Spectro-Perfectionism: An Algorithmic Framework for Photon Noise-Limited Extraction of Optical Fiber Spectroscopy

ADAM S. BOLTON[1,2] AND DAVID J. SCHLEGEL[3]

ABSTRACT. We describe a new algorithm for the "perfect" extraction of one-dimensional (1D) spectra from two-dimensional (2D) digital images of optical fiber spectrographs, based on accurate 2D forward modeling of the raw pixel data. The algorithm is correct for arbitrarily complicated 2D point-spread functions (PSFs), as compared to the traditional optimal extraction algorithm, which is only correct for a limited class of separable PSFs. The algorithm results in statistically independent extracted samples in the 1D spectrum, and preserves the full native resolution of the 2D spectrograph without degradation. Both the statistical errors and the 1D resolution of the extracted spectrum are accurately determined, allowing a correct χ^2 comparison of any model spectrum with the data. Using a model PSF similar to that found in the red channel of the Sloan Digital Sky Survey spectrograph, we compare the performance of our algorithm to that of cross-section based optimal extraction, and also demonstrate that our method allows coaddition and foreground estimation to be carried out as an integral part of the extraction step. This work demonstrates the feasibility of current and next-generation multifiber spectrographs for faint-galaxy surveys even in the presence of strong night-sky foregrounds. We describe the handling of subtleties arising from fiber-to-fiber cross talk, discuss some of the likely challenges in deploying our method to the analysis of a full-scale survey, and note that our algorithm could be generalized into an optimal method for the rectification and combination of astronomical imaging data.

1. INTRODUCTION

Optical fibers offer a compelling advantage for astronomical survey spectroscopy. By allowing the light from targets over a wide field of view on the sky to be rearranged into a compact format and fed to any number of spectrographs in parallel, they can provide a vast multiplex advantage over imaging spectrographs. For this reason, fiber technology has been adopted for a number of major survey programs such as the Las Campanas Redshift Survey (LCRS: Shectman et al. 1996), the Two-degree Field Survey (2dF: Colless et al. 2001), the Sloan Digital Sky Survey (SDSS: York et al. 2000), and the recently initiated Baryon Oscillation Spectroscopic Survey (BOSS: Schlegel et al. 2009a) of the SDSS3 project. The use of fiber spectrographs for faint-object spectroscopy has however been restrained by concerns over throughput and systematic limitations on the quality of subtraction of night-sky emission foregrounds. The spectroscopic extractions of the SDSS multifiber instrument have established a high standard, but significant systematic shortcomings remain. Methods to characterize and partially remove sky-

subtraction residuals in SDSS fiber spectra can mitigate the problem somewhat (e.g., Bolton et al. 2004; Wild & Hewett 2005), but do not substitute for a formally correct sky-subtraction model. The faintest spectroscopic galaxy surveys have tended to make use of multislit imaging spectrographs (e.g., Cowie et al. 1996; Steidel et al. 2003; Davis et al. 2003; Abraham et al. 2004; Le Fèvre et al. 2005), along with sky-subtraction techniques such as nod-and-shuffle (Cuillandre et al. 1994; Glazebrook & Bland-Hawthorn 2001) or B-spline–based modeling of the 2D sky spectra in the slits (e.g., Kelson 2003). Nod-and-shuffle techniques in particular are ill-suited to most fiber spectrographs, require at least double the detector area, and furthermore reduce the background-limited signal-to-noise ratio (S/N) by a factor of $1/\sqrt{2}$ due to their subtraction of data from data.

This paper outlines an algorithmic framework for the modeling and extraction of optical and near-infrared astronomical fiber spectroscopy to the limit of photon noise and native instrumental resolution. By comparing our method with the standard techniques of optimal extraction currently in wide use, we identify and resolve key systematic barriers to "perfect" extraction. As a result, multifiber spectroscopy emerges as a clear and compelling technique for current and future generations of faint-galaxy spectroscopic surveys, even well below the brightness of the night sky at all wavelengths. This algorithm will deliver significant benefits to the reanalysis of the original SDSS (hereafter SDSS1) archive and to the ongoing analysis of the BOSS survey, both for core redshift-survey goals and for projects that

---

[1] Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112;bolton@physics.utah.edu.

[2] Beatrice Watson Parent Fellow, Institute for Astronomy, University of Hawai'i, Honolulu, HI 96822.

[3] Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720-8160; djschlegel@lbl.gov.

aim to select rare emission-line objects from within the regions of the spectrum dominated by OH line emission from the night sky (e.g., Bolton et al. 2004; Willis et al. 2005; Bolton et al. 2006; Willis et al. 2006; Bolton et al. 2008).

In considering this subject, we will make a clear distinction between the problems of "calibration" and "extraction": calibration is the description of the way in which any set of astronomical and environmental stimuli translate into the responses of the digital detectors (which we assume here to be pixellated charge-coupled devices, or CCDs); extraction is the reconstruction of particular stimuli from particular responses. More strictly speaking, we view an "extracted spectrum" not as a model for the flux of the observed source itself, but rather as a properly calibrated 1D compression of the instrumental response to an observation of that source. When executed and reported correctly, an extracted spectrum permits a statistically valid χ^2 test of an input model spectrum against the extracted pixels. This article specifically illustrates a method for carrying out this sort of perfect extraction assuming that perfect calibration is available. We do not mean to trivialize the problem of calibration, and in our concluding remarks we will discuss the relationship of our extraction method to current and future spectroscopic calibration regimes.

The article is organized as follows. Section 2 frames the problem of extraction in terms of image modeling, lays out the first part of our algorithm, and compares its performance with that of standard extraction techniques in terms of the quality of their respective models to the raw 2D data. Section 3 addresses the issue of resolution and covariance in the extracted spectra, and presents the second part of our method, which establishes optimal properties in both these regards. Section 4 presents a more realistic demonstration of our algorithm on simulated data, illustrating several additional strengths and subtleties of the method. Finally, Section 5 provides a discussion and conclusions.

We will observe the following conventions in this article. Without loss of generality, we will assume that spectroscopic traces run roughly parallel to CCD columns (i.e., "vertically"), with wavelength increasing with row number and cross-sectional position increasing with column number. We will denote vectors in lowercase bold type ($\mathbf{f}$) and matrices in uppercase bold type ($\mathbf{A}$). We assume that all errors have a Gaussian distribution, and we assume no formal priors on fitted model parameters.

2. EXTRACTION AS IMAGE MODELING

The current standard of quality for the extraction of optical fiber spectroscopic data from digital images to 1D spectra is the optimal technique described by Hewett et al. (1985), Horne (1986), and Robertson (1986), and subsequently implemented in countless forms (e.g., Marsh 1989; Kinney et al. 1991; Hall et al. 1994; Bacon et al. 2001; Piskunov & Valenti 2002; Cushing et al. 2004; Bershady et al. 2005; Zanichelli et al. 2005; Dixon et al. 2007; Bolton & Burles 2007; Cui et al. 2008). This algorithm models the 2D spectrum image row by row with a minimum-χ^2 scaling of the cross-sectional profile. The fitted amplitude is taken as an optimal estimator of the 1D spectrum at the wavelength corresponding to that row. Optimal extraction as currently understood and practiced has significant advantages, most notably in the increased S/N of the extracted spectrum in comparison with the boxcar-aperture summing technique. It also allows for model-based flagging of pixels afflicted by cosmic-ray hits. Optimal extraction furthermore leads to statistically uncorrelated extracted spectrum values, although this property is not preserved under the rebinning that is required for the combination of multiple exposures that are not precisely aligned at the raw pixel level.

The most significant (and underappreciated) shortcoming of the optimal extraction method is that it is only correct in the case where the *two-dimensional* image I of a monochromatically illuminated fiber (the point-spread function, or PSF) is a separable function of column x and row y offset from the PSF centroid:

$$I(x, y) = I_x(x)I_y(y). \tag{1}$$

If this condition does not hold—and it does not hold for nearly all PSF models beyond a single Gaussian ellipsoid whose principal axes are aligned with the CCD rows and columns—then the cross-sectional profile of the 2D spectrum is no longer fixed, but rather depends upon the details of the input spectrum itself. This is of limited concern if the spectrum is smoothly varying with wavelength. However if the spectrum involves many sharp features such as the bright OH rotational emission lines in the near-infrared spectrum of the night sky, then the output of the row-by-row optimal extraction algorithm becomes ill-defined. For example, the cross section through the core of an emission line may have a different shape than the cross section through the wing of the line, yet the extraction will attempt to model them both with a single average cross-sectional shape. This will degrade the resolution and S/N of the extracted spectrum, and may introduce a bias into the spectrum estimation.

The shortcomings of the row-wise optimal extraction method can be overcome by modeling the 2D spectroscopic data image in a manner that more accurately reflects the way in which an input spectrum translates into photon counts in the CCD image. For the case of a linear CCD detector (which we will assume), the system calibration can be described as a matrix $\mathbf{A}$, whose elements $A_{i\ell}$ describe the (noise-free) counts in CCD pixel i for a unit of monochromatic input at wavelength λ_ℓ. Note that we are suppressing the natural two-dimensionality of the CCD by allowing i to index all column/row combinations. We are also suppressing any possible multifiber dependence; it is straightforward to incorporate this sort of dependence into the ℓ-indexed dimension, and we will in fact do so in § 4. The calibration matrix $\mathbf{A}$ incorporates all the effects of wavelength calibration, spectrum trace position, PSF shape and its dependence

upon position, flat-fielding effects, spectrophotometry, and exposure time. Note that $\mathbf{A}$ will generally be a sparse matrix, since an input at a given wavelength in a given fiber will only influence a relatively small number of the full CCD pixels. For the case of an input spectrum vector $\mathbf{f}$, the observed CCD pixel count vector $\mathbf{p}$ will be given by

$$\mathbf{p} = \mathbf{Af} + \mathbf{n}, \qquad (2)$$

or in indexed notation

$$p_i = \left(\sum_\ell A_{i\ell} f_\ell \right) + n_i, \qquad (3)$$

where $\mathbf{n}$ is a pixel noise vector. Extraction is then the reconstruction of $\mathbf{f}$ from $\mathbf{p}$ given a knowledge of $\mathbf{A}$. Although extraction is nontrivial, it is a linear problem.

Since the wavelength coordinate is continuous, equation (2) should most properly be written in integral rather than matrix form. However our interest is in *solving* for $\mathbf{f}$: inversion for an infinite number of differentially-spaced wavelength amplitudes from a finite number of pixels and resolution elements is neither possible nor desirable. We therefore restrict our attention to a discrete set of wavelength sampling positions $\{\lambda_\ell\}$. The means of determining the most appropriate sampling density for these positions will be addressed in more detail at the end of this section; in general they may be either more closely or more widely spaced than the pixel rows in the raw data. However, if they are too closely spaced, then the solution for $\mathbf{f}$ becomes degenerate.

Assuming Gaussian noise, the minimum-χ^2 solution for the input spectrum vector $\mathbf{f}$ from the data vector $\mathbf{p}$ is

$$\mathbf{f} = (\mathbf{A^T N^{-1} A})^{-1} \mathbf{A^T N^{-1} p}. \qquad (4)$$

Here, $\mathbf{N}$ is a pixel noise matrix, with $N_{ij} = \langle n_i n_j \rangle$. We treat raw pixel errors as statistically independent, and thus the noise matrix is diagonal. Our model for the 2D raw pixel data is simply given by $\mathbf{m} = \mathbf{Af}$. Again, this requires that the sampling points not be too closely spaced in wavelength, so that $\mathbf{A^T N^{-1} A}$ does not become noninvertible.

To demonstrate how this extraction method improves upon row-by-row optimal extraction based on a fixed cross-sectional profile, we will adopt the following illustrative fiber spectrograph PSF:

$$I(x,y) = \frac{(1-b)}{\sqrt{2\pi}\sigma} \exp\left[\frac{-r_{ell}^2}{(2\sigma^2)}\right] + \frac{b \exp(-r/r_0)}{2\pi r_0 r}. \qquad (5)$$

Here, σ (in pixels) controls the size of a Gaussian core to the profile (first term), and r_0 sets the characteristic size of the profile wings (second term). The parameter b controls the fraction of the total flux in the wing component. The form of the wing

component is taken from a parameterization of the near-IR scattering seen within the SDSS1 spectroscopic CCDs (Gunn, J. E., 1997, private communication). The coordinate $r = \sqrt{x^2 + y^2}$ is the radial offset in pixels on the CCD from the center of the PSF spot. The form r_{ell} in the argument of the Gaussian term indicates that the core profile can have an elliptical shape as determined by

$$r_{ell} = \sqrt{qx'^2 + y'^2/q}, \qquad (6)$$

where q is a minor-to-major axis ratio, and (x', y') are related to the CCD column/row coordinates (x, y) by a possible rotation and translation. For the wings, we adopt the values $b = 0.1$ and $r_0 = 36$ pixels, roughly appropriate to the SDSS1 CCDs at a wavelength of 8500 Å. For the core, we take $\sigma = 1.1$ pixels, roughly the median characteristic value for the SDSS1, and an ellipticity of $q = 0.75$ with a major axis position angle inclined at 45° relative to the column/row axes. This core ellipticity is somewhat more exaggerated that what is seen in well-focused SDSS1 spectrograph exposures, but will provide a good test of our method in the presence of the astigmatic aberrations that can arise in wide-field camera systems.

Generally speaking, the calibration matrix $\mathbf{A}$ is determined by the combination of the fiber PSF functional form, the dependence of its parameters upon wavelength and position (which can be derived from calibration frames illuminated by gas-discharge lamps that have multiple discrete emission lines at known wavelengths), and the wavelength- and pixel-dependent throughput of the system (which can be derived from calibration frames uniformly illuminated by incandescent lamps with approximately flat continuum spectra). For simplicity, we will use CCD pixel rows as a proxy for wavelength, neglect any variation of the PSF shape parameters with position, and neglect any pixel-to-pixel sensitivity variations. Assuming that all these effects can be calibrated successfully, they can be incorporated into the approach that we describe. We also neglect the problem of cosmic rays here, although a model-based extraction such as ours is ideally suited to the iterative masking of cosmic-ray pixels in the raw data based upon their statistical disagreement with the best extraction model. Finally, we ignore the possibility of a large-scale scattered-light component in the image, but the incorporation of this phenomenon (as well as those of imperfectly subtracted instrumental bias and dark-current levels) into our 2D modeling would be straightforward.

The most stringent test of any extraction is provided by unresolved emission lines, which show the sharpest real features possible in the data. To test our method alongside the standard row-wise extraction, we therefore create a noise-free spectrograph-like image consisting of 11 emission lines of equal flux spaced vertically over 80 pixels within a 31×101 pixel image. We include a slight tilt of the spectral trace relative to the vertical, and incorporate a mild rate of change in the spacing of the lines so as to ensure a range of subpixel positions for the

line centers. Each pixel value is computed by integrating the profile with 4×4 subsampling. We incorporate a small core radius in the denominator of the scattering wing term to avoid numerical difficulties with the (integrable) singularity. This simulated image is shown in the leftmost panel of Figure 1.

Next we perform an extraction of this model image using both the standard row-wise, cross-section based optimal extraction and our 2D modeling method. For the 2D extraction, we compute a basis set of 101 PSFs centered on the trace position in each of the 101 image rows. Note that the "emission lines" of the simulated image will in general not be coincident with these basis functions, but will instead be offset from the nearest basis PSF by some subpixel amount in y. For the row-wise extraction, we sum all 101 basis functions with equal weight to generate the image of a flat continuum spectrum, and take the normalized cross-sectional shape of this spectrum in each row as the basis profile for the extraction of that row. Thus both the 2D and row-wise extractions will have an equal number of free extraction parameters, with the same sampling interval between them. Giving equal weight to all pixels, we compute the linear least-squares set of amplitudes that multiply the basis functions of each extraction method to best reproduce the simulated image. The results are shown in Figure 1. The difference between the quality of the two models can be seen immediately, with the row-wise extraction model unable to capture the variation of the cross-sectional shape with spectrum that comes from the ellipticity of the PSF core and the presence of the scattering wings acting in concert with the sharp spectral features. Figure 1 also shows the residuals of subtraction of the two extraction models from the simulated image. The residuals of the 2D method have

a scale roughly 200 times smaller than the residuals of the row-wise method.

We can make a quantitative comparison of the relative quality of the two extraction models to the simulated image in terms of the following two figures of merit:

$$D_{\rm sq} \equiv \left(\sum_i (p_i - m_i)^2 \Big/ \sum_i p_i^2 \right)^{1/2}, \qquad (7)$$

$$D_{\rm ab} \equiv \sum_i |p_i - m_i| \Big/ \sum_i |p_i|. \qquad (8)$$

Here, the p_i are the raw pixel data values, and the m_i are the 2D extraction model values in those pixels. Both of these figures will be equal to zero in the case of a perfect model, and equal to 1 for a model that is identically zero (i.e., no model at all). For the case described here, we find $D_{\rm sq} = 0.0006$ and $D_{\rm ab} = 0.0008$ for the 2D model, and $D_{\rm sq} = 0.12$ and $D_{\rm ab} = 0.20$ for the row-wise model, again confirming a factor of ≈ 200 improvement with the 2D model.

Before proceeding further, we address the question of appropriate sampling. While the sampling of the row-wise extraction is naturally limited to one sampling point per row, the sampling of the 2D extraction method can easily be adjusted to be either finer or more coarse than this. The residuals seen in the case given after subtracting the 2D model from the simulated emission-line image, although very small compared to those of the row-wise model, nevertheless exceed the numerical precision of the calculation, and are due to the misalignment between the extracting basis and the input emission lines in the "wavelength" direction. We might
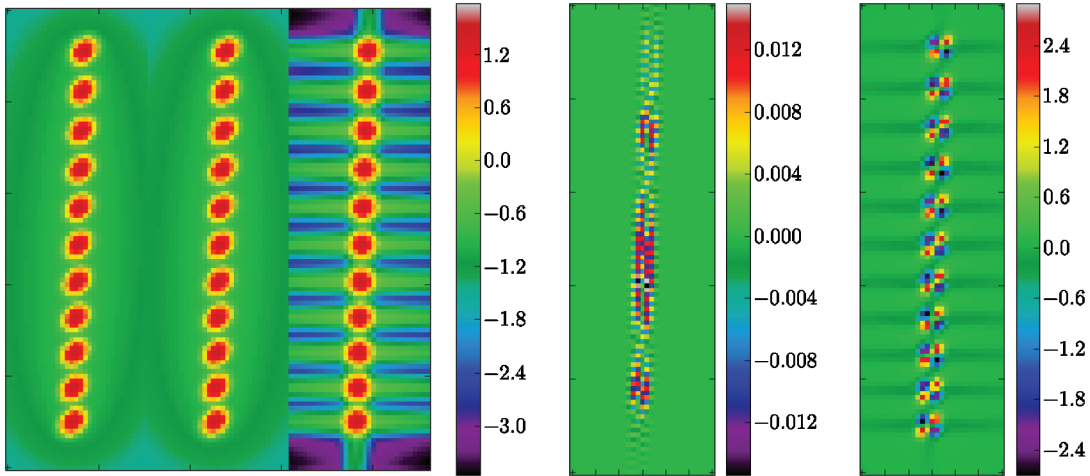

FIG. 1.—*Left, from left to right:* simulated noise-free emission-line image, 2D extraction model of simulated image, and row-wise extraction model of simulated image. Color scaling is in units of the base-10 logarithm of the pixel value, with the pixel values themselves scaled to have an average value of unity across the entire image. *Right:* residuals resulting from the subtraction of the 2D extraction model (*left*) and row-wise extraction model (*right*) from the simulated emission-line image. Color scale is in (nonlogarithmic) units of residual counts. Image regions are 31×101 pixels in size in all cases.

reasonably improve upon the extraction model by adopting a more finely-spaced basis set. If the typical fiber spectrograph PSF were purely band-limited, then the appropriate sampling could be chosen via the Nyquist criterion. However, fiber spectrograph PSFs are primarily determined not by a diffraction limit, but rather by the convolution of a sharp "top-hat" fiber image with a set of optical aberrations, and thus the determination of an appropriately critical extraction sampling must be made according to more empirical factors. For the particular case under consideration here, we have tested the quality of the 2D model to the simulated image for sampling densities ranging from 0.5 to 2 basis functions per row in the 2D image. Both D_{sq} and D_{ab} improve with increasingly fine sampling until about 1.3 basis functions per row, where they attain a value of a few times 10^{-5}. With finer sampling beyond this point, the model becomes worse as the matrix to be inverted approaches singularity. In practice, the appropriate sampling density will be determined by the details of the spectrograph PSF and dispersion, and consequently by the spectral resolution as defined in the following section.

3. RESOLVING THE RESOLUTION

We have demonstrated the fidelity of our 2D extraction model to the 2D pixel data, but what of our extracted spectrum? Naively, $\mathbf{f}$ from equation (4) is our 1D spectrum estimator, and the matrix $\mathbf{C} \equiv (\mathbf{A^T N^{-1} A})^{-1}$ is the covariance matrix of its individual values. But all is not well: equation (4) not only extracts the spectrum, it deconvolves the spectral resolution. The instability of the deconvolution manifests itself in a significant "ringing" in the extracted spectrum, along with large correlations and anti-correlations between extracted pixels. Both of these are decidedly undesirable features in a spectrum. The way to remedy this situation is the second key element of our extraction method: to *reconvolve* our deconvolved spectrum back to the same resolution as the raw data.

Consider the inverse covariance matrix $\mathbf{C}^{-1}$ of the deconvolved spectrum basis:

$$\mathbf{C}^{-1} = \mathbf{A^T N^{-1} A}. \tag{9}$$

This matrix is symmetric and band-diagonal, and all of its entries are nonnegative (assuming that there is no way for any input spectrum to *subtract* counts from the CCD). Now take the unique nonnegative matrix square root of $\mathbf{C}^{-1}$ to find a symmetric matrix $\mathbf{Q}$ such that

$$\mathbf{C}^{-1} = \mathbf{QQ}. \tag{10}$$

This may be done by determining the eigenbasis of $\mathbf{C}^{-1}$, taking the element-wise square root of the diagonal matrix of its eigenvalues (which will all be positive since $\mathbf{C}^{-1}$ is positive definite), and transforming this new diagonal matrix back using the unitary matrix that relates the eigenbasis to the original basis. Equation (10) would still hold if we applied any arbitrary set of

sign-flips to the square root-eigenvalue matrix (leading to other matrix square roots in addition to the unique nonnegative one), but these would introduce undesired negativity into $\mathbf{Q}$.

Next, define a normalization vector $\mathbf{s}$ through

$$s_{\tilde{\ell}} = \sum_{\ell} Q_{\tilde{\ell}\ell}, \tag{11}$$

a matrix $\mathbf{R}$ through

$$R_{\tilde{\ell}\ell} = s_{\tilde{\ell}}^{-1} Q_{\tilde{\ell}\ell} \text{ (no sum)}, \tag{12}$$

and a diagonal matrix $\widetilde{\mathbf{C}}^{-1}$ with entries given by

$$\widetilde{C}_{\tilde{\ell}\tilde{\ell}}^{-1} = s_{\tilde{\ell}}^2. \tag{13}$$

By construction, we now have

$$\mathbf{C}^{-1} = \mathbf{R^T} \widetilde{\mathbf{C}}^{-1} \mathbf{R}, \tag{14}$$

and consequently

$$\widetilde{\mathbf{C}} = \mathbf{RCR^T}. \tag{15}$$

Note the complete analogy between equations (9) and (14). The matrix $\mathbf{R}$ defines a transformation from the deconvolved spectrum to a "reconvolved" spectrum for which (1) the extracted 1D pixels are statistically independent of one another (since the matrices $\widetilde{\mathbf{C}}^{-1}$ and $\widetilde{\mathbf{C}}$ are diagonal), and (2) the blurring of input spectra in wavelength from the deconvolved basis is statistically equivalent to the blurring imposed by the actual 2D spectrograph matrix $\mathbf{A}$ on the physical input spectrum. It is in this sense that we can reconvolve with the "same" resolution as was inherent in the 2D data. Our extracted 1D spectrum is then

$$\widetilde{\mathbf{f}} = \mathbf{Rf}, \tag{16}$$

with uncorrelated errors described by the diagonal covariance matrix $\widetilde{\mathbf{C}}$ and undegraded resolution described by the matrix $\mathbf{R}$. We may think of this reconvolved spectrum as a model for what *would have been observed* by a truly 1D spectrograph with the same resolution as our 2D CCD spectrograph. Since the resolution of the extracted spectrum is precisely characterized by the matrix $\mathbf{R}$, we can convolve any theoretical model for the input 1D spectrum with this matrix in order to compare to the data and compute a statistically correct χ^2 value. The sense of the normalization imposed by equation (12) is to describe the elements of $\widetilde{\mathbf{f}}$ as a weighted sum over the elements of the theoretical input spectrum, with the sets of weights individually summing to 1.

Figure 2 illustrates the deconvolved and reconvolved spectra that result from the extraction of the simulated emission-line
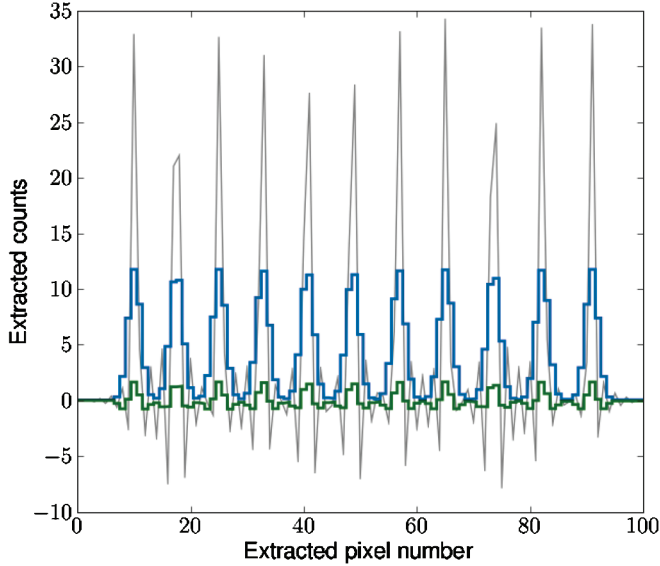
FIG. 2.—Extracted spectra of simulated noise-free emission-line image. *Thin line:* deconvolved spectrum from the 2D modeling extraction method. Note significant ringing. *Thick blue line* (upper of the two thick lines, rendered with steps): deconvolved spectrum from 2D modeling extraction reconvolved to the native spectrograph resolution using the resolution matrix **R** defined in the text. *Thick green line* (lower of the two thick lines, rendered with steps): $10\times$ the difference between the reconvolved 2D-model extracted spectrum and the row-model extracted spectrum. "Upward peakiness" at the position of the emission lines indicates that the 2D-extracted spectrum has higher resolution than the row-extracted spectrum.

spectrum of § 2. Note that the ringing in the deconvolved spectrum is completely absent from the reconvolved spectrum. The peaks in the reconvolved spectrum represent the true undegraded resolution of the 2D spectrographic data. The values in the deconvolved spectrum are highly correlated with one another, whereas the values in the reconvolved spectrum are completely uncorrelated. Also shown is a representation of the difference between the extracted spectra using the 2D and row-wise methods. Although the two spectra are sufficiently similar that their differences cannot be seen in a direct plot, they differ at the few percent level, which is significant for application to the extraction of faint-galaxy spectra in the presence of strong night-sky line emission. Also note that the resolution of the 2D-extracted spectrum is higher than the resolution of the row-extracted spectrum.

In defining the resolution with which to reconvolve our deconvolved extracted spectrum, we have explicitly chosen a form that results in an exactly diagonal sample covariance matrix in the reconvolved basis. Additionally, as a consequence of the band-diagonal nature of $\mathbf{C}^{-1}$, our **R** matrix is also essentially band-diagonal, corresponding to a localized "line-spread function" in the extracted 1D spectrum as seen in Figure 2. It is nevertheless worth remembering that other choices for **R** are available. For instance, there may be applications for which it is desirable to have a simpler parameterized form for the res-

olution than will in general result from the procedure outlined here. The parameters of this form could be optimized so as to make the off-diagonal correlations in the reconvolved basis sufficiently small while still nonzero.

4. A MORE INTERESTING TEST

Before taking up the longer-term challenge of implementing our algorithm on actual survey spectroscopy, we illustrate its power with a more realistic simulation than the simple case in §§ 2 and 3. In particular, we now include the following effects, all of which will be found in real survey data:

1. Noise,
2. A varying 2D fiber PSF shape,
3. Overlap ("cross talk") between the 2D images of neighboring spectra on the CCD detector of a multifiber spectrograph,
4. Multiple exposures with relative movement ("flexure") of the fiber images on the CCDs between them, and
5. A night-sky foreground that may vary between exposures and must be modeled and subtracted to reveal a much fainter object spectrum of interest.

We use nearly the same PSF model as described in § 2, but now generate four "spectra" on the same image, separated by 5.7 pixels from one another in the horizontal direction. This separation is somewhat smaller than the ≈ 6.2 pixel separation between neighboring spectra in the SDSS1, and therefore leads to greater cross talk. The Gaussian core of each fiber is given a different ellipticity (ranging from 0.9 to 0.6) and position angle (ranging between $15°$ and $60°$ from the x pixel direction), to mimic optical variations and fiber heterogeneities. The central two spectra are taken as "object" spectra, and the outside two are "sky" spectra. We also generate three "exposures" of these four spectra, shifting the fiber images relative to the CCD grid in each case to simulate spectrograph flexure. We generate a "sky" spectrum of 15 emission lines at increasing separation over the range of the spectra. In each of the three exposures, the amplitudes of these individual sky lines are randomized, and the sky realization for each exposure is projected through all four "fibers" in that image. The positions of the sky lines in wavelength are fixed across all exposures. Next, two "object" spectra are projected through the central two fibers of all three exposures and added to the image of the sky spectra. The relative fluxes are taken so that, in the fibers containing both object and sky, the total sky counts exceed the total object counts by a factor of 20. We use an extraction sampling density of 1.2 basis functions per CCD row. The noise level is set by assigning 10^4 statistical sky counts (i.e., photons) per spectrum per sampling point in total across all three exposures. This implies 500 total object counts per sampling point, and an approximate background-limited S/N of 5 per extraction pixel for the objects. We also include a "read-noise" term of 5 rms counts per pixel. We use the statistical noise level to generate a noise-image realization that we
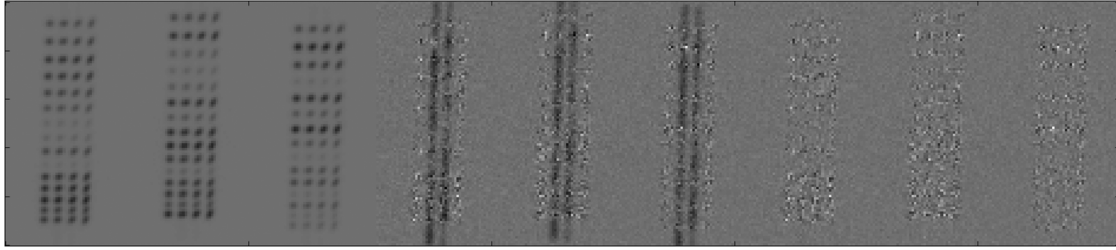
Fig. 3.—Simulated multifiber, multiexposure spectroscopic data, including noise, flexure, a nonuniform PSF, and "sky." *Left:* Three "exposures" of four fiber spectra, including simulated flexure and sky-spectrum variability. Object spectra are included in central two fibers in each set, but are too faint to see directly. *Center:* same as left, after subtraction of extraction model for the sky component in each exposure, with gray-scale stretched by a factor of 40 to reveal the traces of the object spectra. *Right:* As in center, but after subtraction of extraction model for the object spectra as well. When scaled by the pixel errors, these residuals are consistent with a reduced χ^2 of unity. Each "exposure" is 51×101 pixels in size.

add to the sky and object data. The resulting three sets of four spectra are seen in the leftmost third of Figure 3.

The full power of the 2D modeling extraction now becomes apparent. We construct a generalized calibration matrix, involving input spectra of both the night sky (one for each exposure) and the science target objects (one for each object). In each exposure, the sky basis projects into the images of all fibers—sky and object alike—while the object bases project only into the images of the object fibers. The multiple exposures of each object spectrum are extracted in a single step: there is no need for the separate steps of registration, extraction, and coaddition of individual frames. Similarly with the sky spectra, we extract a single sky spectrum from all fibers of a given exposure. Within the object fibers, we extract the sky and object spectra in sum together. There is no explicit sky-subtraction step, but rather a sky–object decomposition that is an integral part of the extraction into individual 1D spectrum components. All at once, we model and extract three skies—one for each exposure—and two object spectra. In summary: extraction, co-addition, and sky subtraction are all subsumed into a single image-modeling operation. The results of this modeling can be seen in the second and third parts of Figure 3, illustrating the sky–object decomposition and the "photon-noise" limited quality of the sky + object model to the three 2D exposures.

We note that the all-in-one sky modeling and extraction is necessary in order to model and decompose the sky spectrum from the object spectra *in the deconvolved frame*, upstream from the fiber-to-fiber PSF variations that would lead to systematic errors in a traditional modeling and subtraction of the sky from extracted and resolution-convolved spectra. If necessary, an accounting for any spatial variations of the sky brightness over the telescope focal plane could be built in by way of additional linear sky components. The combination of multiple object spectra at the time of extraction, on the other hand, is not strictly necessary. Sky-subtracted object spectra from the individual exposures could be combined together in a subsequent step, to allow for the nonlinear determination of spectrophotometric variations between the frames. However, once these variations are determined, the proper combination of these multiple

frames would resemble a second extraction, with the concatenation of their individual $\mathbf{R}$ matrices forming the new $\mathbf{A}$ matrix (in the notation of § 2).

The presence of cross talk between neighboring spectra, and between the object spectra and the skies, makes the determination of the 1D resolution matrix somewhat more challenging. Consider the full covariance matrix $\mathbf{C}$ and inverse covariance matrix $\mathbf{C}^{-1}$ in the deconvolved extracting basis: they consists of multiple blocks, with each block describing the statistical coupling between different extracted spectra. The matrix $\mathbf{C}^{-1}$ is band diagonal in all blocks, but it has nonzero elements in the off-diagonal blocks coupling different extracted spectra to one another. Assuming the spectral cross talk on the CCD is from spectra that are otherwise unassociated with one another, we do not simply want to take the matrix square root of $\mathbf{C}^{-1}$ to define our resolution, since the resolution defined in this way would mix extracted spectra with one another. A possible solution which we adopt here is to invert $\mathbf{C}^{-1}$ to obtain $\mathbf{C}$, zero out all the entries in the off-diagonal blocks of $\mathbf{C}$, reinvert, and define the resolution in terms of the square root of the resulting matrix. To make this explicit with block matrix notation for a case of two extracted spectra

$$\mathbf{C} = (\mathbf{C}^{-1})^{-1} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}, \quad (17)$$

$$\mathbf{C} \equiv \begin{bmatrix} \mathbf{C}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{22} \end{bmatrix}, \quad (18)$$

$$\mathbf{C}^{-1} = \begin{bmatrix} [\mathbf{C}_{11}]^{-1} & \mathbf{0} \\ \mathbf{0} & [\mathbf{C}_{22}]^{-1} \end{bmatrix} = \mathbf{Q}\mathbf{Q}, \quad (19)$$

and proceed to define the resolution matrix $\mathbf{R}$ in terms of this $\mathbf{Q}$ as indicated by equations (11) and (12). Note that $[\mathbf{C}_{11}]^{-1} \neq \mathbf{C}_{11}^{-1}$ and $[\mathbf{C}_{22}]^{-1} \neq \mathbf{C}_{22}^{-1}$ due to the nonzero off-diagonal blocks in the constructed inverse covariance matrix $\mathbf{C}^{-1}$.

It can be shown that by this definition, the resolution matrix will diagonalize **C** *within its diagonal sub-blocks* so that extracted samples will be statistically independent of one another within each spectrum. There will in general be nonzero correlations between the extracted samples of different spectra with one another, but this is unlikely to be important for spectra that are otherwise unrelated. Note that if the problem were rather to extract the spectra of an integral-field spectrograph, with cross talk between fibers that were also adjacent to one another on the sky, then it would perhaps be desirable instead to have a resolution matrix that mixes spectra with one another but provides full statistical independence of samples both within a single spectrum and among neighboring spectra.

Computing the resolution of our simulated multifiber, multiexposure set in the manner outlined here, and reconvolving the extracted object and sky spectra, we obtain the results shown in Figure 4. The sky spectra are scaled down by a factor of 20 for display purposes. When weighted by the error estimates, the extracted spectra have a reduced χ^2 of approximately unity relative to the resolution-convolved input spectra. This demonstrates the suitability of our approach for the extraction of faint-galaxy spectra in the presence of high and sharply wavelength-dependent night-sky foregrounds.

5. DISCUSSION AND CONCLUSIONS

The extraction algorithm that we have described does not come without a price. In the presence of fiber-to-fiber cross talk, the standard row-wise optimal extraction will couple extracted amplitudes between fibers in a single row, leading to a banded matrix of dimension equal to the number of fibers that must be inverted; this process must then be repeated for each row in each exposure. Our 2D modeling extraction, on the other hand, couples extracted amplitudes between fibers, wavelengths, and exposures. Thus the matrix to invert for the solution set of spectra has sides of dimension equal to

$$N_{\text{spectra}} \times N_{\text{samples per spectrum}} \times N_{\text{exposures}}. \quad (20)$$

For one SDSS1 pointing, this would correspond to 320 fibers (one of the two spectrographs), approximately 4000 sampling points, and three exposures: i.e., a square matrix, nearly 4 million to a side. With a brute-force approach this matrix could never be stored, let alone inverted, with any hardware of the present or foreseeable future. The way forward to determining the extracted spectra will no doubt lie in exploiting the sparseness of the inverse covariance matrix to reduce storage and computation, and to apply an iterative method such as the conjugate gradient to solve for the extracted spectra. To determine the resolution with which to reconvolve the extracted spectrum, which formally requires the inversion of the full inverse covariance matrix $\mathbf{C}^{-1}$, the practical solution will be to invert a sufficient subspace of $\mathbf{C}^{-1}$ surrounding each spectrum (or subsegment of a spectrum) to define an acceptably accurate approximation to
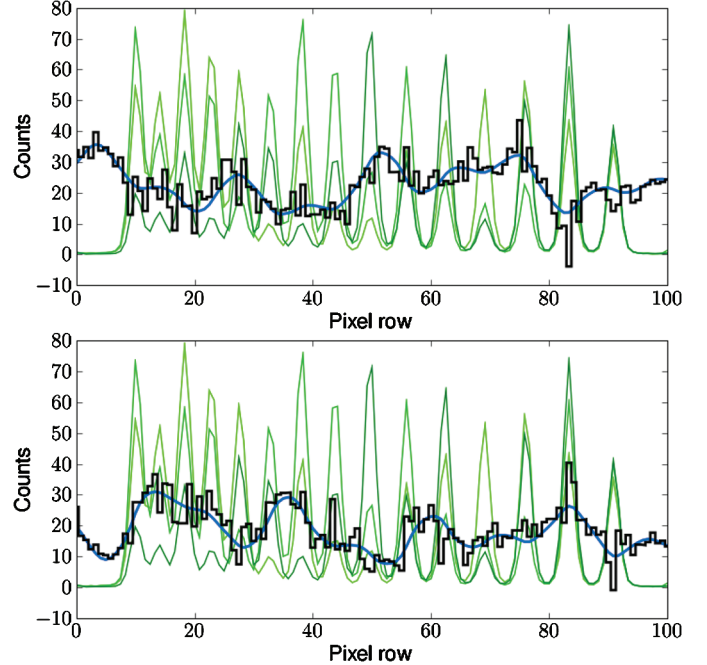

FIG. 4.—Extracted spectra of two simulated objects from multiple exposures as described in § 4 and depicted in Fig. 3, together with extracted "sky" spectra. *Black lines* (rendered with steps) show the extracted spectra, while *blue lines* (rendered smoothly and tracing the black lines) show the input spectra convolved with the 1D resolution. *Green lines* (of varying shades, and identical in both plots) indicate the extractions of the three different realizations of the "sky" spectrum in the three individual exposures, divided by a factor of 20 for display purposes.

the desired resolution matrix. The exact requirements will depend upon the degree of cross talk between neighboring fibers in the spectrograph under consideration. Even with these strategies, a usable implementation of our algorithm for real multifiber survey data may require high-performance parallel computing, depending on the computational expense of the necessary matrix-element calculations.

Most immediately, we plan to implement the strategy outlined in this article to the extraction of spectra from the BOSS Survey. We also plan to conduct a reanalysis of the SDSS1 archive, to provide the definitive version of this important spectral database with the best possible extracted resolution, S/N, and foreground subtraction. These techniques also offer promise for the upcoming Apache Point Observatory Galaxy Evolution Experiment (APOGEE: Allende Prieto et al. 2008) of the SDSS3, a high-resolution, near-infrared multifiber survey of red giant stars in our Galaxy that aims to constrain their evolutionary history as traced by multiple chemical abundances. This extraction strategy will also form a key part of the technical feasibility of the Big-BOSS survey (Schlegel et al. 2009b), which proposes to use a 5000 fiber spectrograph fed by a 3° field-of-view focal plane positioner system on a 4 m class telescope to measure the baryon acoustic scale and redshift-space distortions over the redshift range $0.2 < z < 3.5$. Although the implementations for these

different surveys will differ in detail, we believe that the software engine for the core extraction computations can be written in a general-purpose form. The application of this technique to slit spectroscopy may be possible as well, although the preservation of object spatial information by the slit makes the problem substantially more complicated.

In all cases, the full power of this extraction can only be realized with sufficiently accurate calibration. The current standard calibrations for fiber spectroscopy are exposures of uniform spatial illumination by flat-spectrum incandescent lamps ("flats") and multi–emission-line gas-discharge lamps ("arcs"). Within the framework of our extraction algorithm, the former will determine the relative sensitivity of the individual fibers and pixels, and the latter will determine the spectrograph PSF shape and position as a function of illuminating wavelength in each fiber. Assuming there are no systematic offsets between the calibrations and the science exposures, and assuming that the variation of the spectrograph PSF is smooth enough with wavelength to be well-sampled by the arc frames (which are sparsely distributed in wavelength), these calibrations should contain sufficient information to determine the calibration matrix $\mathbf{A}$. We may proceed by extracting the arcs and flats together, with each one described by a single spectrum projected through an initial guess for $\mathbf{A}$, and then optimize the parameters of $\mathbf{A}$ by nonlinear iteration so as to improve the quality of the 2D extraction models to convergence. A more direct and ambitious approach to the determination of the calibration matrix would be to illuminate the facility calibration screen with either a high-wattage monochromater or a tunable laser (cf. Stubbs & Tonry 2006; Cramer et al. 2009), and to step the illumination source through wavelength in subsequent exposures so as to map out $\mathbf{A}$ explicitly. In practice, with the exception of spectrograph systems that are very stable thermally and mechanically, there is likely to be some shifting of the fiber positions and focus on the CCDs between the calibration and science frames. In this case, "tweaks" to the parameters of $\mathbf{A}$ will be derived from the shapes and positions of the fiber traces and night-sky emission lines in the science frames. Finally, we note that the calibration may be significantly improved by incorporating all available knowledge about the optical design of the telescope and instrument, rather than treating the system as a black box to be specified entirely by empirical calibration data.

Putting aside the computational challenges that arise from the consideration of continuous 2D input data, the method we have described can also be generalized into an optimal recipe for the rectification and combination of multiple CCD imaging exposures: i.e., taking the $\mathbf{f}$ of § 2 to be a *two-dimensional* image model to be extracted from the data. As with the spectroscopic application, the resulting extractions would have optimal resolution, statistically independent extracted image samples, and a definition in terms of the optimization of a well-motivated scalar objective function describing the quality of a model fit to all of the individual exposures. The implementation would be nontrivial, but the benefits could be great. A significant challenge on the calibration side is that the details of the imaging PSF, which must be known, will depend upon the spectral energy distributions of the imaged objects, which will in general be unknown.

In summary, we have demonstrated a new spectrum extraction algorithm for optical and near-infrared astronomical fiber spectroscopy. Given sufficiently accurate calibration, this method can extract spectra to the statistical noise limit in the presence of arbitrarily complicated point-spread functions and arbitrarily high and wavelength-varying foregrounds. The extracted spectra have optimal and precisely quantified resolution and S/N, along with statistically uncorrelated pixels, for any number of subexposures in combination together. As such, statistically accurate χ^2 tests may be made between the extracted data and theoretical models of the input object spectrum. This algorithm represents a fundamental improvement upon the current state-of-the-art methods in use for the extraction of fiber spectroscopy, and thus motivates a serious and positive reevaluation of the promise of fiber-fed spectrographs for next-generation ground-based faint-object surveys.

REFERENCES

Abraham, R. G., et al. 2004, AJ, 127, 2455

Allende Prieto, C., et al. 2008, Astron. Nachr., 329, 1018

Bacon, R., et al. 2001, MNRAS, 326, 23

Bershady, M. A., Andersen, D. R., Verheijen, M. A. W., Westfall, K. B., Crawford, S. M., & Swaters, R. A. 2005, ApJS, 156, 311

Bolton, A. S., & Burles, S. 2007, New J. Phys., 9, 443

Bolton, A. S., Burles, S., Koopmans, L. V. E., Treu, T., & Moustakas, L. A. 2006, ApJ, 638, 703

Bolton, A. S., Burles, S., Schlegel, D. J., Eisenstein, D. J., & Brinkmann, J. 2004, AJ, 127, 1860

Bolton, A. S., et al. 2008, ApJ, 682, 964

Colless, M., et al. 2001, MNRAS, 328, 1039

Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839

Cramer, C. E., Brown, S., Caldwell, N., Dupree, A. K., Korzennik, S. G., Lykke, K. R., & Szentgyorgyi, A. 2009, in OSA Technical Digest Series, Conference on Lasers and Electro-Optics/International Quantum Electronics Conference, JThE85

Cui, B., Ye, Z. F., & Bai, Z. R. 2008, Acta Astron. Sinica, 49, 327

Cuillandre, J. C., et al. 1994, A&A, 281, 603

Cushing, M. C., Vacca, W. D., & Rayner, J. T. 2004, PASP, 116, 362

Davis, M., et al. 2003, Proc. SPIE, 4834, 161

Dixon, W. V., et al. 2007, PASP, 119, 527

Glazebrook, K., & Bland-Hawthorn, J. 2001, PASP, 113, 197

Hall, J. C., Fulton, E. E., Huenemoerder, D. P., Welty, A. D., & Neff, J. E. 1994, PASP, 106, 315

Hewett, P. C., Irwin, M. J., Bunclark, P., Bridgeland, M. T., Kibblewhite, E. J., He, X. T., & Smith, M. G. 1985, MNRAS, 213, 971

Horne, K. 1986, PASP, 98, 609

Kelson, D. D. 2003, PASP, 115, 688

Kinney, A. L., Bohlin, R. C., & Neill, J. D. 1991, PASP, 103, 694

Le Fèvre, O., et al. 2005, A&A, 439, 845

Marsh, T. R. 1989, PASP, 101, 1032

Piskunov, N. E., & Valenti, J. A. 2002, A&A, 385, 1095

Robertson, J. G. 1986, PASP, 98, 1220

Schlegel, D., White, M., & Eisenstein, D. 2009a, in Astro2010: The Astronomy and Astrophysics Decadal Survey, 314

Schlegel, D. J., et al. 2009b, ArXiv e-prints (arXiv:0904:0468)

Shectman, S. A., Landy, S. D., Oemler, A., Tucker, D. L., Lin, H., Kirshner, R. P., & Schechter, P. L. 1996, ApJ, 470, 172

Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Dickinson, M., & Giavalisco, M. 2003, ApJ, 592, 728

Stubbs, C. W., & Tonry, J. L. 2006, ApJ, 646, 1436

Wild, V., & Hewett, P. C. 2005, MNRAS, 358, 1083

Willis, J. P., Hewett, P. C., & Warren, S. J. 2005, MNRAS, 363, 1369

Willis, J. P., Hewett, P. C., Warren, S. J., Dye, S., & Maddox, N. 2006, MNRAS, 369, 1521

York, D. G., et al. 2000, AJ, 120, 1579

Zanichelli, A., et al. 2005, PASP, 117, 1271