A Bayesian Analysis of SDSS J0914+0853, a Low-mass Dual AGN Candidate

Adi Foord1, Kayhan Gültekin2, Mark T. Reynolds3, Edmund Hodges-Kluck2,3, Edward M. Cackett4, Julia M. Comerford5, Ashley L. King6, Jon M. Miller7, and Jessie C. Runnoe8

1 Department of Astronomy and Astrophysics, University of Michigan, Ann Arbor, MI 48109, USA
2 Department of Astronomy, University of Maryland, College Park, MD 20740, USA
3 Code 662, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
4 Department of Physics and Astronomy, Wayne State University, Detroit, MI 48202, USA
5 Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder, CO 80309, USA
6 Department of Physics, Stanford University, Stanford, CA 94305, USA

Received 2019 February 22; revised 2019 April 8; accepted 2019 April 10; published 2019 May 17

Abstract

We present the first results from Bayesian AnalySis of Multiple AGN in X-rays (BAYMAX), a tool that uses a Bayesian framework to quantitatively evaluate whether a given Chandra observation is more likely a single or dual point source. Although the most robust method of determining the presence of dual active galactic nuclei (AGNs) is to use X-ray observations, only sources that are widely separated relative to the instrument’s point-spread function are easy to identify. It becomes increasingly difficult to distinguish dual AGNs from single AGNs when the separation is on the order of Chandra’s angular resolution (<1″). Using likelihood models for single and dual point sources, BAYMAX quantitatively evaluates the likelihood of an AGN for a given source. Specifically, we present results from BAYMAX analyzing the lowest-mass dual AGN candidate to date, SDSS J0914+0853, where archival Chandra data shows a possible secondary AGN ≈0″3 from the primary. Analyzing a new 50 ks Chandra observation, results from BAYMAX shows that SDSS J0914+0853 is most likely a single AGN with a Bayesian factor of 13.5 in favor of a single point source model. Further, posterior distributions from the dual point source model are consistent with emission from a single AGN. We find a very low probability of SDSS J0914+0853 being a dual AGN system with a flux ratio f > 0.3 and separation r > 0″3. Overall, BAYMAX will be an important tool for correctly classifying candidate dual AGNs in the literature, as well as studying the dual AGN population where past spatial resolution limits have prevented systematic analyses.

Key words: galaxies: active – galaxies: interactions – galaxies: statistics – X-rays: galaxies

1. Introduction

Given that almost all massive galaxies are thought to harbor nuclear supermassive black holes (SMBH) (Kormendy & Richstone 1995) and that classical hierarchical galaxy evolution predicts that later stages of galaxy evolution are governed by mergers (e.g., White & Rees 1978), galaxy mergers provide a favorable environment for the assembly of active galactic nuclei (AGNs) pairs (Volonteri et al. 2003). The role galaxy mergers play in triggering AGN and/or AGN pairs remains unclear (e.g., Hopkins & Quataert 2010; Kocevski et al. 2012; Schawinski et al. 2012; Hayward et al. 2014; Villforth et al. 2014, 2017; Capelo et al. 2015), but both observations and simulations agree that AGN activity should increase with decreasing galaxy separation (e.g., Koss et al. 2012; Blecha et al. 2013; Ellison et al. 2013; Barrows et al. 2017; Capelo et al. 2017; Goulding et al. 2018).

“Dual AGNs” are usually defined as a pair of AGNs in a single galaxy or merging system (with typical separations of ∼1 kpc), while a “binary AGN” is a pair of AGNs that are gravitationally bound with typical separations ≤100 pc (see Begelman et al. (1980) for a summary of the main merging phases of SMBHs). Understanding the specific environments where dual AGNs occur provides important clues about black hole growth during the merging process. Additionally, as progenitors to SMBH-binary mergers, the rate of dual AGNs is intimately tied to gravitational wave events detectable by pulsar timing arrays and space-based interferometry (see Mingarelli (2019) and references within). Thus, dual AGNs offer a critical way to observe the link between galaxy mergers, SMBH accretion, and SMBH mergers.

The frequency of galaxy mergers in our observable universe implies that dual AGN should be relatively common. In particular, assuming a dynamical friction timescale of ∼1 Gyr, we expect the galaxy merger fraction with separations ≤1 kpc by z = 0.1 to be between ∼6% and 10% (Hopkins et al. 2010). However, these estimated merger fractions do not take into account the AGN duty cycle, and observations of nearby AGN have shown that, at separations ≤1 kpc, the fraction of dual AGN may be much higher (Barrows et al. 2017). However, very few dual AGNs with separations <1 kpc have been confirmed. Such systems become difficult to resolve with Chandra beyond z ≥0.05, where separations on the order of 1 kpc approach Chandra’s angular resolution (where the half-power diameter is ≈0″8 at ≈1 keV). For example, the closest dual AGN candidate identified using two resolved point sources with Chandra is NGC 3393 (Fabbiano et al. 2011), which has a projected separation of ≈150 pc (∼0″6; however, see Koss et al. (2015) for a critical analysis of the X-ray emission). Thus, many indirect detection techniques have been developed to search for evidence of dual and binary AGNs, primarily relying on optical spectroscopy and photometry.

Perhaps the most popular method of finding dual AGN candidates is via double-peaked narrow line emission regions (e.g., Zhou et al. 2004; Gerke et al. 2007; Comerford et al. 2009, 2012, 2013; Liu et al. 2010; Wang et al. 2009; Ge et al. 2012; Fu et al. 2012). Double-peaked narrow lines can be a result of a dual AGN system during the period of the merger when their narrow line regions (NLR) are well-separated in
velocity. The system can display two sets of narrow line emission regions, such as [O III], where the separation and width of each peak will depend on parameters such as the distance between the two AGN. However, the optical regime alone is insufficient to confirm a dual AGN candidate, because of ambiguity in interpretation of the observed double-peaked NLRs. For example, bipolar outflows and rotating disks can also produce the double-peaked emission feature (e.g., Greene & Ho 2005; Rosario et al. 2010; Müller-Sánchez et al. 2011; Smith et al. 2012; Nevin et al. 2016). Indeed, follow-up observations using high-resolution imaging and spatially resolved spectroscopy have found that many double-peaked dual AGN candidates are most likely single AGNs (Shen et al. 2011; Fu et al. 2012; Comerford et al. 2015). Dual or binary AGN candidates can be confirmed using high-resolution radio imaging (Rodriguez et al. 2006; Fu et al. 2015; Müller-Sánchez et al. 2015; Kharb et al. 2017); however, an absence of radio emission does not necessarily mean an absence of AGNs (as only ~10% of AGN are radio-loud), while a detection of two radio nuclei can have multiple physical explanations (such as star-forming nuclei). Nuclei can only be classified as AGN at radio frequencies if they are compact and have flat or inverted spectral indices (e.g., Burke-Spolaor 2011; Hovatta et al. 2014).

1.1. X-Ray Observations of Dual AGN Candidates

The most robust method of confirming the presence of dual AGNs is to use X-ray observations. Due to the relatively few possible origins of emission above $10^{36}$ erg s$^{-1}$, apart from accretion onto an SMBH (Lehmer et al. 2010), X-rays are one of the most direct methods of finding black holes, especially with Chandra’s superb angular resolution. Unlike the optical regime, X-rays are less sensitive to absorption from the dusty environments of merger remnants. Currently, many analyses searching for dual AGN candidates using Chandra observations implement the Energy-Dependent Subpixel Event Respositioning (EDSER) algorithm (Li et al. 2004). EDSER improves the angular resolution of Chandra’s Advanced CCD Imaging Spectrometer (ACIS) by reducing photon impact position uncertainties to subpixel accuracy; in combination with Chandra’s dithering, this method can resolve subpixel structure down to the limit of the Chandra High Resolution Mirror Assembly (HRMA). However, thus far it has only been used to make images and qualitatively analyze them for dual point sources. In the absence of corroborating evidence from other data, the reliance on visual interpretation of dual AGNs with separations comparable to Chandra’s resolution leads to both false negatives and false positives. This issue is worse in the low-count regime (<200 counts), where even dual AGN with larger separations (>0.5") but low flux ratios are not clearly distinct.

We have developed a PYTHON tool, Bayesian AnalYsis of AGNs in X-rays (BAYMAX), that allows for a quantitative and rigorous analysis of whether a source in a given Chandra observation is more likely composed of one or two point sources. This is done by taking calibrated events from a Chandra observation and comparing them to the expected distribution of counts for single or dual source models. The main component of BAYMAX is the calculation of the Bayes factor (BF), which represents the ratio of the plausibility of observed data, given two different models. Values $>1$ or $<1$ signify which model is more likely (see Section 2 for explicit details). Further, BAYMAX returns the maximum likelihood values for the parameters of each model. In this paper, we introduce BAYMAX and present its analysis of the Chandra observations of the lowest-mass dual AGN candidate, SDSS J0914+0853. Here, we specifically highlight BAYMAX’s capabilities with respect to the Chandra observations of SDSS J0914+0853. We are using a subset of BAYMAX’s full capabilities, i.e., analyzing an on-axis source, assuming identical spectra for both the primary and secondary AGN, and deeming the background contribution to be negligible. Furthermore, false positives are only analyzed for regions in parameter space (such as count number, separation, and flux ratio) that are specific to SDSS J0914+0853. Our following paper (A. Foord et al. 2019, in preparation) will expand upon the explicit details of BAYMAX, including its capability to correctly identify dual AGN as a function of observed flux, angular separations, off-axis angle, and flux ratios. In this paper, we restrict our discussion to BAYMAX’s abilities with regard to our observations of SDSS J0914+0853.

1.2. SDSS J0914+0853

SDSS J0914+0853 was originally identified by Greene & Ho (2007) as one of ~200 low-mass SMBHs, based on “viral” black hole mass estimates where the velocity dispersion and radius of the broad line region were estimated from H0 emission line characteristics. The system is at $z = 0.14$ ($D_L = 661$ Mpc and $D_A = 509$ Mpc for a ΛCDM universe, where $H_0 = 69.6$, $\Omega_M = 0.286$, and $\Omega_\Lambda = 0.714$) and is a low-mass ($M_{BH} = 10^{5.5} M_d$), low-luminosity AGN. SDSS J0914+0853 was observed by Chandra as part of a Cycle 13 program targeting low-mass AGNs (Proposal ID:13858, PI: Gültekin). These data were taken to investigate the fundamental plane in the low-mass regime, and thus are on-axis (Gültekin et al. 2014). Analyzing the 15 ks Chandra data with EDSER, the archival Chandra exposure shows a possible secondary source 0.3" away from the primary. Possible contamination from an ultraluminous X-ray source (ULX) is very low; following the methodology in Foord et al. (2017a) we calculate the number of expected ULXs with $L > 10^{35}$ erg s$^{-1}$ to be $<10^{-3}$ within a radius of 0.3" from the center of the galaxy. If the emission is found to most likely originate from two point sources, it will be the lowest-mass dual AGN discovered to date, and analysis of this system would pave the way toward a better understanding of the role of mergers and AGN activity in low-mass systems. In particular, dual AGNs in low-mass galaxies with low luminosities are the perfect test bed for discerning the best among the competing models for the connection between galaxy mergers and AGN activity. It has been argued that mergers can trigger high-luminosity AGNs but not low-luminosity AGNs, which are triggered by stochastic processes (Hopkins & Hernquist 2009; Treister et al. 2012). A competing hypothesis is that there is no correlation between AGN luminosity and mergers (e.g., Villforth et al. 2014). Because dual AGNs most likely arise from mergers, the presence of a low-luminosity dual AGN in SDSS J0914+0853 would show that low-luminosity AGNs can arise from mergers. However, effects due to pileup and artifacts from the point-spread function (PSF) cannot be ruled out at a high statistical confidence for the low-count (∼250 counts between 2 and 7 keV) image. At 10%, the pile-up fraction is relatively small, but combined with asymmetries in the Chandra PSF (Jud & Karovska 2010), it could produce a
spurious dual AGN signature. Thus, a statistical analysis is necessary before a discovery can be confirmed.

We aim to unambiguously determine the true nature of SDSS J0914+0853. As stated above, this cannot be accomplished using only existing Chandra data, because (i) the pileup introduces systematic uncertainties in the EDSER processing, (ii) the existing exposure is relatively shallow (15 ks), and (iii) potential PSF artifacts can produce spurious dual AGN signatures. To help determine the true nature of SDSS J0914+0853, we received a new observation (Proposal ID:19464, PI:Gültekin) that addresses all three of the above points. In particular, the observation: (i) uses the shortest possible frame time with a subarray, thereby eliminating pileup; (ii) goes 3 times deeper, with a 50 ks exposure; and (iii) uses a substantially different roll angle, so that any PSF artifacts will not appear in the same location on the sky. With a total of ∼723 counts between 2 and 7 keV (combining both data sets), BAYMAX is able to statistically analyze the likelihood that SDSS J0914+0853 is a dual AGN for separations >0.3 arcsec and flux ratios >0.1.

The remainder of the paper is organized into five sections. Section 2 introduces Bayesian inference, focusing on the specific components of the BF and how BAYMAX calculates the likelihood and prior densities. In Section 3, we analyze the Chandra observations of SDSS J0914+0853, including both a photometric and spectral analysis. In Section 4, we present our results when running BAYMAX on the Chandra observations of SDSS J0914+0853. In Section 5, we discuss the sensitivity and limitations of BAYMAX across parameter space and how they affect our results. Finally, we summarize our findings in Section 6. Please see Table 1 for a list of symbols used throughout this paper.

### 2. Methods

#### 2.1. Bayesian Inference

BAYMAX is capable of statistically and quantitatively determining, on the basis of a Bayesian framework, whether a given observation is better described by a model composed of one or two point sources. A Bayesian approach combines all available information (using prior distributions and likelihood models) to infer the unknown model parameters (posterior distributions). Bayes’ Theorem implies:

\[
\frac{P(M_2|D)}{P(M_1|D)} = \frac{P(D|M_2)}{P(D|M_1)} \times \frac{P(M_2)}{P(M_1)},
\]

where the posterior odds represent the ratio of the dual point source model (M2) versus the single point source model (M1) given the data D, the BF quantifies the evidence of the data for M2 versus M1, and the prior odds represent the prior probability ratio of M2 versus M1. Specifically, the BF is the ratio of the marginal likelihoods:

\[
BF = \frac{\int P(D|\theta_2, M_2)P(\theta_2|M_2)d\theta_2}{\int P(D|\theta_1, M_1)P(\theta_1|M_1)d\theta_1},
\]

which represent the ratio of the plausibility of observed data D, given two different models, and are parameterized by the parameter vectors \(\theta_2\) and \(\theta_1\). Values >1 or <1 signify whether M2 or M1 is more likely (see Jeffreys (1935) and Kass & Raftery (1995) for the historic interpretations of the strength of a BF value; we analyze our data to define a “strong” BF value in Section 5). In this paper, we assume that M2 and M1 are equally probable, such that \(P(M_2) = P(M_1) = 0.5\) and the BF directly represents the posterior odds. Thus, calculating the BF can be broken into two components: the likelihood density, \(P(D|\theta_j, M_j)\), and the prior density, \(P(\theta_j|M_j)\).

#### 2.2. Data Structure and Modeling the PSF

In this section, we will focus on the likelihood density implemented in BAYMAX. Each reprocessed Chandra level-2 event file tabulates the directional coordinates \((\mu, \nu)\) and energy \(E_i\) for each detected photon, where \(i\) indexes each detected photon (see Table 1 for a summary of notation). The detector itself records the pulse height amplitude (PHA) of each event, which is roughly proportional to the energy of the incoming photon. In the reprocessed files, the energy \(E_i\) is calculated from the event’s PHA value, using the appropriate gain table. Thus, BAYMAX takes calibrated events \((\mu, \nu, E_i)\) from reprocessed Chandra observations and compares them to new simulations based on single and dual point source models.

---

**Table 1**

| Symbol | Definition |
|--------|------------|
| \((x_i, y_i)\) | Sky coordinate of photon \(i\) |
| \(E_i\) | Energy of photon \(i\), in keV |
| \(\mu\) | Total flux (counts) of given source |
| \(k\) | Central position of given source in sky coordinates (2D: \(\mu = [\mu_x, \mu_y]\)) |
| \(\Delta x_k\) | Number of Chandra observations being modeled |
| \(\Delta y_k\) | Translational astrometric shift in \(x\) (\(K = [1, ..., k - 1]\)) |
| \(\Delta \phi_k\) | Translational astrometric shift in \(y\) |
| \(r\) | Radial astrometric shift \((\Delta x_k)^2 + (\Delta y_k)^2\) |
| \(f\) | Rotational astrometric shift |
| \(M_j\) | Flux ratio between secondary and primary source |
| \(\theta_j\) | Given model being analyzed by BAYMAX |
| \(\theta_j\) | Parameter vector for \(M_j\), i.e., \([\mu, r, \Delta x_k, \Delta y_k, \Delta \phi_k]\). |

**Note.** Column 1 lists the symbols, and Column 2 provides the corresponding definitions.
We characterize the properties of the Chandra PSF by simulating the PSF of the optics from the High Resolution Mirror Assembly (HRMA) via ray-tracing simulations. The two primary methods to simulate the HRMA PSF are SAOTrace\(^7\) and the Model of AXAF Response to X-rays (MARX) (Davis et al. 2012). While the MARX model uses a slightly simplified (and faster) description of the HRMA, differences between SAOTrace and MARX simulations are minimal for on-axis simulations. For our PSF analysis below, we find consistent parametric fits between a PSF generated by SAOTrace and one generated by MARX—in particular, the rms error between the two fits is on the order of \(\sim 0.1\%\).

Thus, our likelihood models for single and dual point sources are created by parametrically modeling the Chandra PSF using high-count simulations created by MARX-5.3.3. To translate the PSF model to an event file, the HRMA ray-tracing simulations are projected on to the detector plane via MARX. Ray-tracing simulations generated by both MARX and SAOTrace will have roughly the correct total intensity but small deviations in the overall shape. Specifically regarding MARX—the PSF wings are broader than observations, while the PSF core is narrower than observed (Primini et al. 2011). These discrepancies can be reduced by blurring the PSF when projecting it to the detector plane via the AspectBlur parameter. This parameter is used to account for the uncertainty in the determination of the aspect solution (such as effects from pixel quantization and pixel randomization), as well as the uncertainty in the instrument and dither models within MARX.

The best value should be considered carefully for each unique observation.\(^8\) For MARX-generated simulations on ACIS-S, we expect the AspectBlur parameter to have values between 0\(^\circ\)25 and 0\(^\circ\)28. For our PSF analysis, we set AspectBlur to 0\(^\circ\)28. We note that the value used for AspectBlur does not represent the accuracy at which we can centroid.

For a given observation, a user-defined source model is input to MARX to generate X-ray photons incident from a single point source centered on the observed central position of the AGN (\(\mu_{\text{drag}}\) defined as the set of coordinates where the hard X-ray emission from the AGN is estimated to peak). Because we do not model the spectral parameters of the system (see Section 1), we are only interested in modeling the spatial distribution of a photon due to its energy \(E\); and our PSF does not depend on the spectral shape of our model. Each simulation uses the observation-specific detector position (RA_Nom, Dec_Nom, Roll_Nom) and start time (TSTART). We set the number of generated rays (NumRays) to \(1 \times 10^6\), and the readout strip is excluded by setting the parameter ACIS_Frame_Transfer_Time to 0.

We model the PSF as a summation of 2D Gaussians, where the amplitude and standard deviation of each Gaussian is energy-dependent. In general, the PSF may be any function that is unique to a given observation and can be quickly evaluated. For both the 15 and 50 ks observations, we fit a variety of possible functions to the PSF. Using the Bayesian Information Criterion (BIC) as a diagnostic for model comparison, we find that a summation of three circular, concentric 2D Gaussians yields the best fit for both Chandra observations (specifically, we find the PSF wings are modeled best by the broadness of a Gaussian component versus the addition of a Lorentzian component). We model the PSF for each observation individually, but we find that the best-fit parameters for each PSF model are consistent with one another within the \(1\sigma\) error bars (which is not surprising, given that both sources were observed on-axis for ACIS-S, albeit in different Chandra cycles).

Each photon is assumed to originate from a single or dual point source system. For example, for a single point source, the probability that a photon observed at location \(x, y\) on the sky with energy \(E\) is described by the PSF centered at \(\mu = (x, y, E)\), i.e., the energy-dependent PSF. For \(n\) total events, the total probability is the product of the probabilities for each detected photon, i.e., the likelihood density is:

\[
\mathcal{L} = P(x, y | \mu, E) = \prod_{i=1}^{n} P(x_i, y_i | \mu, E_i)
\]

\[
= \prod_{i=1}^{n} \frac{M_{i,j}(\theta_j)^{D_i}}{D_i!} \exp(-M_{i,j}(\theta_j)), \tag{3}
\]

where we use the Poisson likelihood—which is appropriate given that Chandra registers each event individually. Here, \(M_{i,j}(\theta_j)\) is the probability for event \(i\), given our PSF model, and \(D_i\) is the data value for event \(i\). For a dual point source, the total probability is \(P(x, y | \mu_P, \mu_S, E, n_S/n_P)\), where \(\mu_P\) and \(\mu_S\) represent the locations of the primary and secondary AGNs. The ratio of the fluxes (or total counts) between the secondary and primary is represented by \(n_S/n_P = f\), where \(0 \leq f \leq 1\). We note that our analysis on SDSS J0914+0853 does not include fitting for the spectral models. Using the archival data, we find consistent hardness ratios between the candidate primary and secondary AGNs, where we use circular and nonoverlapping apertures centered on their apparent locations. Thus, we assume that the spectra are the same spectral shape as that for the entire system, but with different normalizations. Future analyses with BAYMAX will include fitting for different spectral shapes.

Because BF represents the ratio of likelihood densities and we use the same data across both models, our calculations become simpler. We are left with:

\[
\ln \mathcal{L}' = \sum_{i=1}^{n} D_i \ln M_{i,j}(\theta_j) + \text{constant}, \tag{4}
\]

where \(M_{i,j}(\theta_j)\) is calculated for either a single \((j = 1)\) or dual \((j = 2)\) point source model via our parametrically fit PSF, for each detected event \(i\).

2.3. Prior Distributions

BAYMAX requires user input regarding two points: (i) the number of data sets and (ii) the prior distributions for each parameter. Regarding point (i), SDSS J0914+0853 has \(k = 2\) observations, and thus the parameter vector \(\theta_i = [\mu, \Delta x_1, \Delta y_1]\) while \(\theta_2 = [\mu_P, \mu_S, \log f, \Delta x_1, \Delta y_1]\). Here, \(\mu = (\mu_P, \mu_S)\) is the central sky \(x, y\) positions of the AGN; \(\Delta x_1\) and \(\Delta y_1\) account for the translational components of the relative astrometric registration for the \(k - 1\) observation; and \(\log f\) is the log of the flux ratio, where \(f = n_S/n_P\). The relative astrometric registration adds an uncertainty that must be taken into account in order to avoid spurious dual AGN signals that can be generated from slight mismatches between two or more observations. We take this into account by including the astrometric registration of multiple observations as a set of

---

\(^7\)http://cxc.harvard.edu/cal/Hrma/Raytrace/SAOTrace.html

\(^8\)http://cxc.harvard.edu/ciao/why/aspectblur.html
parameters to be marginalized over. For SDSS J0914+0853, we find that the rotational component of the relative astrometric registration is expected to be very small ($\Delta \phi_1 < 1^\circ$) and including the parameter does not affect our results. Thus, we only include $\Delta x_1$ and $\Delta y_1$, which are analyzed for the shallower observation (i.e., relative to the 50 ks exposure).

Regarding point (ii), BAYMAX can incorporate any user-defined function to describe the prior distributions for each parameter. For SDSS J0914+0853, the prior distributions of $\mu$ for both $M_1$ and $M_2$ are described by a continuous uniform distribution:

$$\mu = U(a, b),$$

(5)

where we constrain all $\mu$ values to be between $a = \mu_{\text{obs}} - 2$ and $b = \mu_{\text{obs}} + 2$. Thus, the 2D parameter space for possible $\mu_x$ and $\mu_y$ is a $4 \times 4$ sky-pixel box ($\approx 1.998 \times 1.998$) centered on the observed central X-ray coordinates of SDSS J0914+0853. Further, the prior distributions of $\Delta x_1$ and $\Delta y_1$ are also described by a uniform distribution with $a = \delta \mu_{\text{obs}} - 3$ and $b = \delta \mu_{\text{obs}} + 3$, where $\delta \mu$ represents the difference between the observed central X-ray coordinates of the two observations (in practice, $\delta \mu$ is expected to be small; however, because the most recent observation of SDSS J0914+0853 was taken in a subarray mode, the difference between the aimpoints of the two observations is $\approx 15$ sky pixels). For $M_2$, the prior distribution for $\log f$ is also described by a uniform distribution (and thus $f$ is described by a log uniform distribution), where $a = -2$ and $b = 0$. The range for the prior distribution of $\log f$ covers possible values expected for “major mergers” (with mass ratios $>1/3$), while accounting for a large range of possible Eddington fractions between the two black holes. In general, informative priors can be incorporated if prior information is available. For example, we might set the prior distributions of $\mu_P$ and $\mu_S$ to Gaussian distributions centered on coordinates that are better constrained by other observations (such as spectroastrometric [O III] observations or complementary IR photometry).

2.4. Calculation of the Bayes Factor

Bayesian inference can be divided into two categories: model selection and parameter estimation. In this section, we review how we address each component in BAYMAX.

Computing the marginal likelihood is challenging, as it involves a multidimensional integration over all of parameter space. Only over the last $\approx 20$ yr have advances in computational power allowed Bayesian inference to become a more common technique for model selection. In addition, general numerical methods based on Markov chain Monte Carlo (MCMC) (e.g., Metropolis et al. 1953) have been developed, allowing one to conduct Bayesian inferences in an efficient manner, with few constraints on dimensionality or analytical integrability.

To calculate the marginal likelihood, BAYMAX implements a sampling technique called nested sampling (Skilling 2004). In nested sampling, the marginal likelihood is rebranded as the “Bayesian evidence,” denoted by $Z$. Here, $Z = \int P(D|\theta, M)P(\theta|M)d\theta$. Nested sampling transforms the multidimensional integral into a one-dimensional integral by introducing the prior mass $X$, defined as $X = \int_{f_r(\theta) > \lambda} P(\theta|M)d\theta$. Here, the integral extends over the regions of parameter space contained within the isolikelihood contour $f_r(\theta) = \lambda$, and at any given time has a value $0 < X < 1$. For a step-by-step explanation of nested sampling, we refer the reader to Skilling (2004), Shaw et al. (2007), Feroz & Hobson (2008), and Feroz et al. 2009. A direct effect of the nested sampling methodology is sparse sampling in low-likelihood regions and dense sampling where the likelihood is high. To calculate the evidence, BAYMAX uses the PYTHON package nestle.10 The package provides a pure PYTHON implementation of nested sampling, where prior mass space can be sampled via different techniques. In particular, we use multi-ellipsoidal sampling by setting method = “multi” (see Mukherjee et al. 2006; Shaw et al. 2007; Feroz & Hobson 2008).

The two parameters that affect the accuracy of $Z$ are the number of “active points” and the stopping criterion $d\log Z$. The number of active points represent how many points in prior mass space one is sampling at a given time (roughly analogous to the number of walkers in an MCMC run). The stopping criterion determines when the nested sampling loop terminates (when the current largest sampled likelihood does not increase by more than the stopping criterion value, the sampling will end).11 For our analysis of SDSS J0914+0853, we use 500 active points (generally, a lower limit on the number of active points is $2N_{\text{dim}}$, and due to using the multi-ellipsoidal method, we add additional points to characterize each mode well) and set the stopping criterion $d\log Z = 0.1$. By continually increasing the number of active points and decreasing the stopping criterion, the estimated evidence and its accuracy should converge. We find no significant difference in results when increasing our active points above 500 and using $d\log Z < 0.1$, so we conclude that these values have properly sampled the likelihood space.

Finally, for parameter estimation, BAYMAX uses PyMC3 (Salvatier et al. 2016), which uses gradient-based MCMC methods for sampling. Specifically, we use PyMC3’s built-in Hamiltonian Monte Carlo (HMC) sampling method. HMC uses the gradient information from the likelihood to converge much more quickly than normal Metropolis–Hastings sampling. In general, HMC is more powerful for high dimensionality and complex posterior distributions (e.g., Betancourt et al. 2014).

In Figure 1, we show the results when we analyze two simulations with BAYMAX. The simulations have been reprocessed using EDSER and binned by two-thirds of the native pixel size. We simulate a single and a dual AGN system via MARX using the same telescope configuration as our new 50 ks observation. Both simulations have $n = 700$ photons between 2 and 7 keV, as well as the same 2–7 keV spectrum as SDSS J0914+0853. We do not include a background contribution in our simulations (see Section 3). For the dual AGN simulation, each AGN has the same spectra but with normalizations such that the flux ratio $f = n_2/n_1 = 0.8$ while the separation between the two AGN is 0$''$.4. As is evident, it is difficult to visually distinguish whether a given simulation is actually composed of one or two sources. Using the

9 The continuous uniform distribution $U(a, b)$ is a probability distribution where all values between the minimum $a$ and maximum $b$ are equally probable.

10 https://github.com/kbarbary/nestle

11 Specifically, at a given iteration $i$ where the current evidence is $Z_i$ and the estimated remaining evidence in the likelihood landscape is $Z_{\text{est}}$, if $\log(Z_i + Z_{\text{est}}) - \log(Z_i) < d\log Z$, the sampling will terminate
methodology presented above, BAYMAX favors the correct model for both simulations: for the single AGN simulation, BAYMAX estimates a BF of $17 \pm 1.6$ in favor of the single point source model, while for the dual AGN simulation, BAYMAX estimates a BF of $25 \pm 1.5$ in favor of the dual point source model (error bars have been determined by running BAYMAX multiple times on each simulation; see Section 4).

Although the hard X-ray emission appears quite similar between the two simulations, the joint posterior distributions are significantly different from one another. Specifically, for the dual AGN simulation, the joint posterior distribution is more tightly concentrated around the true values. BAYMAX is able to recover the true separation and flux ratio within the 68% credible interval. However, for the single AGN, the separation and flux ratio are consistent with zero at the 99.7% confidence level. We note that the shape of this particular joint distribution (namely, an “L” shape) is consistent with a single AGN, where at very large flux ratios, the dual AGN candidate is likely to have $r = 0$, and at very large separations, the dual AGN candidate is likely to have $\log f = -2$. More specifically, the dual point source model places each AGN at the same location and/or with arbitrarily low $f$, effectively consistent with one point source.

3. Data Analysis

3.1. X-Ray Data

SDSS J0914+0853 was originally targeted in order to study low-mass AGNs and their relation to the plane of black hole accretion. The quasar was placed on the back-illuminated S3 chip of the ACIS detector, with an exposure time of 15 ks (Obs ID: 13858). We received a new 50 ks exposure at a roll angle significantly different from the previous observation, and used the smallest subarray (1/8) on a single chip in order to get the shortest standard frame time (Obs ID: 19464). This was done for three reasons: (1) to place the PSF artifact in a different location, (2) to avoid pileup, and (3) to receive ~2–3 times more counts. We re-reduced and re-analyzed the archival data in order to ensure a uniform analysis between the two data sets.
We follow a data reduction similar to those described in previous X-ray studies analyzing AGN (e.g., Foord et al. 2017a, 2017b), using the Chandra Interactive Analysis of Observations (CIAO) v4.8 (Fruscione et al. 2006). Both data sets are analyzed with EDSER (Li et al. 2004), which can be included in the standard CIAO reprocessing command chandra_repro with the parameter pix_adj=EDSER. For each observation, we first evaluate the aspect solutions of the reprocessed level-2 event files in order to ensure the Kalman lock was stable at all times. Further, we inspect the event detector coordinates as a function of time and find that they followed the instrument’s dither pattern, indicating no aspect-correction–based degradation of the PSF.

We then correct for astrometry, crossmatching the Chandra-detected point-like sources with the Sloan Digital Sky Survey Data Release 9 (SDSS DR9) catalog. The Chandra sources used for crossmatching are detected by running wavdetect on the reprocessed level-2 event file. We require each matched pair to be less than 2″ from one another and have a minimum of three matches. The 15 ks observation meets the criterion for an astrometrical correction; we find eight matches between the Chandra observation and the SDSS DR9 catalog, resulting in a shift of less than 0.5″. The 50 ks observation was taken in a subarray, and thus does not meet these criteria. However, because BAYMAX takes into account astrometric shifts between observations, this step will not affect our final results; see Section 4. Background flaring is deemed negligible, as neither data set contain intervals where the background rate is 3σ above the mean level.

We then rerun wavdetect on filtered 0.5–7 keV data to generate a list of X-ray point sources. We use wavelets of scales 1, 1.5, and 2.0 pixels with a 1.5 keV exposure map, and set the detection threshold significance to 10^−5 (corresponding to one false detection over the entire S3 chip). We identify the quasar as an X-ray point source 0″4 (Obs ID: 13858) and 0″7 (Obs ID: 19464) from the SDSS-listed optical center (2″ corresponds to 95% of the encircled energy radius at 1.5 keV for ACIS). Counts are extracted from a 2″ radius circular region centered on the X-ray source center, where we use a source-free annulus with an inner radius of 20″ and outer radius of 30″ for the background extraction. We compare the estimated background contribution from the data sets to the Chandra blank-sky files. Here, the blank-sky files are properly scaled in exposure time and have matching WCS coordinates, dimensions, and energies. We find consistent results, where within 2–7 keV we expect ≤1 and 1.5 background counts within a 4 × 4 sky-pixel box (≈1″98 × 1″98) centered on the quasar for the archival and new data set, respectively.

Our reduced data are shown in Figure 2. Here, both exposures have been reprocessed using the EDSET algorithm and are binned by a tenth of the native pixel size. The archival data appear to have subpixel structure, indicating a possible secondary AGN ~0″3 away from the primary, but our new observation is inconsistent with this picture. Although the X-ray emission may be slightly extended in the east–west direction (north is up, while east is left), we find minimal photometric evidence supporting the presence of a secondary AGN.

3.2. Spectral Fitting

The quasar’s net count rate and flux value are determined using XSPEC, version 12.9.0 (Arnaud 1996). All errors evaluated in this section are at the 95% confidence level, unless otherwise stated, and error bars quoted are calculated with MCMC via the XSPEC tool chain. We implement the Cash statistic (cstat) (Cash 1979) in order to best assess the quality of our model fits.

Both spectra have an excess of flux at soft X-ray energies (<1 keV) with respect to the power-law continuum, while the 15 ks data appear to catch the source in a higher flux state in the soft X-ray band (see Figure 3). Both of these behaviors are observed (e.g., Lohfink et al. 2012, 2013) in AGNs with a “soft excess” component (see Miniutti et al. 2009) and Ludlam et al. (2015) for examples of soft X-ray excess in low-mass AGN candidates), an excess in emission above the extrapolated 2–10 keV flux that is detected in over 50% of Seyfert 1 galaxies (Halpern 1984; Turner & Pounds 1989; Piconcelli et al. 2005; Bianchi et al. 2009; Scott et al. 2012). The physical origin of the soft excess remains uncertain; the shape is suggestive of a low-temperature, high-optical-depth Comptonization of the inner accretion disk, but the temperature of this component appears to be constant over a wide range of black hole masses (and thus inferred accretion disk temperatures; see Gierliński & Done 2004 and Crummy et al. 2006). The two most popular explanations for the soft excess are blurred ionized reflection from the inner parts of the accretion disk (e.g., Fabian et al. 2002, 2005; Gierliński & Done 2004; Crummy et al. 2006) and Comptonization components (such as partial covering of the source by cold absorbing material; see Boller et al. 2002 and Tanaka et al. 2004).

Indeed, we find a statistically better fit when including an absorbed redshifted blackbody component to account for this soft excess (phabs * xphabs * (zpow + zbody)). We fix the Galactic hydrogen column density (the photoelectric absorption component phabs) to 4.21 × 10^20 cm^−2 (Kalberla et al. 2005) and the redshift to z = 0.14. In Figure 3, we show the X-ray spectra of both observations, along with the best-fit XSPEC models. For the archival data set, we find best-fit values for intrinsic N_H = 3.38±0.10 × 10^20 cm^−2, power-law component G = 2.01±0.07, and blackbody component kT = 0.10±0.11 keV. For our new data set, we find best-fit values for intrinsic N_H = 4.07±0.10 × 10^20 cm^−2, power-law component G = 2.51±0.11, and blackbody component kT = 0.11±0.04 keV.

However, because our analysis with BAYMAX is restricted to the 2–7 keV photons from SDSS J0914+0853, our results are not affected by the soft emission component in the spectrum. In particular, although we detect variability between the two observations in the low-energy band, the 2–10 keV fluxes are consistent with one another (at the 99.7% C.L.) when we independently fit each spectra between 2 and 7 keV with an absorbed redshifted power law. For the 15 ks observations, we calculate a total observed 2–10 keV flux of 3.20+0.23−0.19 × 10^−13 erg s^−1 cm^−2, while for the 50 ks observations, we calculate a total observed 2–10 keV flux of 2.23±0.19 × 10^−13 erg s^−1 cm^−2 s^−1. This corresponds to respective rest-frame 2–10 keV luminosities of 1.83±0.30 × 10^43 erg s^−1 and 1.25±0.23 × 10^43 erg s^−1 at z = 0.14 (assuming isotropic emission).

4. Results

Analyzing the 15 ks Chandra data with the EDSET option enabled, SDSS J0914+0853 appears to be an interesting dual
AGN candidate. When binned, the data show a possible secondary source 0\"3 away from the primary (see Figure 2). Although a potentially interesting result, classifying the source based on a qualitative analysis runs the risk of a false positive. A statistical analysis is necessary before discovery can be confirmed. With an abundance of photons and a robust model of the Chandra PSF, we now aim to unambiguously determine the true nature of SDSS J0914+0853. We must first individually analyze each observation using BAYMAX, and then combine the two (yielding a total of \( n = 723 \) counts between 2 and 7 keV).

We restrict our analysis to photons that (i) have energies between 2 and 7 keV and (ii) are contained within a 4 × 4 sky-pixel box (1\"98 × 1\"98) centered on the nominal X-ray coordinates of the AGN. This corresponds to \(~95\%\) of the encircled energy radius for the 2–7 keV photons. Because we anticipate respective background counts of \( \lesssim 1 \) and 1.5 within this region for the archival and new data sets, each photon is assumed to originate from either one (\( M_1 \)) or two (\( M_2 \)) point sources, with no background contamination. The asymmetric PSF feature is within this extraction region, and sits approximately 0\"7 from the center of the AGN (see Figure 2). Within 2–7 keV, there are 6 and 12 photons that spatially coincide with the feature for the 15 and 50 ks observations, respectively. We mask the feature in both exposures before running BAYMAX.

We run BAYMAX with the initial conditions for the parameter vectors \( \theta_1 \) and \( \theta_2 \) as stated in Section 2. When running BAYMAX on our 15 and 50 ks observations individually, \( k = 1 \), and thus we exclude the \( \Delta x_1 \) and \( \Delta y_1 \) from \( \theta_1 \) and \( \theta_2 \). Further, we run BAYMAX with the initializations for nestle as described in Section 2, with 500 active points and \( d\log Z = 0.1 \).

Our 15 ks observation has a total of \( n = 251 \) counts between 2 and 7 keV, while our 50 ks observation has a total of \( n = 472 \) counts between 2 and 7 keV. Using BAYMAX, we calculate BF values (defined as the ratio of the evidence for the dual point source model to the single point source model) of \( \frac{Z_2}{Z_1} = 0.154 \) and \( \frac{Z_2}{Z_1} = 0.102 \) for the 15 and 50 ks observations, respectively. This represents respective BF values of \( \approx 6.5 \) and \( \approx 9.8 \) in favor of a single point source. The relative magnitudes of the BF values are not surprising—because the 15 ks observation has fewer counts than the 50 ks observation, we expect there to be less evidence in favor of a given model. Indeed, using the definitions presented in Kass & Raftery (1995), both of these BF values are considered “positive” against the dual point source model. Further, the posterior distributions for \( \theta_2 \) are consistent across both data sets: the best-fit locations for \( \mu_p \) and \( \mu_s \) are consistent with one another at the 95\% confidence interval, and the joint posterior distributions have shapes consistent with a single point source (i.e., consistent with the “L” shape seen in Figure 1).

Given that the individual analyses on each data set favor the same model, and that we can treat the two spectra as the same between 2 and 7 keV, we increase our statistical power and combine both data sets. This yields a total of \( n = 723 \) counts...
between 2 and 7 keV. Although we analyze the two observations jointly, we emphasize that the likelihoods for each observation are calculated independently of one another and are a function of their respective PSF models. Here, \( k = 2 \), and \( \Delta x_1 \) and \( \Delta y_1 \) are included in parameter vectors for each model. We use BAYMAX to calculate a BF of \( \frac{7.40 \times 10^{-2}}{7} \) = 1.35 in favor of a single point source.

To test the impact of the MCMC nature of nested sampling, we run BAYMAX multiple times on the combined data sets. We find consistent results, with a spread in ln BF space well-described by a Gaussian distribution centered at ln BF = 2.6 with standard deviation of 0.2. This BF strongly indicates that the single point source model best describes the X-ray emission from SDSS J0914+0853. In Table 2, we list the best-fit values (defined as the median value of the posterior distributions) for parameter vector \( \theta_2 \).

We examine the posterior distributions for \( \theta_2 \) to better understand our results. In Figure 4, we show the combined 2–7 keV data set (≈65 ks, where the photons associated with the 15 ks exposure have been spatially shifted by the most likely \( \Delta x_1 \) and \( \Delta y_1 \)) with the best-fit \( x \) and \( y \) sky coordinates for the primary and secondary AGN (\( \mu_{px} \) and \( \mu_{py} \)), as well as the joint posterior distribution for the separation, \( r \), and log flux ratio, \( \log f \).

The joint posterior distribution has a shape consistent with a point source—the median values of the marginal posterior distributions are

\[
r = 0.15 \pm 0.5 \quad \text{and} \quad \log f = -1.6 \pm 0.4,
\]

the separation is consistent with zero at very large flux ratios (\( \log f \to 0 \)), and the flux ratio is consistent with zero at very large separations (\( r \to 2^\mu \)). The best-fit values for all the parameters in parameter vector \( \theta_2 \) are listed in Table 2.

We investigate the influence of our prior distributions on our results. In particular, the Bayesian evidence automatically implements Occam’s razor—the simpler model will be more easily favored than the more complicated one, unless the latter is significantly better at explaining the data. For our analysis, this means that the dual point source model needs enough data to overcome the inherent bias that BAYMAX has toward favoring the single point source model. Whenever the prior distribution is relatively broad compared with the likelihood function, the prior has fairly little influence on the posterior. Thus, we re-run BAYMAX with Gaussian prior distributions for \( \mu_x, \mu_y, \) and \( \mu_S \):

\[
\mu = \mathcal{N}(\mu, \sigma^2),
\]
where $\mu_m$ and $\sigma^2$ represent the mean and variance of the distribution. For $\mu_P$ and $\mu_S$, we set $\mu_m$ to the nominal X-ray positions of the potential primary and secondary AGN and set $\sigma$ to the observed separation between the two ($\sim0.8$), given the 15 ks archival observation. For $\mu_f$, we set $\mu_m$ to the nominal X-ray position of the AGN and similarly set $\sigma$ to 0.8. BAYMAX calculates a BF of 10.8 ± 1.2, consistent within the errors of our previous measurement. Further, the posterior distributions returned by BAYMAX are consistent with those listed in Table 2. We conclude that using sharper priors (comparable to the sharpness of our likelihoods) has no effect on our results.

5. Discussion

Our results support the hypothesis that the low-mass dual AGN candidate SDSS J0914+0853 is instead a single AGN. Individually, we find BF values of 6.5 and 9.8 in favor of a single point source for the 15 and 50 ks observations. When we combine the two data sets for a joint analysis, we find a BF $\sim 13.5$ in favor of a single point source, and the posterior distributions are consistent with this model. Further, the prior distributions do not appear to have a great influence on our posteriors, indicating that the data should be sufficient to favor the correct model even when accounting for the Bayesian bias. In this section, we discuss the significance of our results by analyzing BAYMAX’s capabilities across a range of parameter space for both the single and dual point source models. Assuming that SDSS J0914+0853 is indeed a dual AGN system, we investigate how the BF determined by BAYMAX depends on parameters $r$ and $f$. In particular, we aim to understand where in parameter space BAYMAX loses sensitivity for simulations with a number of counts comparable to our observations.

5.1. BAYMAX’s Sensitivity across Parameter Space

The first step is to investigate how well BAYMAX can classify a sample of simulated single AGN, i.e., our frequency of false positives. This measurement will allow us to better define a range of BFs that we can consider “strongly” in support of the dual point source model. We simulate 100 single AGN via MARX, assuming the same telescope configuration and spectrum as our new data set. Further, each simulation has 700 photons between 2 and 7 keV. We analyze each simulation with BAYMAX and find that only two are misclassified as a dual AGN with BF $> 3$ (with the largest BF = 3.5). Thus, we define a BF $> 3$ in favor of a dual AGN as “strong evidence,” while anything below this cut is classified as inconclusive.

We then run BAYMAX on a suite of simulated dual AGN systems, generated via MARX. The simulations were created with the same assumptions as listed above. Each simulation has 700 photons between 2 and 7 keV. Each simulated AGN has the same 2–7 keV spectrum as SDSS J0914+0853, but with normalizations proportional to their flux ratio. We simulated systems with separations that range between 0.8 and 0.2 and flux ratios that range between 0.1 and 1.0. For each $r$–$f$ point in parameter space, we evaluate 100 simulations with randomized position angles between the primary and secondary. Our results are shown in Figure 5, where we plot the logarithm of the mean BF for each point in parameter space. Consistent with expectations, BAYMAX favors the dual point source model more strongly as the separation and flux ratio of a given dual AGN simulation increases, where we can expect BF on the order of $\approx 10^4$ for systems with $r \approx 0.5$. We enforce a cut of BF $> 3$, where only BF above this value are classified as strongly in favor of the dual point source model. We find that we are sensitive to most flux ratios where $r \gtrsim 0.8$.

5.2. A Quasi-frequentist Approach

Our analysis is intended to be a fully Bayesian inference, but some readers may find a frequentist interpretation more intuitive. In the following section, we describe a potential interpretation of our results using a quasi-frequentist perspective.
Figure 5. Bayes factor (defined as $Z_f/Z_i$) for simulated dual AGN with varying separation ($r$, in arcseconds) and flux ratios ($f$). For each point in parameter space, we evaluated 100 simulations with randomized position angles (0°–360°) between the primary and secondary AGN. Here, we plot the logarithm of the mean BF for each point in parameter space. We enforce a cut of BF > 3, where above this value the Bayes factor is classified as strongly in favor of the dual point source model. Points in parameter space with a BF below this value are above this value the Bayes factor is classified as weakly in favor of the dual point source model. For a frequentist perspective, we add a contour that SDSS J0914+0853 is a dual AGN, we can reject the null hypothesis (with $p < 0.03$) at $f > 0.2$ for separations as low as 0.3″.

On average, for separations below 0.35″, BAYMAX will not necessarily favor the correct model for a dual AGN system. For SDSS J0914+0853, we estimate a possible separation of 0.3″, given the shallower Chandra observation. However, the strength of the BF in favor of a single AGN has its own significance. From a frequentist perspective, we may ask what is the probability of measuring a BF $\geq 13.5$ in favor of a single AGN if the system is dual AGN. In this specific scenario, our “null hypothesis” is that SDSS J0914+0853 is a dual AGN system and our $p$-value represents the probability of measuring a BF $\geq 13.5$ in favor of a single AGN. Using our suite of dual AGN simulations, we analyze the probability of measuring a BF $\geq 13.5$, as a function of $r$ and $f$. Across all of parameter space, we find $p \leq 0.05$ and thus reject the null hypothesis at a 95% confidence level. If we set our $p$-value threshold to $p < 0.03$, we find that only for the smallest flux ratios ($f < 0.2$) can the null hypothesis not be rejected for $r < 0.3$ (see Figure 1). Thus, there is a very low probability of SDSS J0914+0853 being a dual AGN system with a (1) flux ratio $f > 0.3$, (2) separation $r > 0.3″$, and (3) measured BF $= 13.5$ in favor of a single AGN.

We find that, for observations with 700 counts, BAYMAX is sensitive to a large region of $r$-$f$ parameter space, such that we would expect different results if SDSS J0914+0853 were a dual AGN. Our results and discussion highlight the importance of a robust, quantitative analysis of dual AGN candidates that are classified by their X-ray emission. Most candidate dual AGN are discovered via indirect detection methods, such as narrow-line optical spectroscopy or optical/IR photometry. However, directly detecting the X-ray emission unambiguously associated with a AGN is necessary for confirmation. For candidate AGNs with separations on the order of Chandra’s resolution (<1″), receiving observations with sufficient counts, paired with a robust model of the Chandra PSF, will allow for the most accurate analysis. In particular, we may expect that most dual AGN candidates should have separations <1″, as the physical-to-angular scale becomes 1.0 kpc arcsec$^{-1}$ at a distance of 200Mpc ($z \approx 0.05$). Given the small number of currently confirmed dual and binary AGN, tools such as BAYMAX will be important because they permit precise measurement of the dual AGN rate—and as a result, an improved physical understanding of the evolution of SMBHs and their activity.

6. Conclusions

In this work, we present the first analysis by BAYMAX, a tool that uses a Bayesian framework to statistically and quantitatively determine whether a given observation is described best by one or two point sources. BAYMAX takes calibrated events from a Chandra observation and compares them to simulations based on single and dual point source models. BAYMAX determines the most likely model by the calculation of the BF, which represents the ratio of the plausibility of the observed data $D$, given the model $M_i$ and parameterized by the priors. We present the results of a BAYMAX analysis of the lowest-mass dual AGN candidate SDSS J0914+0853, which was originally targeted as a dual AGN based on shallow archival Chandra imaging. The 15 ks exposure appears to have a secondary AGN ~0.3″ from a primary AGN. We received a new 50 ks Chandra exposure with a shorter frame time (to avoid pileup) and a different roll angle, with the aim of unambiguously determining the true accretion nature of the AGN. The main results and implications of this work can be summarized as follows:

1. Analyzing our new 50 ks observation, we find (by visual analysis) that the 2–7 keV emission more closely resembles that of a single point source. Both spectra have an excess of flux at soft X-ray energies (<1 keV) with respect to the power-law continuum, while the 15 ks observation appear to catch the source in a higher flux state in the soft X-ray band. Both of these behaviors are seen in AGN with a “soft excess” component, and we fit our spectra with an absorbed redshifted power law and blackbody (phabs x phabs x (zpow + zbbbody)). For the archival data set, we find best-fit values for intrinsic $N_H = 3.38^{+0.10}_{-0.10} \times 10^{20}$ cm$^{-2}$, power-law component $\Gamma = 2.01^{+0.11}_{-0.12}$, and blackbody component $kT = 0.10^{+0.03}_{-0.05}$ keV. For our new data set, we find best-fit values for intrinsic $N_H = 4.07^{+0.10}_{-0.10} \times 10^{20}$ cm$^{-2}$, power-law component $\Gamma = 2.51^{+0.10}_{-0.12}$, and blackbody component $kT = 0.11^{+0.05}_{-0.04}$ keV.

2. We find that the 2–7 keV emission is consistent between the two observations, and fit the spectra in this energy-range with an absorbed redshifted power law. For the 15 ks observations, we calculate a total observed 2–10 keV flux of $3.20^{+0.90}_{-0.80} \times 10^{-13}$ erg s$^{-1}$ cm$^{-2}$. For the 50 ks observations, we calculate a total observed 2–10 keV flux of $2.23^{+1.0}_{-0.99} \times 10^{-13}$ erg s$^{-1}$ cm$^{-2}$. This corresponds to a rest-frame 2–10 keV luminosity of $1.83^{+0.31}_{-0.46} \times 10^{43}$ erg s$^{-1}$ and $1.25^{+0.35}_{-0.21} \times 10^{43}$ erg s$^{-1}$ at $z = 0.14$ (assuming isotropic emission).

3. We use BAYMAX to analyze the 15 and 50 ks observations, both individually as well as combined, restricting our analysis to photons with energies between 2 and 7 keV. Using BAYMAX, we calculate respective BF in favor of the single point source model of $\approx 6.5$ and 9.8 for the 15 ks and 50 ks observations. When combining the two observations, we calculate a BF of 13.5 in favor of a
single point source. To test the impact of the MCMC nature of nested sampling, we run BAYMAX multiple times on the combined data sets. We find consistent results, with a spread in ln BF-space well-described by a Gaussian distribution centered at ln BF = 2.6 with standard deviation of 0.2.

4. Our posterior distributions for both the single and dual point source model further support that SDSS J0914 +0853 is a single AGN. Spatially, the best-fit locations from $\mu_\beta$ and $\mu_\delta$ are consistent with one another within the 68% error level. Further, the joint posterior distribution has a shape expected from a single point source—the median values of the marginal posterior distributions are $r = 0.15 \pm 0.5$ and $\log f = -1.6 \pm 0.4$, at very large flux ratios ($\log f \to 0$), the separation is consistent with 0, and the flux ratio is consistent with zero at very large separations ($r \to 2^\circ$).

5. We investigate the influence of our prior distributions by running BAYMAX with Gaussian prior distributions for $\mu$, $\mu_\beta$, and $\mu_\delta$. BAYMAX calculates a BF in favor of a single point source of 10.8 ± 1.2, consistent within the errors of our initial measurement. Further, the posterior distributions returned by BAYMAX are consistent with those listed in Table 2.

6. We investigate how the BF determined by BAYMAX depends on the separation and flux ratio of a given dual AGN system. For Chandra observations with at least 700 counts between 2 and 7 keV, we find that BAYMAX is capable of strongly favoring the correct model for most flux ratios when $r \geq 0.035$. For the smallest separations ($r \leq 0.03$), BAYMAX is capable of identifying the correct model when the flux ratio $f \geq 0.8$.

7. From a quasi-frequentist perspective, we estimate the probability of measuring a BF $\geq 13.5$ in favor a single AGN, using a null hypothesis that SDSS J0914+0853 is actually a dual AGN. Across all of parameter space ($0.03 < r < 0.5$ and $0.1 < f < 1.0$), we find $p \leq 0.05$ and can reject the null hypothesis at a 95% confidence level. Thus, there is a very low probability of SDSS J0914+0853 being a dual AGN system with a (1) flux ratio $f > 0.3$, (2) separation $r > 0.03$, and (3) measured BF = 13.5 in favor of a single AGN.

We have shown through various analyses that there is a dearth of evidence supporting SDSS J0914+0853 as a dual AGN system. Specifically, BAYMAX estimates a BF strongly in favor of a single AGN, and the posterior distributions for a possible separation and flux ratio between a primary and secondary AGN are consistent with zero. Moving forward, statistical analyses with BAYMAX on Chandra observations of dual AGN candidates will be important for a robust measurement of the dual AGN rate across our visible universe. Finally, our Bayesian framework will eventually be capable of more general analyses, such as evaluating binary active star systems.

A.F. and K.G. acknowledge support provided by the National Aeronautics and Space Administration through Chandra Award Number TM8-19007X, issued by the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of NASA, under contract NAS8-03060. We also acknowledge support provided by NASA through Chandra Award Number GO7-18087X, issued by the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of NASA, under contract NAS8-03060. A.F. thanks Abderahmen Zoghbi for helpful discussion regarding PyMC3. This research has made use of NASA’s Astrophysics Data System.

Software: CIAO (v4.8; Fruscione et al. 2006), XSPEC (v12.9.0; Arnaud 1996), nestle (https://github.com/kbarbary/nestle), PyMC3 (Salvatier et al. 2016), SAOTrace (http://cxc.harvard.edu/cal/Hrma/Raytrace/SAOTrace.html), MARX (v5.3.3; Davis et al. 2012).

ORCID iDs
Adi Foord https://orcid.org/0000-0002-1616-1701
Kayhan Gültekin https://orcid.org/0000-0002-1146-0198
Mark T. Reynolds https://orcid.org/0000-0003-1621-9392
Edmund Hodges-Kluck https://orcid.org/0000-0002-2397-206X
Edward M. Cackett https://orcid.org/0000-0002-8294-9281
Jessie C. Runnoe https://orcid.org/0000-0001-8557-2822

References
Arnaud, K. A. 1996, in ASP Conf. Ser. 101, Astronomical Data Analysis Software and Systems V, ed. V. Systems & J. Jacoby (San Francisco, CA: ASP), 17
Barrows, R. S., Comerford, J. M., Greene, J. E., & Pooley, D. 2017, ApJ, 838, 129
Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, Natur, 287, 307
Betancourt, M. J., Byrne, S., Livingstone, S., & Girolami, M. 2014, arXiv:1410.5110
Bianchi, S., Guainazzi, M., Matt, G., Fonseca Bonilla, N., & Ponti, G. 2009, A&A, 495, 421
Blecha, L., Loeb, A., & Narayan, R. 2013, MNRAS, 429, 2594
Boller, T., Fabian, A. C., Sunyaev, R., et al. 2002, MNRAS, 329, L1
Burke-Spolaor, S. 2011, MNRAS, 410, 2113
Capelo, P. R., Dotti, M., Volonteri, M., et al. 2017, MNRAS, 469, 4437
Capelo, P. R., Volonteri, M., Dotti, M., et al. 2015, MNRAS, 447, 2123
Cash, W. 1979, ApJ, 228, 939
Comerford, J. M., Gerke, B. F., Newman, J. A., et al. 2009, ApJ, 698, 956
Comerford, J. M., Gerke, B. F., Stern, D., et al. 2012, ApJ, 753, 42
Comerford, J. M., Pooley, D., Barrows, R. S., et al. 2015, ApJ, 806, 219
Comerford, J. M., Schluns, K., Greene, J. E., & Cool, R. J. 2013, ApJ, 777, 64
Crummy, J., Fabian, A. C., Gallo, L., & Ross, R. R. 2006, MNRAS, 365, 1061
Davis, J. E., Bautz, M. W., Dewey, D., et al. 2012, Proc. SPIE, 8443, 84431A
Ellison, S. L., Mendel, T. J., Scudder, J. M., Patton, D. R., & Palmer, M. J. D. 2013, MNRAS, 430, 3128
Fabbiano, G., Wang, J., Elvis, M., & Risaliti, G. 2011, Nat, 477, 431
Fabian, A. C., Ballantyne, D. R., Merloni, A., et al. 2002, MNRAS, 331, L35
Fabian, A. C., Miniutti, G., Iwasawa, K., & Ross, R. R. 2005, MNRAS, 361, 795
Ferroz, F., & Hobson, M. P. 2008, MNRAS, 384, 449
Feruz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Foord, A., Gallo, E., Hodges-Kluck, E., et al. 2017a, ApJ, 841, 51
Foord, A., Gültekin, K., Reynolds, M., et al. 2017b, ApJ, 851, 106
Fruscione, A., McDowell, J. C., Allen, G. E., et al. 2006, Proc. SPIE, 6270, 62701V
Fu, H., Myers, A. D., Djorgovski, S. G., et al. 2015, ApJ, 799, 72
Fu, H., Yan, L., Myers, A. D., et al. 2012, ApJ, 745, 67
Ge, J.-Q., Hu, C., Wang, J.-M., et al. 2012, ApJS, 201, 31
Gerke, B. F., Newman, J. A., Lotz, J., et al. 2007, ApJL, 660, L23
Gierliński, M., & Done, C. 2004, MNRAS, 349, L7
Goulding, A. D., Greene, J. E., Bezanzer, R. et al., 2018, PASI, 70, S37
Greene, J. E., & Ho, L. C. 2005, ApJ, 627, 721
Greene, J. E., & Ho, L. C. 2007, ApJ, 670, 92
Gültekin, K., Cackett, E. M., King, A. L., Miller, J. M., & Pinkney, J. 2014, ApJL, 788, L22
Halpern, J. P. 1984, ApJ, 281, 90
Hayward, C. C., Torrey, P., Springel, V., Hernquist, L., & Vogelsberger, M. 2014, MNRAS, 442, 1992
Hopkins, P. F., Bundy, K., Croton, D., et al. 2010, ApJ, 715, 202
Hopkins, P. F., & Hernquist, L. 2009, ApJ, 694, 599
