exposure is thus a linear sum of these two components that can be fit to source-masked pixels of the exposure with, e.g., a least-squares fit.

As mentioned above, the pipeline up-the-ramp fit of the WFC3/IR exposures produces a linear fit to the source-plus-background count rate for each pixel throughout an exposure. The linear fit clearly assumes that the source and background count rates are constant in time, whereas the time-variable He-line flux can result in highly non-linear ramps (WFC3 ISR 2014-03). Because the 2D structure for the two components differ significantly, the linear ramp fit can result in a final background structure that is no longer a simple linear sum of the two components separately. For this reason, WFC3/IR exposures that suffer a significant time-variable component from the He-line background should be reprocessed with a modified version of the ramp-fitting algorithm, or, e.g., with the up-the-ramp CR-rejection step turned off (CRCORR=OMIT). Additional algorithms are being explored, such as fitting the background at each non-destructive read of an IR exposure or otherwise identifying the time-variable, non-linear component of the IR ramps.

An example of the sky-subtraction process for a WFC3 G141 exposure is shown in Figure 4. The flat-field is another source of structure in the WFC3/IR exposures (multiplicative, as opposed to additive backgrounds) that will be described in more detail below (§**Error! Reference source not found.**). Residuals after subtracting the multi-component master sky images are of the order 0.01 e^-/s/pix, whereas the typical background count rates are of order 1.0 e^-/s/pix. Nevertheless, systematic residuals at the level of a few hundredths of an electron per second are likely the factor limiting the ability to extract faint continuum spectra of, e.g., high-redshift galaxies.

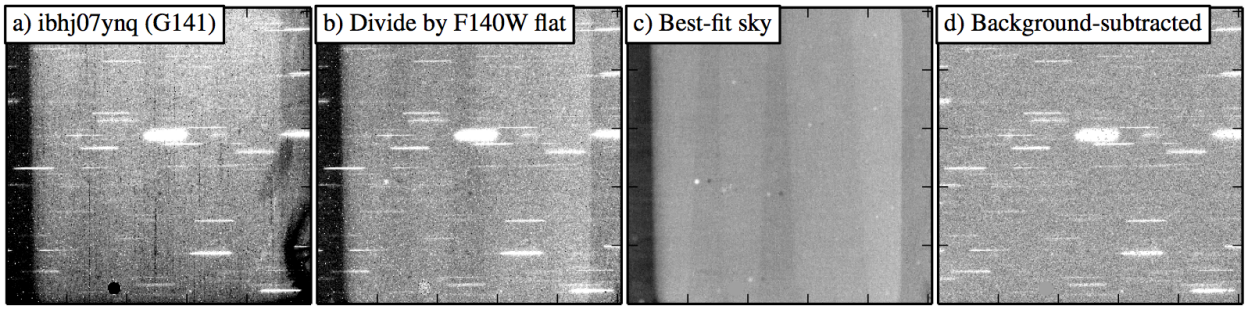

Figure 4: Sky-subtraction of grism exposures with master sky images. Panel b) represents a trivial grey flat-field correction using the imaging flat-field reference file. The best-fit sky image in panel c) is a composite of zodiacal continuum and He 1.083 μm line components; the structure results from overlapping vignetted grism orders.

4.3. Source Identification (pre-imaging)

To define astronomical objects for which to extract slitless spectra, the aXe and 3D-HST pipelines typically work with externally defined reference images and source catalogs derived from them with the SExtractor software (Bertin & Arnouts 1996). The reference images may be the direct images taken concurrently with the grism exposures, or alternatively they could be (deeper) astrometrically aligned images from separate programs (the validity of using external images as a grism reference would require negligible proper motions of target objects between the observations epochs). The reduced spectra are usually accompanied by cutouts of the objects from the direct imaging reference, as it is the non-trivial morphologies (in the case of resolved objects) that effectively define the line-spread function in the 2D grism spectra. Image cutouts themselves are often preferable to analytical (e.g., Gaussian) descriptions of the object morphologies as the 2D spectra share the diffraction-limited image quality of the broad-band imaging filters.

The SExtractor software provides a catalog of parameters for each object, such as size and brightness, as well as an image of "segmentation" regions that indicate which *pixels* in the reference image are associated with a given object. The software provides some capability of deblending neighboring sources, but the format of the segmentation images used by the grism pipelines assigns single objects to a given pixel assuming no overlap. The *HST* slitless analysis pipelines use the segmentation polygons to define which pixels are dispersed together as a single object. §6.7 below describes techniques for identifying emission-line objects from observations taken at multiple roll angles largely independent of the catalog + segmentation requirements described here.

4.4. Astrometric alignment

The calibration of the *HST* slitless spectroscopic traces is defined with respect to positions of objects in "direct" images through normal imaging filters[2] (see §4.6 and §4.7). In a *HST* grism visit, it is current policy to require that programs obtain at least one pair of direct + grism exposures without offsetting the telescope between them. The direct images can be aligned to astrometric reference images and catalogs using software such as the TweakReg function of the DrizzlePac software package. The astrometric alignment of the direct image can then be propagated to the paired grism exposures.

Most programs obtain at least one pair of direct + grism exposures per orbit of a multi-orbit visit and many programs obtain pairs of direct + grism exposures at all dither positions within a visit (e.g., §5.5). If the direct images are not needed for science analysis (i.e., if deep images in the same filters are available from a different program), the stability of *HST* pointing and dithering is usually sufficiently precise (2–20 mas for small dithers) so that only a single pair is required.

Alignment of the grism exposures alone using compact features such as the zeroth order spectra may be possible, but has not been explored in great detail for the *HST* grisms.

4.5. Geometric Rectification

Geometric rectification is handled in two distinct cases: when creating the direct image catalog and when extracting spectra. When creating a source catalog, multiple distorted broadband images are combined to produce deeper, higher signal to noise images. During that process, the images are rectified/undistorted and proper celestial coordinates for each source are computed. When extracting spectra, and since spectral extraction takes place in the distorted frame, the R.A./Dec. coordinates of the master source catalog must be used with knowledge of the field distortion to compute accurate pixel coordinates for each source and for each of the available grism images.

4.6. The Spectral Trace

In practice the spectral trace of the primary dispersion order is never perfectly aligned with either the rows or the columns of the detector. Furthermore, the slope of the spectral trace, measured with respect to the row direction on the detector can vary over the field; this is the case for the NICMOS, ACS and now WFC3 grisms. In the cases of ACS and WFC3, the variation of the relative position of the trace and its angle are parameterized in such a way that they can be computed as a function of position on the detector. The calibration parameterization described below was defined for the aXe software package (§5.3); to make

[2] Due to small astrometric shifts between imaging filters, the trace calibration was performed and is therefore most appropriate for the following direct/grism pairs: ACS F814W/G800L, WFC3 F105W/G102, and WFC3 F140W/G141.

use of the existing calibrations, the parameterization has been adopted by other more recent implementations of *HST* slitless spectroscopy analysis software (§5.5 and §6.3).

The relative y-displacement, dx, with respect to a fiducial reference point (taken to be the position of the source in the undispersed image; Figure 5) dx is a polynomial of the form:

$$dy(dx) = t_0(i,j) + t_1(i,j)dx + t_2(i,j)dx^2 + \cdots + t_n(i,j)dx^n,$$

where the individual coefficients t_k are themselves 2D polynomials evaluated at pixel (i,j) in the direct image in detector coordinates:

$$t_k(i,j) = u_0 + u_1 i + u_2 j + u_3 ij + u_4 i^2 + u_5 j^2 + \ldots.$$

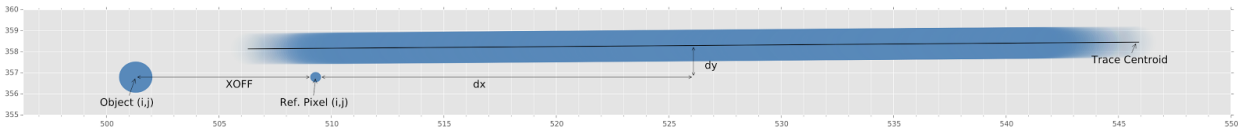

Figure 5: Demonstration of the aXe trace calibration, which is defined as an offset of dy as a function of the distance dx, along pixel rows, from a reference position in the direct image.

4.7. Wavelength Calibration

The aXe wavelength calibration, defined at a distance dp *along* the spectral trace (i.e. in the reference frame of the rectified trace described in §4.6; Figure 6), is also parameterized by a polynomial

$$\lambda(dp) = l_0(i,j) + l_1(i,j)\, dp + \cdots + l_n(i,j)\, dp^n,$$

where, again, the coefficients l_k can vary across the detector field of view:

$$l_k(i,j) = m_0 + m_1 i + m_2 j + m_3 ij + m_4 i^2 + m_5 j^2 + \ldots.$$

The parameterization of the wavelength calibration along the spectral trace was a design decision made for aXe but that is not necessarily optimal for future systems. Specifically, computing dp requires an integration along the trace that is not analytical and is potentially computationally intensive for nonlinear functions of dy/dx along the trace. In principle, any variable that is single-valued along the trace can be used, defined in such a way to be more computationally efficient.

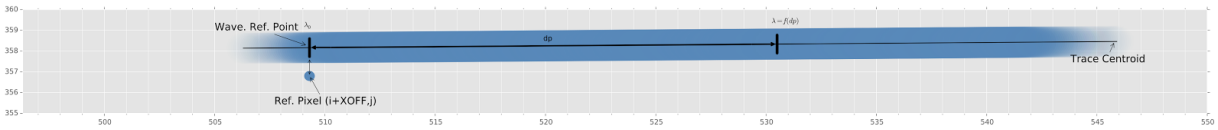

Figure 6: Demonstration of the aXe wavelength solution, which is defined as a wavelength, λ, as a function of the position along the spectral trace.

4.8. Field-dependent calibration

Current software such as aXe relies on a polynomial parameterization of the field dependence of the trace and wavelength calibrations across the field of view (see the coefficients $t_k(i,j)$ and $l_k(i,j)$ above). Polynomials may not be optimal for describing the field dependence of future missions such as WFIRST. In particular, a piecewise/local approach could provide an improved description of the much larger field of view of WFIRST. This could divide the larger field of view in smaller sections (e.g., each of 18 detectors) that can be calibrated independently to a much higher accuracy than if single (high-order) polynomials are required to define the dependence over the entire field of view.

4.9. Flat-field Calibration

Flat fielding of grism data must be performed before information from different detector pixels corresponding to the same wavelength are combined. Each pixel (i,j) in the spectrum has an assigned wavelength and a given signal in e^-/s. The value of each pixel is flat fielded by dividing these pixels by a wavelength dependent flat-field of the form $F(i,j,\lambda)$. In the case of NICMOS, ACS and WFC3, the function $F(i,j,\lambda)$ is derived from a linear combination of available narrow and broad band filters with

$$F(i,j,x) = a_0 + a_1 x + a_2 x^2 + \cdots,$$

where the coefficients a_k are 2D matrices defined at each pixel position (i,j) on the detector. The variable $x(\lambda)$ is normalized with parameters λ_0 and λ_1 defined for each grism and

$$x = (\lambda - \lambda_0)/(\lambda_1 - \lambda_0).$$

Note that a given pixel can have different flat-field values depending on the sources that contribute to it (i.e., sky and neighboring sources) at effective wavelengths λ determined as in §4.7.

4.10. Creating 1D Spectra

In aXe, 1D spectra are generated from the 2D distribution of pixels belonging to a given spectrum by collapsing/summing pixels in the cross dispersion of the trace (Figure 7). Without slits, the user must specify an ``effective slit'' that defines the cross dispersion axis; implemented definitions for the effective slit include the source major axis or simply the axis perpendicular to the spectral trace (both of these definitions would be roughly equivalent for the example shown in Figure 7.

While constant weights can be added, it is sometimes better to apply different weights to different pixels, akin to an “optimal extraction” (Horne 1986), using the known profile of the object in the cross dispersion direction to create a set of weights. This way, pixels that are closer to the trace are more heavily weighted than pixels farther away from the trace when they are added together to create a 1D spectrum.

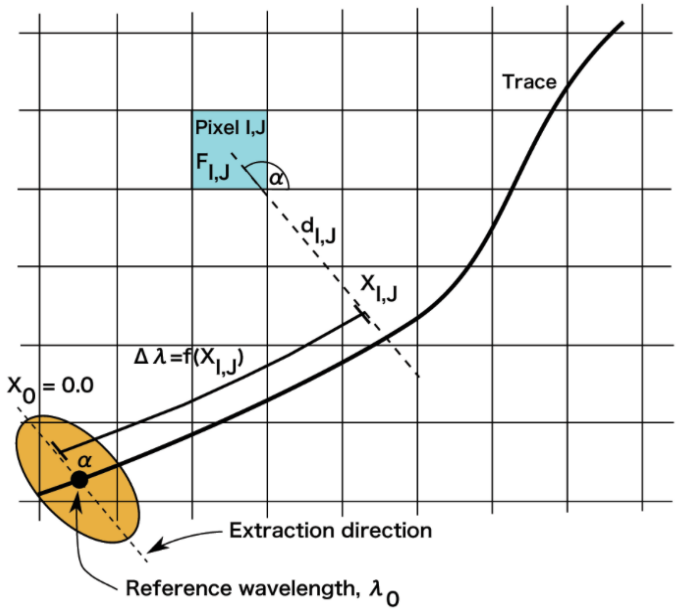

Figure 7: Conventions for extracting a 1D spectrum along a user-defined "extraction direction", which may be, but isn't required to be, perpendicular to the spectral trace.

As an alternative to this approach, aXe version 1.4 implemented "aXeDrizzle," which remaps the original grism data onto a rectified grid of wavelength versus distance from the trace in the cross-dispersion direction. This remapping is done for each available observation (taken at the same PA). The resulting rectified 2D grism spectra are then combined pixel-to-pixel, allowing for cosmic ray and bad pixel rejection. Creating a 1D spectrum from the stacked 2D grism stamp is then performed by simply summing up the data in the cross dispersion direction (which is now aligned along columns). 1D spectra can be extracted with "optimal weighting" with weights determined from the cross-dispersion profile of the source measured from the available direct imaging.

4.11. Flux Calibration

The flux calibration is performed once all the pixels of a spectrum with the same wavelength bin have been combined (either simply added, or optimally weighted using varying weights in the cross dispersion direction). In aXe flux calibration is done by dividing the uncalibrated spectrum that is in units of e^-/s/pixel by a response function which has units of erg/s/cm^2/Å –per– e^-/s/pix, accounting for the width of each bin $\Delta\lambda$.

5. Existing Software Systems

5.1. Historical systems

Several wide field extraction software systems have been designed over the years. As the NICMOS instrument provided IR grism slitless spectroscopy over a moderate field of view, ESA ST-ECF took on the task of calibrating these modes and of designing and implementing

software to extract fully calibrated 1D spectra from the 2D *HST* NICMOS data. The "NICMOSLook" software was an interactive IDL package that allowed a user to mark an object's direct image and would then identify, extract and sum up the appropriate pixels in the associated grism image to produce 1D spectra that were both flux and wavelength calibrated (Figure 8). NICMOSLook was interactive and designed specifically for NICMOS and its moderate field of view, allowing for a few hundred spectra at most to be manually extracted. A pipeline version of the code CALNICC, based on the IDL code of NICMOSLook, was also created. While meant to be a "pipeline-able" addition of the NICMOS pipeline tasks CALNICA and CALNICB, its use was limited and it was never added to the *HST* NICMOS data pipeline.

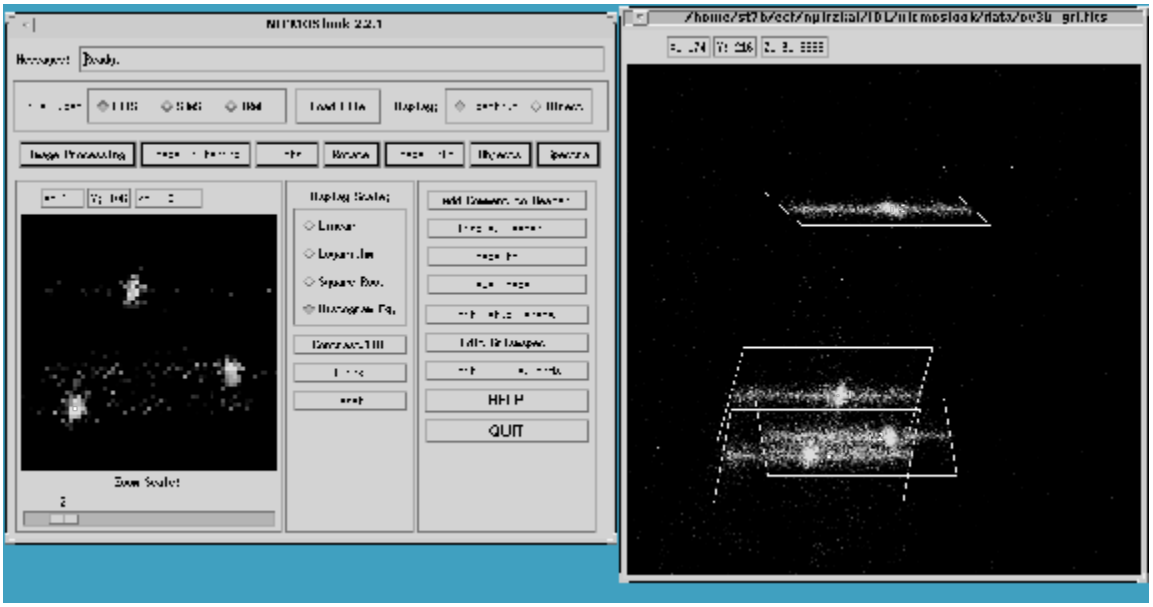

Figure 8: Screenshot of the NICMOSLook interface.

5.2. ESA ST-ECF aXe system

The installation of ACS on *HST* brought a larger set of detectors with slitless spectroscopic capabilities. ST-ECF designed a new extraction package, which, while based loosely on the capabilities of NICMOSLook, differed in its approach as it was non-interactive and able to deal with several thousands of spectra at once.

The ESA ST-ECF aXe was originally designed in 2000 to be a modular system where the process of extracting ACS slitless grism spectra was split into well-defined and independent steps that could then be scripted into a pipeline like operation (Figure 9). Version 1.2 of aXe dealt with preparation of extraction catalogs, background subtraction, extraction of grism pixels for each spectrum, flat fielding, 1D spectra generation, and finally conversion into fully flux calibrated 1D spectra using C-based programs. The interface of this task was a configuration file that described the instrument dispersion properties such as the trace and wavelength calibration (§4.6 and 4.7). Input object catalogs defined which objects to extract and separate tasks produced intermediate products such as Aperture Catalogs, and Pixel Extraction Tables

that subsequent tasks read and further processed. This system was conceived to allow for flexibility and possible changes to tasks such as, e.g., background subtraction.

Version 1.4 of aXe added an interface into Pyraf by wrapping up the original aXe C executables inside Python so that the required number of steps could be reduced in order to extract ACS spectra using pre-determined options. This reduced the flexibility to the user but simplified the use of aXe. Additionally, aXe added the aXeDrizzle option that allowed the combination of multiple spectra taken during the same *HST* visit. This introduced a Drizzle like process that removes cosmic rays and detector defects in data obtained using small dither offsets. Also, the later versions of aXe enable more quantitative modeling of spectral contamination by allowing users to use multiple band direct images of all the sources in the field. With this it obtains realistic pre-estimates of the dispersed data in terms of the spectral slope and object morphology, which could be used to generate a quantitative estimate of the level of contamination/overlap between spectra of neighboring sources.

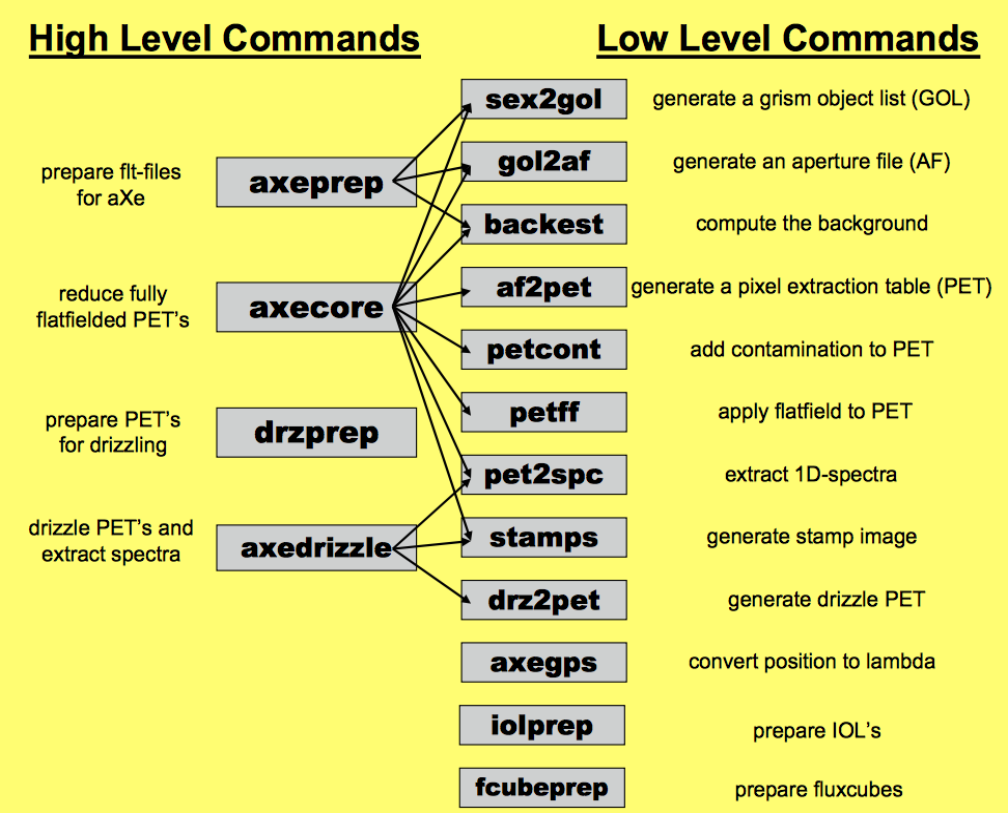

Figure 9: Flowchart of the high- and low-level aXe commands. For more details on the commands themselves, see the aXe User Manual (Kümmel et al. 2011; http://axe.stsci.edu/axe/manual/html/).

5.3. STScI aXe

The aXe package was delivered to STScI/OED in 2011. It has since been maintained at STScI by OED with the help of the WFC3 Grism group. No new features have been added to the aXe package but the software was updated to remove its dependence on Multidrizzle (both for input data products and for internal computation of coordinates). With the accumulation of science

and calibration data, the WFC3 team has updated the WFC3 grism configuration files periodically to improve the spectral extraction and calibration of WFC3 grism data. Ongoing monitoring programs indicate no significant evolution of the WFC3 grism sensitivity or geometric calibration (e.g., the trace polynomials) since the installation of the instrument onboard *HST* in 2009.

5.4. User enhancements to aXe (GRAPES, PEARS, WISPS)

Science teams exploiting ACS and WFC3 slitless spectroscopy data have made a variety of extensions beyond the core aXe capabilities.

GRAPES and PEARS: The GRAPES ("Grism-ACS Program for Extragalactic Science"; GO 9793; PI: Malhotra) and PEARS ("Probing Evolution and Reionization Spectroscopically"; GO 10530; PI: Malhotra) projects were the two largest programs using slitless spectroscopy on *HST* with the ACS instrument at optical wavelengths. The GRAPES projects relied on the ST-ECF aXe 1.2 software and ACS calibration files. Additional processing was performed by the team using Python scripts to remove the background in ACS grism images using a Super Sky frame. This bypassed the built-in background subtraction task of aXe 1.2 and reduced the systematic errors in the background subtraction to better than 0.05 e^-/s. PEARS also used the ST-ECF aXe ACS calibration products but switched to the newer aXe 1.4 version and added the use of aXeDrizzle to produce both 1D and 2D spectra. As for GRAPES, new techniques were developed to properly subtract the background light from the PEARS images and a new background subtraction task was added into the PEARS pipeline, adding a 2D fitted model of the general distribution of the background light in the ACS grism images in addition to the Super Sky subtraction. This step reduced the residuals by a factor of 2, greatly improving the quality of the spectra of faint sources.

WISPS (WFC3 Infrared Spectroscopic Parallel Survey): The WISPS programs (PI: Malkan) obtained WFC3/IR G102 and G141 exposures *in parallel* with primary COS and STIS spectroscopic observations. While the WISPS team used the aXe calibration products and configuration files, as well as the aXe software, the team had to add additional steps in the process. The main modification was a way to deal with the impact of cosmic rays and hot pixels in the WFC3 data as their data consisted mainly of undithered short exposures. These modifications were IDL scripts used in addition to aXe version 1.2, and later version 1.4 tasks. The process is fully outlined by Atek et al. (2010).

5.5. 3D-HST

The "3D-HST" survey is a 248-orbit Cycle 21 Treasury program that covered roughly two-thirds of the CANDELS *HST* imaging area with slitless spectroscopy from the WFC3 G141 grism and the ACS/WFC G800L grism in parallel (GO 12177&12328, PI: van Dokkum; Brammer et al. 2012). The goal of the 3D-HST program was to provide uniform, unbiased redshift measurements for the galaxies detected in the CANDELS imaging and photometry, and this goal requires more dedicated pipeline processing of every object in the field, as opposed to a procedure relying on visual inspection of spectra and subjective selection strong line-emitters

"by eye". The full 3D-HST survey covers 625 arcmin^2: 60% of a single pointing of the WFIRST-AFTA wide-field camera as described in the final SDT report (Spergel et al. 2015).

The relatively narrow wavelength coverage of the G141 grism (1.1–1.7 μm) typically provides only a single emission line species (if any) that can be difficult to identify and interpret from the spectrum alone. The rich ground- and space-based supporting datasets in the CANDELS fields can be used to help to, first, constrain line identifications based on photometric redshift fits to the broad-band spectral energy distributions. Second, the ancillary measurements can aid the detection and interpretation of lower equivalent width lines that are somewhat more difficult to see from visual inspection but that are found at wavelengths and spatial line morphologies consistent with the photometric redshift constraints and deep imaging data (Figure 10). For this reason, the 3D-HST team devoted considerable effort to constructing *photometric* catalogs of the fields covered by the grisms to aid the interpretation of the slitless spectra (see Skelton et al. 2014). The photometric redshift precision for these catalogs reaches ~1% for fields covered with up to 40 photometric bands and ~3% for fields with more sparsely sampled photometric SEDs. These photometric catalogs therefore provide priors both on where objects physically lie in the direct images and also their broad-band colors.

While the aXe software is suitable for modeling a grism image and extracting individual spectra of relatively isolated objects, the 3D-HST survey goals require more direct interaction with the two-dimensional slitless spectra. Specifically, the extracted one-dimensional spectra of well-resolved and/or asymmetric objects often contain spurious features that result from the convolution of the object morphology with the grism sensitivity function, which contain sharp edges in the case of the two WFC3/IR grisms. Fitting spectra and determining redshifts in the presence of these systematic features proved difficult. Therefore, 3D-HST opted to perform (e.g., redshift) fits to the full two-dimensional spectra themselves, and this requires a mechanism to quickly generate synthetic two-dimensional model spectra from arbitrary input spectral templates. These fits, along with the data reduction and spectral extraction, are performed with a custom pipeline independent of aXe. While the basic product of the pipelines are similar—2D model spectra based on the direct images and "dispersed" using the grism configuration polynomials (§4.6)—the 3D-HST pipeline offers much more flexibility for (quickly) generating model spectra well-suited for iterative fitting algorithms.

Each of 124[3] separate $2' \times 2'$ WFC3/IR pointings of 3D-HST (adjacent but with minimal overlap) was covered in a two orbit visit divided between four dithered paired exposures in the F140W direct filter (t_{exp}~800 s) and the G141 grism (t_{exp}~4700 s). Another way that the 3D-HST pipeline differs from aXe is that these exposures are combined by *interlacing* the half-pixel dither positions in the distorted frame rather than *drizzling* them to a rectified frame. A drizzled output pixel typically contains input from multiple input pixels, and, alternatively, a given input pixel is drizzled to multiple (neighboring) output pixels. Therefore, the noise in the drizzled spectra can

[3] An additional 28 pointings from GO-11600 (PI: Weiner) cover the GOODS-N field with a similar observing strategy.

be correlated between pixels and compact correlated noise features can easily be mistaken for emission lines given the relatively coarse pixel sampling of the WFC3 PSF. In the case of interlacing, however, the template fits to the 2D spectra are therefore fits to the original pixels themselves, along with their relatively well-behaved instrumental noise model, and the resulting fits to weak emission lines are more robust.

The 3D-HST team recently (2014 November) extracted spectra for all objects in the survey fields down to a magnitude H_{F140W}< 24 (AB) (Momcheva et al., in prep). Redshifts and emission line strengths are determined automatically for the extracted spectra (Figure 10), along with the information in the photometric catalogs. Among some 20,000 galaxies, a comparison of roughly 2,000 galaxies with ground-based spectroscopic redshift measurements suggest a G141 redshift precision of 0.3%, a factor of 3–10 improvement over the photometric redshift measurements alone. While the precision depends somewhat on the line equivalent width and object morphology (with compact morphologies resulting in more localized emission lines), precise redshifts can be measured for galaxies with weak emission lines or even from *continuum* features alone (see also van Dokkum & Brammer 2010; Whitaker et al. 2013). Again, these represent automatic extractions of the survey field with visual inspection performed simply to rule out 5–10% of the objects with the worst data quality failures (see also *§6.1*).

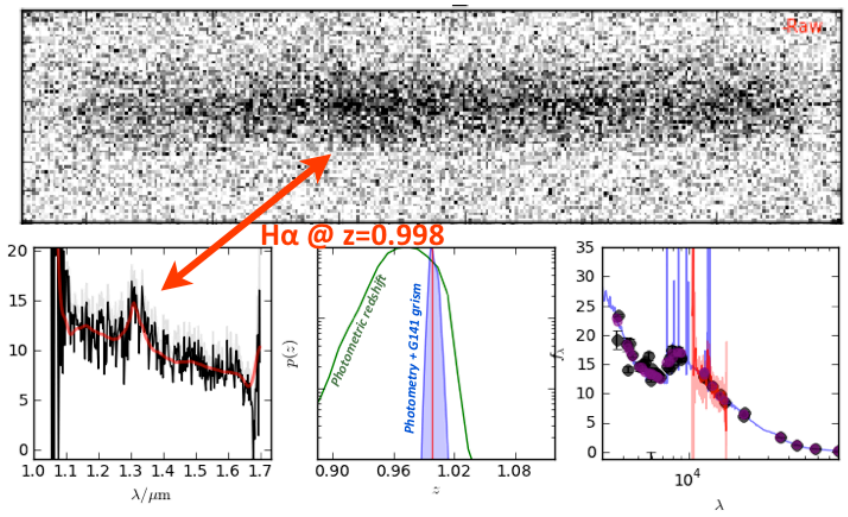

Figure 10: Example of a spectral extraction and fit from the 3D-HST survey. The galaxy shown has an Hα emission line at *z*=0.998, easily visible in the 1D extraction but more difficult to see in the 2D spectrum. The red curve in the lower-right panel shows a template fit to the spectrum, indicating that the line morphology modeled from the spatial information in the direct image closely follows the observed line shape. The redshift constraints from the combined fit to the spectrum + photometry are in excellent agreement (bottom-middle panel) with the ground-based spectroscopic measurement shown in the vertical red line. Note that the 1D spectrum and model are shown for display only; the template fits are performed on the full 2D spectrum shown in the top panel.

5.6. Special Cases

The WFC3 grisms have been used for certain observations not suitable for analysis with the spectral extraction procedures outlined above (e.g., aXe). These observations include use-

cases that may also be relevant to WFIRST-AFTA: time-resolved exoplanet transit observations and observations of extremely bright sources. Observations of exoplanet transits and eclipses have been very successful with *HST* (e.g., Stevenson et al. 2014); the salient property of these observations is their differential nature. The most extreme implementation has been the intentional motion of the telescope during the observation in the cross-dispersion direction (i.e. "spatial scanning") to maximize photon collection efficiency and to average over limitations of the detector calibration (flat-field, bad pixels, etc.). While some elements of the general-purpose systems have been used (e.g., the wavelength calibration), such observations have been analyzed using purpose-built software optimized for the differential experiment.

A second example of non-standard observations is exposures of extremely bright stars: STScI has obtained spectra of Vega for use as an absolute flux standard. These observations are only possible by taking advantage of the factor of ~100 decreased throughput in the -1st order because the primary +1st order saturates in all of the available WFC3/IR detector readout sequences. The -1st order can be modeled with the existing general-purpose software, but it isn't extracted for analysis.

Both of these examples illustrate observational cases not readily handled by the standard software analysis tools and raise questions as to the desired scope for such software. Another (not yet explored) analysis problem would be to devise means of dealing with highly extended sources (e.g. a face-on galaxy covering most of the detector field of view).

5.7. How is Information Embedded/Modeled

Coming from the instruments, *HST* grism exposures are FITS files similar to standard imaging exposures but using the grism spectral element. The FITS headers contain many header keywords pertaining to the exposure, among them the spectral element, read-out sequence, total exposure time, nominal WCS, etc. In practice (and by current policy), *HST* grism exposures are accompanied in the same visit by direct images in imaging filters that are used for astrometric alignment to reference images and catalogs (see §4.4). The astrometric alignment of the direct images can be propagated to the WCS headers of the grism exposures.

The aXe and 3D-HST spectral extraction and modeling pipelines use static configuration files for each instrument (+chip) + grism element combination to define how the multiple beams, or spectral orders, are dispersed by the grism element. The configuration files define the grism trace and wavelength calibration as polynomials whose coefficients can vary in both the *x* and *y* dimensions across the instrumental field of view (§4.6–**Error! Reference source not found.**). The polynomials are defined in the nominal distorted frame of the *HST* science instruments (i.e., FLT images produced by the ACS and WFC3 calibration pipelines).

The flux calibration of the spectral orders is defined in separate sensitivity files with the sensitivity and its corresponding uncertainty provided as a function of wavelength. Variation of the sensitivity across the field of view of the *HST* instruments is incorporated into the grism-specific flat-field calibration files (§**Error! Reference source not found.**).

In the reduction process of extracting 2D and 1D spectra from the instrumental images, relevant parameters from the exposure headers and source catalogs are propagated to the FITS headers of the extracted spectra. Ideally these keywords would provide sufficient geometric information for mapping directly between the frames of the original instrumental exposures and the reduced extracted spectra. In practice this is achieved to only a modest degree of success, in part, due to limitations imposed by the polynomial description of the slitless spectral traces.

5.8. Use of existing datasets/catalogs/archive

Many of the past and current users of *HST*-based grism spectroscopy have relied on existing archival data. The use of existing deep multi-band broad band images makes it possible to create deep catalogs of objects to extract before grism observations are even taken. This also allows for highly realistic simulation of a given astronomical field, which is important if one wants to understand the effect of source crowding/overlapping in planned observations. Creating these deep mosaics, usually done with Multidrizzle/Astrodrizzle also requires that these data be easily identified, which is usually done with the *HST* MAST Archive.

6. Future Systems

6.1. Pipeline versus user tools

The large-scale slitless spectroscopy surveys with *HST*, the WFC3 3D-HST and WISPS programs combined, comprise fewer IR pixels (and sources) than will a *single* exposure from WFIRST-AFTA. The ~24,000 sources with good signal-to-noise in 3D-HST ($H_{AB} < 24$) compelled that team to develop a pipeline solution (§5.5) for modeling/extracting source spectra from the pipeline-reduced WFC3 grism exposures delivered by STScI. In this subsection, we explore the pros and cons of a pipeline solution versus a user-tools solution for WFIRST-AFTA slitless spectroscopy.

Currently, the slitless spectroscopy modes of *HST* instruments (STIS, ACS, and WFC3) are pipeline-processed at STScI only through the initial frame reductions (§4.1). Further analysis of these exposures to extract object-specific information (1D or 2D spectra) is not done via a pipeline, but is instead supported by STScI via user tools (aXe; §5.3). Much of the rationale for STScI opting to support a user-tool solution for slitless spectroscopy is relevant to the case of WFIRST-AFTA. Conversely, if a pipeline solution is ultimately deemed the best solution for WFIRST-AFTA slitless spectroscopy, the obstacles that STScI has encountered in that regard will need to be overcome.

One advantage of a user-tool solution is to maximize flexibility of reductions and extractions for "corner cases". By placing the software in the hands of the observer and allowing a large degree of latitude in configuration tweaking, the user-tool solution allows observers to best perform extractions matched to their particular needs and their particular observing setup. For example, a pipeline optimized for total-source spectral extraction would poorly serve observers seeking to investigate galaxies' core emission, or independent spectra of resolved galaxy substructures, etc.

A corollary benefit of the user-tool solution is allowing the best characterization of the limitations on extracted spectra. Not unlike the case of source extraction in direct imaging, the accurate estimation of detection limits, contamination handling, etc., can be difficult if not impossible to obtain via a generic pipeline extraction. The proper estimation of these uncertainties typically involves modeling of the data by the investigator, which lends itself much more to a user-tool solution than an external pipeline. Along with the post-observation analysis and characterization, these same tools can also be used to aid observation planning, for example, enabling the investigator to determine appropriate roll angles that would minimize contamination for specific targets of interest.

A potential further advantage of the user-tools solution is the multiplexing of a large user base with distributed computing resources. In the event that the WFIRST-AFTA Guest Observer slitless spectroscopy program comprises many separate, modestly sized investigations by many independent groups worldwide, their computational needs may well be served by a user-tool solution on local computing resources. This paradigm also would likely expand the time domain for the required computational resources, as contrasted with a centralized pipeline solution where the user community would expect a rapid turnaround of extractions on each successive newly-completed dataset.

A key disadvantage of the user-tools model is the comparably high cost in human capital to amass the necessary expertise in spectral extractions from slitless spectroscopy. This disadvantage is certainly a contributing factor in the relatively modest exploitation of *HST* slitless spectroscopy assets. There is the danger that this high bar of expertise may also tend to reduce diversity among slitless spectroscopy programs accepted by a TAC, which may favor proposals from the small number of research teams who have already built this expertise via prior approved programs. In this regard, the pipeline paradigm is perhaps more egalitarian to the user community by centralizing that expertise and making it available to all proposers.

Another major disadvantage is the greater difficulty in creating a uniform archive of extracted spectra. The ESA ST-ECF undertook to extract most of the pre-SM4 ACS G800L spectra using aXe with manual support. This resulted in a catalog of approximately 48,000 spectra now available in the Hubble Legacy Archive (see Kummel et al. 2009).

The vast data volumes of the High Latitude Survey (HLS) will almost certainly require an unsupervised pipeline approach (§7) as opposed to the “user-tools”. Balancing the needs of the core mission surveys and the GO program, it seems likely that WFIRST-AFTA will be required to develop both approaches. It is worth considering to what extent a pipeline approach developed specifically for the High Latitude Survey’s science objectives would support other large projects.

6.2. The Next Generation of Algorithms

Here we describe two related algorithms that we have begun testing for extracting spectral information from a set of dispersed images. An important aspect of both approaches is that they consider the full context of the two-dimensional grism exposures, as opposed to “reduction” to one-dimensional extractions of individual objects more typical of the aXe products and long-slit spectral analysis. This allows for a more natural accounting for irregular object morphologies, contamination, and incorporation of exposures taken at multiple position angles. The algorithms in preparations share a set of assumptions, the most fundamental of which is that a given collection of pixels shares a common spectrum. This "separability assumption" implies that the three-dimensional flux cube can be written as:

$$F(x, y, \lambda) \; = \; p(x, y)\, f(\lambda)$$

where $p(x, y)$is the two-dimensional profile of the source (i.e., a direct image, more details below) and $f(\lambda)$ is the spectrum of interest. At present, these assumptions are meant to guide the first phase of development and implementation benchmarks, and may be relaxed in the future. Below we further discuss the importance of the separability assumption and the potential limitations of relaxing it.

6.3. Linear Reconstruction

As a general rule, we seek to extract the optimal spectrum for a collection of pixels that are assumed to represent a single astrophysical source. In the limiting case of a single position angle (or orient in the case of *HST*), this requires careful understanding of the possible sources of contamination that may disperse onto the spectrum of interest. However in a more data-rich environment where multiple (perhaps many) position angles are available, then it is possible to more precisely determine the amount of contamination since the spectra of a set of objects may only overlap under certain orientations. We formulate the extraction of multiple-orient data by noting that the flux in a given dispersed-image pixel is the sum of all sources (at various wavelengths), weighted by their projected area onto the pixel, instrumental response, and spatial profiles. Of course many sources have a small angular sizes (compared to the detector), and so typically only a few sources will contribute to a given dispersed-image pixel.

This situation sets up a system of many linear equations (that is, $N_{pix}{\times}N_{src}{\times}N_{orient}$ equations, where N_{pix} is the number of pixels in a dispersed image, N_{src} is the number of sources in the scene, and N_{orient} is the number of realizations of the scene—whether in the form of multiple orientations or dithers). For typical extragalactic fields, this system of linear equations is very sparse and therefore specialized algorithms can expedite the inversion of the system.

6.4. Forward Modeling

The Linear Reconstruction paradigm described above is designed to optimally extract the information from the dispersed images. By contrast, the Forward Modeling approach is motivated to find the best model description of the data. As with the Linear Reconstruction, here we make the usual assumption that a collection of pixels share a common spectrum, but

now we characterize the collection of pixels as a set of model parameters. By comparing the prediction of the dispersed image(s) to the data, we can determine which spectra can best reproduce the observations. The model spectra can be described in a host of ways, such as stellar population synthesis models, empirical libraries, or even piecewise functions (essentially a nonparametric spectrum).

6.5. Comparisons of Methods

We argue that these methods are not incompatible, but rather complement each other with a unique set of strengths and weaknesses. In Table 3, we briefly describe some of the advantages and limitations of these schemes. As the algorithms for reduction and extraction of slitless spectroscopy continue to mature, we expect hybrid approaches may provide the best results, such as using the results from the Linear Reconstruction as initial conditions for the Forward Modeling optimizers (such as a Markov chain).

Table 3: Comparison of Next-Generation of Algorithms

Method	Strengths	Weaknesses
Linear Reconstruction	1. optimal use of the spectroscopic data 2. pure, data-driven problem with solution from "first principles" 3. naturally include sky signal	1. presently assumes a great deal is known to high degree of precision (e.g. spacecraft pointing) 2. potentially challenging for objects that cover a large fraction of the detector (e.g. large nebulosities or extended galaxies)
Forward Modeling	1. include and marginalize over known unknowns (e.g. astrometry) 2. many science aims include modeling of 1D spectra, as a direct path to science 3. naturally introduce additional information (e.g. broadband data) and/or astrophysical priors	1. deeply model dependent 2. minimization techniques (e.g. Markov chains) are likely very slow and scale in unclear ways 3. other numerical algorithm issues (e.g. initial conditions, convergence, monitoring/training, or uncertainty analysis)

6.6. Options in the Absence of Direct Imaging

To be clear, we refer to direct imaging as data that exists in advance and largely independent of the slitless spectroscopy. In many cases, such data motivated the spectroscopic campaign (such as 3D-HST opting to point at CANDELS fields) and are seen as value-added in terms of science productivity. On the other hand, we refer to pre-imaging as data taken concurrent to the slitless spectroscopy with the primary purpose of calibrating the spectroscopy—most often by establishing the astrometric reference. With these distinctions in mind, pre-imaging is often a relatively inexpensive cost in observing time (compared to direct imaging), since one simply needs to identify a few bright, well-localized sources to refine the pointing of the spacecraft.

Whereas direct imaging is often as deep as required to study some interesting sources, such as emission-line galaxies. Consider a hypothetical survey where no deep, direct imaging exists however shallow pre-imaging is taken for calibrations. It is entirely plausible that a strong emission-line source with weak continuum can go undetected in the pre-imaging but leave a strong signal in the deeper dispersed data (such as the so-called naked emission-line sources). Of course a single dispersed image is unlikely to be sufficient for a unique extraction or modeling of such a source (since one cannot uniquely know the wavelength zeropoint without knowing the intrinsic position of the source), however with multiple position angles it is possible to triangulate and determine the intrinsic position of the source and wavelengths for the spectrum (e.g. Straughn et al. 2008, Pirzkal et al. 2012). As it pertains to the two approaches described above, the absence of direct imaging requires relaxing the separability assumption described above and solving for the entire flux cube simultaneously. This introduces additional complications and unknowns, and our team is currently evaluating options for altering these new algorithms.

6.7. Spectra without initial images for source identification

While spectral extraction requires a-priori knowledge of the source content of a field observed using a grism, there is at least one alternative in the case of emission line targets. This may be situation for cases where either 1) no direct imaging is available at all or 2) the available direct image is too shallow to identify sources whose spectrum appears in the grism exposure, be it in the continuum and/or emission line(s). Recovering this information requires that grism exposures be obtained at multiple position angles (which will be the case for the baseline plan of the WFIRST-AFTA HLS; Spergel et al. 2015).

Emission lines can be identified directly in the distorted 2D ACS or WFC3 exposures, which is usually done by first smoothing the dispersed image and subtracting the smoothed image from the original image (similar to unsharp masking or high-pass filtering). A program like SExtractor (Bertin & Arnouts 1996) is then used to look for point-like feature in the residual 2D grism image. When an emission line is identified, and with knowledge of the spectral trace of the grism (§4.6), one can compute the coordinates of the spectral trace passing though the emission line. These points are then converted to RA and DEC on the sky, essentially projecting the trace passing through the emission line back onto the sky. With multiple PAs, the traces will project differently onto the sky; the point on the sky where the projected spectral traces intersect is the source of the emission line and the RA and DEC of the inferred source of the emission can be determined (Figure 11). Separately in each PA the wavelength of the emission line can be determined by computing the distance of the emission line to the source on the detector and a consistency check can be made at this point by checking that the same wavelength is derived for the same line at each of the available PAs.

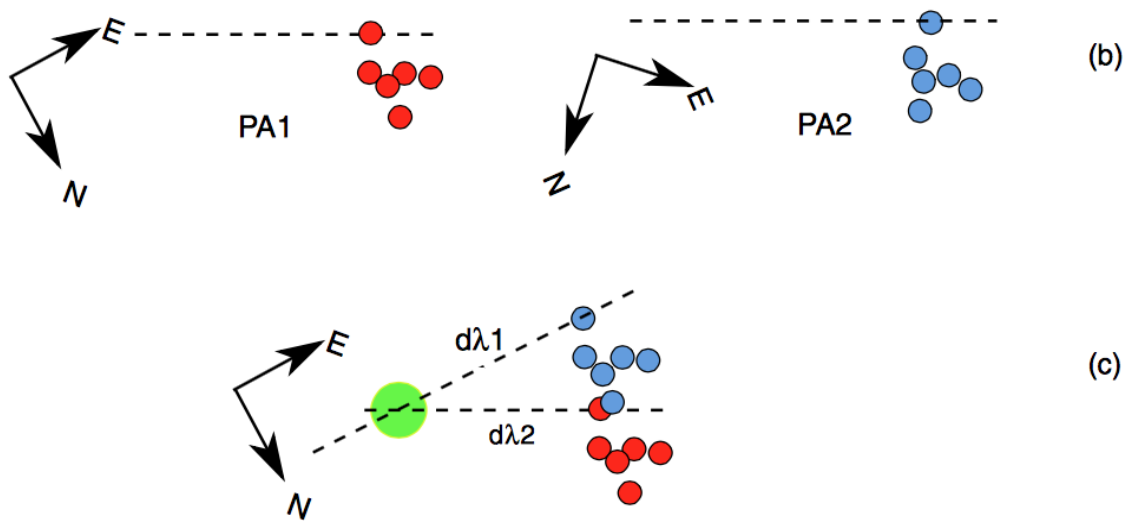

Figure 11: Triangulation method for identifying faint emission line galaxies without prior knowledge of where galaxies were in the field. This method developed by the PEARS team (Pirzkal 2012) enables the detection of "naked" high equivalent width emission lines where the host galaxies where too faint to be detected or when the emission line was generated far on the outskirt of an extended galaxy.

7. Scaling to very large datasets

The prospect of WFIRST-AFTA scale datasets poses multiple issues beyond those raised by *HST* or *JWST* observations. The most obvious difference is that an approach requiring significant human interaction is simply not feasible: WFIRST-AFTA will produce more than three orders of magnitude more slitless spectroscopy data than *HST* or *JWST* and will thus need to develop a far more automated approach regardless of the underlying spectral extraction methodology (i.e. an aXe type approach or some combination of the next generation algorithms considered in §6). Therefore considerable attention to data organization and tracking is necessary together with (potentially) the need for quite powerful computing facilities. However, there are several main reasons for optimism outlined below.

7.1. Parallelism

The spectral extraction from large-scale grism surveys is inherently parallelizable. The appropriate tile size is certainly no larger than the WFIRST-AFTA field of view and is probably as small as a 2x2 detector tile. Very little information actually needs to be passed between tiles (e.g. sky levels). that the application of highly compute intensive approaches (e.g. forward modeling) is practical even with large surveys using current computing technology. The necessary investment is the number of cores not their performance or interconnect speed.

7.2. Opportunities for self-calibration

Past experience has primarily relied upon calibrations obtained independently from the science observations. This links the stability of these calibrations to the obtainable accuracy. Large-scale surveys offer several avenues to self-calibration. The inclusion of standard fields that are

re-observed at a cadence compatible with calibration stability provides an internal self-calibration. These fields can be calibrated either directly with WFIRST-AFTA to existing standards or via ground observations. Alternately, a ground or space calibration of a subset of the primary science target population can serve as a calibration basis. Self-calibration may involve constructing calibration products over time as science data are accumulated (for example, the “master sky” images created for WFC3/IR described in §4.2), so efficient reprocessing of the extractions and analysis with the best available calibrations should be incorporated in the pipeline design.

7.3. Machine learning

One complexity that may occur when scaling up to much larger datasets is that the manual handling of exceptions could become impractical. Exceptions range from the impact of bright stars, glints and stray light, moving solar system objects, to instrumental anomalies. While manual screening is possible, WFIRST will likely present an opportunity (or necessity) to apply machine-learning techniques to exclude defective data samples. These techniques would reduce observer bias (and costs) while also providing a traceable record of which artifacts were flagged for exclusion.

7.4. Statistical analysis

A final, but potentially critical, item for consideration is the resulting statistical understanding of the extracted spectral data. That is, an understanding of how effects that operate on small angular scales—such as spectral contamination and detection completeness and how these depend on source density on the sky—can affect the large-scale statistical properties of the entire survey which themselves are key inputs to the cosmological experiments of the WFIRST-AFTA mission.

8. Conclusions and Suggestions for Future Work

Slitless spectroscopic observations have played an important role in astronomy, represent a key HST and JWST scientific capability, and are expected to be essential for key WFIRST-AFTA science objectives. Successful exploitation of such data will require significant efforts in both software systems and calibrations (both ground and flight). The desire to maximize the scientific returns and the need to careful understanding of the statistical properties of the extremely large datasets expected from WFIRST-AFTA poses significant challenges. Current or near-term software systems are probably inadequate for analysis beyond very small WFIRST-AFTA observation programs. Large scale programs will require major advances in dealing with overlapping spectra, using the information inherent in observations at multiple position angles, book-keeping to combine multiple tiled observations, and the optimal inclusion of priors obtained from WFIRST-AFTA (or other) imaging observations.

We make two distinct recommendations. First, that efforts to explore improved approaches for use with HST data be encouraged and mechanisms established to capture lessons learned. Second, that WFIRST-AFTA should develop a detailed data analysis plan including software

algorithms and calibration plans for slitless spectroscopic observations. The WFIRST-AFTA project should recognize both the magnitude of this task and the need to make key decisions regarding mission design and ground calibration based upon this data analysis plan.

9. Acknowledgements

This work was carried out under contract with the WFIRST-AFTA Study Office at the NASA Goddard Space Flight Center (GSFC), as part of joint pre-formulation science center studies by the Space Telescope Science Institute (STScI) and the Infrared Processing and Analysis Center (IPAC). We are grateful for feedback and helpful comments from our colleagues at GSFC, IPAC, and STScI; especially Lee Armus , James Colbert, Alaina Henry, Davy Kirkpatrick, Seppo Laine, and Swara Ravindranath.

10. References

Atek, H. et al. 2010, *ApJ,* 723, 104
Barger, A et al. 2012, *ApJ*, 749, 106
Bertin, E. & Arnouts, S. 1996, *A&AS*, 117, 393
Brammer, G. et al. 2012, *ApJS*, 706, 173
Cowie, L. et al. 2010, *ApJ*, 711, 928
Horne, K 1986, *PASP*, 98, 609
Kummel, M. et al. 2009, ADASS XVIII, ASP Conf. Series 411, 430
McCarthy, P. et al. 1999, ApJ 520, 548
Morrissey, P. et al. 2007, *ApJS*, 173, 682
Pirzkal, N. et al. 2012, *ApJ*, 748, 122
Rodney, S. et al 2012, ApJ 746,5
Skelton, R. et al. 2014, *ApJS*, 214, 24
Salzer, J. et al., 2000, *AJ*, 120, 80
Spergel, D. et al. 2015, *arXiv*/1503.03757
Stevenson, K. et al. 2014, *Science*, 346, 838
Straughn, A. N. et al. 2008, *AJ*, 135, 1624
van Dokkum, P. & Brammer, G. 2010, *ApJ*, 718, 73
Whitaker, K. et al. 2013, *ApJ*, 770, 39