The Astronomical Journal, 159:22 (15pp), 2020 January https://doi.org/10.3847/1538-3881/ab5aeb

Covariance-regularized Reconstruction of Data Cubes in Integral Field Spectroscopy and Application to MaNGA Data

Dou Liu[1], Michael R. Blanton[1], and David R. Law[2]
[1] Center for Cosmology and Particle Physics, Department of Physics, New York University, 4 Washington Place, New York, NY 10003, USA
[2] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
Received 2019 June 25; revised 2019 November 18; accepted 2019 November 19; published 2019 December 17

Abstract

Integral field spectroscopy can map astronomical objects spatially and spectroscopically. Due to instrumental and atmospheric effects, it is common for integral field instruments to yield a sampling of the sky image that is both irregular and wavelength dependent. Most subsequent analysis procedures require a regular, wavelength-independent sampling (for example a fixed rectangular grid), and thus an initial step of fundamental importance is to resample the data onto a new grid. The best possible resampling would produce a well-sampled image, with a resolution equal to that imposed by the intrinsic spatial resolution of the instrument, telescope, and atmosphere, and with no statistical correlations between neighboring pixels. A standard method in the field to produce a regular set of samples from an irregular set of samples is Shepard's method, but Shepard's method typically yields images with a degraded resolution and large statistical correlations between pixels. Here we introduce a new method, which improves on Shepard's method in both these respects. We apply this method to data from the Mapping Nearby Galaxies at Apache Point Observatory survey, part of Sloan Digital Sky Survey IV, demonstrating a full width at half maximum close to that of the intrinsic spatial resolution (and ∼16% better than Shepard's method) and low statistical correlations between pixels. These results nearly achieve the ideal resampling. This method can have broader applications to other integral field data sets and to other astronomical data sets (such as dithered images) with irregular sampling.

Unified Astronomy Thesaurus concepts: Observational astronomy (1145); Galaxies (573); Astronomy data reduction (1861); Spectroscopy (1558)

1. Introduction

Integral field spectroscopy (IFS) yields a rich set of information about extended objects such as galaxies. Modern facilities for performing IFS include the Multi Unit Spectroscopic Explorer (Laurent et al. 2006), the Keck Cosmic Web Imager (Morrissey et al. 2012), and several efforts to create large samples of nearby galaxies, such as the Calar Alto Legacy Integral Field Area Survey (CALIFA; Sánchez et al. 2012), the Sydney-AAO Multi-object IFS (SAMI; Croom et al. 2012) and Mapping Nearby Galaxies at Apache Point Observatory (MaNGA; Bundy et al. 2015).

Here, we consider the MaNGA survey, one of three core programs in the fourth-generation Sloan Digital Sky Survey (SDSS-IV; Blanton et al. 2017). MaNGA uses integral field units (IFUs) consisting of optical fiber bundles to obtain spectra across the face of a sample of 10,000 low redshift galaxies, making it possible to map the spectroscopic properties of galaxies and to interpret this spectroscopy in terms of two-dimensional maps of stellar age, gas phase and stellar phase elemental abundances, star formation histories, and kinematics.

The MaNGA fiber bundles are arranged in a hexagonal grid with a separation of 151 μm. Each fiber has a cladding and buffer in its outer annulus and has an active core size of 120 μm, corresponding to 2″ in the Sloan Foundation Telescope focal plane. This configuration results in an effective filling factor for the active cores of 56% (Law et al. 2015). With a typical seeing FWHM of 1.″5 at Apache Point Observatory (APO), a single observation undersamples the sky image considerably. Thus, to increase the sampling of the point-spread function (PSF) and to avoid sampling irregularities, MaNGA uses dithered observations. Ideally, each observation consists of three dithered exposures, performing a set of subsequent exposures on each side of an equilateral triangle. This pattern of observations leads to a finer, but still hexagonal, pattern on the sky at the guiding wavelength. In the ideal case when all observations are taken at the same hour angle, the dithers produce a hexagonal pattern at all wavelengths, with an overall wavelength-dependent shift of the hexagonal pattern on the sky due to chromatic differential refraction by the atmosphere. If the observations are taken at different hour angles, the set of dithers will not produce a perfect hexagonal pattern of samples at all wavelengths. Thus, generally speaking, the MaNGA data produce an irregular sampling of fluxes on the sky that varies with wavelength.

Most analysis techniques are designed to handle regular, usually rectangular, grids. Except for analyses involving a full "forward modeling" of the MaNGA data, resampling the fluxes onto such a regular grid spatially is important. The resulting product is referred to as a data cube. These data cubes are the primary product of the MaNGA data reduction pipeline (DRP; Law et al. 2016). The DRP extracts the individual fiber spectra and calibrates them, producing row-stacked spectra (RSS)—one spectrum per fiber all interpolated onto the same wavelength grid. At each wavelength, MaNGA then uses a modified version of Shepard's method (Shepard 1968) to resample the RSS data for that wavelength onto a rectangular grid. This method, which is a flux-conserving variation of Shepard's interpolation method, is widely adopted because of its simplicity and robustness (Yang et al. 2004), for example in the CALIFA pipeline (e.g., Sánchez et al. 2012).

In this paper we reconsider this choice of resampling method and propose an alternative. Before beginning, we ask: what is the best performance we should expect from a resampling

algorithm? Here we define this "best performance" as being equivalent to actually sampling the real sky with the real instrument at the location of every grid point in the desired resampling. This process would produce an image with a resolution equivalent to the intrinsic spatial resolution of the instrument, telescope, and atmosphere, and zero statistical correlations between the errors in each pixel. This definition means we are not attempting to deconvolve the intrinsic resolution—but to resample with as little loss of resolution or introduction of statistical correlations as possible.

Relative to this ideal, Shepard's method suffers two problems. First, it broadens the final reconstructed PSF substantially. In the cases we consider in this paper, the FWHM of the final PSF can be 20% broader than the intrinsic resolution. Second, the errors in pixels in the reconstruction are highly correlated with each other. In the cases we consider in this paper, for Shepard's method pixels separated by $\sim$2 spaxels (1″) can have correlation coefficients between their errors as high as 0.70, complicating any correct analysis of the images that accounts for errors. These deficiencies drive us to find a way to improve the performance of the resampling.

To do so, we make use of the techniques developed in different contexts by Hamilton (1997); (for decorrelation of galaxy power spectra) and Bolton & Schlegel (2010); (for decorrelation of extraction of optical spectra). The latter term their method "spectroperfectionism," alluding to the fact that they are performing a spectroscopic extraction. However, we prefer the name "covariance-regularized reconstruction" as more generally applicable, since the approach of Bolton & Schlegel (2010) is appropriate for a number of other contexts beyond spectroscopic extraction or (as applied here) IFS image reconstruction. As far as we are aware, this general approach has not been previously applied in the context that we explore here of two-dimensional spatial reconstruction. We would like to point out that neither Shepard's method nor our method intend to decorrelate in the spectral dimension.

We will make some important distinctions between different terms in this paper to avoid confusion:

1. *Kernel*: The response $K(x, y)$ to a delta-function source of the atmosphere, telescope, and instrument combined. In the specific case of MaNGA, the kernel is the atmospheric seeing convolved with the telescope's optical response convolved with the fiber profile (which we approximate as a 2″ diameter top hat).
2. *Kernel-convolved image*: The actual or model image on the sky convolved with the kernel.
3. *Shepard's image*: The resampling resulting from Shepard's method. Referred to as a vector of pixel values $\boldsymbol{S}$.
4. *Deconvolved reconstruction*: A model of the image on the sky with the kernel deconvolved. Referred to as a vector of pixel values $\boldsymbol{F}$ (never used directly in the final result).
5. *Covariance-regularized Reconstruction (CRR)*: The resampling resulting from our method. Referred to as a vector of pixel values $\boldsymbol{G}$.
6. *Point-spread function (PSF)*: In our usage, the PSF will refer to the response to a delta function on the sky of the output image from the analysis, either Shepard's image or the CRR. In an ideal reconstruction, the PSF would be identical to the kernel values at each pixel (using the delta-function location as the center of the kernel function).

This article is organized as follows. Section 2 describes the resampling methods we consider here. Section 3 presents a series of tests of this method for a simulation. Section 4 presents a demonstration of the method for real data. Section 5 summarizes and discusses the results.

2. Image Resampling Methods

2.1. Shepard's Method, the MaNGA DRP Standard

A specialized version of Shepard's method is described in Sánchez et al. (2012) in the context of CALIFA and in Law et al. (2016) in the context of MaNGA. The input data from the IFU survey consists of the flux intensity $f[i]$ and the variance of the error $N[i] = \sigma[i]^2 = \langle \Delta f[i]^2 \rangle$ in each fiber. For a system with $N_{\rm fiber}$ fibers, with which we have taken $N_{\rm exp}$ exposures, for each wavelength channel there are $N_0 = N_{\rm fiber} \times N_{\rm exp}$ values, each corresponding to a different location on the sky. We will refer below to each such observation as a fiber-exposure. For the MaNGA DRP, the output grid is rectangular with a pixel size 0.″5 per spaxel, and we write the total number of pixels as M_0.

The transformation from intensities at the irregularly located fiber locations to the Shepard's image is:

$$S[j] = \sum_{i=1}^{N_0} W[j, i] f[i] \tag{1}$$

where $\boldsymbol{S}$ is Shepard's image, and the $M_0 \times N_0$ matrix $W[j, i]$ is the weight of each fiber location i contributing to the output grid point j. In Shepard's method, the weight function is a circularly symmetric Gaussian that depends on the distance $r[i, j]$ between the fiber location i and the grid point j:

$$W[j, i] = \frac{b[i]}{W_0[j]} \exp\left(-\frac{r[i, j]^2}{2\sigma_0^2}\right), \tag{2}$$

for $r[i, j] < r_{\rm lim}$, and zero otherwise. σ_0 defines the width of the Gaussian function, and for the MaNGA DRP is set to 0.″7. The MaNGA DRP takes $r_{\rm lim} = 1.''6$. The normalization parameter $W_0[j]$ is defined as the sum of the N_0 weights for each output grid point j, to guarantee the conservation of flux:

$$W_0[j] = \sum_i b[i] \exp\left(-\frac{r[i, j]^2}{2\sigma_0^2}\right), \tag{3}$$

over all pixels i for which $r[i, j] < r_{\rm lim}$. In Equations (2) and (3), $b[i]$ is zero if the inverse variance $N[i]^{-1} = 0$ and is unity otherwise.

2.2. Covariance-regularized Reconstruction

Our method can be written in the same form as Shepard's method (Equation (1)), but with a different choice of $W[j, i]$. The choice is motivated in the following manner, which explicitly designs the weights so that the final reconstruction remains consistent with the samples and so that its covariance matrix has small off-diagonal entries. We refer to an image using our method as a CRR.

Consider fitting a linear model to reconstruct the observables $f[i]$. Our model consists of a set of delta functions with fluxes $F[j]$, distributed on a regular grid (the same grid we want to use for the reconstruction). Each fiber-exposure i observes this function convolved with the kernel appropriate for that

particular observation, $K_i(x, y)$. In the case of MaNGA, the kernel is the convolution of the atmospheric seeing, the telescope optical response, and the fiber. This model for the observables can be written as

$$m[i] = \sum A[i, j]F[j], \tag{4}$$

where:

$$A[i, j] = K_i(x_i - X_j, y_i - Y_j) \tag{5}$$

is the value of the kernel at the separation between the model pixel j at (X_j, Y_j) and the fiber-exposure i at (x_i, y_i). The kernel is a known property of the observations based on the estimated seeing at the time of each observation.

We can fit for the model parameters $F[j]$ by minimizing the χ^2 error:

$$\chi^2 = \sum_i \frac{(f[i] - m[i])^2}{\sigma[i]^2}. \tag{6}$$

Because the model is linear in $F[j]$, the solution can be written:

$$\boldsymbol{F} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1}\boldsymbol{f}, \tag{7}$$

where $\boldsymbol{N}$ is the data covariance matrix, which in our case is diagonal with diagonal entries $\sigma[i]^2$.

Therefore, we can estimate $\boldsymbol{F}$, which is analogous to a deconvolved image based on the fiber samples, because it is the sky image before convolution with the kernel. As one expects for a deconvolution, the error covariance for $\boldsymbol{F}$ is highly nondiagonal—there are strong correlations and anticorrelations in the errors of neighboring pixels. The covariance matrix is

$$\boldsymbol{C} = \langle \Delta\boldsymbol{F}\Delta\boldsymbol{F}^{\mathrm{T}} \rangle = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1}\boldsymbol{A})^{-1}. \tag{8}$$

Using $\boldsymbol{F}$ for science is very undesirable because of these correlations, which would complicate any error analysis but also lead to large fluctuations among the values of $\boldsymbol{F}$.

A common technique is to regularize the values of $\boldsymbol{F}$, either under a Tikhonov regularization, a maximum entropy criterion, or something else (e.g., Warren & Dye 2003). However, here we take a different approach, which is to regularize the covariance such that it is diagonal, which turns out to be similar to reconvolving $\boldsymbol{F}$ to a resolution similar to the kernel.

The covariance matrix can be whitened through a linear transformation of $\boldsymbol{F}$ that can be found by taking a square root of the inverse covariance matrix $\boldsymbol{C}^{-1} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1}\boldsymbol{A}$. For this symmetric and positive definite matrix, we can take its square root by finding its eigensystem:

$$\boldsymbol{C}^{-1} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1} = (\boldsymbol{P}\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{P}^{-1})(\boldsymbol{P}\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{P}^{-1}) = \boldsymbol{Q}\boldsymbol{Q}, \tag{9}$$

where $\boldsymbol{D}$ here is a diagonal matrix of eigenvalues and $\boldsymbol{P}$ is the matrix of eigenvectors. For $\boldsymbol{D}^{\frac{1}{2}}$, the positive root is always chosen.

Another path to finding the same matrix $\boldsymbol{Q}$ is through the singular value decomposition (SVD):

$$\boldsymbol{N}^{-\frac{1}{2}}\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}. \tag{10}$$

where $\boldsymbol{\Sigma}$ is diagonal, $\boldsymbol{V}$ is orthogonal ($\boldsymbol{V}\boldsymbol{V}^{\mathrm{T}} = 1$), and $\boldsymbol{U}$ is close to orthogonal ($\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}$ is diagonal with ones for dimensions i with $\Sigma_i \neq 0$ or zeros for dimensions i with $\Sigma_i = 0$). In this case:

$$\boldsymbol{F} = \boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^{T}\boldsymbol{N}^{-\frac{1}{2}}\boldsymbol{f} \tag{11}$$

if $\Sigma_i \neq 0$ for all i. We can also use the standard Moore–Penrose inverse technique and set $\Sigma_i^{-1} = 0$ for $\Sigma_i = 0$, which allows us to handle truly degenerate cases smoothly. We can show that

$$\boldsymbol{C}^{-1} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}} \tag{12}$$

and therefore,

$$\boldsymbol{Q} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}. \tag{13}$$

Thus, we can find $\boldsymbol{Q}$ without ever constructing the full covariance matrix or its inverse, or explicitly finding its eigenvectors, which may be useful in cases when the covariance matrix is ill conditioned.

The linear transformation of $\boldsymbol{F}$ that we will use is not quite $\boldsymbol{Q}$, because we want the transformation to conserve the total flux in $\boldsymbol{F}$. We can achieve this goal by normalizing over each row to give the transformation matrix:

$$R[i, j] = \frac{1}{\sum_j Q[i, j]} Q[i, j]. \tag{14}$$

$\boldsymbol{R}$ is then a linear transformation under which $\boldsymbol{C}$ can be diagonalized:

$$\boldsymbol{C}_G = \boldsymbol{R}\boldsymbol{C}\boldsymbol{R}^{\mathrm{T}} \tag{15}$$

with the entries of the diagonal matrix $\boldsymbol{C}_G$ given by

$$C_G[i, i] = \left(\sum_j Q[i, j] \right)^{-2}. \tag{16}$$

(In detail, the resulting covariance matrix is not precisely diagonal if any $\Sigma_i = 0$.) This result means that if we define:

$$\begin{aligned} \boldsymbol{G} &= \boldsymbol{R}\boldsymbol{F} \\ &= \boldsymbol{R}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1}\boldsymbol{f} \\ &= \boldsymbol{R}\boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{N}^{-\frac{1}{2}}\boldsymbol{f} \\ &\overset{\mathrm{def}}{=} \boldsymbol{W}\boldsymbol{f} \end{aligned} \tag{17}$$

then $\boldsymbol{C}_G$ is the covariance matrix of $\boldsymbol{G}$. Multiplying by the matrix $\boldsymbol{R}$ turns out to be similar to a reconvolution of $\boldsymbol{F}$ with the kernel, although it is not strictly speaking a convolution.

$\boldsymbol{G}$ will be our reconstruction, which by design should have close to a diagonal covariance matrix. We will show later that in the context of MaNGA it also has a sharper PSF than Shepard's method. Like Shepard's method it is just a linear combination of the input fluxes and so can be written in the same form as Equation (1).

There are some adjustments we will make to this method. The first adjustment is that we will apply a regularization term to handle singular values. This adjustment will make the procedure more numerically robust but have almost no impact on the results. We can think of this regularization term in terms of an adjustment to the function we are minimizing in Equation (6) with a quadratic term, the simplest version of Tikhonov regularization:

$$\chi^2 = \sum_i \frac{[f[i] - m[i])]^2}{\sigma[i]^2} + \lambda^2 \sum_j F[j]^2. \tag{18}$$

The solution $\boldsymbol{F}$ is altered under this regularization to:

$$\begin{aligned} \boldsymbol{F} &= (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1}\boldsymbol{A} + \lambda^2\boldsymbol{I})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{N}^{-1}\boldsymbol{f} \\ &= \boldsymbol{V}\boldsymbol{\Sigma}_*^{-1}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{N}^{-\frac{1}{2}}\boldsymbol{f} \end{aligned} \tag{19}$$

where $\boldsymbol{\Sigma}_*$ is diagonal and

$$\Sigma^{-1}_{*,ii} = \frac{\Sigma_{ii}}{\Sigma^2_{ii} + \lambda^2}, \tag{20}$$

and the solution $\boldsymbol{G}$ becomes

$$\boldsymbol{G} = \boldsymbol{R}\boldsymbol{V}\boldsymbol{\Sigma}^{-1}_* \boldsymbol{U}^{\mathrm{T}}\, \boldsymbol{N}^{-\frac{1}{2}}\boldsymbol{f}. \tag{21}$$

The second adjustment is that we will not use the actual variance matrix $\boldsymbol{N}$ created from the observed fiber spectra. If we did so, this would introduce a correlation between the brightness of an object (which for bright sources dominates the true variance vector) and the values of the corresponding column of the weight matrix $\boldsymbol{W}$. This correlation would lead to an undesirable dependence of the reconstructed PSF on the brightness of the source that would greatly complicate analysis of our data cubes. Additionally, it is well known that using an inverse variance–based weight function can produce a systematic underestimate of the derived fit when the variances are derived from the signal itself. (see, e.g., Section 5.1 of Mighell 1999; Law et al. 2016). For these reasons, we use $\tilde{\boldsymbol{N}}$ in our calculations, which equals unity where $\boldsymbol{N}^{-1}$ is not zero, and zero where $\boldsymbol{N}^{-1}$ is zero.

A third adjustment we make is to slightly alter the kernel to remove contributions for which the radius is larger than $4''$. This choice makes little difference in the final result but makes the calculation considerably faster.

The fourth adjustment we make is to remove from the SVD calculation any model pixels j that are further than some distance r_{lim} (which we here set to $1\farcs6$, the same as DRP's choice) from any fiber—basically, any pixels not well-constrained by the data—which lowers the condition number of $\boldsymbol{A}$ and makes the SVD more stable.

The final method then can be written as follows:

$$\boldsymbol{G} = \boldsymbol{W}\boldsymbol{f}, \tag{22}$$

where

$$\boldsymbol{W} = \boldsymbol{R}\boldsymbol{V}\boldsymbol{\Sigma}^{-1}_* \boldsymbol{U}^{\mathrm{T}} \tilde{\boldsymbol{N}}^{-\frac{1}{2}}, \tag{23}$$

in which $\boldsymbol{V}$, $\boldsymbol{U}$, and Σ are from the SVD of $\tilde{\boldsymbol{N}}^{-1/2}\boldsymbol{A}$, Σ_* is defined in Equation (20), and $\boldsymbol{R}$ is defined by Equations (13) and (14). The covariance matrix of $\boldsymbol{G}$ is then:

$$\boldsymbol{C}_G = \langle \boldsymbol{W}\Delta\boldsymbol{f}\Delta\boldsymbol{f}^{\mathrm{T}}\boldsymbol{W}^{\mathrm{T}}\rangle = \boldsymbol{W}\boldsymbol{N}\boldsymbol{W}^{\mathrm{T}}. \tag{24}$$

Here, $\boldsymbol{N}$ is the diagonal variance matrix from the individual fibers. This covariance matrix is not guaranteed to be diagonal, but because of how we have constructed the weights it will prove to be much closer to diagonal than Shepard's method.

When operating on the resulting data cube for science, it is important to identify unreliable spaxels about which there is little information provided by the fiber data. The MaNGA pipelines use two masks, `NOCOV` to indicate that there is no information about the pixel, and `LOWCOV` to indicate that there is a low amount of information about the pixel. These masks are based on the fiber-level maskbits, which record the effect of hot pixels, cosmic ray hits, and other effects that can make the data from a fiber unusable at some or all wavelengths. Our approach to defining these masks for the data cube spaxels is different from the standard MaNGA pipeline. The MaNGA pipeline counts the total contribution of fibers to each pixel, based on the $\boldsymbol{W}$ matrix for Shepard's method, to identify poorly constrained spaxels. We use a different approach, which is to use the diagonal elements of the covariance matrix when we assume constant unit noise in each fiber that is not masked:

$$\tilde{\boldsymbol{C}} = \boldsymbol{W}\tilde{\boldsymbol{N}}\boldsymbol{W}^{T}. \tag{25}$$

This covariance matrix is not the same as the actual covariance $\boldsymbol{C}_G$. We set spaxels with a variance more than twice the median variance to `LOWCOV`, so that they may be ignored in scientific analysis. This procedure appropriately masks spaxels too near the edge of the fiber bundle or that are affected by bad fibers.

There are several free parameters in the CRR method as applied to MaNGA, which we now discuss. First, there is the shape of the kernel, which is determined by our estimate of the observing conditions at each exposure. Second, there are the conditions for dropping edge pixels. Third, there is the regularization parameter λ. Fourth, there is the pixel scale of CRR. The first three parameters prove to have no significant effect on the results for MaNGA, as we show later in this paper. We will explore below the effect of the reconstruction pixel scale.

Some general aspects of the method, compared to Shepard's method, are worth noting before describing those tests. Although our method is also just a linear combination of the fluxes, the weights are not determined by a stationary function as they are for Shepard's method. They are also not restricted to be nonnegative. These properties, particularly the latter, are essential to reducing the off-diagonal covariances and producing a PSF close to the kernel resolution.

Although we motivated the method based on the model fit to the parameters $\boldsymbol{F}$ expressed in Equation (6) and the diagonalization of their covariance matrix $\boldsymbol{C}$, we never need to explicitly determine either $\boldsymbol{F}$ or $\boldsymbol{C}$. We will nevertheless calculate these quantities below in order to demonstrate their properties.

Our approach is mathematically identical to the approach proposed by Bolton & Schlegel (2010) in the different context of spectroscopic extraction. Our notation differs somewhat from theirs. Specifically, their $\boldsymbol{p}$ is our $\boldsymbol{f}$, their $\boldsymbol{f}$ is our $\boldsymbol{F}$, and their $\tilde{f}$ is our $\boldsymbol{G}$.

3. Tests on Simulated Data

In this section, we use simulations to characterize the performance of Shepard's method, which is what the MaNGA DRP uses, and our CRR method.

3.1. General Information

MaNGA uses a hexagonal pattern of fibers to detect flux as a function of position on the sky. The hexagonal pattern is dithered in position between different exposures, and varies on the sky with wavelength. Our goal is to resample the fluxes onto a rectilinear grid. In Figure 1, the blue points show the fiber locations at $\lambda = 5500$ Å for the plate-IFU 8720-1901, shown as a function of X and Y position in the focal plane (corresponding to R.A. and decl.). The red points show the pixels we are using in the reconstruction. As explained above, we exclude pixels when their distance to all the fibers is larger than $1\farcs6$. In this figure, we show both $0\farcs5$ and $0\farcs75$ pixel^{-1} scales. We will examine the effect of difference choices of pixel scales on our reconstruction later.

We will simulate observations for a point source to test the methods. Since the methods we test are linear, the point source responses at different locations can fully characterize the

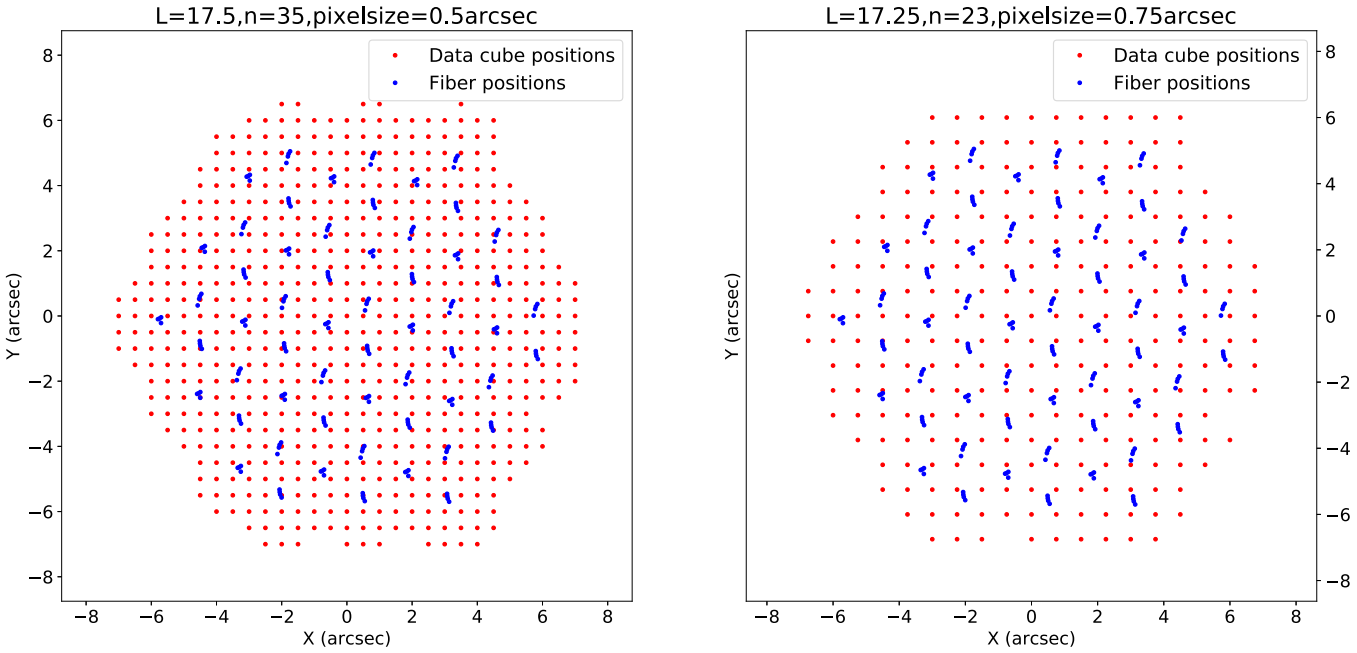

Figure 1. Configuration of plate-IFU 8720-1901 with two different pixel scales and for wavelength $\lambda = 5500$ Å. Blue points are IFU fiber locations, red points are pixels in the reconstruction. The left panel shows pixel scale $0\farcs5$ pixel^{-1}, and the right panel shows pixel scale $0\farcs75$ pixel^{-1}.

performance. The flux for each fiber is sampled from the kernel function K as the response of a point source at some location (in most cases, we use $X = Y = 0$).

Panel (a) of Figure 2 shows our adopted kernel for wavelength $\lambda = 5500$ Å. We use a double Gaussian function to simulate the atmospheric seeing, and convolve with the fiber profile (a $2''$ top hat) to generate the kernel function. The normalized double Gaussian is defined by the standard deviation σ_1 for the inner Gaussian, the ratio σ_2/σ_1 of the outer to inner Gaussian standard deviation, and the ratio A_2/A_1 of the central values of the outer and inner Gaussians. We use a typical pair of ratios that approximates the inner parts of a Moffat-like atmospheric PSF (Law et al. 2015; J. Gunn 2019, private communication):

$$\frac{\sigma_2}{\sigma_1} = 2$$
$$\frac{A_2}{A_1} = \frac{1}{9}. \quad (26)$$

The FWHM of the resulting model PSF is:

$$\mathrm{FWHM} = 1.05\,\mathrm{FWHM}_1 = 2.473\sigma_1. \quad (27)$$

The MaNGA DRP reports the FWHM of atmospheric seeing for each exposure at the guider wavelength $\lambda_0 = 5400$ Å. For other wavelengths, we will assume the seeing varies as $\lambda^{-1/5}$ (Yan et al. 2016).

For each simulation, the fiber locations are provided by a MaNGA observation with several exposures, for which positions accounting for a variety of observational effects (bundle metrology, dithering, chromatic and field differential refraction, etc.) are stored in the MaNGA data products, for example those shown in Figure 1. Normally the number of MaNGA exposures ranges from 6 to 21 in order to achieve uniformity and the required signal-to-noise ratio (S/N).

Each pixel in the output image is a weighted sum of the values sampled for each fiber-exposure. Figure 3 shows an example for the central pixel. For Shepard's weights (right panel), the weights for each fiber-exposure are just a decreasing function of distance from the pixel. For the CRR, the weights are both positive and negative and have an oscillatory nature, similar to that found in sinc-interpolation methods.

Whereas the fluxes sampled from the kernel are noiseless, the fluxes of the actual observations are not. The main sources of noise are Poisson noise in the number of electrons due to the object, sky, and dark current, plus the read noise from the amplifiers. For the bulk of locations and wavelengths in MaNGA, the noise is object dominated. When we include simulated noise in our tests, we concentrate on this regime, so that the noise is proportional to the square root of the flux. We characterize the S/N of the simulations based on that of the fiber with the maximum simulated (noiseless) flux $f_{\max}$. For a chosen S/N we then define a scale factor s converting flux to number of photons $N_p = fs$, so that:

$$S/N_p = \frac{f_{\max}\cdot s}{\sqrt{f_{\max}\cdot s}} = \sqrt{f_{\max}\cdot s}\,. \quad (28)$$

Then for each fiber we apply Poisson noise based on the resulting N_p for each fiber.

3.2. Nominal Case

We start our tests with the nominal case at $0\farcs5$ pixel^{-1}, in order to compare directly at the MaNGA DRP pixel scale. Figure 2 shows the response of a point source in the center, averaged over the g-band wavelength range. Panels (b)–(d) show, respectively, MaNGA's DRP data cube (which uses Shepard's method), the CRR image, and our implementation of Shepard's image, all at $0\farcs5$ pixel^{-1}. Panels (e) and (f) show the CRR image and Shepard's image at $0\farcs75$ pixel^{-1}. As we said previously, panels (b) and (d) are almost the same because they use the same method. We show the DRP result in (b) to confirm that we are analyzing the data consistently.

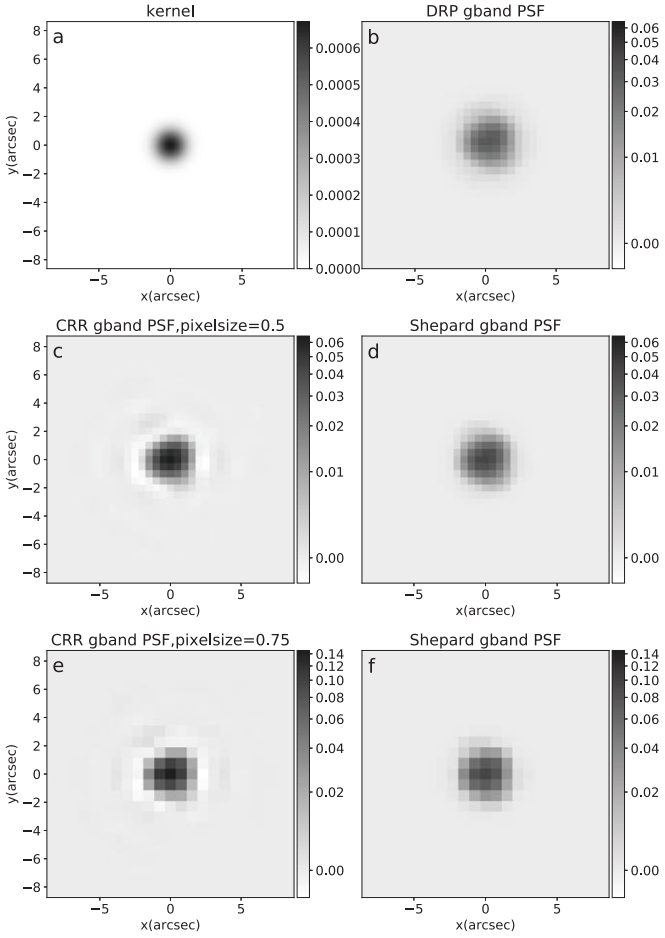

Figure 2. Data cube slices from simulated observations of a point source generated for plate-IFU 8720-1901. The flux is averaged over the g band, and we assume seeing at 5500 Å is around $1\farcs19$. Panel (a): the kernel for wavelength $\lambda = 5500$ Å. Panel (b): the MaNGA DRP implementation of Shepard's method ($0\farcs5$ pixels). Panel (c): reconstruction using the CRR method ($0\farcs5$ pixels). Panel (d): our implementation of Shepard's method ($0\farcs5$ pixels). Panel (e): reconstruction using the CRR method ($0\farcs75$ pixels). Panel (f): our implementation of Shepard's method ($0\farcs75$ pixels). The image scaling is linear for Panel (a), but is arcsinh-scaled for Panels (b)–(f) to show the behavior in the PSF wings.

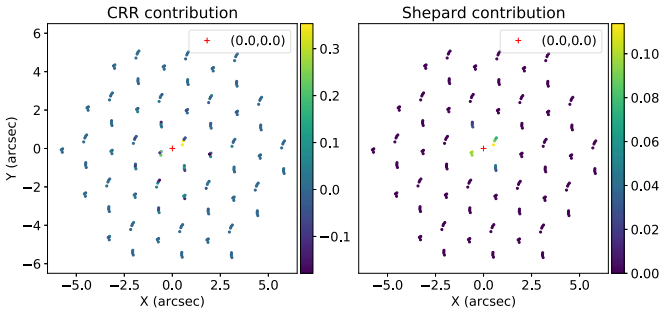

Figure 3. The contribution of all fibers on the pixel at (0, 0) for plate-IFU 8720-1901, which is a row of weights matrix $\boldsymbol{W} \cdot \lambda = 5500$ Å. Left panel: our CRR image. Right panel: Shepard's method. The red cross is the central pixel evaluated in the graph.

We compare the performance of the methods quantitatively by measuring the FWHM and also a pseudo-Strehl ratio for the resulting PSF. We define the pseudo-Strehl ratio as the ratio of the peak image data cube slice intensity compared to the center of the kernel-convolved point source image. This is the ratio of

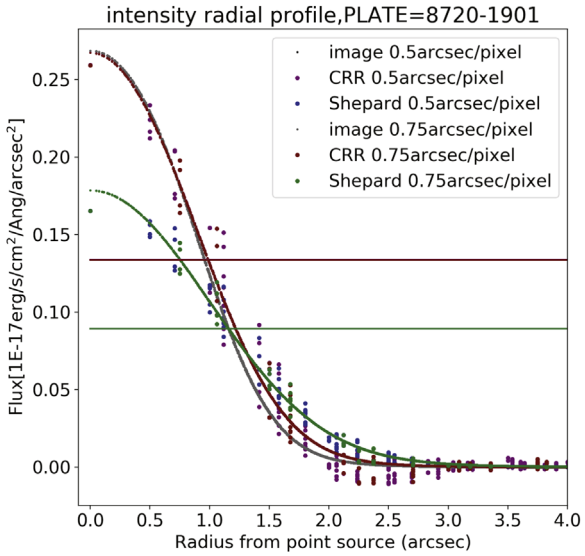

Figure 4. Profile of the PSF in the output images from the simulated data in Figure 2. Points are flux values per pixel throughout the image at $0\farcs5$ and $0\farcs75$ pixel^{-1}, for wavelength $\lambda = 5500$ Å. The horizontal lines indicate the half maximum for each radial profile fit.

Table 1
Sharpness Measures of PSF

Analysis Item	Pixel Scale Per Pixel	Kernel	Shepard's	CRR
FWHM(arcsec)	0.5	1.905	2.316	1.973
Strehl ratio	0.5	1	0.664	0.996
FWHM(arcsec)	0.75	1.905	2.317	1.973
Strehl ratio	0.75	1	0.664	0.996

the peak intensity to the largest possible given the resolution of the instrument. We show here a single typical wavelength slice at $\lambda = 5500$ Å with seeing FWHM around $1\farcs19$, but we find similar results throughout the whole spectrum and various plate-IFUs.

Figure 4 shows the profile of the intensity of the resulting images. To determine a FWHM reliably, we fit a model to the pixel values. In particular, we use the kernel function K as our model, which is parameterized by a choice of seeing. We use least-squares to find the best-fitting seeing value, and infer the FWHM and the pseudo-Strehl ratio from the corresponding kernel. In analyzing the MaNGA DRP results, Law et al. (2016) used a Gaussian fit instead; we find that this is adequate for a broad PSF but not for the CRR image PSF. The results for plate-IFU 8720-1901, including the FWHM and the Strehl ratio for the kernel, for Shepard's method, and for the CRR method, are listed in Table 1.

For consistency, given a pair of Strehl ratios or FWHM values v_1 and v_2 for the Shepard and CRR images, we define the improvement in this quantity as $\frac{2\,|\,v_2 - v_1\,|}{v_2 + v_1}$. The radial profile shows that there is a 16.0% improvement in the FWHM between Shepard's method and CRR, which turns out to be a typical level of improvement. Denoting the kernel FWHM as v_0, the maximum possible improvement of the FWHM would be $\frac{2(v_2 - v_0)}{v_2 + v_0} = 19.5\%$. The pseudo-Strehl ratio increase is around 40%. Our results are very close to the best possible performance without performing a deconvolution (i.e., very

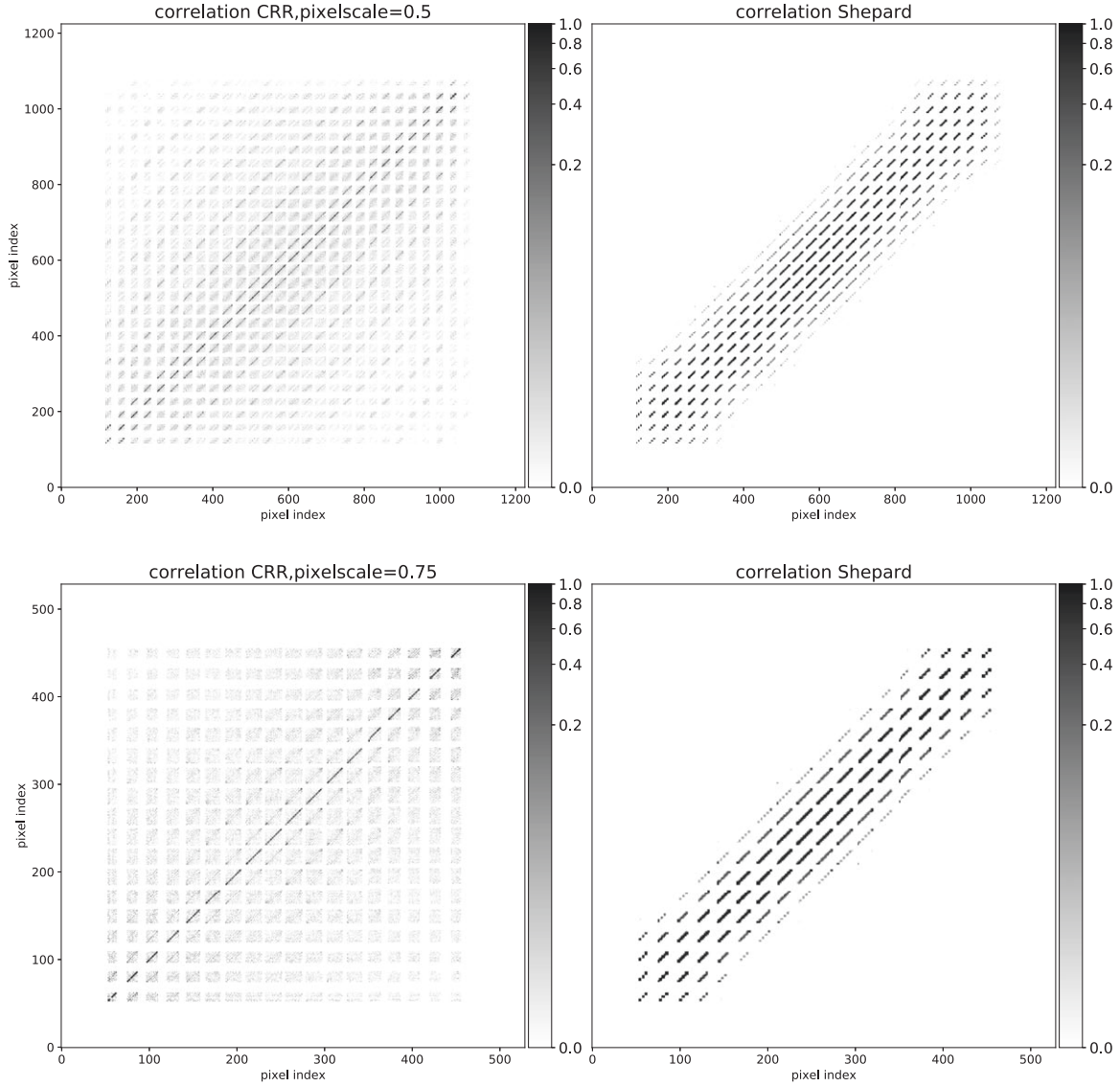

Figure 5. The correlation matrix of CRR's and Shepard's result for plate-IFU 8720-1901, both shown at $0\farcs5$ and $0\farcs75$ pixel^{-1}, for wavelength $\lambda = 5500$ Å. The image scaling is arcsinh to better see the small off-diagonal elements. Left panel: CRR result. The correlation coefficients between values separated by two or more pixels are typically of the order 10^{-2} or less. Right panel: Shepard's result.

close to the kernel). We note in passing that we were surprised that the kernel had a FWHM slightly less than $2''$, given that the fiber top-hat FWHM is precisely $2''$; however, this result is correct—slightly blurring the top hat with a Gaussian slightly reduces the FWHM.

Next we compare the covariance between pixels for both methods. The covariance between pixels of the final image is expressed in Equation (24), for both our method and Shepard's method, depending only on which weights $\boldsymbol{W}$ are used. In order to factor out the magnitude of the covariance along the diagonal, we will examine the correlation matrix between pixels, defined in the usual way as $\rho_{jk} = C_{jk} \Big/ \sqrt{C_{jj} C_{kk}}$.

Figure 5 shows the resulting correlation matrix between pixels. We omit masked pixels, which are mostly at the corners. The image scaling is arcsinh to best examine the off-diagonal components. Shepard's output image has a broad covariance matrix, while our method is nearly diagonal, as expected from the previous section. We show pixel scales of $0\farcs5$ and $0\farcs75$ pixel^{-1}, and for the CRR covariances the difference in the level of correlation between these cases is clear. We consider the choice of pixel size and its effect on covariance more quantitatively in the next subsection.

3.3. Choice of Pixel Size

Here we discuss the choice of pixel size. The primary considerations are the sharpness of the resulting PSF (as quantified by FWHM and pseudo-Strehl ratio), the covariance between pixels, and the sampling of the image. The results shown in this section for plate-IFU 8720-1901 and $\lambda = 5500$ Å are representative of what we find at other wavelengths and in other plate-IFUs.

The sharpness of the PSF differs very little between the two pixel scales, as Table 1 demonstrates.

The correlation coefficients do tend to depend on pixel scale, as shown in Table 2, based on the central pixel of plate-IFU 8720-1901 at 5500 Å. In this table, we list the quadratic means of the correlation coefficients between the central pixel and the pixels which are separated from it by 1 pixel or 2 pixels. We examine the correlations more fully in Figure 6, which shows

Table 2
Correlation Coefficient between Pixels with Different Separations (Quadratic Mean of Values between the Central Pixel and Its Surrounding Pixels), for Shepard's Method and CRR

Separation Method	Pixel Scale arcsec/pixel	1 pixel	2 pixels
Shepard's	0.5	0.876	0.703
Shepard's	0.75	0.767	0.465
CRR	0.5	0.303	0.094
CRR	0.75	0.050	0.024

the full correlation between the central pixel and all others for both methods with the two pixel scales. Each image corresponds to a row of the correlation matrix. The CRR image outperforms the Shepard's image in either case. The CRR image has some ringing in its correlation matrix that is greatly reduced at a pixel scale of $0\farcs75\,\mathrm{pixel}^{-1}$. Statistical independence is desirable so that the fluxes in the cube can be used to fit models and to propagate errors without tracking a broad covariance matrix, which is complex and burdensome, or relying on approximated error-scaling relations (e.g., Law et al. 2016).

The final major consideration for pixel size is whether the pixels provide sufficient sampling of the image. At $0\farcs75\,\mathrm{pixel}^{-1}$, the FWHM of the images are about 2.5 pixels, which is above the usual rough guide for critical sampling. Although for unusual PSFs of the sort produced by this instrument there may yet be some poorly sampled power at $0\farcs75\,\mathrm{pixel}^{-1}$, sampling more densely leads to an increase in covariance between pixels, as shown above.

These considerations of sampling and covariance lead us to $0\farcs75\,\mathrm{pixel}^{-1}$ as the best choice for our application. This pixel scale provides decent sampling without inducing excess correlations between pixels or ringing behavior.

3.4. PSF as a Function of Source Position

The PSF response for both CRR and Shepard's is sensitive to the relative position on the sky of the point source and the fibers. We therefore need to characterize the PSF response across the face of the IFU and verify that we satisfy MaNGA's requirement that the PSF FWHM varies by less than 10%

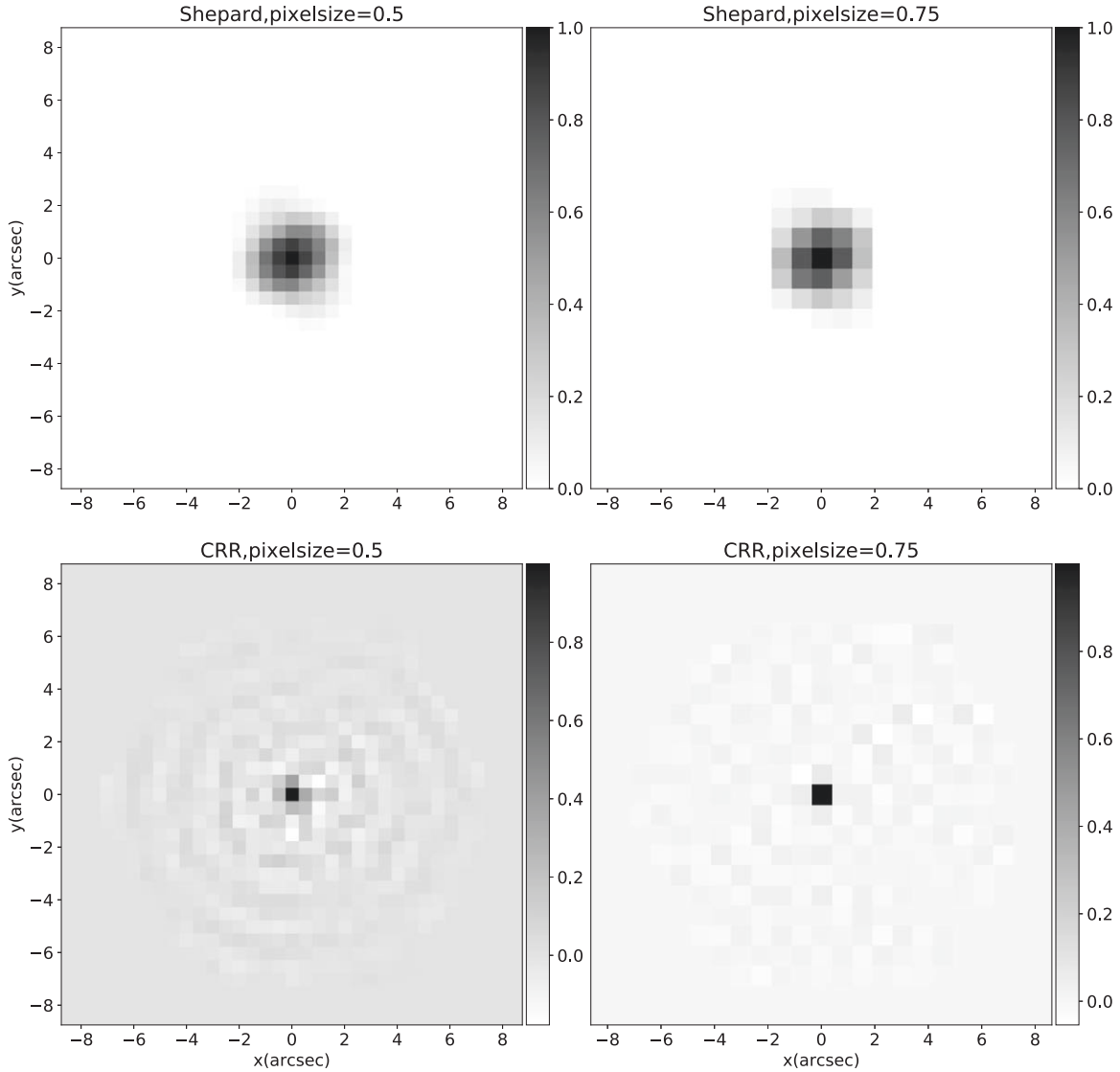

Figure 6. The correlation between the central pixel flux and that of other pixels for images produced from plate-IFU 8720-1901, at wavelength $\lambda = 5500$ Å. Upper panels: Shepard's method at $0\farcs5$ and $0\farcs75\,\mathrm{pixel}^{-1}$. Lower panel: our method at $0\farcs5$ and $0\farcs75\,\mathrm{pixel}^{-1}$.

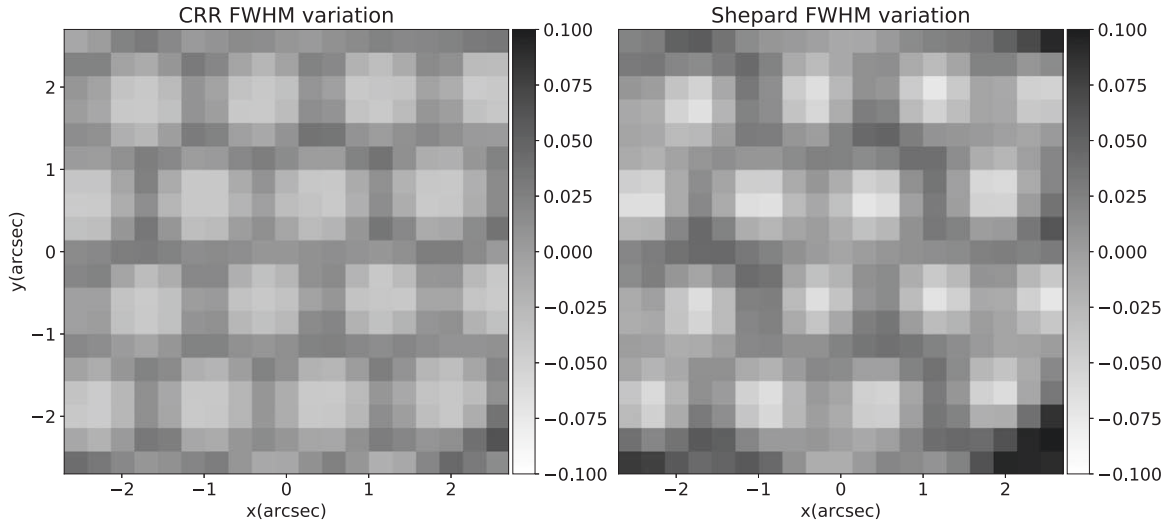

Figure 7. Fractional FWHM variation as a function of point source location for plate-IFU 8720-1901 at wavelength $\lambda = 5500$ Å. Left panel: CRR image, using $0\farcs75$ pixel^{-1}. Right panel: Shepard's image, using $0\farcs5$ pixel^{-1} to simulate the MaNGA DRP performance.

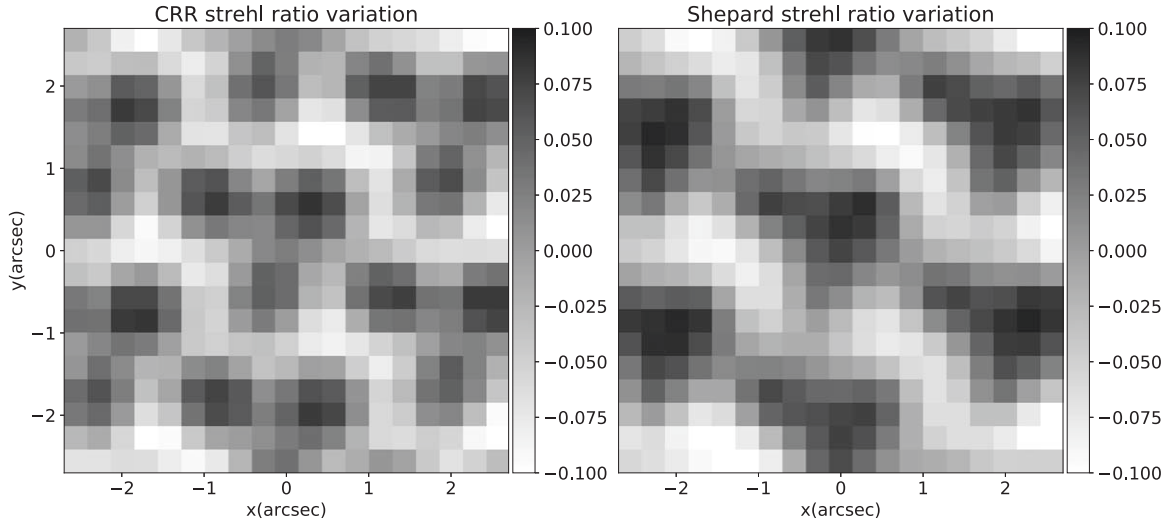

Figure 8. Similar to Figure 7, for the fractional Strehl ratio variation.

across the IFU. In Figure 7 we consider the FWHM homogeneity and in Figure 8 we consider the pseudo-Strehl ratio homogeneity. Below, we also discuss the homogeneity of the axis ratio b/a of the PSF.

Each figure uses a grid of locations of the point source, using the $0\farcs75$ pixel^{-1} CRR and the $0\farcs5$ pixel^{-1} Shepard's method results (to match MaNGA's implementation). We quantify the variation of each quantity relative to its median:

$$\delta\eta = \frac{\eta - \eta_0}{\eta_0}. \tag{29}$$

Here, η is the FWHM or the pseudo-Strehl ratio, and η_0 is its median value. As noted above, the CRR image is consistently higher resolution than the Shepard's image for point sources regardless of where they are relative to the fibers. In addition, in this case the CRR image also shows less fractional variation in resolution than Shepard's method, especially in the FWHM. In the example plate-IFU 8720-1901, the fractional standard deviation of the FWHM is 0.033 for Shepard's method, compared to 0.023 for our method. Meanwhile, the fractional standard deviation of the pseudo-Strehl ratio is 0.056 in Shepard's method, compared to 0.048 in our method. Thus, the CRR image slightly outperforms Shepard's image in terms of PSF homogeneity for this fiber bundle. This comparison depends on the location of the fiber-exposures; for the other fiber bundles we have tested, the CRR images always give similar or better homogeneity than Shepard's images for the FWHM and pseudo-Strehl ratio.

However, for the PSF roundness we find that Shepard's method outperforms CRR. We quantify the roundness with the axis ratio b/a of the PSF. For each point source location within a radial distance to the center equaling to the FWHM, we fit the PSF with a 2D Gaussian. The parameters of this fit yield the minor-to-major axis ratio. Following Law et al. (2015), we quantify the performance with $(b/a)_{99}$, which is the axis ratio for which 99% of the point source locations yield a PSF with $b/a > (b/a)_{99}$. This quantity is a conservative lower limit on the PSF roundness and its variation. In order to compare more directly to Law et al. (2015), we implemented our test with constant seeing. In our analysis, Shepard's image has $(b/a)_{99} = 0.91$, whereas the CRR image has $(b/a)_{99} = 0.85$.

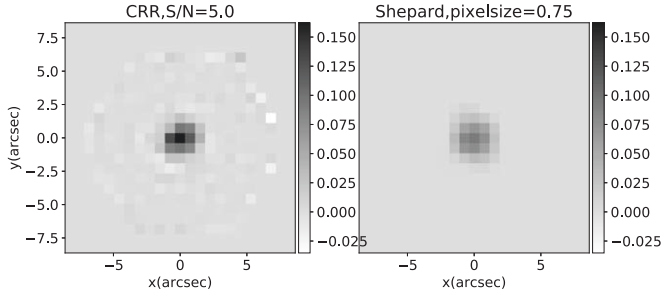

Figure 9. Simulated data cube slice with noise at $0\farcs5$ pixel^{-1} and S/N = 0.75 for plate-IFU 8720-1901 at wavelength $\lambda = 5500$ Å. Left panel: CRR method. Right panel: Shepard's method.

We were unable to reproduce the results of (Law et al. 2015) for Shepard's method, who found a substantially higher value of $(b/a)_{99} = 0.96$ using a very similar methodology to ours. This is partly because the coordinates we adopt are from an actual observational fiber bundle. If we use a three-point dithered hexagonal grid with seeing FWHM $= 1\farcs4$ to do the simulation, we get $(b/a)_{99} = 0.942$ with Shepard's reconstruction method. The homogeneity statistics parameters are consistently better than with the real observational coordinates. But we also found Shepard's method outperforms our result in the homogeneity.

3.5. Adding Noise

The simulated data we test on above is noiseless. However, we need to test our performance in the presence of realistic noise. The sources of noise in MaNGA are outlined in Section 3.1, but as noted there we will only consider the effect of Poisson noise due to the object signal, which is the most problematic (increasing with the signal rather than staying constant) and which is usually dominant.

Figure 9 shows a point source with noise, quantified by the S/N of peak flux in the reconstruction image, for both the CRR and Shepard's image. Obviously the image exhibits some noise. For Shepard's image, the result looks smoother, due to the larger off-diagonal covariances—i.e., the pixel-to-pixel fluctuations are reduced because neighboring pixels are correlated.

Figure 10 examines the FWHM and pseudo-Strehl ratios under a range of noise conditions. The CRR image remains sharper and brighter in the center regardless of the noise level.

3.6. Effect of an Inaccurate PSF Model

We build our analysis based on the assumption that the kernel is the same as our assumed one—that is that our double Gaussian model for the seeing with a given width, convolved with the fiber profile, is correct. However, the seeing width is an estimate and its FWHM can deviate from reality as much as 20%; other aspects of the kernel model may be incorrect as well.

Shepard's method does not use any information about the kernel, and thus its behavior is independent of how accurately we know the kernel. However, our CRR method depends on a kernel model to determine the weights. Therefore we need to test whether the CRR image varies under reasonable assumptions about the inaccuracy of our kernel model.

To test the behavior of the algorithm under these conditions, we varied the FWHM of the assumed seeing from the actual

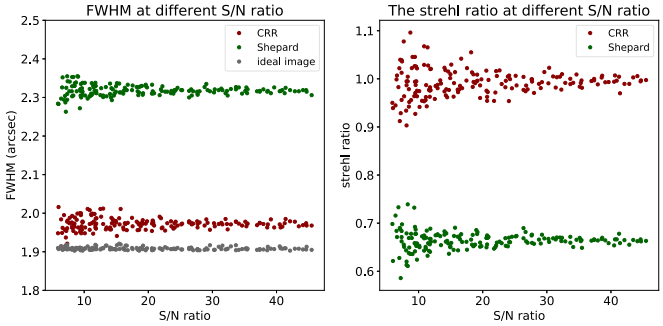

Figure 10. Comparison of Shepard's method and CRR method at different S/N for plate-IFU 8720-1901 and wavelength $\lambda = 5500$ Å. The pixel size $= 0\farcs75$. Left panel: FWHM of the PSF. Right panel: the pseudo-Strehl ratio of the PSF.

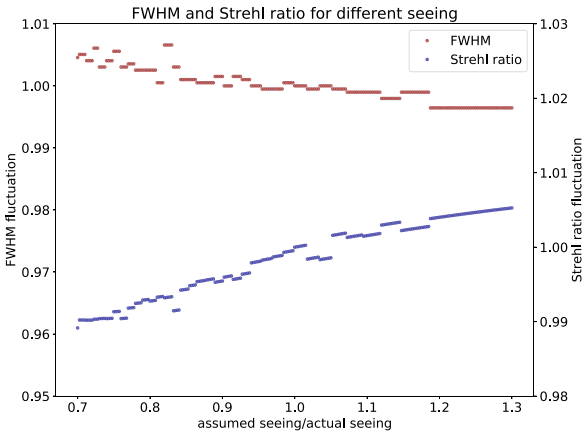

Figure 11. The FWHM and Strehl ratio of the CRR image when the assumed kernel deviates from actual (simulated) kernel, for plate-IFU 8720-1901 at wavelength $\lambda = 5500$ Å, with pixel size $0\farcs75$. The x-axis indicates the ratio of the assumed FWHM of kernel to the average of the actual seeing values for the observations ($1\farcs19$). The piecewise discontinuities are an artificial result of the fact that our kernel is constructed at a set of discrete FWHM values.

value in our simulations, and compared the CRR image with the case that the assumed seeing was the same as the actual seeing. The results are shown in Figure 11. We varied the seeing by as much as 30% from the observational estimate for this plate of $1\farcs19$. The FWHM and pseudo-Strehl ratio for this case are $1\farcs973$ and 0.996, when the seeing is correctly estimated. For a 10% error in our assumed seeing, the FWHM of CRR image changes by around 0.30%; for a 20% error in our assumed seeing, the FWHM of CRR image changes by around 0.40%. The fractional variation of the Strehl ratio is similarly small. Therefore our results are relatively insensitive to whether or not we know the kernel exactly. We have checked that this variation is small independent of the actual seeing in the simulation (e.g., if the actual seeing is poor rather than the relatively good value of $1\farcs19$).

3.7. Regularization Parameter

One of the free parameters in the method is the regularization parameter λ. Here we examine the effect of this parameter on the results. Figure 12 shows results without regularization in the upper panels and with regularization in the lower panels, for comparison. The left panels show the deconvolved reconstruction $\boldsymbol{F}$; the upper left panel (using $0\farcs75$ pixel^{-1}) for a simulated point source shows significant ringing, with large correlations and anticorrelations between pixels. The upper center panel

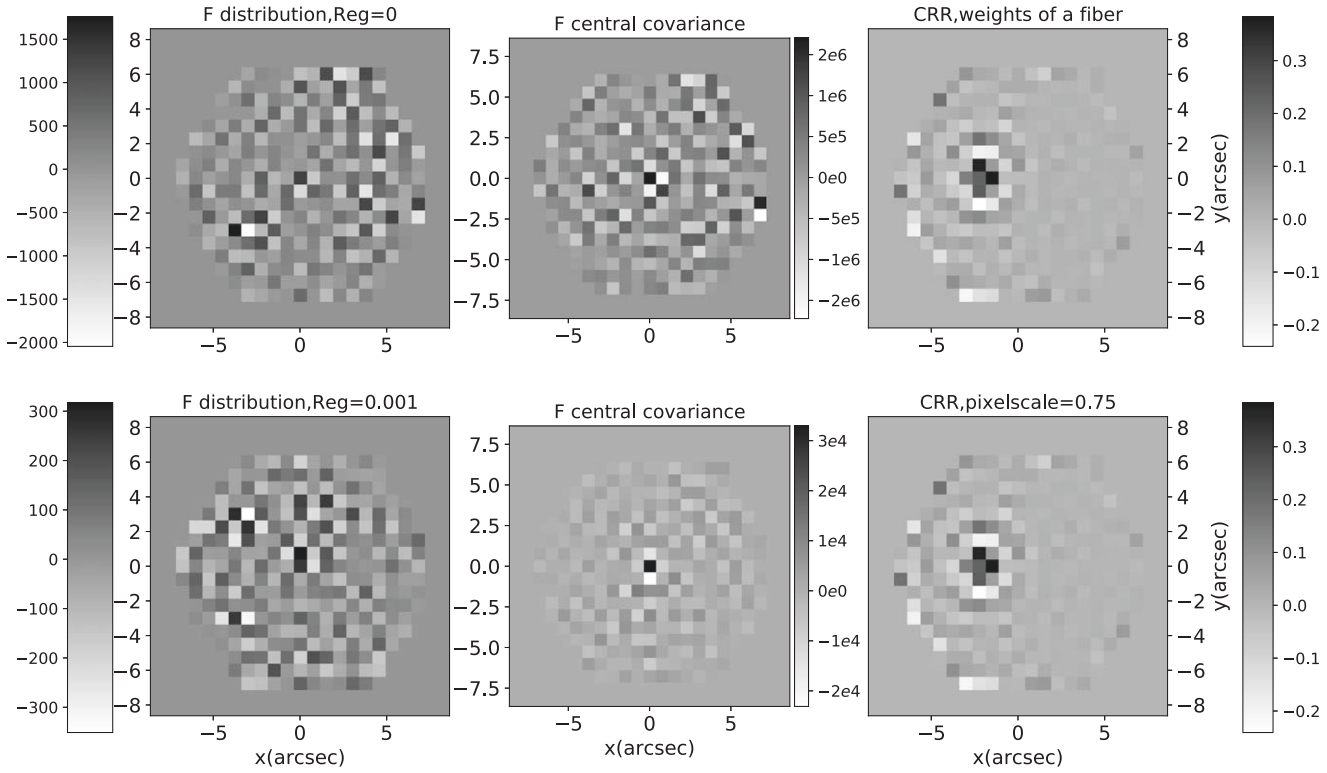

Figure 12. The effect of regularization for our method for plate-IFU 8720-1901. Upper panels are without regularization and lower panels are with regularization 10^{-3}. We used pixel size $0\farcs75$ pixel^{-1}. Left panels: the distribution of the deconvolved solution. Middle panels: the covariance between the central pixel and other pixels in the deconvolved solution. Right panels: the contribution of a particular fiber to all pixels in the CRR image.

quantifies the covariance between the central pixel of $\boldsymbol{F}$ and all the other pixels (the central row of the covariance matrix), showing large anticorrelations between the central pixel and its immediate neighbors.

These correlations and anticorrelations are large and can introduce numerical instability to our analysis, since $\boldsymbol{F}$ results from the last four factors in Equation (23). However, with appropriate regularization, the behavior of $\boldsymbol{F}$ (and therefore those last four factors) is much more constrained. While improving the numerical stability, this regularization will not affect the extraction of our final reconstruction from the fiber data, due to the first factor $\boldsymbol{R}$ in Equation (23).

The right panels show the contribution of a particular fiber to all the pixels on the grid, equivalent to a column of the $\boldsymbol{W}$ matrix. Unlike the Shepard's method weights, the contributions are not all positive and exhibit some ringing. This behavior is characteristic of accurate interpolation kernels, such as sinc-interpolation, so it is expected. These weights remain identical even under regularization, meaning that for the values of λ we use (or smaller), it does not affect the final results.

3.8. Consistency Across Wavelength

The locations of fibers and the kernel shape are a function of wavelength, which causes a necessary variation with wavelength of the weights $\boldsymbol{W}$ in our method. Here we examine whether the CRR method produces spectra that are consistent across wavelength under this variation.

We simulate a constant f_λ point source for all wavelengths and use the meta-data (positions, atmosphere conditions) for plate-IFU 8720-1901. We compare the output spectra for CRR and Shepard's method in Figure 13 for the total flux within the bundle. In Figure 14, we perform the same comparison for just the spectrum in the central pixel.

Figure 13 shows that the variation of the sum of all the pixels in the simulation results is within about 3% of constant. Once the overall difference in amplitude of ∼1% between the CRR and Shepard's method is accounted for, the relative variation is consistent between the methods to about 0.5%. The variation in both methods is due to not having enough dithers of exposures, which inevitably leads to a variation in the flux due to the changing position of the point source relative to the fibers as a function of wavelength. Artificially adding in more dithers can reduce this variation for the simulations (though of course we cannot do that for the observations!).

Figure 14 shows the variation of the central pixel intensity across wavelength. The variation of intensity is 6% for the CRR result, compared to 3% for Shepard's result. Therefore, for an individual pixel there is a slightly larger spectrophotometric inconsistency in the CRR method than in Shepard's method.

For the real case, besides the nonconstant flux and extended fiber configuration, we need to consider the bad fibers identified by the MaNGA DRP for each plate-IFU, leading to the "low coverage" or "no coverage" masks. These fiber masks occur because of cosmic ray events, bad flat fields, CCD defects, broken fibers, or 2D extraction problems (Law et al. 2016). They will result in zero weights in the $\boldsymbol{W}$ matrix. But we have verified that they do not affect the spectrophotometric consistency as a function of wavelength, because of the normalization over the contributions from all the good fibers.

We applied our method on 140 fiber bundles in order to demonstrate that the example plate-IFU (8720-1901) is typical. In each case, we simulate a point source at the center of the

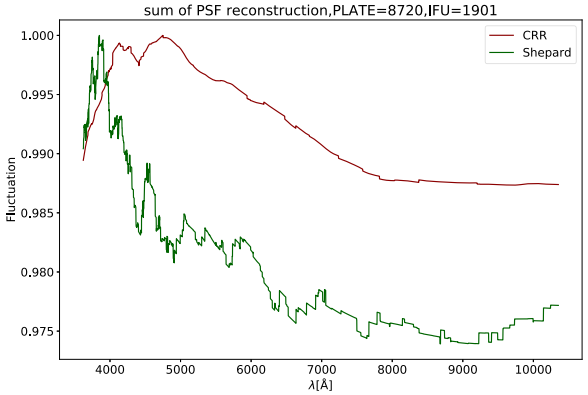

Figure 13. Spectrum of the sum of all pixels for a constant f_λ point source with Shepard's method and CRR method. The y-axis is normalized by its maximum value.

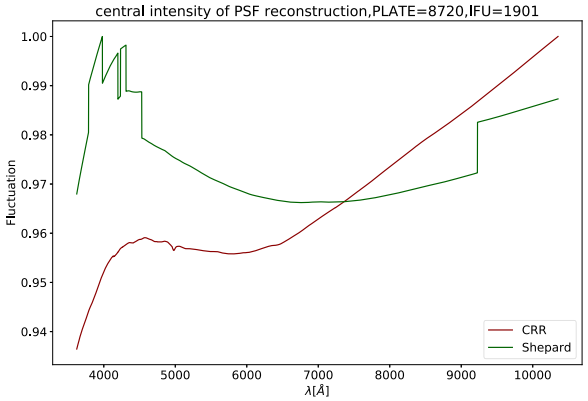

Figure 14. Similar to Figure 13, for the spectrum of the central pixel.

cube. We compare the PSF FWHM and central intensity between CRR and Shepard's image at several different wavelengths. As shown in Figures 15 and 16, the reconstruction in our method is generally narrower than Shepard's result and with higher pseudo-Strehl ratio.

4. Demonstration with Actual Data

The tests in the previous section used fiber sampling locations and weather conditions taken from real observations, but with artificially injected fluxes from a simulated point source profile. In this section we apply the method to several real sets of MaNGA data as a demonstration and to verify that it is working as expected.

Figure 17 shows four test galaxies at pixel size $0\farcs75$ pixel^{-1}. As the *irg*-band color images in the left column show, two are elliptical galaxies and two are spiral galaxies. The first two rows (plate-IFUs 8720-1901 and 8143-6101) are active galactic nuclei (AGNs), and the second two (plate-IFUs 8247-6101 and 9183-12702) are non-AGN. We use these different types of galaxies just to show the variety of possible cases.

The remaining columns, from left to right, show the Hα emission in the CRR method and in Shepard's method, the [O III] 5007 Å emission in both methods, and the mean continuum of the 5300–6000 Å region in both methods. The Hα and [O III] line emission fluxes are estimated by subtracting a continuum estimate using side bands around each line, and an unweighted integration of the line flux in a fixed rest frame wavelength range around each line. This relatively crude

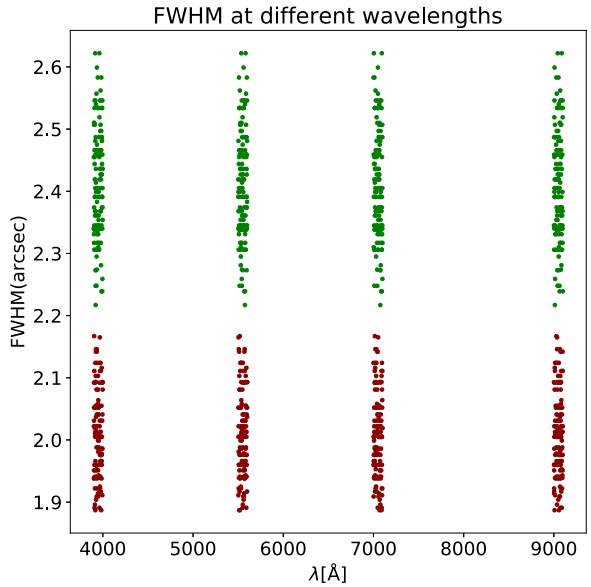

Figure 15. PSF FWHM of point source simulation results at several different wavelengths. Each set of points represents a fiber bundle simulation. Dark red points are for CRR image, and green points are for Shepard's image. The points are randomly offset slightly in wavelength for clarity. The four wavelength slices we use are $\lambda = [3900, 5500, 7000, 9000]$ Å.

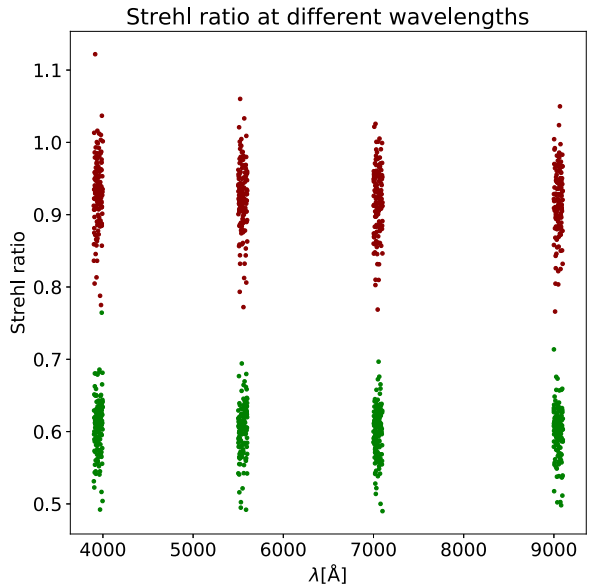

Figure 16. Similar to Figure 16, for the pseudo-Strehl ratio of the point source simulation results.

method is adequate to characterize the image quality, and we defer to the next phase of this work a more careful analysis with the MaNGA Data Analysis Pipeline (Westfall et al. 2019).

In each image, the sharper nature of the CRR image is clear. This clarity is most dramatically shown in the Hα emission in the bottom row and in the [O III] emission in the second row. It is also clear that the off-diagonal covariance in the CRR images is lower, manifesting as a noisier-looking image. For example, in the [O III] images in the bottom two rows, the regions without significant emission show uniform white noise in the CRR image, but show the characteristic mottling of correlated noise in Shepard's method.

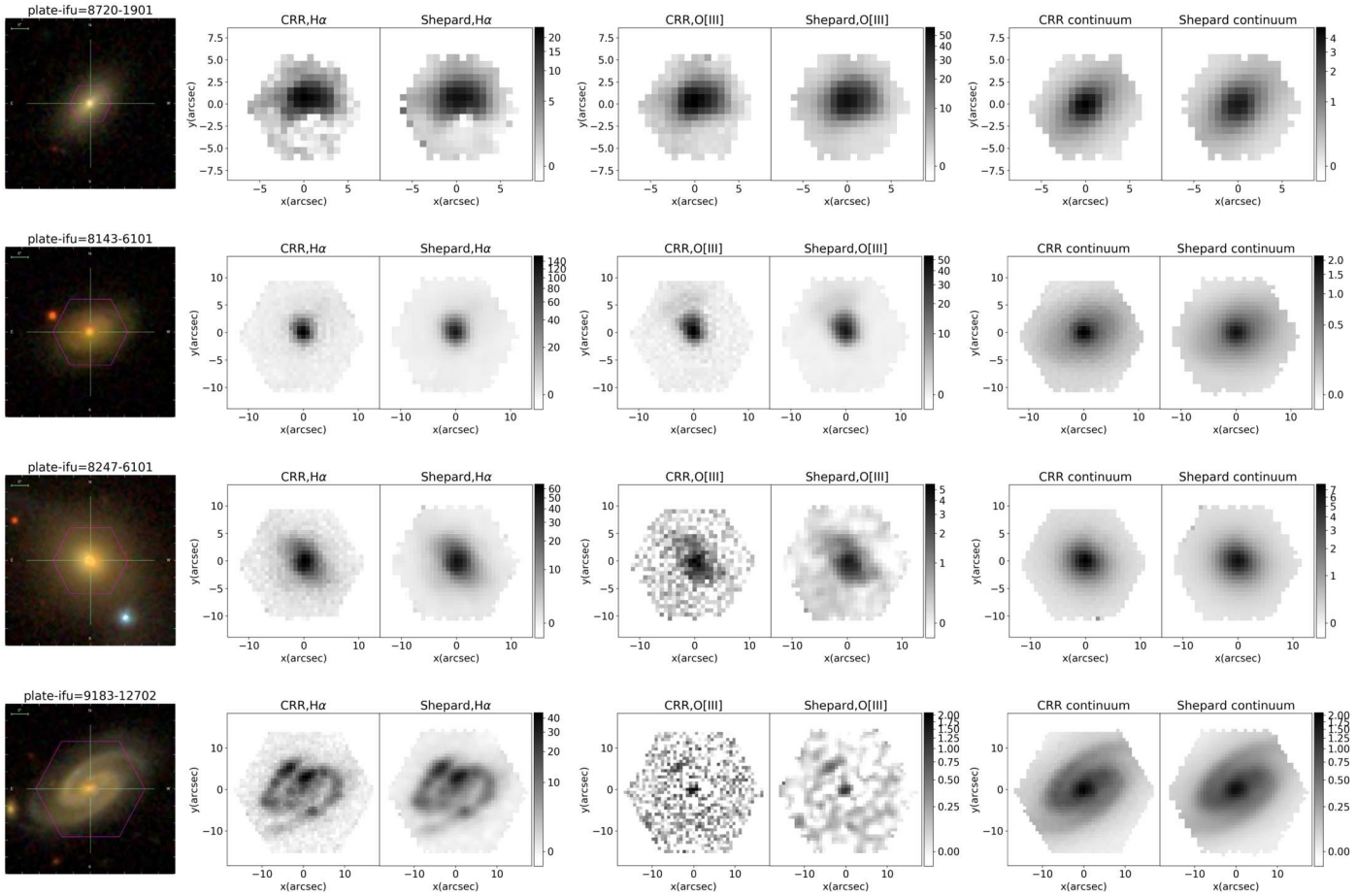

Figure 17. Real galaxy extraction slices for four different plate-IFUs at $0\farcs75$ pixel^{-1}. From top to bottom, the rows correspond to plate-IFUs 8720-1901, 8143-6101, 8247-6101, and 9183-12702. From left to right the columns are the *irg*-band color image in MaNGA, the continuum subtracted CRR images in the Hα 6563 Å region, the O[III] 5007 Å region, and a continuum 5300–6000 Å region. In each case, we show our method on the left and Shepard's method on the right.

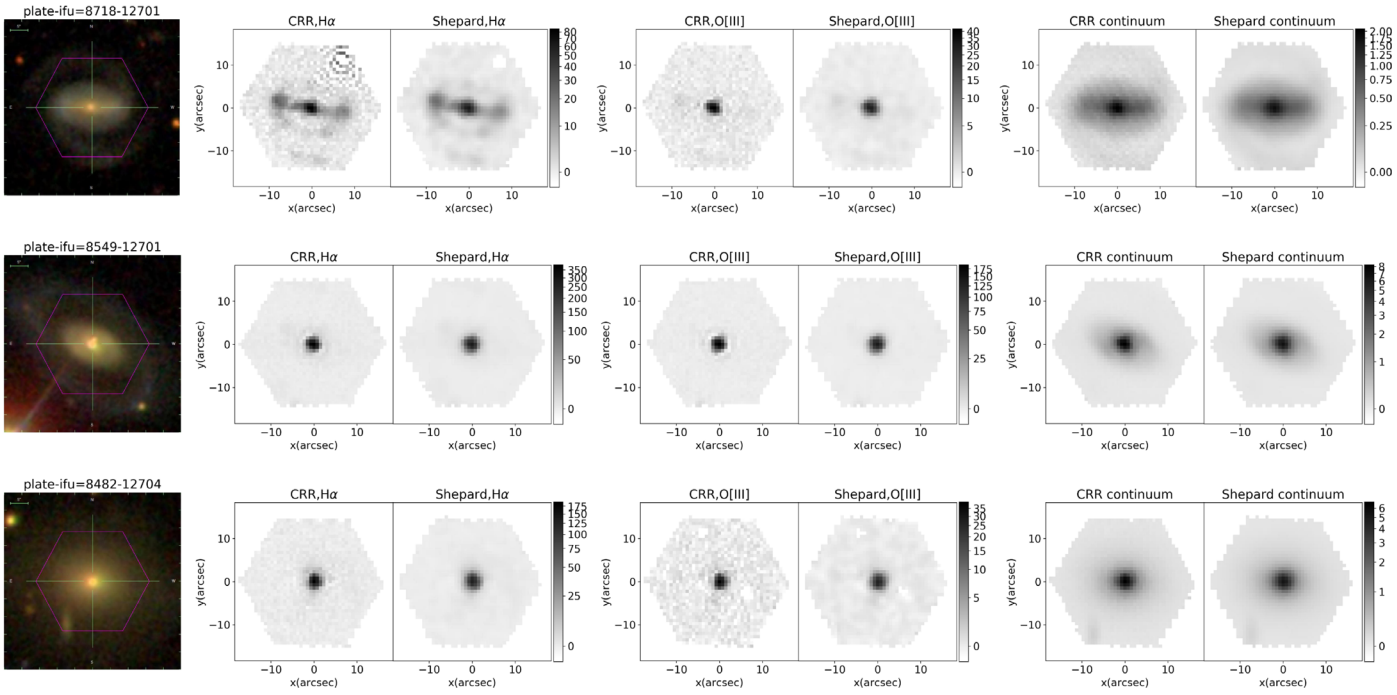

Figure 18. Similar to Figure 17, for three AGN-like plate-IFUs: 8718-12701, 8549-12701, and 8482-12704.

Table 3
Radial Profile Measurements for Real Galaxies Whose Morphologies Are Dominated by AGN

Plate	Method	Hα 6563 Å		O[III] 5007 Å	
		FWHM	Center	FWHM	Center
8718-12701	CRR	2.511	79.5	2.181	42.9
	Shepard's	2.729	62.7	2.644	29.0
8549-12701	CRR	2.193	405	2.184	201.6
	Shepard's	2.640	278	2.591	202
8482-12704	CRR	2.283	181	2.264	37.8
	Shepard's	2.729	128	2.729	26.2

Note. FWHM is in units of arcsec, while the central flux is in units of 10^{-17} erg s^{-1} cm^{-2} Å^{-1} arcsec^{-2}.

We also check the profile as a function of radius from the center for real images. Considering that the morphology of the galaxy will affect the radial profile, it is best to choose point-like sources for the comparison. In Figure 18, we select three targets identified as AGN by Rembold et al. (2017). Using the same methods we use for the PSF fitting, we measure the FWHM and central flux. Table 3 shows the results. Only the comparison of FWHM and central intensity for different methods has meaning, not the absolute value of these quantities, since the galaxies are more extended than a PSF and may have more irregular morphologies than a PSF. In addition, we expect for extended sources the change in their size to be smaller than for a PSF. The CRR image reveals a 28.3% brighter center and a 16.6% smaller spatial size on average based on these three cases, showing that the method is working as we expect from the simulation tests in Section 3.2.

Finally, we check the full spectra of the CRR and Shepard's image results for real data. Figure 19 shows the ratio of the spectrum from CRR to the spectrum from Shepard's image. We consider the full spectrum summed over all pixels, and the intensity of the central pixel, as labeled, for plate-IFU 8720-1901. The thick solid curve overlaid on each spectral ratio is the running mean for the ratio smoothing over 100 wavelength slices or 84 Å. For the sum of all pixels of the running mean, the ratio of Shepard's image and our reconstruction is 0.9729 on average with 0.0033 for the standard deviation, meaning it is almost constant in the running mean. The central intensity of our reconstruction is 29.9% brighter than Shepard's result, with a standard deviation of 0.0439. These results indicate that we have not introduced major wavelength-dependent artifacts relative to what might exist in Shepard's method.

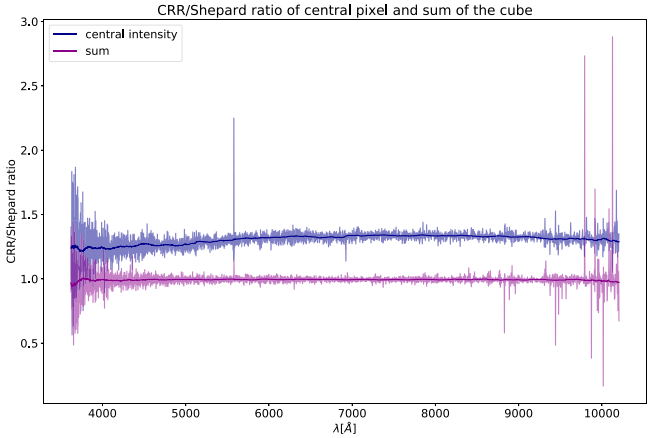

Figure 19. Ratio of the spectrum from CRR method to the spectrum from Shepard's method for plate-IFU 8720-1901. The pixel scale = 0.″75. The blue line is for the central pixel, and the magenta line is for the sum of all pixels. The solid curves are running averages for the ratio smoothing over 100 wavelength slices.

5. Discussion

Shepard's method is one of the most commonly used techniques to spatially interpolate irregularly sampled data onto a regular grid (see Franke & Nielson 1980; Dell'Accio et al. 2016). The largest IFU surveys to date are all using a variant of it. MaNGA and CALIFA both used the form of Shepard's method described here, and SAMI uses a differently motivated technique, but one which is very similar to Shepard's method with a different kernel.

However, Shepard's method produces images whose flux errors are correlated with one another; i.e., they have a very nondiagonal covariance matrix (Law et al. 2016). In addition, as we found in this investigation, it unnecessarily broadens the PSF of the resulting image.

Other techniques exist in the statistical literature for interpolating from irregularly distributed samples, for example radial basis function techniques, Wiener interpolation, "kriging," and Gaussian processes (Krige 1951; Wiener 1964; Schaback 1995; Hartkamp et al. 1999; Rasmussen & Williams 2006; Press 2007). However, these methods are not designed to produce a consistent and tight PSF, and generally lead to highly off-diagonal covariances.

Motivated by a desire to avoid off-diagonal covariances, we examined the techniques of Bolton & Schlegel (2010), and in this paper have adapted them to the imaging context. The result is a technique that successfully reduces the off-diagonal covariances to a very small level and also provides a final image PSF that is better than other methods by around 16% in the FWHM, for the MaNGA example we consider.

The off-diagonal covariances produced by other methods can heavily affect the subsequent analysis. For simple measurements such as aperture fluxes, the correct propagation of errors becomes cumbersome and complex. For more complicated measurements such as maximum likelihood model fitting, even determining the best-fit parameters depends on accounting for the covariance accurately. The MaNGA data, and IFU data generally, is often used for such measurements. Our reconstruction represents a way to simplify these measurements up front.

One may wonder why it is possible for us to obtain near-zero off-diagonal covariances, considering that dithered samples are spaced every 1.″44 while our pixel scale is 0.″75. First, it is worth noting that the covariance between adjacent neighbors is not entirely eliminated, but remains at the few percent level for this pixel size. Second, the differences in seeing and even slight differences in coordinates for different exposures allow the method to create nearly independent data points in neighboring pixels.

The improved image PSF is substantial. It is equivalent to building an instrument with 30%–40% more fibers. It provides a greater ability to resolve structures within galaxies and to measure gradients accurately, as well as to find fainter point-like features (e.g., AGN).

A critical aspect of obtaining diagonal covariances is that some of the weights are negative, whereas Shepard's method has all positive weights (see Figure 3). All positive weights will always produce correlated errors. Although at first glance the fact that a positive fiber flux can contribute negatively to a pixel flux may be non-intuitive, this same feature exists in all accurate image interpolation techniques on a grid (e.g., those based on a sinc kernel) and is not in itself a cause for concern.

We plan to apply this method to the entire MaNGA sample and to test the MaNGA Data Analysis Pipeline code on the revised cubes. Since our method is still experimental (we have only applied it to around 100 or so cubes) we cannot determine yet if it will yield a practical improvement to the results, but we view its prospects as promising. A drawback of this approach is that it will take about 4–5 times longer to build a cube than with Shepard's method, mainly in the SVD process.

So far as we can determine, although our technique was derived from that of Bolton & Schlegel (2010), this application is entirely new, and represents a new method for scattered image interpolation. Since its main function is to control the covariance matrix of the result, rather than to regularize the result by enforcing a notion of smoothness, we refer to it as CRR. It is relevant when the input samples are noisy and irregular, when result of the interpolation is meant to be a specific grid of values, and when there is a natural resolution (in our case, the kernel) in the sampled image that is meant to be preserved. These conditions apply in other IFU data sets (e.g., those that will be provided by the *James Webb Space Telescope*), as well as to a number of ground-based and space-based imaging data sets. Thus, this method may provide an alternative and improved method to analyze those data sets.

We thank Yacine Ali-Haïmoud, Stephen Bailey, Adam Bolton, Cristina Mondino, and David W. Hogg for useful discussions regarding this work. M.R.B. was supported in part by National Science Foundation grant NSF-AST-1615997.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS website is www.sdss.org.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

ORCID iDs

Dou Liu https://orcid.org/0000-0002-4069-6415
Michael R. Blanton https://orcid.org/0000-0003-1641-6222
David R. Law https://orcid.org/0000-0002-9402-186X

References

Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, AJ, 154, 28
Bolton, A. S., & Schlegel, D. J. 2010, PASP, 122, 248
Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, ApJ, 798, 7
Croom, S., Lawrence, J. S., Bland-Hawthorn, J., et al. 2012, MNRAS, 421, 872
Dell'Accio, F., Tommaso, F. D., & Hormann, K. 2016, IJNA, 36, 359
Franke, R., & Nielson, G. 1980, IJNME, 15, 1691
Hamilton, A. J. S. 1997, MNRAS, 289, 285
Hartkamp, A. D., De Beurs, K., Stein, A., & White, J. W. 1999, Interpolation Techniques for Climate Variables (Mexico, D.F.: CIMMYT) http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1017.9308&rep=rep1&type=pdf
Krige, D. G. 1951, SAIMM, 52, 119
Laurent, F., Henault, F., Renault, E., et al. 2006, PASP, 118, 1564
Law, D. R., Cherinka, B., Yan, R., et al. 2016, AJ, 152, 83
Law, D. R., Yan, R., Bershady, M. A., et al. 2015, AJ, 150, 19
Mighell, K. J. 1999, ApJ, 518, 380
Morrissey, P., Matuszewski, M., Martin, C., et al. 2012, Proc. SPIE, 8446, 844613
Press, W. H. 2007, Numerical Recipes: The Art of Scientific Computing (3rd ed.; Cambridge, New York: Cambridge Univ. Press)
Rasmussen, C. E., & Williams, C. K. I. 2006, Gaussian Processes for Machine Learning (Cambridge, MA: MIT Press)
Rembold, S. B., Shimoia, J., Storchi-Bergmann, T., et al. 2017, MNRAS, 472, 4382
Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., et al. 2012, A&A, 538, A8
Schaback, R. 1995, in Approximation Theory VIII, Vol. 1, Approximation and Interpolation, ed. C. K. Chui & L. L. Schumacker (Singapore: World Scientific), 491
Shepard, D. 1968, in Proc. 1968 ACM National Conf. , ed. R. B. Blue, Sr. & A. M. Rosenberg (New York: ACM), 517
Warren, S. J., & Dye, S. 2003, ApJ, 590, 673
Westfall, K., Cappellari, M., & Bershady, M. A. 2019, AJ, 158, 231
Wiener, R. 1964, Extrapolation, Interpolation, and Smoothing of Stationary Time Series (Cambridge, MA: MIT Press)
Yan, R., Tremonti, C., Bershady, M. A., et al. 2016, AJ, 151, 8
Yang, C. S., Kao, S. P., Lee, F. B., & Hung, P. S. 2004, in Proc. XXth ISPRS Conf., Geo-imagery Bridging Continents, ed. M. Orhan Altan (Hannover: ISPRS), 778