School of Physics and Astronomy

Senior Honours Project
Astrophysics

Deep Stacking of AGN Hypervariables

Alexander S. Wheaton
April 9, 2020

Abstract

In this project, I examine the unusual luminosity of the galactic nucleus SDSS J094511 located at redshift $z = 0.758$ using data gathered with the Liverpool Telescope on Las Palmas between 2012 and 2018. I develop several algorithms for centroiding on the object, for resampling to obtain higher resolution, and for stacking images to increase the S/N ratio in the frame. I extract radial and linear cross sections from the object and compare these to the point-spread-function of the Liverpool Telescope CCD to corroborate the hypothesis that linear increases in the luminosity over large timescales are *not* due to instability of the viscous AGN accretion disk, but a gravitational microlensing event by a star in an intervening galaxy or dwarf galaxy.

Declaration

I declare that this project and report is my own work.

Signature:  Date: April 9, 2020

Supervisor: Professor A. Lawrence, FRSE, FRaS 12 Weeks

Contents

1 Introduction

The luminosity of most galaxies is dominated by stellar emission and emission from hot gas and dust in the interstellar medium, which is heated by those stars. As such, it is mostly confined to the wavelength range of typical stars—the ultraviolet through the infrared. Between 1 and 50% of galaxies in the local universe, however, show strong emission in a wider range of bands, from gamma rays to the radio spectrum. The source of these emissions is typically confined to a compact nuclear region, and such objects are therefore called active galactic nuclei (AGN).[8]

AGN are distinguished from one another by their luminosities, spectra, and ratio of nuclear to galactic luminosity into four main groups: Seyfert galaxies, quasars, radio galaxies, and optically-violent variables. The aforementioned properties vary among these broad catagories, but most AGN exhibit (among other common characteristics):

- luminous and compact appearance
- broad spectral energy distribution which is not explicable by thermal (blackbody) emission
- variability over short timescales of minutes to days

which are relevant to this project.[8]

1.1 The Viscous Accretion Disc

The underlying mechanism for this luminosity is widely understood to be the accretion of matter onto a supermassive black hole.[8] The luminosity, compactness and sprectra of AGN are well described by a model in which matter that falls into a black hole with some amount of angular momentum and forms an accretion disk. Gravitational potential energy is changed for kinetic energy (and therefore angular momentum), and viscous processes between particles on different orbits create drag, transferring angular moment outward and heating the disk. This energy is then radiated thermally.[7]
Small adjustments to this model for surrounding clouds and "atmosphere" closely fit observed spectral energy distributions.[7]. What this model fails to explain, however, is the extreme variability of AGN. A viscosity parameter which explains observed luminosities should also result in a very stable disk, for which changes in the optical spectrum take place only on the order of 10^3 years.[7]

The accretion disk model was briefly rescued by the idea of X-ray reprocessing **Citation Needed**, but recent observations of "hypervariable," high-luminisity AGN by the Sloan Digital Sky Survey (SDSS), the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) have once again cast uncertainty onto the model. These objects vary on timescales comparable to the time for information to propagate over the Scharzchild radius of a supermassive black hole[8][7]:

$$\Delta t \geq \frac{R_{Sch}}{c} = \frac{2GM}{c^3} \tag{1}$$

That such extreme changes should happen on these timescales and in the optical spectrum suggests that the disk undergoes incredible physical change and that this change propagates relativistically through the disk—a strong argument for instability.

1.2 The Pan-STARRS Survey

In 2016 Lawrence et. al identified fifteen transiently luminous, blue AGN in the Pan-STARRS1 Survey by comparison with earlier data from the SDSS. These targets were subsequently monitored with the Liverpool Telescope (LT) on Las Palmas, and found to be hypervariable AGN over a timescale of decades.[5]

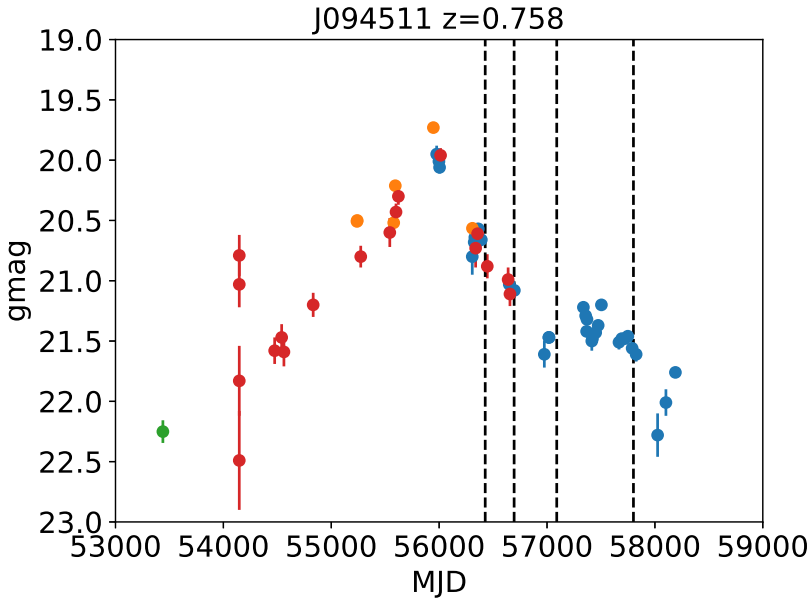

Figure 1: Magnitude of SDSS J094511 from SDSS, CRTS, PanSTARRS, and LT observations, over ten years.

One such transiently luminous, blue hypervariable AGN is SDSS J094511, a **insert quasar classification here** in the constellation Leo and at redshift $z = 0.758$.[5][2] Fig. 1 shows the evolution of magnitide of J094511 in the Sloan g' band between 2004 and 2014, using combined data from the SDSS, from the Catalina Real-Time Transit Survey (CRTS), from Pan-STARRS1, and the LT. The oject evolves at ~ 0.4 mag/yr from $g' \sim 22.2$ to 19.7 and back again, during which time its luminosity changes by a factor of ~ 2.[2]. This magnitude of evolution over a period of ten years is not explicable by the accretion disk model alone.

Lawrence et. al therefore propose several possible explanations for these events, with the most likely candidates being i) changes in the accretion state of the disk (i.e. matter falling into an otherwise quiet AGN) and ii) large amplitude microlensing events by stars in intervening galaxies or dwarf galaxies.[5] Microlensing is a well established explanation for AGN variability, and can take the form of either macro-lensing due to the gravitational of the entire intervening galaxy or microlensing by a single or few constituent stars. The latter event is more likely to be the case for small galaxies, for which the macroscopic gravitational potential does not dominate the lensing effect.[5]

Possibly more information from Lawrence et. al p.16 here

There are several signatures which may indicate that this is the case for a particular object. If the AGN is modeled as point source and intervening star modeled as a lens (henceforth: simple microlensing model), it is possible to compute the lightcurve $F(t) = \mu(t)F_s$ which describes the evolution of the lensed AGN with respect to its unlensed flux, as described by Bruce et. al (2017). This model contains seven free parameters, one of which is a flux contribution F_b from the host galaxy of the lensing star(s). Although a simplification, this model is particularly useful for understanding hypervariables like J094511, whose light curves evolve smoothly.[2]

The parameter F_b is key to demonstrating a lensing event, as background flux may be detectable in the cross sectional luminosity curve of the object. In this project, I develop methods for quantifying this parameter by stacking frames from the Liverpool Telescope to improve signal to noise ratio and depth of field, and for extracting the luminosity profile of objects in the frame for comparison to the point spread function of the Liverpool Telescope.

2 Liverpool Telescope Data

There are several technical problems associated with the LT data. In many frames, astronomical seeing is very poor. All frames, particularly in the U-band, are littered with hot pixels or cosmic rays detections. Each of these factors obfuscates the already faint luminosity profile of J094511, and I have employed several statistical methods to mitigate them.

2.1 Stacking to Increase S/N Ratio

For astronomical frames in which the noise is background limited (i.e. thermal, source, and read noise are negligible compared to the sky) the signal to noise ratio goes as[6][8]:

$$\frac{S}{N} = \frac{n_\star t}{\sqrt{n_{pix} n_s t}} \propto t^{1/2} \tag{2}$$

Where t is the exposure time of the frame. S/N ratio increases with exposure time, although with diminishing marginal returns. Very long exposures, however, are typically limited by the saturation saturation point of the CCD. A typical approach to increasing the S/N ratio of observations, then, is to "stack" the frames of many short exposures, summing up the counts recieved at each point in the sky in each frame.[8] As the S/N ratio increases, faint objects in the frame like J094511 become more distinguishable from the background (increased depth of field), and cosmic ray detections disappear into the background, both without saturating the CCD.

2.2 Astronomical Seeing

Atmospheric turbulence creates a differential density of air above ground-based telescopes, distorting light wavefronts and obfuscating astronomical objects with "speckles." This phenomena is called atmospheric seeing, and is typically quantified by θ_s, the hypothetical full width at half maximum of an "airy disc" which distorts incoming wavefronts.

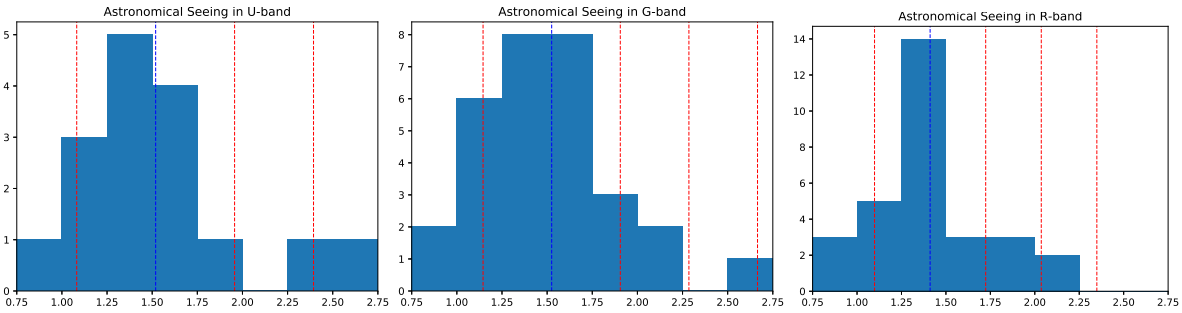

Figure 2: Astronomical seeing of the LT data in Sloan *u'g'r'* bands.

Astronomical seeing values of $\theta_s = 1''$ are typical for optical wavelengths, where the Sloan bands lie. Values below this are considered exceptionally good seeing.[8][6] It is desirable to

discard frames with higher seeing values, but this must be balanced against the need for additional frames to increase the S/N ratio in the final stack. Fig. 2 shows the distribution of seeing values, in arcseconds, in each of the Sloan *u'g'r'* bands, with the mean (blue) and standard deviations (red) plotted vertically. It is clear that discarding frames with $\theta_s > 1''$ would result in substantially reduced dataset, and a resultingly low S/N ratio for an already faint object. For this reason, I have elected to discard all frames for which $\theta_s > \bar{\theta_s} + \sigma_\theta$, where the seeing parameter is more than one standard deviation above the mean parameter. This includes the most of the available data in each band, while exluding the worst seeing parameters and is a more rigorous and reproducible approach to ensuring the best possible stack.

A single image of J094511 in the U-band has a seeing of $r_0 = 999.0''$, which is likely an error or the result of exceptional atmospheric turbulence. Including this value in the statistics derived from the seeing distribution would badly skew $\bar{\theta_s}$ and σ_θ For this reason, the image has been discarded and its seeing parameter excluded from the calculation of the mean and standard deviation of the seeing distribution.

2.3 FITS World Coordinate Data

Finally, individual frames in the dataset are not all centered on the same equatorial coordinates, and must therefore be aligned before they can be stacked. Individual frames are stored using the NASA FITS standard, which contains a series of header data units (HDUs), each of which contain a data cube and metadata associated with that cube.[9] The LT HDUs contain a single, 2-dimensional arrays of counts per pixel for each frame. The metadata of each HDU contains the astronomical seeing, roation, and crucially, the World Coordinate System (WCS) associated with each frame.[4] The WCS data contains several key pieces of information for aligning individual frames.

Header Key	Value	Meaning
CTYPE1	'RA—TAN'	WCS projection
CTYPE2	'DEC–TAN'	WCS projection
CRPIX1	512.	WCS reference pixel coordinate
CRPIX2	512.	WCS reference pixel coordinate
CRVAL1	146.295274426	[degrees] World coordinate at the ref pix
CRVAL2	17.76338016	[degrees] World coordinate at the ref pix
CDELT1	-7.7508973E-05	[degrees/pixel]
CDELT2	7.7508973E-05	[degrees/pixel]
CROTA1	90.361783	[degrees]
CROTA2	90.361783	[degrees]

Table 1: Example FITS header keys and values from the LT.

While the WCS projection and platescale (CDELT1/CDELT2) are constant across all frames, other WCS values—most importantly the world coordinate at the reference pixel—vary from frame to frame. Random frames are rotated by multiples of $\pi/4$ from the $(x, y) \rightarrow (-\alpha, \delta)$ mapping. Additionally, frames taken after 20?? have a twice the field of view of previous frames, with a reference pixel of (1024,1024), rather than (512,512). There are also systematic errors in (α, δ) coordinate of the reference pixel (see Sec. 4.2).

3 The AstroPy and FITS-Utils Modules

To correct for these factors, I utilised the AstroPy module for Python 3.5.2. The AstroPy module allows one to load in FITS compliant files as HDU objects which contain all the data and metadata associated with a particle file. Performing math on the data stored in these objects, however, can modify the original FITS file[1][10], so I modified and extended the functionality of the fits-utils module, written for aligning and stacking on the Telescope Group Project course at the University of Edinburgh.[1] HDU data cubes are loaded into NumPy arrays and stored in an "image-dictionary." Relevant metadata for each image are also stored in this dictionary, which behaves similarly to the AstroPy HDU object.

I then algorithmically sort frames by seeing value according to the standards defined in Sec. 2.2, rotate each array if necessary so that $(x, y) \rightarrow (-\alpha, \delta)$, crop the array if necessary to 1024x1024 pixels, and reassign the coordinates of reference pixel to (512,512) if the array is cropped. The image-dictionaries are then ready to be aligned and stacked.

4 Centroiding on J094511

In order to align each image, its offset, in pixels, with respect to every other image in the stack must be determined.The location of J094511 varies around the reference pixel in each frame, and the relative offset of each image can be computed by comparing the location of the object in each.[2]

Determining the precise location of the object, however, is a non-trivial task. J094511 has a spread of ~ 10 pixels in each frame. A naive approach to determining its centerpoint would be to take the pixel coordinates of the maximum value of the object, but doing this algorithmically quickly proves futile and innaccurate. J094511 is by far one of the faintest objects in the frame, and there is no obvious way to distinguish pixels which contain the object from pixels which do not. Even a hard coded cutout from the array, which was manually verified to always contain the object, frequently contains sufficient noise and cosmic ray detections to confuse any maximum value finding algorithm. Even if this were possible, poor atmospheric seeing distorts the object so that the brightest pixel is commonly *not* at the center of the object (dubiously assuming that the "above atmosphere" brightest pixel is even at the center, or the same across frames).

4.1 Two Dimensional Weighted Mean

A better method for computing the centroid of the object is to calculate the flux weighted mean in two dimensions:

$$(\bar{x}, \bar{y}) = \left(\frac{1}{n_{tot}} \sum_i n(x_i) x_i, \frac{1}{m_{tot}} \sum_j m(y_j) y_j \right) \tag{3}$$

Whereby all the counts are first summed along the y-axis and the flux weighted mean of position, $\bar{x}$, calculated, and then the inverse—summing along the x-axis and calculating $\bar{y}$—is performed. This is a more reliable method of calculating the object centroid, which tends to downweight

[1] `github.com/aswheaton/fits-utils`

[2] Note that from here on, the phrases "pixel coordinates" or "image coordinates" will be used to refer to the (x, y) coordinates of pixels in the image, which are mapped to $(-\alpha, \delta)$ as defined by the FITS file standard[9]. The phrase "array coordinates" will be used to refer to the $(row, column)$ indices of the NumPy array containing the image. This is an important distinction, as pixel coordinates are mapped to array coordinates by $(x, y) \rightarrow (column, row)$.

individual speckles from atmospheric seeing and cosmic ray detections. Unfortunately, performing a 2D weighted mean on the entire frame invariably yields the reference (center) pixel coordinates for $(\bar{x}, \bar{y})$. No source in the LT frames is sufficiently bright to reliably pull this average away from the center pixels, even when rescaled exponentially. Least of all is J094511 able to do this.

The weighted average calculation, it turns out, is only sensitive to objects which nearly fill the sample region. There is enough variation in the location of J094511 between frames that a hard coded cutout from the array which will always contain the object is large enough to suffer the same problem. A smaller cutout region is need. Thus, a technique for accurately and algorithmically making an initial guess at the location of the object, around which a small cutout region can be made dynamically, was developed using the WCS data from the LT FITS headers.

4.2 Initial Guess: WCS Centroiding

By comparing the known coordinates of J094511, (α,δ) = (9:45:11.08,+17:45:44.78) to the world coordinates at the reference pixel, we can compute an initial guess for the location of J094511. Some trivial linear algebra gives that:

$$x_\star = p_\alpha(\alpha_\star - \alpha_{ref}) + x_{ref} \tag{4}$$

$$y_\star = p_\delta(\delta_\star - \delta_{ref}) + y_{ref} \tag{5}$$

Where (x, y) are pixel coordinates, (α, δ) are equatorial coordinates in degrees, $\star$ denotes values associated with the object of interest and ref denotes values associated with the reference pixel. The values p_α and p_δ are the platescale values from the FITS header, corresponding respectively to keys CDELT1 and CDELT2 in Table 1.

In developing this method, however, I found that the LT WCS data is consistently off by ~ 25 pixels along the x/α-axis, or about $1''$. This is easily corrected by an extra term in Eq. 5, so I wrote the WCS centroiding function to accommodate a user specified correction factor.

Once corrected, the pixel coordinatess of the intitial guess are rounded to the nearest whole pixel, mapped back to array coordinates and a cutout of 10x10 pixels taken around them. I perform the weighted mean calculation for this smaller array, round the result to the nearest whole pixel value, and map this centroid back to the array coordinates of the entire image.

This process produces reasonably good centroid values in almost every frame, but has some disadvantages. First, there are two points in the process where a non-integer pixel coordinate are necessarily rounded: once when calculated the offset from the reference pixel, and again after performing the weighted mean. Each of these introduce an error into the centroid calculation. For larger objects, this might be negligible, but for objects such as J094511 which are scarcely 10 pixels across on the LT CCD, each error of ± 0.5 pixels is considerable. Second, this approach still limits resolution to the number of pixels in the objects' footprint on the CCD, while the true centroid is, of course, not guaranteed to be at an integer pixel value in any frame.

Once these centroid values are obtained, I carry out the alignment of each frame and stack them accordingly. This process is elaborate, and will only be described in summary here.[3] First, the maximum displacements in array coordinates Δr_{max} and Δc_{max} are calculated from all the centroid values in the stack. These values are used to create an array of zeros, with dimensions $(1024 + \Delta r_{max})$x$(1024 + \Delta c_{max})$, which are sufficient to hold all the images, once aligned. For each image, then, the displacement in array coordinates with respect to the *minimum* centroid

[3]Full implementation of the algorithm at `github.com/aswheaton/lt-stacking-project`.

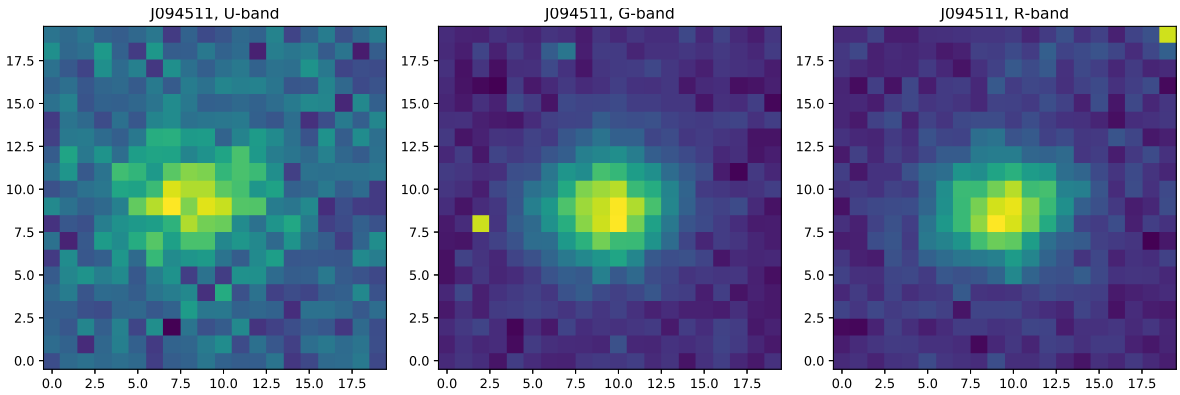

Figure 3: Image stack of J094511 in the Sloan $u'g'r'$ bands.

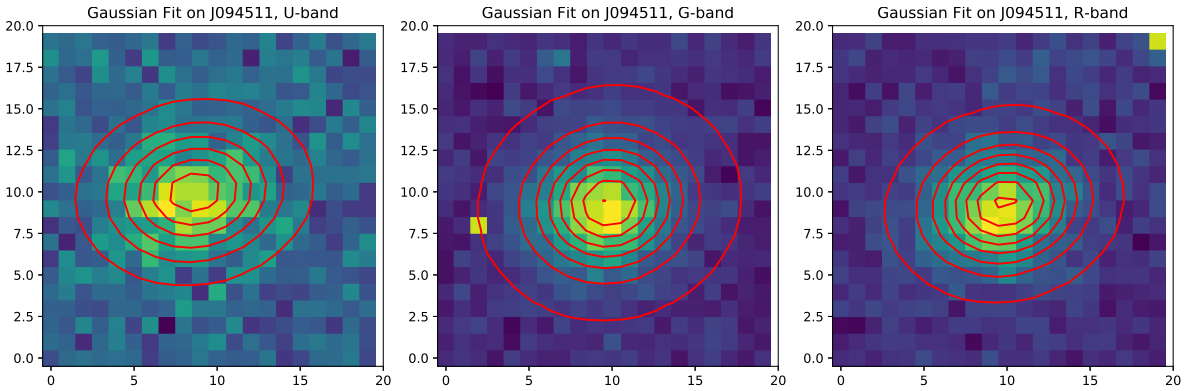

Figure 4: Gaussian fit on J094511 in the Sloan $u'g'r'$ bands.

values in the stack, r_{min} and c_{min}, and these values, Δr and Δc are used to cast the unaligned image into its aligned position in the array of zeros. This is done such that:

$$(r_{align}, c_{align}) = (r_{unalign} + \Delta r, c_{unalign} + \Delta c) \tag{6}$$

For each pixel in the array. The resulting stack of aligned images is summed and the resulting array saved to a new image dictionary. This produces a much better defined object, as shown in Fig. 3. By fitting a two dimensional gaussian to this data, I begin to quantify the properties of J094511 and the sky background on which it lies. Table 2 shows the fitted parameters for the resulting stack in each of the Sloan $u'g'r'$ bands.

Band	Amplitude	x_0	y_0	σ_x	σ_y	θ
u'						
g'						
r'						

Table 2

Discuss fitted parameters and errors here!

4.3 Refinement: Gaussian Fitted Centroiding

Gaussian fitting, however, presents a solution to some of the shortcomings of the flux weighted mean centroid described in Sec. 4.1. Once a gaussian function is fitted to the existing data, the

fitted parameters can be used to interpolate the number of counts at non-integer pixel values. This is called resampling, and can be used to obtain higher resolution

In two dimensions, the power to which e is raised in the Gaussian function is any negative-definite quadratic form. Consequently, the level sets of the Gaussian will always be ellipses.[3] A particular example of a two-dimensional Gaussian function is:

$$f(x,y) = A\exp\left(-\left(\frac{(x-x_0)^2}{2\sigma_x^2} + \frac{(y-y_0)^2}{2\sigma_y^2}\right)\right) \tag{7}$$

Where A is the amplitude, (x_0, y_0) are the center in pixel coordinates, σ_x, σ_y are the spread of gaussian in the x and y directions.[11] Generally, a two-dimensional elliptical Gaussian function is expressed as:

$$f(x,y) = A\exp\left(-\left(a(x-x_0)^2 + 2b(x-x_0)(y-y_0) + c(y-y_0)^2\right)\right) \tag{8}$$

The coefficients:

$$a = \frac{\cos^2\theta}{2\sigma_x^2} + \frac{\sin^2\theta}{2\sigma_y^2} \tag{9}$$

$$b = \frac{\sin 2\theta}{4\sigma_x^2} - \frac{\sin 2\theta}{4\sigma_y^2} \tag{10}$$

$$c = \frac{\sin^2\theta}{2\sigma_x^2} + \frac{\cos^2\theta}{2\sigma_y^2} \tag{11}$$

Allow one to specify a counterclockwise angle θ.[3]

5 Cross Sections of J094511

5.1 Linear and Radial Cross Sections

5.2 Excess Luminosity in the Wings

5.3 Point Spread Function Subtraction

5.4 Spectrum of J094511

Compare the magnitude in each band to typical AGN spectra (see AGN Zoo notes)

6 Conclusion

References

[1]Astropy Collaboration, T. P. Robitaille, and E. J. Tollerud, "Astropy: A community Python package for astronomy", A&A **558**, A33, A33 (2013).

[2]A. Bruce, A. Lawrence, C. MacLeod, M. Elvis, M. J. Ward, J. S. Collinson, S. Gezari, P. J. Marshall, M. C. Lam, R. Kotak, C. Inserra, J. Polshaw, N. Kaiser, R.-P. Kudritzki, E. A. Magnier, and C. Waters, "Spectral analysis of four 'hypervariable' agn: a microneedle in the haystack?", Monthly Notices of the Royal Astronomical Society **467**, 1259–1280 (2017).

[3]A. Developers.

[4]E. W. Greisen and M. R. Calabretta, "Representations of world coordinates in fits", Astronomy & Astrophysics **395**, 1061–1075 (2002).

[5]A. Lawrence, A. G. Bruce, C. MacLeod, S. Gezari, M. Elvis, M. Ward, S. J. Smartt, K. W. Smith, D. Wright, M. Fraser, P. Marshall, N. Kaiser, W. Burgett, E. Magnier, J. Tonry, K. Chambers, R. Wainscoat, C. Waters, P. Price, N. Metcalfe, S. Valenti, R. Kotak, A. Mead, C. Inserra, T. W. Chen, and A. Soderberg, "Slow-blue nuclear hypervariables in panstarrs-1", Monthly Notices of the Royal Astronomical Society **463**, 296–331 (2016).

[6]A. Lawrence, *Astronomical measurement : a concise guide*, eng (Springer, Heidelberg ; London, 2014).

[7]A. Lawrence, "Quasar viscosity crisis", eng, Nat Astron **2**, 102–103.

[8]R. McClure, *Lecture notes in astrophysics: active galaxies*, 2019.

[9]Pence, W. D., Chiappetti, L., Page, C. G., Shaw, R. A., and Stobie, E., "Definition of the flexible image transport system (fits), version 3.0", A&A **524**, A42 (2010).

[10]A. M. Price-Whelan, B. M. Sipőcz, and A. Contributors, "The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package", AJ **156**, 123, 123 (2018).

[11]E. W. Weisstein, *"gaussian function." from mathworld–a wolfram web resource.* (2020) `https://mathworld.wolfram.com/GaussianFunction.html`.

Appendix A: Alignment Algorithm