– Independent Work Report Spring 2020 –

Model-Based Search for Extended Emission-Line Regions in Astronomical Images

Charles Zhao*
Advisor: Peter Melchior†

Abstract

The recently developed method SCARLET *by Melchior et al. (2018) allows for the morphologies and spectra of individual astronomical sources to be extracted from large data sets even when they are partially or completely overlapping. In this work, we search for fluorescent emissions of gas known as extended emission-line regions (EELRs), which are energized by active galactic nuclei (AGN) at the centers of galaxies. EELRs have so far rarely been observed cleanly because their host galaxies tend to dominate the emission, but* SCARLET *is designed to separate multiple sources of emission. We develop an approach to extract EELRs from multi-band images without the need for targeted spectroscopic measurements. Our approach uses Gaussian Process regression to generate samples of likely EELR spectra, and then computes a likelihood-weighted model average of each EELR's morphology and spectrum as obtained from* SCARLET. *This model-based search approach is especially useful for finding known or suspected physical processes in the growing data volumes of future large astronomical surveys.*

1. Introduction

With the development of more advanced telescopes that allow us to see deeper into space, accurately and computationally efficiently distinguishing overlapping astronomical objects in images has become increasingly important. Morphologies and spectra are basic characteristics of galaxies, and with the massive amounts of data being gathered by modern telescopes, determining these characteristics in the presence of overlaps is essential for large-scale astronomical studies. Much

*Department of Computer Science, Princeton University, Class of 2021
†Department of Astrophysical Sciences, Princeton University

of today's astronomical studies are done not on astronomical images themselves, but on catalogs of objects detected in these images, so source extraction and separation impacts all downstream studies.

Melchior et al. [12] have recently developed a model-based framework for source separation in multi-band images called SCARLET. One particularly useful application of SCARLET that takes advantage of SCARLET's model-based nature is the deblending of fluorescent emissions of gas known as extended emission-line regions (EELRs), which are produced by active galactic nuclei (AGN) at the centers of galaxies. AGN feedback processes, the same processes that produce EELRs, are thought to explain several phenomena in galaxy evolution, such as the lack of massive star-forming galaxies in the nearby universe [7]. However, the mechanisms underlying AGN are still poorly understood. Many proposed mechanisms exist, but the details as well as relative importances of these mechanisms are unclear. Narrow-line regions (NLRs), fluorescent regions of gas that emit narrow spectral lines, are one of the defining features of AGN. EELRs are surprisingly large NLRs, spanning tens of kiloparsecs.[1] Due to their large extents, EELRs have been used to help understand AGN feedback, variability over time, and obscuration by dusty tori (see Sun et al. [19] for a list of related studies). Therefore, an important step towards a better understanding of AGN is to have more and higher quality observations of EELRs.

Past studies have required targeted observations of EELRs that require access to the world's most sensitive telescopes and thus limit the number of known systems and related studies. In this work, we develop a method for imaging EELRs in large-scale astronomical surveys without any kind of targeted observations. By allowing for large-scale search of EELRs, we hope our method will be able to image more EELRs and thus enable future studies to better understand AGN.

[1] A parsec is defined as the distance at which one astronomical unit (roughly the radius of Earth's orbit around the Sun) subtends an angle of one arcsecond. 1 pc $\approx$ 3.26 light-years.

2. Background and Related Work

2.1. Deblending

The problem of deblending is to model an image as the sum of contributions of multiple, often overlapping astronomical sources affected by a known blurring, called a point spread function (PSF).[2] This blurring might, for example, be caused by diffraction by the telescope, and by Earth's atmosphere (which is what causes stars to twinkle). The goal is then to fit a model to each source so that the resulting reconstructed image is as close to the original image as possible. Much of the following two paragraphs is adapted from Lupton [11].

Stars are relatively easy to deblend since they can be modeled as point sources. In the case of an image composed only of stars, the full model is

$$I = S + \sum_r F_r \delta(x - x_r) * \phi + n \tag{1}$$

where I is the observed intensity, S is the sky level, F_r is the flux of the rth star, x_r is the 2D coordinate of the rth star, δ is the Dirac delta function, ϕ is the PSF, $*$ denotes a convolution, and n is noise. Fitting this model is a minimization problem in $3r+1$ unknowns, since we need to fit S, F_r, and x_r.

Galaxies, however, are much harder. Their shapes and light intensities can be irregular, sometimes with multiple peaks (see Figure 1). In other words, their morphology is not a simple δ function. For example, we cannot be sure whether a multi-peaked observation is made up of multiple separate galaxies, a single blobby irregular galaxy, or an elliptical galaxy occluded by a dust cloud. For deblending algorithms, these uncertainties translate to degeneracies. We will now review several approaches to deblending.

SExtractor [5] was published in 1995 and designed especially for processing large amounts of astronomical imaging data. It is achromatic, i.e. it only takes in a single-band image. Its

[2] To be more precise, an imaging system's PSF is its response to a point source, as can be seen in Equation 1.

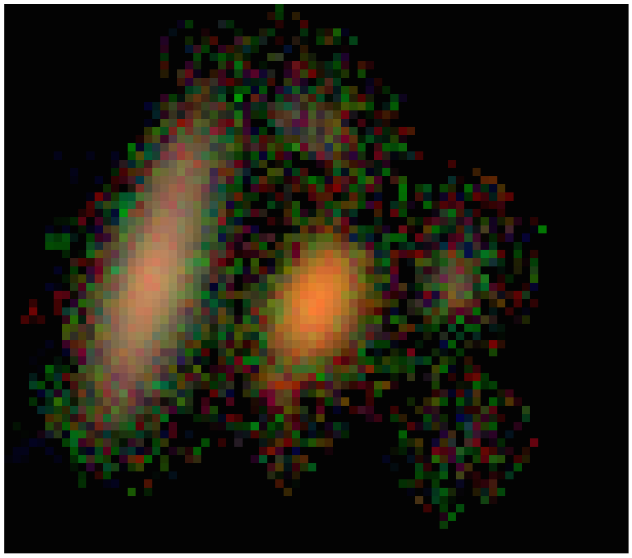


Figure 1: A multi-peaked observation [11].

approach to deblending is to threshold an image at exponentially spaced intensity levels and search for sets of pixels that are connected at a lower threshold but split into multiple connected regions at a higher threshold. This approach is simple and fast. However, one downside is that it can only associate each pixel with a single peak, which is unrealistic and can lead to larger galaxies being incorrectly shredded into small unphysical chunks due to overlapping galaxies.

SCARLET [12], the recently developed deblending framework used in this work, models an observed scene as a mixture of components with compact spatial support and uniform spectra over their support. One significant advantage SCARLET has over many other approaches is that it directly operates on multi-band images.[3] This is important since color often serves as a key discriminator between overlapping objects. In addition, while many previous approaches perform pixel-object association one object at a time, SCARLET performs this association simultaneously. This allows one to explore the degeneracies arising from overlapping objects.

For each component, SCARLET estimates its morphology (i.e. its spatial intensity variation) and its spectrum. SCARLET's approach is a generalization of non-negative matrix factorization (NMF). Given B images (one for each band) with N pixels per image, and representing each image as a

[3] A multi-band image is an image consisting of observations through multiple filters, each one covering a different range of wavelengths. For example, a regular consumer camera has red, green, and blue filters.

flattened N-element array, SCARLET constructs a model $\mathsf{M} \in \mathbb{R}^{B \times N}$ as a sum of K components,

$$\mathsf{M} = \sum_{k=1}^{K} \mathsf{A}_k \otimes \mathsf{S}_k = \mathsf{AS} \tag{2}$$

where $\mathsf{A}_k \in \mathbb{R}^B$ is the amplitude of component k across all bands, also known as its spectral energy distribution (SED), and $\mathsf{S}_k \in \mathbb{R}^N$ is the morphology of that component. $\otimes$ denotes the outer product. The A_k form the columns of $\mathsf{A} \in \mathbb{R}^{B \times K}$, and the $\mathsf{S}_k^\top$ form the rows of $\mathsf{S} \in \mathbb{R}^{K \times N}$. Note how, compared to the simple model in Equation 1, this model allows for arbitrary morphologies, not just point sources.

SCARLET maximizes the likelihood with respect to the parameters A_k and S_k under constraints. Assuming a homoscedastic Gaussian error, the negative log-likelihood is $f(\mathsf{A},\mathsf{S}) = \frac{1}{2}\|\mathsf{Y} - \mathsf{AS}\|_F^2$ where Y is the observed image. This is a simplified version of the negative log-likelihood minimized by SCARLET.[4]

Unlike NMF, whose only constraint is that all entries are nonnegative, SCARLET can enforce an arbitrary number of quite general constraints simultaneously. In particular, SCARLET maximizes the likelihood function using a proximal gradient method, so it can enforce any constraint that can be expressed as a proximal operator.[5] For example, two such constraints on a component's morphology are monotonic decrease and 180° degree rotational symmetry, both defined with respect to the component's center. Enforcing such constraints can help resolve the aforementioned degeneracies, causing SCARLET to "prefer" certain models over others.

Astro R-CNN [8] is another example of a recent deblending framework, this one using a deep learning approach. The authors developed a Mask Region-based Convolutional Neural Network

[4]The actual negative log-likelihood minimized by SCARLET accounts for the PSF, allows for heteroscedastic errors, and includes terms for constraints.

[5]The basic idea behind proximal gradient methods is to alternate between stepping in the direction of the negative gradient and projecting onto the constrained manifold. For example, to minimize a closed smooth proper convex function $f(\boldsymbol{x})$ under a convex constraint $g(\boldsymbol{x})$ with $\boldsymbol{x} \in \mathbb{R}^n$, the update step is

$$\boldsymbol{x}^{\mathrm{it}+1} \leftarrow \mathrm{prox}_{\lambda^{\mathrm{it}} g}(\boldsymbol{x}^{\mathrm{it}} - \lambda^{\mathrm{it}} \nabla f(\boldsymbol{x}^{\mathrm{it}}))$$

where λ^{it} is the step size. g is never directly evaluated and can be non-differentiable. Instead, it is only accessed through its proximal operator, which can be thought of as a projection.

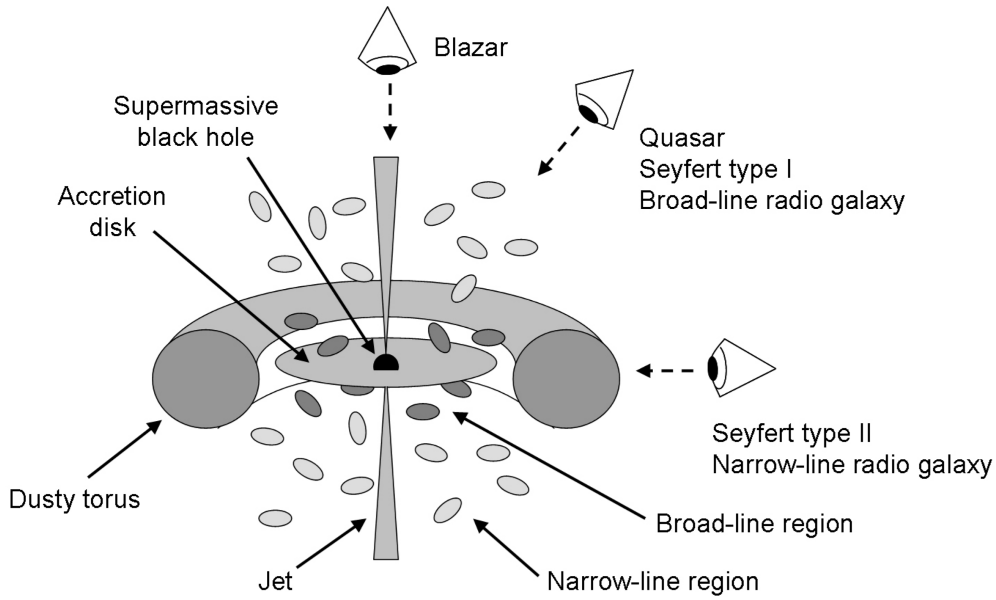


Figure 2: Unified AGN model [24].

(Mask R-CNN) to perform source detection, classification, and deblending. Like SCARLET, Astro R-CNN takes in multi-band images and is able to deblend crowded images quite well. However, since it uses Mask R-CNN to perform instance segmentation, Astro R-CNN only provides object masks, not morphologies or spectra. That is, it assigns pixels to objects without coming up with a model for each object, although, unlike SEXTRACTOR, a pixel can be assigned to multiple overlapping objects. In addition, since it requires a large training set, Astro R-CNN is trained on simulated images and catalogs, which can systematically differ from real images. This is an issue in our case since EELRs are poorly understood, so we certainly cannot simulate them well. We believe this issue essentially prohibits us from taking a deep learning approach. Instead, we employ a model-based approach in which we can enforce physical priors instead of having the system learn from large amounts of data.

2.2. Active Galactic Nuclei and Extended Emission-Line Regions

An EELR is a spatially extended narrow-line region, which is a fluorescent emission of gas from a galaxy. EELRs have widely varying spatial scales and typically have high-excitation spectra. They are found around many AGN, which are highly luminous compact regions at the centers of galaxies fueled by accretion of matter onto a black hole. AGN have anisotropic radiation fields that photoionize the surrounding interstellar medium (ISM), so the morphologies of EELRs are likely determined by the interaction of the AGN radiation with the surrounding ISM [17]. According to unified AGN models, AGN are comprised of a central black hole, a rotating disk of material surrounding and falling into the black hole, and jets shooting out from the poles of the disk. Thus, the prototypical morphologies of EELRs surrounding AGN, although diverse, tend to be jet-like and coaxial with the radio source. In addition, a dusty torus surrounds the inner parts of the AGN, which can obscure parts of the AGN depending on the viewpoint of the observer and leads to their diverse appearances. Unified AGN models posit that different classes of AGN are actually a single type of object observed from different angles (see Figure 2). He et al. [9] provide observational evidence for the unified AGN model, finding that 81% of the AGN in the MaNGA survey have bi-conical or bi-polar narrow-line region morphology, and also finding significant evidence that the major axis of the host galaxy disk and the AGN ionization cones tend to be orthogonal to each other.

However, these unified models are intensely debated [3, 21]. Furthermore, EELRs can also be produced by intrinsic properties of the ISM, meaning that EELRs do not always lie along the radio source axis or within AGN ionization cones. In fact, some EELRs have been observed lying almost perpendicular to the radio axis [10]. Villar-Martín et al. [20] analyzed the radio galaxy PKS 193246 and discovered a giant EELR that extends well beyond the ionization cones of the AGN. Their analysis suggests that, rather than being emitted by the AGN, the giant structure is a star-forming halo associated with the debris of the merger that triggered the activity. Indeed, emissions from star formation and from AGN both tend to produce similar spectra.

Thus, AGN and EELRs are very active areas of study. As mentioned in the introduction of this paper, a better understanding of EELRs can provide valuable insights into galaxy evolution and other areas of cosmology.

2.3. EELR Imaging

Our work is similar in spirit to that of Sun et al. [19]. EELRs generally have strong [O III][6] emission lines that can be observed in broad-band images, which make these lines excellent for studying AGN. Sun et al. develop a broad-band imaging technique that can reconstruct images of the [O III] line, using the Subaru Hyper Suprime-Cam (HSC) Survey for broad-band images and the Sloan Digital Sky Survey (SDSS) for spectra. Their technique uses spectra from SDSS to carefully subtract out the galactic stellar continuum from broad-band images, thus isolating the narrow-line emissions. This technique does not require targeted observations, and is thus able to cover much larger samples than traditional targeted techniques. They use their technique to image the NLRs around 300 obscured AGN, finding 8 EELRs that extend beyond 10 kpc from the nucleus.

Although this technique already requires much less targeted data than previous work to image EELRs, allowing it to search around hundreds rather than just dozens of AGN, it still requires prior spectroscopic measurements. This dependency on spectroscopic measurements limits the objects searchable by Sun et al.'s technique for two reasons. First, SDSS is less sensitive than HSC, and Sun et al. were only able to use objects observed in both surveys. Second, spectroscopic measurements are observationally resource intensive. Astronomical objects are first identified in the SDSS imaging survey, and then a small fraction of these, which have to be relatively bright, are selected to have their spectra measured by having individual fibers pointed at each object. 640 spectra can be observed at a time, with a total integration time of 45-60 minutes, depending on observing conditions [22].

The Legacy Survey of Space and Time (LSST) will be the leading imaging survey in the near future. It will be similarly as sensitive as HSC but cover a much wider area, greatly increasing the available data. But since the spectroscopic coverage of LSST will be limited, there is a need for EELR imaging techniques that do not rely on spectroscopic inputs. This is the role our method aims to fill.

[6]This notation refers to the spectral line corresponding to the second ionization state of oxygen.

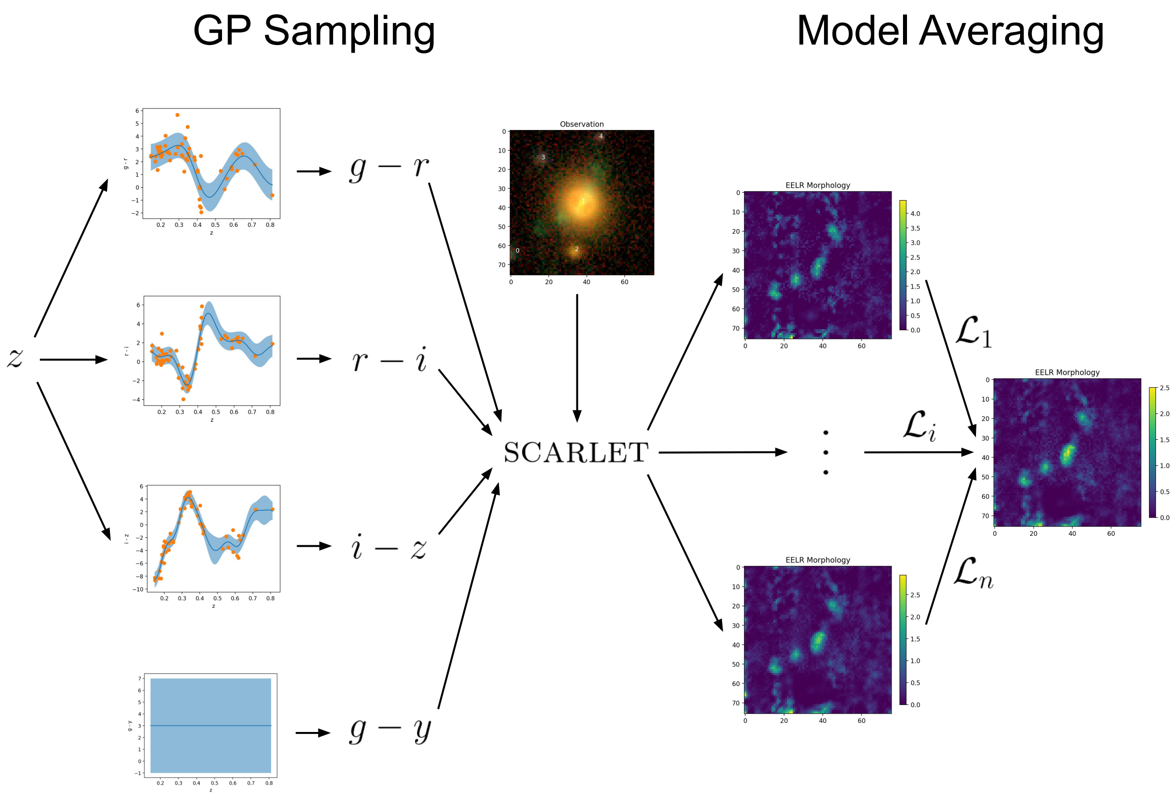


Figure 3: Summary of our inference procedure. Gaussian Process sampling is used to generate EELR spectra (see Section 3.1), which are used by SCARLET as constraints when deblending, and the SCARLET models are then combined using likelihood-weighted model averaging (see Section 3.2).

3. Approach

We develop a probabilistic method for automatic EELR search that does not require spectroscopic measurements. By using SCARLET's constraint system to represent prior information about expected spatial and spectral properties of objects (see Section 4.3 for details), we believe this approach can compensate for the lack of precise spectra. Figure 3 is a schematic of our approach.

3.1. EELR Spectrum Sampling

SDSS provides high quality measurements of both redshift and spectra. Although EELRs have varying compositions, they all tend to have strong [O III] lines that dominate their spectra. However, the redshift z (not to be confused with the photometric band z) of the galaxy moves all the

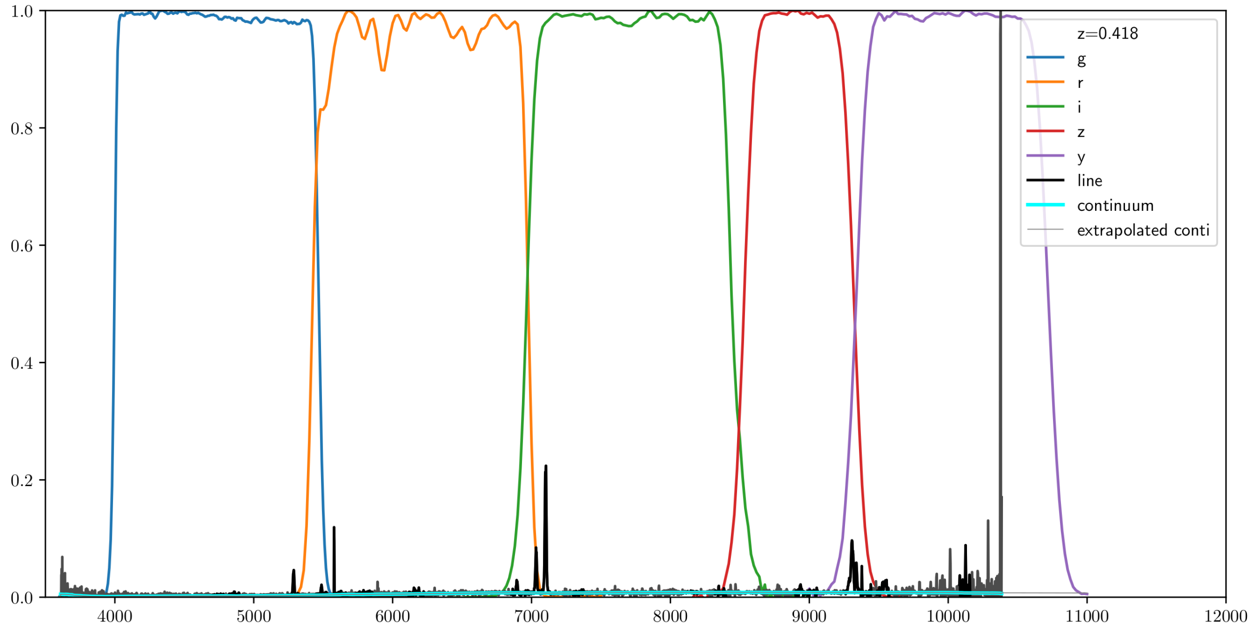


Figure 4: SDSS spectrum of one of the obscured AGN in our sample, and the filter response functions of the five HSC bands [19]. The horizontal axis is wavelength in Å, and the vertical axis is the normalized wavelength flux density.

lines by a factor of $1+z$, so we observe them at different wavelengths. Therefore, we use Gaussian Process (GP) regression to fit a distribution of functions which map redshift to colors, defined as the difference between magnitudes in two bands.[7] Namely, we regress $g-r$, $g-i$, $g-z$, $r-i$, $r-z$, and $i-z$ on redshift. We can use these fitted GPs to generate samples of EELR colors based on galaxy redshift, and then use SCARLET to try fitting models in which the EELR component is constrained to having the given set of colors. Furthermore, having multiple samples allows us to assess the uncertainty of our estimates.

Note that we do not use GP regression to fit any colors involving the y band. This is because we noticed that the SDSS measurements in the y band tend to be unreliable. Looking at Figure 4, we can see that there is a "forest" of abnormally strong lines in the y band, which we believe is due to Earth's atmosphere rather than the AGN. Although SDSS performs processing to subtract out the Earth's atmosphere, it may have failed to do so cleanly in the y band due to the atmosphere's spatial variability. This renders the y band measurements essentially meaningless. Fortunately, our

[7] The reason for using colors rather than magnitudes directly is that we do not want to fit the correlation between redshift and overall EELR brightness. Because magnitudes are defined on a log scale, we only care about the relative differences (not ratios) between bands.

sampling approach provides a natural solution to this issue, which is to uniformly randomly choose y band magnitudes within a reasonable range when generating color samples. This allows us to still utilize the y band in determining the EELR morphology despite our lack of trust in the y band magnitudes.

3.2. Model Averaging Procedure

Model averaging is a technique that uses an ensemble of models to produce more robust predictions than any single model by itself. It does this by taking into account each model's ability to explain the observation and weighting the models accordingly. In our implementation, we sample from the GP priors to obtain a set of EELR colors. We also initialize the EELR morphology as a uniform random pixel vector. SCARLET then optimizes the likelihood under constraints and converges to a local maximum. We define our final estimator as

$$\widehat{\mathbb{E}}(M \mid D) = \frac{1}{Z}\sum_i M_i \cdot \mathscr{L}(D \mid M_i) \tag{3}$$

where M_i are the optimized models and $\mathscr{L}(D \mid M_i)$ are their respective likelihoods. Z is a normalization factor. This estimator weights each optimized model by its likelihood (as computed by SCARLET), giving more weight to "good" models, i.e. those that are able to produce a close reconstruction of the original image, and less weight to "bad" models. By combining models in this way and essentially marginalizing over our uncertainty about the exact EELR spectra and morphologies, we seek to robustly detect and characterize EELRs.

Furthermore, interpreting the likelihoods as reliability weights, we can compute the unbiased model variance as

$$\widehat{\sigma^2}(M \mid D) = \frac{\sum_i w_i (M_i - \widehat{\mathbb{E}}(M \mid D))^2}{V_1 - (V_2/V_1)} \tag{4}$$

where $w_i := \mathscr{L}(D \mid M_i)$, $V_1 := \sum_i w_i = Z$, and $V_2 := \sum_i w_i^2$. This provides a pixel-wise measure of uncertainty about our model average.

Another benefit of our randomized approach is that it should be robust to noisy redshift measurements. This is important because our goal is to image EELRs based solely on broad-band images, and we therefore want our method to work well with *photometric* redshift measurements rather than *spectroscopic* redshift measurements. Whereas spectroscopic redshift techniques estimate redshift by observing the wavelengths of specific characteristic spectral lines, photometric redshift techniques come up with much rougher estimates of redshift by observing the source emission in a few broad-band filters of an imaging survey. Fortunately, as increasingly larger-scale astronomical surveys allow us to photometrically observe fainter and fainter galaxies but not their spectra, photometric redshift techniques have been an active area of research and are becoming increasingly accurate [6, 18].

4. Implementation

4.1. Data

We use HSC for broad-band images and SDSS for spectra and redshift, so our dataset consists of objects detected in both surveys. The HSC survey is part of the Subaru Strategic Program [2] and provides high-resolution g, r, i, z, y broad-band images from the 8.2m Subaru telescope in Hawaii. We use the HSC S18A-wide data release [1], which covers an area of $305\,\mathrm{deg}^2$ at full depth in all five bands. The HSC's median seeing, defined as the full width at half maximum of the PSF (smaller is better), in the i band is about 0.6". SDSS [22] uses a telescope in New Mexico and is much wider but shallower than HSC, and has a typical seeing of 1.4" in the i band. Importantly, SDSS also provides spectra for some of the objects detected in its imaging survey, allowing us to fit the GPs as described in Section 3.1.

Following the procedure of Sun et al. [19], our targets are selected from four SDSS spectroscopically identified obscured type 2 AGN samples—Zakamska et al. [25]; Reyes et al. [16]; Mullaney et al. [13]; and Yuan, Strauss, and Zakamska [23]. We exclude type 1 AGN from our dataset since their bright nuclei can completely dominate observations and thus interfere with the imaging of EELRs (see Figure 2 for a schematic comparison of type 1 and type 2 AGN according to unified

AGN models). Mullaney et al. target low redshift objects ($z \leq 0.3$), while Zakamska et al. and Yuan et al. primarily focus on higher redshifts ($z \sim 0.3-0.7$), and Reyes et al. have both.

We crossmatch this SDSS sample with the HSC S18A-wide sample, using the same target selection criteria as Sun et al. and ending up with a sample of 444 observations with a host galaxy centered in each image. Note that our sample is larger than that of Sun et al. since we use the updated HSC S18A data release, which covers a greater area than the S16A release.

4.2. Gaussian Process Regression

The first step in our method is to generate SED samples given a redshift. To do this, we use scikit-learn's [15] `GaussianProcessRegressor` to fit a GP for each color involving the g, r, i, or z bands. Each GP uses as its kernel the sum of a radial-basis function (RBF) kernel and a white noise kernel. Since the RBF kernel is infinitely differentiable, it produces smooth functions. The white noise kernel captures the noise in the data so that the GP does not overfit, and so that we have an estimate of uncertainty at each redshift. The GPs, fit on SDSS colors and spectroscopic redshift, are shown in Figure 5.

As a sanity check, we can see in Figure 5 that $r-i$ increases sharply around $z \approx 0.4$. This aligns with the fact that the [O III] line moves from the r band into the i band around this redshift. We can verify this by looking at Figure 4. This AGN is at a redshift of $z = 0.418$, and we can see that its [O III] line, which is the strongest line in the spectrum not including the noisy y band, is right at the transition between the r and i bands.

Note that these GPs are actually redundant. Since we only care about the relative differences between band magnitudes, we can set an arbitrary constant for one of the bands, e.g. $g = 25$. Then, we only need three of the six colors to determine the magnitudes of the other three bands. We decided to use $g-r$, $r-i$, and $i-z$ since their GPs had relatively small variance.

By the definition of a GP, for a given redshift, we have a Gaussian distribution for each of these three colors. We can therefore sample from these Gaussian distributions to produce samples of the g, r, i, and z bands. As explained in Section 3.1, we do not have reliable measurements of the y

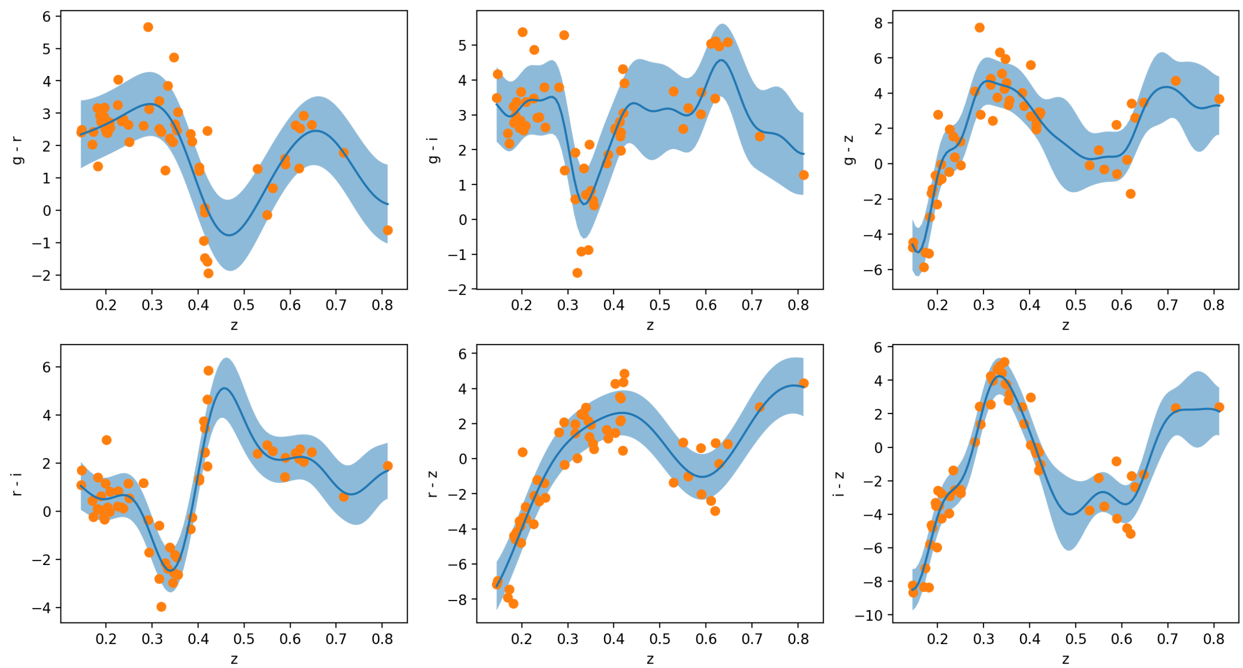
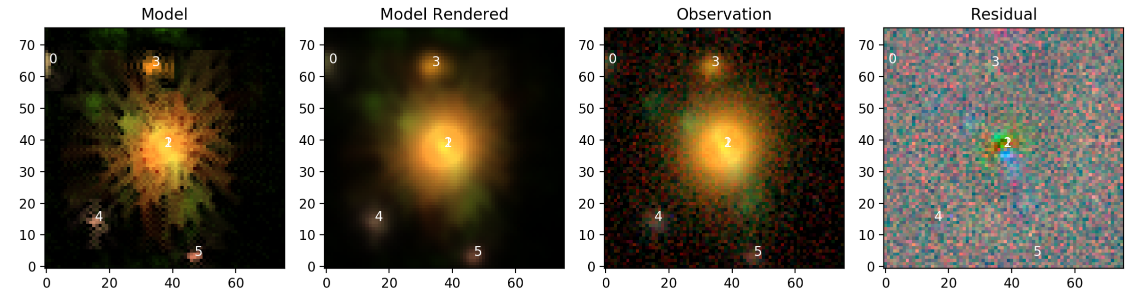


Figure 5: Gaussian Processes regressing colors on redshift z (not to be confused with the z band). The blue lines are the GP mean functions, and the light blue bands are within-1σ confidence intervals.

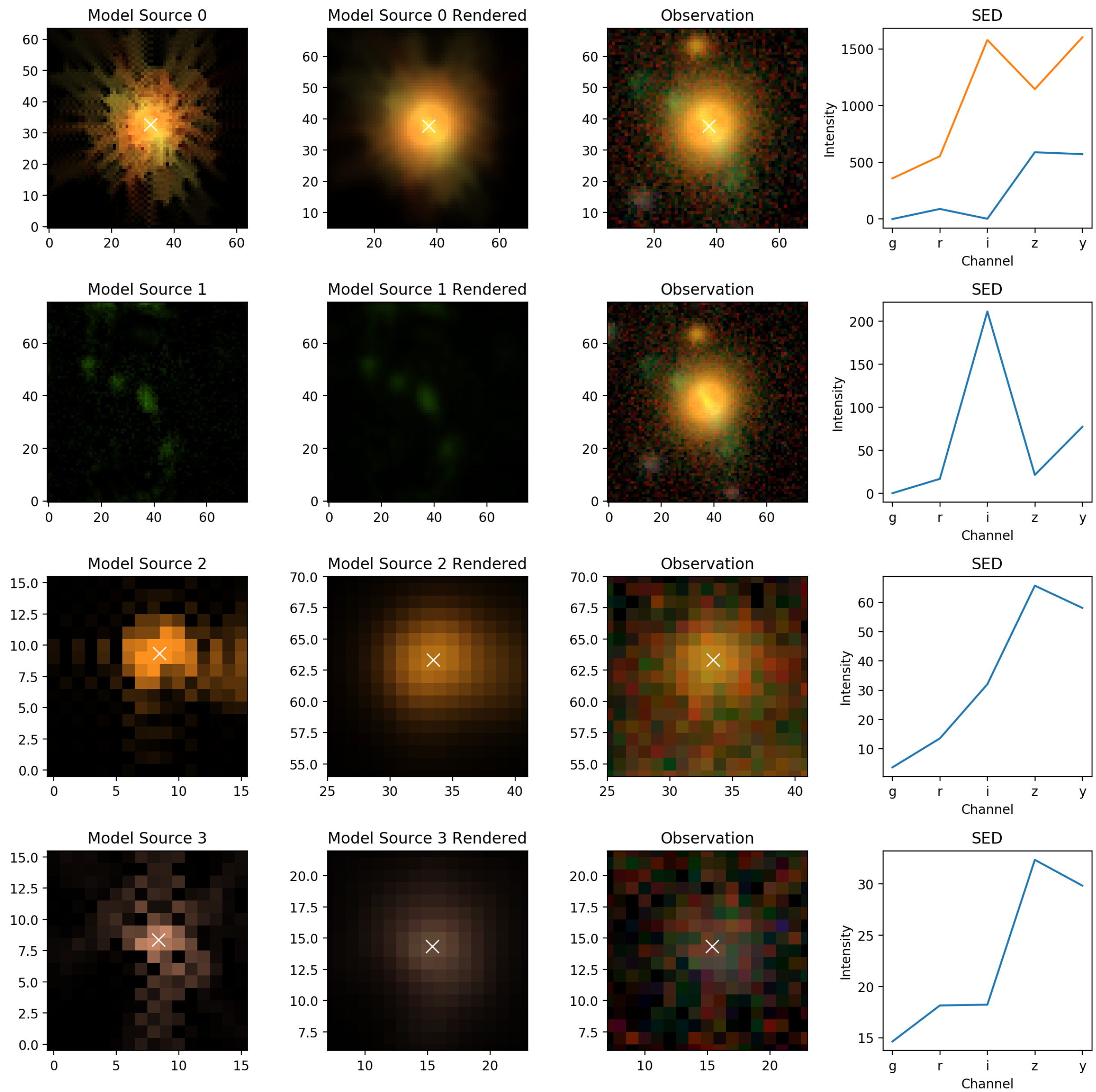
band, so we uniformly randomly sample y band magnitudes between 18 and 26.

4.3. SCARLET Configuration

Given a multi-band (g, r, i, z, y) observation, we first use SEP [4], an implementation of SEXTRACTOR, to generate a catalog of objects and their coordinates within the image. This provides a starting point for SCARLET to then determine the precise morphologies and spectra of each object. Given their highly variable morphologies, we model the EELR as a SCARLET `RandomSource` initialized with uniform random morphology, but we constrain its SED, i.e. its relative magnitudes in each band. We model the host galaxy as a `MultiComponentSource`, comprised of two `ExtendedSources` which are each constrained to have monotonically decreasing flux from the center, with one on top of the other. The rationale for modeling the host galaxy as two components is that, in SCARLET, each component has a uniform spectrum over its spatial support, which is not very realistic for galaxies. Most extended galaxies have a redder inner core containing older stars, and a bluer outer part with younger stars, so two stacked components of uniform spectra better capture this property. We model all other objects as `ExtendedSources`.



(a) Scene model.



(b) Source models. Not all sources are shown here.

Figure 6: An example model fit by SCARLET using the configuration in Section 4.3.

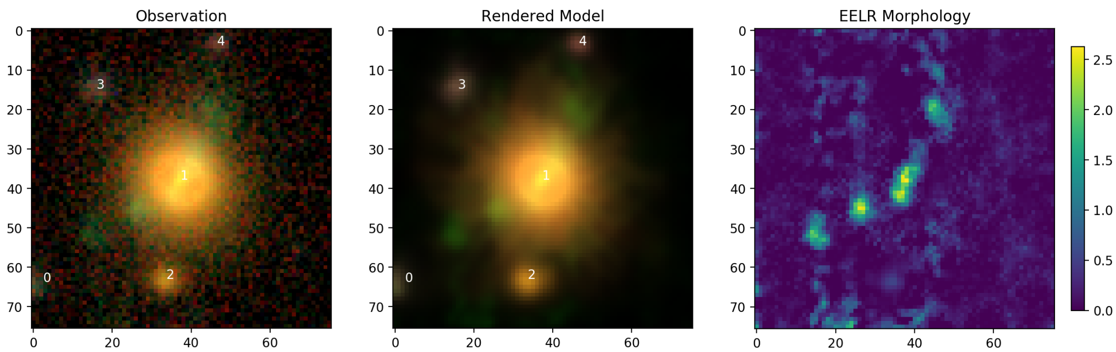
Figure 6 shows an example output of SCARLET using this configuration. In Figure 6b, Source 0 is the two-component host galaxy model and Source 1 is the EELR model. Sources 2 and 3 are two other objects detected in the scene. The "rendered" models are simply the SCARLET models convolved with the PSF. Note that these figures are from a single run of SCARLET, i.e., without model averaging.

4.4. Model Averaging

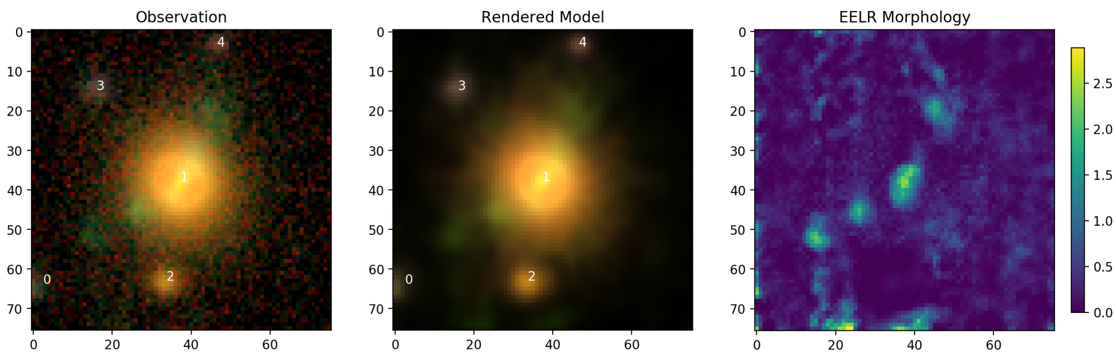
Using the procedure described in Section 4.2, for a given observation, we generate a sample of 50 EELR spectra. For each of these samples, we deblend the observation using SCARLET and the configuration described in Section 4.3, constraining the EELR's SED to the sampled spectrum. SCARLET returns a scene model as well as the model's log-likelihood. Finally, we combine all these models by computing the model average using Equation 3.

One implementation issue is arithmetic underflow. Often, different model samples for the same observation have vastly different likelihoods due to some models fitting the observation very poorly. This causes underflow when Equation 3 is applied directly. Therefore, we first drop any samples with log-likelihood less than the highest log-likelihood minus 7 (i.e., with likelihood more than $\exp(7) \approx 1097$ times smaller than the highest likelihood). We then subtract the largest log-likelihood from all the remaining log-likelihoods so that the normalized likelihoods are all between 0 and 1, making the computation of Equation 3 well-conditioned.

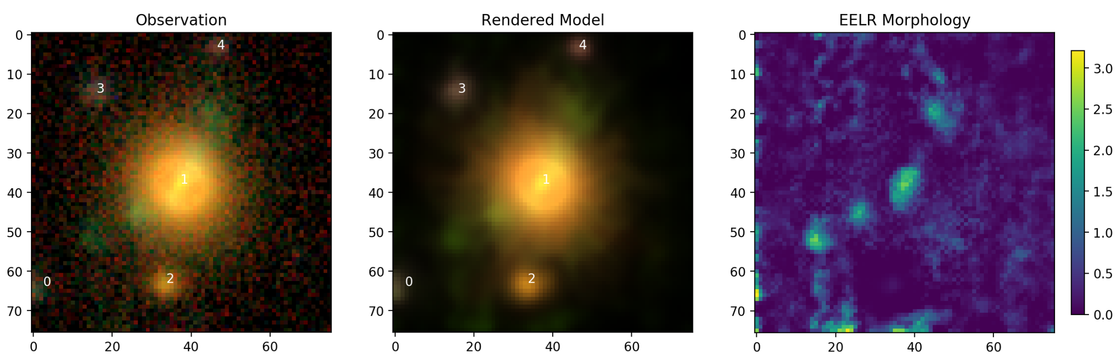
We compute the likelihood-weighted model average of the entire SCARLET model (i.e. the entire scene), as well as the likelihood-weighted model average of just the EELR morphology. Figure 7b shows an example result.



(a) Model averages using SDSS spectra, i.e., the baseline method.



(b) Model averages using the GP method.



(c) Model averages using the GP method with noisy redshift measurement.

Figure 7: Example likelihood-weighted model averages of a full scene (middle column) and of just the EELR morphology (right column).

5. Results

5.1. Metrics

Assuming a homoscedastic Gaussian error, a simplified version of the negative log-likelihood minimized by SCARLET is

$$f(\mathsf{A},\mathsf{S}) = \frac{1}{2}\|\mathsf{Y} - \mathsf{A}\mathsf{S}\|_F^2 \tag{5}$$

	Med. MSE	Med. MPIV
Baseline	**5.258**	**0.0098**
GP	5.276	0.0108
Noisy GP	5.281	0.0100

Table 1: Median performance metrics.

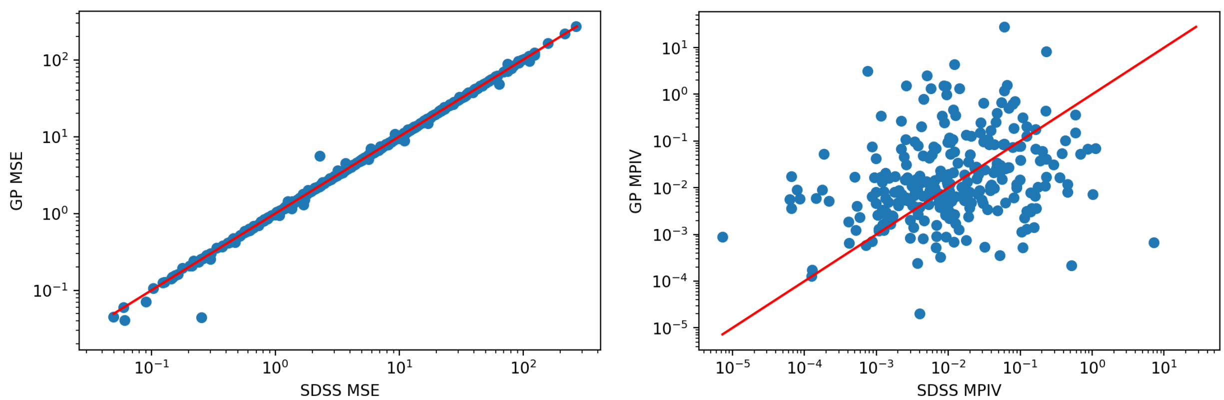
where A is the amplitude/SED matrix, S is the morphology matrix, and Y is the observed image. Therefore, one performance metric we use is the mean squared L2 error (MSE), which is proportional to the mean of the simplified negative log-likelihood given above.

Another metric we use is the mean pixel intensity variance (MPIV) of the EELR model. This is done by applying Equation 4 to get the pixel-wise model variance, and then computing the mean over all pixels. This provides a measure of uncertainty about our model average EELR morphology, which we would generally like to minimize.

Since these metrics do not mean much by themselves, we define a baseline method for comparison. The baseline method is the same as our full method except it uses spectra from SDSS to constrain the EELR spectra, so only the y band is randomly sampled. Note that SCARLET's random initialization of the EELR morphology (see Section 4.3) is also a source of randomization. For all methods, we use 50 samples per observation to compute model averages. We expect the baseline method to perform best since it has the most precise data. Therefore, our goal is to get performance close to that of the baseline method, not to beat it. We compare the baseline method with our full method using GPs to sample spectra, which we call the "GP method." Table 1 gives a summary of the results, and Figure 7 shows an example output using each method.

5.2. Spectroscopic Redshift

In this comparison, the GP method takes in spectroscopic redshifts from SDSS. A per-observation performance comparison between the baseline method and the GP method is shown in Figure 8. In terms of MSE, the GP method beats (i.e. has lower MSE) the baseline method in 49.8% of the observations. In terms of MPIV, the GP method beats (i.e. has lower MPIV) the baseline method in 42.7% of the observations. These comparative metrics as well as the median metrics in



(a) MSE comparison. The red line is the 1:1 line.

(b) MPIV comparison. The red line is the 1:1 line.

Figure 8: Comparison of performance metrics between the baseline method and the GP method.

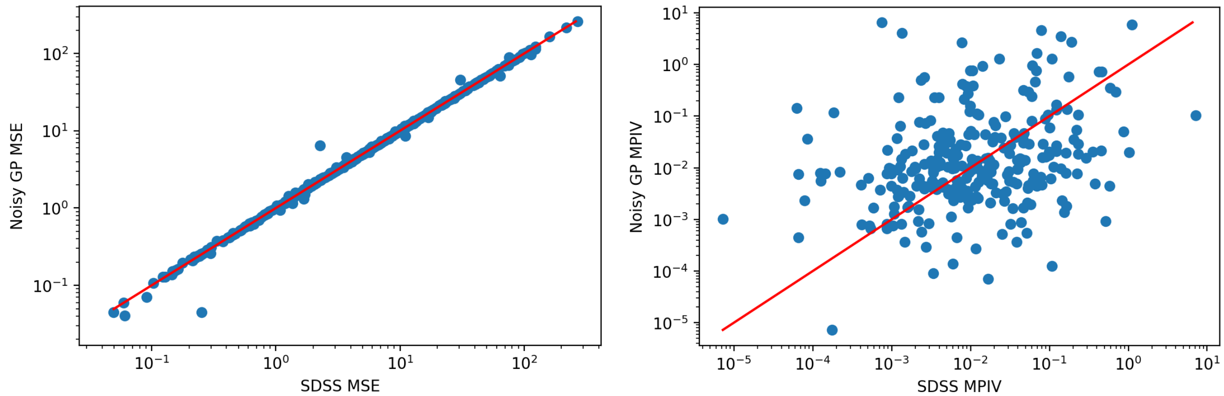
Table 1 suggest that, as desired, using redshift rather than direct spectrum measurements seems to negligibly degrade performance.

5.3. Noisy Redshift

In this comparison, the GP method takes in *noisy* redshifts to simulate the loss of precision when using photometric rather than spectroscopic redshift measurements. In particular, for each observation, we add uniform random noise between -0.05 and 0.05 to the SDSS spectroscopic redshift. This range of noise values roughly corresponds to the range of noise from modern photometric redshift techniques [18]. A per-observation performance comparison between the baseline method and the GP method with noisy redshift is shown in Figure 9. In terms of MSE, the GP method beats the baseline method in 46.2% of the observations. In terms of MPIV, the GP method beats the baseline method in 47.3% of the observations. These comparative metrics as well as the median metrics in Table 1 suggest that our method is robust to noisy redshift measurements.

6. Conclusion

AGN have been theorized to play an important role in several cosmological phenomena, yet AGN are poorly understood. EELRs emitted by AGN are a promising source of information since they represent large-scale interactions of AGN with their host galaxies. However, as astronomical surveys grow larger and see deeper, we can no longer rely on targeted observations or spectroscopic



(a) MSE comparison. The red line is the 1:1 line.

(b) MPIV comparison. The red line is the 1:1 line.

Figure 9: Comparison of performance metrics between the baseline method and the GP method with noisy redshift.

measurements if we want to take advantage of this growing amount of data. Therefore, we have developed a method of imaging EELRs automatically and accurately using only photometric data.

Our method takes in two pieces of data to search for an EELR: redshift and a multi-band image. Redshift can be estimated from the broad-band image using photometric redshift techniques, so in reality only the multi-band image is needed. We use Gaussian Process regression to fit Gaussian distributions of EELR SED to redshift. After generating a sample from these GPs, we then use SCARLET's constraint system to constrain the EELR's SED as well as other physical characteristics of the objects in the observation. SCARLET produces a model and a likelihood for each of these sampled spectra. We then compute the likelihood-weighted model average to produce a final, combined model.

Our method seems to perform quite well despite its lack of precise input data, performing similarly to a method which is given spectra measurements from SDSS rather than having to infer these from redshift. Thus, we believe this method holds promise for detecting and imaging EELRs in future large-scale surveys, enabling future studies on EELRs and AGN.

7. Future Work

After 10 years of observation, the Legacy Survey of Space and Time, currently under construction, will provide observations over a much larger area and greater depth than HSC. This will be an important proving ground for our method in the future.

Although we have examined a very specific application in this work, our approach to search in multi-band images is quite general. For the same reasons our approach is useful for EELR imaging, it could potentially be useful for imaging dust lanes in galaxies, or the bars of spiral galaxies. In both cases, our sampling approach to marginalizing over uncertain SED and morphology constraints could be applied.

There has also been recent work in the area of *multi-output* Gaussian Process modeling (e.g., Parra and Tobar [14]). In our case, this would amount to not only modeling correlations between bands and redshift, but also modeling between-band correlations, and could thus produce better samples. This may in turn increase the computational efficiency of our method by reducing the number of samples needed to produce a good model average.

8. Acknowledgments

I would like to thank my advisor Peter Melchior for his mentorship and guidance throughout this project. I would also like to thank Andy Goulding from the Department of Astrophysical Sciences at Princeton University for gathering the HSC and SDSS data used in this project.

9. Code

All of this project's code is available at `https://github.com/pmelchior/hsc_eelr`.

References

[1] H. Aihara *et al.*, "Second data release of the Hyper Suprime-Cam Subaru Strategic Program," *Publications of the Astronomical Society of Japan*, vol. 71, no. 6, 2019.

[2] H. Aihara *et al.*, "The Hyper Suprime-Cam SSP Survey: Overview and survey design," *Publications of the Astronomical Society of Japan*, vol. 70, no. SP1, 2017.

[3] C. R. Almeida *et al.*, "Testing the unification model for AGN in the infrared: are the obscuring tori of type 1 and 2 Seyferts different?" *The Astrophysical Journal*, vol. 731, no. 2, p. 92, 2011.

[4] K. Barbary, "SEP: Source extractor as a library," *Journal of Open Source Software*, vol. 1, no. 6, p. 58, 2016.

[5] E. Bertin and S. Arnouts, "SExtractor: Software for source extraction," *Astronomy & Astrophysics Supplement Series*, vol. 117, pp. 393–404, 1996.

[6] M. Bolzonella, J. M. Miralles, and R. Pelló, "Photometric redshifts based on standard SED fitting procedures," *Astronomy and Astrophysics*, vol. 363, pp. 476–492, 2000.

[7] R. G. Bower *et al.*, "Breaking the hierarchy of galaxy formation," *Monthly Notices of the Royal Astronomical Society*, vol. 370, no. 2, pp. 645–655, 2006.

[8] C. J. Burke *et al.*, "Deblending and classifying astronomical sources with Mask R-CNN deep learning," *Monthly Notices of the Royal Astronomical Society*, vol. 490, no. 3, pp. 3952–3965, 2019.

[9] Z. He *et al.*, “Morphology of AGN emission-line regions in SDSS-IV MaNGA survey,” *Monthly Notices of the Royal Astronomical Society*, vol. 478, no. 3, pp. 3614–3626, 2018.

[10] K. Inskip *et al.*, “The origin of extended emission line regions,” *New Astronomy Reviews*, vol. 51, no. 1, pp. 47 – 51, 2007.

[11] R. Lupton, “Deblending and why it is impossible,” 2016.

[12] P. Melchior *et al.*, “scarlet: Source separation in multi-band images by constrained matrix factorization,” *Astronomy and Computing*, vol. 24, pp. 129 – 142, 2018.

[13] J. R. Mullaney *et al.*, “Narrow-line region gas kinematics of 24264 optically selected AGN: the radio connection,” *Monthly Notices of the Royal Astronomical Society*, vol. 433, no. 1, pp. 622–638, 2013.

[14] G. Parra and F. Tobar, “Spectral mixture kernels for multi-output gaussian processes,” in *Advances in Neural Information Processing Systems 30*, I. Guyon *et al.*, Eds. Curran Associates, Inc., 2017, pp. 6681–6690.

[15] F. Pedregosa *et al.*, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[16] R. Reyes *et al.*, “Space density of optically-selected type 2 quasars,” *The Astronomical Journal*, vol. 136, no. 6, pp. 2373–2390, 2008.

[17] A. Robinson, *Extended Emission-Line Regions Around AGN*, ser. Astronomical Society of the Pacific Conference Series, 1997, vol. 113, pp. 280–287.

[18] M. Salvato, O. Ilbert, and B. Hoyle, “The many flavours of photometric redshifts,” *Nature Astronomy*, vol. 3, no. 3, pp. 212–222, 2019.

[19] A.-L. Sun *et al.*, “Imaging extended emission-line regions of obscured AGN with the Subaru Hyper Suprime-Cam Survey,” *Monthly Notices of the Royal Astronomical Society*, vol. 480, no. 2, pp. 2302–2323, 2018.

[20] M. Villar-Martín *et al.*, “The giant star forming halo associated with the radio galaxy PKS 193246,” *Monthly Notices of the Royal Astronomical Society: Letters*, vol. 359, no. 1, pp. L5–L9, 2005.

[21] B. Villarroel *et al.*, “AGN luminosity and stellar age: Two missing ingredients for AGN unification as seen with iPTF supernovae,” *The Astrophysical Journal*, vol. 837, no. 2, p. 110, 2017.

[22] D. G. York *et al.*, “The Sloan Digital Sky Survey: Technical summary,” *The Astronomical Journal*, vol. 120, no. 3, pp. 1579–1587, 2000.

[23] S. Yuan, M. A. Strauss, and N. L. Zakamska, “Spectroscopic identification of type 2 quasars at $z < 1$ in SDSS-III/BOSS,” *Monthly Notices of the Royal Astronomical Society*, vol. 462, no. 2, pp. 1603–1615, 2016.

[24] E. Zackrisson, “Quasars and low surface brightness galaxies as probes of dark matter,” Ph.D. dissertation, Department of Astronomy and Space Physics, Uppsala University, Sweden, 2005.

[25] N. L. Zakamska *et al.*, “Candidate type II quasars from the sloan digital sky survey. i. selection and optical properties of a sample at $0.3 < z < 0.83$,” *The Astronomical Journal*, vol. 126, no. 5, pp. 2125–2144, 2003.