Deep learning model for multiwavelength emission from low-luminosity active galactic nuclei

Ivan Almeida, Roberta Duarte, and Rodrigo Nemmen

1Universidade de São Paulo, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Departamento de Astronomia, São Paulo, SP 05508-090, Brazil

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Most active supermassive black holes (SMBH) in present-day galaxies are underfed and consist of low-luminosity active galactic nuclei (LLAGN). They have multiwavelength broadband spectral energy distributions (SED) dominated by non-thermal processes which are quite different from those of the brighter, more distant quasars. Modelling the observed SEDs of LLAGNs is currently challenging, given the large computational expenses required. In this work, we used machine learning (ML) methods to generate model SEDs and fit sparse observations of LLAGNs. Our ML model consisted of a neural network and reproduced with excellent precision the radio-to-X-rays emission from a radiatively inefficient accretion flow around a SMBH and a relativistic jet, at a small fraction of the computational cost. The ML method performs the fit 10^5 times faster than previous semianalytic models. As a proof-of-concept, we used the ML model to reproduce the SEDs of the LLAGNs M87, NGC 315 and NGC 4261.

Key words: black hole physics – accretion discs – galaxies: active

1 INTRODUCTION

Quiescent galaxies with little or no ongoing star formation dominate the local universe (Bell et al. 2004; Bundy et al. 2006; Faber et al. 2007; Ilbert et al. 2010). The observed quiescence can be related to the activity phase of an active galactic nucleus (AGN). In neighbor galaxies with AGNs, we do not observe a relevant quasar population. Instead the observations show an active galaxy population mainly formed by low-luminosity AGNs (LLAGNs); some examples are SgrA* (Narayan et al. 1995; Yuan et al. 2003), NGC 3115 (Wong et al. 2014), and M87 (Event Horizon Telescope Collaboration et al. 2019).

LLAGNs present different observational signatures compared to more luminous AGNs – i.e., quasars, blazars. All AGNs display a broadband multiwavelength emission, but LLAGNs did not have prominent UV continuum emission – characteristic of an optically thick and geometrically thin accretion disc (Ho 1999, 2008; Nemmen et al. 2006; Wu et al. 2007). LLAGNs show weak and narrow Fe Kα (Terashima et al. 2002) lines in their spectra, which agrees with a thin accretion disc absence. Furthermore, the observed luminosity from these LLAGNs is far below the expectation for a thin disc with an efficiency of 10% (L = 0.1L_*; being L the luminosity and M the accretion rate). These characteristics suggest that LLAGNs are radiatively inefficient systems drastically different from the standard scenario.

The AGN dynamics depends on how the supermassive black hole (SMBH) accrete the available gas in its surroundings. These processes rely strongly on whether the viscously generated thermal energy is radiated away (Abramowicz & Fragile 2013). The LLAGNs in our nearby vicinity are underfed SMBHs with accretion rate M < 10^{-2}M_\text{Edd} (M_\text{Edd} is the Eddington accretion rate), these objects are classified as radiatively inefficient accretion flows (RIAF). RIAFs are extremely hot, optically thin, and geometrically thick accretion discs; for more details, see Yuan & Narayan (2014).

RIAFs have characteristic emission in two primary wavelengths: radio and X-ray. The radio component of the spectral energy density (SED) is related to synchrotron emission in the accretion disc. The elevated temperatures in RIAFs (Yuan & Narayan 2014) indicate an electronic population with very high average energy. These electrons can interact with available photons and transfer energy via inverse Compton scattering (IC). IC is one of the primary sources of X-ray emission. RIAFs can emit bremsstrahlung radiation also, via electron-ion scattering inside the accretion flow. However the radiation field in RIAFs is very faint, and it has negligible effects in the optically thin accretion flow dynamics.

An astrophysical object’s emitted light carries information about the physical processes occurring in the source. With the observed SED, one can model the emitting source. LLAGNs have a broadband emission, and modeling their multiwavelength emission gives us valuable knowledge about the system – e.g., the accretion rate or electronic and ionic distributions in the accretion flow. Several works (Yuan et al. 2003; Nemmen et al. 2006, 2014; Almeida

© 2021 The Authors
et al. 2018; Bandyopadhyay et al. 2019) performed similar analyses. Traditional methods to fit LLAGN SEDs (e.g. Nemmen et al. (2014)) are computationally intensive and require the user to actively monitor and participate in the modelling process.

In the last few years, machine learning (ML) methods have become extremely common in science, especially in dealing with big data problems, like Astronomy. The development of better telescopes gives us a massive amount of data, which is a favorable scenario to implement ML algorithms (George & Huerta 2018). Many authors published works using ML methods in astronomical data, doing object classification (Robde et al. 2005; Banerji et al. 2010), and spectra analysis and modeling (Firth et al. 2003; Valdés & Bonham-Carter 2006; Ball et al. 2007; Marulanda et al. 2020; Liew-Cain et al. 2020), for example.

Neural Network (NN) is a ML method based on biological neurons and how they pass through the information (McCulloch & Pitts 1943). It is widely used in the scientific community since they show results when it comes to classification (Wan 1990; Storrie-Lombardi et al. 1992; Odewahn 1995) and regression (Specht 1991; Comrie 1997). NNs consist of neurons and connections between them. With weights \( w_i \) associated with each connection. Each neuron has an activation function \( f(\sum w_i x_i) \) that is a relation between the inputs \( x_i \) and weights (Leshno et al. 1993). Deep learning (DL) is a very powerful tool, for more details see (Breen et al. 2020; Rodriguez et al. 2018; LeCun et al. 2015).

Other works used NNs to fit observational data (Asensio Ramos & Almeida 2009; Pacheco-Sanchez et al. 2019; Fathivavsari 2020). For instance, Asensio Ramos & Ramos Almeida (2009) shows a technique based on the combination of two ML methods: principal analysis component (PCA) and NNs. It uses a PCA for dimensionality reduction, and the NNs are used to interpolation. The PCA is useful to get faster results since it can shrink the size of the data. In our work, we do not perform a dimensionality reduction, but the speed-up is adequate. Fathivavsari (2020) proposed the use of NNs as a new technique to predict the flux and the shape of Ly\(\alpha\) emission lines in the spectra from quasars. The architecture proposed in Fathivavsari (2020) is a deep NNs with five layers with the input being \( S I_{LV}, C I_{LV}, \text{ and } C I_{II} \) emission lines, and the output is the Ly\(\alpha\) emission line. Similarly to our work, they used hyperparameter tuning methods to find the best architecture. They obtained a NN able to predict Ly\(\alpha\) emission lines with a precision 6% – 12%.

We aimed at calculate LLAGNs’ SEDs using deep learning methods. To model the observational data, we needed to generate a large number of SEDs to find the best fit via traditional methods –like the minimum square error. Creating a large amount of SEDs is challenging; the computational time involved is very high. We used DL method to speed-up our SED generation rate up to 100000 times, depending on the spectrum process. For this work, we generated a training sample of \( \sim 20000 \) SEDs for RIAFs and \( \sim 140000 \) for jets changing the main parameters of the system, and we trained a neural network (NN) with the dataset. The code’s name is AGINES (Active Galactic Nuclei Neural network SED generator).

In section 2, we presented the details of the model, details about the NN construction and training are in section 2.2. In section 3 we presented our results. We compared our results to other similar works in literature in section 4. The summary of the work is in section 5.

2 MODELLING
2.1 Original data
We assumed sub-Eddingston LLAGNs are accreting as radiatively inefficient accretion flow (RIAF). We used a semi-analytical approach to treat the RIAF emitted radiation (Nemmen et al. 2014). Normally LLAGN SEDs are radio-loud (Ho 1999; Sikora et al. 2007). In our model, we also included the synchrotron emission from a relativistic jet, either. Here in this work, we followed the same approach as Nemmen et al. (2014) to generate our SED sample.\(^1\) Henceforth we will call this model as the “fiducial” one.

2.1.1 RIAF
We considered the calculation of an optically thin and geometrically thick two-temperature accretion flow with outer radius \( r_0 = 10^4 R_S \) (\( R_S \) is the Schwarzschild radius), which presents very low radiative efficiency (Narayan et al. 1998). Our assumptions to calculate the system SED were:

(i) The accretion flow is stationary.
(ii) Viscosity is parameterized as Shakura & Sunyaev (1973) \( \alpha \)-viscosity.
(iii) The gravity is described with a pseudo-Newtonian potential (Paczynski & Wiita 1980).

The radiative transfer is treated carefully in more detail. We took into account radiation emission from synchrotron, inverse Compton scattering (IC), and bremsstrahlung processes occurring inside the accretion flow. (Nemmen et al. 2006; Yu et al. 2011; Nemmen et al. 2014).

Outflows, or winds, are an intrinsic feature of RIAFs’s model, e.g. Yuan et al. (2012, 2015); Almeida & Nemmen (2020). In this work, we considered the parameterization proposed by Blandford & Begelman (1999) for the accretion rate as a function of radius following a power-law relationship:

\[
M = M_0 \left( \frac{r}{r_0} \right)^s
\]

the parameter \( s \) is related to the “intensity” of winds, higher values of \( s \) mean stronger winds. \( M_0 \) is the accretion rate at the defined radius value \( r_0 \) in units of \( M_{\text{Edd}} \). In this work, we used \( r_0 \) as the outer radius of the accretion disc, \( r_0 = 10^4 R_S \). Following Nemmen et al. (2014) we assumed that \( s \) is limited to the range \( 0 \leq s \leq 1 \). For \( M_0 \), we considered a sub-Eddington system and limited the possible values to \( M_0 < 0.1 M_{\text{Edd}} \).

The RIAF solution depends on the system’s physical parameters: the black hole mass \( M \), the viscosity parameter \( \alpha \), the adiabatic index \( \gamma \), the parameter \( \beta \) — called modified plasma parameter — defined as the ratio between gas pressure and total pressure \( (\beta = P_{\text{gas}}/P_{\text{total}}) \), and the fraction of turbulence energy dissipated that heats the electronic population of the plasma \( \delta \). Here we followed Nemmen et al. (2006) and adopted \( \alpha = 0.3 \), \( \beta = 0.9 \) and \( \gamma = 1.5 \). For \( \delta \) values, normally they were considered small (0.01 in Narayan & Yi (1995)), but others processes could increase this value (Quataert & Gruzinov 1999; Sharma et al. 2007), like magnetic reconnection in hot plasmas for example, which can allow \( \delta \) to vary between 0.01 \( \leq \delta \leq 0.5 \).

\(^1\) Fiducial model source code can be found in: https://bitbucket.org/nemmen/adsaf-code/src/master/
In the RIAF model, we varied three free parameters: $\delta$, $s$ and $M_0$. We generated $\sim 20000$ RIAF SEDs randomly sampling the parameters in range: $0.01 \leq \delta \leq 0.3$; $0 \leq s \leq 1$; $-5 \leq \log_{10} \left( M_0 / M_{\text{Edd}} \right) \leq -1$.

2.1.2 Jet

The RIAF component does not produce enough radio emission to match the radio observations of LLAGN (Ulvestad & Ho 2001; Nemmen et al. 2006; Liu & Wu 2013). We included in our model the contribution of a jet component modeled based on the scenario of internal shocks (Spada et al. 2001). Following this model in the SMBH surroundings, a portion of the accretion flow gas is transferred to the jet generating a mass outflow rate $M_j$ and a shock wave. The shock wave makes the jet emission dominated by a non-thermal leptonic population.

Our modeled jet has a conical geometry with a half-opening angle of 0.1 rad ($\sim 6^\circ$) and a constant bulk Lorentz factor of 2.9, independent of the distance to the central SMBH. The jet is perpendicular to the accretion disc with an angle with the line of sight of 30$^\circ$. The shocks accelerate the electrons to a power-law with index $p$. The parameters $\epsilon_e$ and $\epsilon_B$ describe respectively the fraction of energy density from the electrons and the magnetic field.

The jet modelling free parameters were: $M_j$, $p$, $\epsilon_e$ and $\epsilon_B$. We allow $p$ to be in the range [2, 3], as the shock theory predicted it. We expect that $M_j < M_0$, here $M_0$ is from RIAF modeling, the mass outflow in the jet should not be higher than the mass inflow in the accretion disc. By definition, both $\epsilon < 1$, since they are fractions of the total energy. We generated $\sim 140000$ jet SEDs. We randomly varied the free parameters in the range: $2 \leq p \leq 3$; $-7 \leq \log_{10} \left( M_j / M_{\text{Edd}} \right) \leq -2$; $-5 \leq \log_{10} \epsilon_e \leq -1$; $-5 \leq \log_{10} \epsilon_B \leq -1$.

2.2 Neural network

Our model is a deep neural network composed of neurons with several layers and weights. In the output layer, we have a function called loss function $L$ which is any error function that gives us how much the prediction and target are different. The learning procedure is given by the backpropagation method that consists of derivatives of $L$ concerning the weight we want to learn (Kelley 1960). In other words, the backpropagation uses gradient descent (Ruder 2016) to find the minimum value in a space given by $L(w_{ij})$, i.e., find which weights $w_{ij}$ gives the minimum of $L$.

The learning part starts when we feed the neural network with data from the training set. We can give the data in batches, which is giving data by parts; this corresponds with faster training. Epoch is passing all batches once. The batch size and epoch are called hyperparameters. They are changed to find which combination of hyperparameters gives better results. In our work, we considered the number of neurons and the number of layers as hyperparameters. In this way, our goal was to find the best architecture for our model. We created a GridSearch to find the best values of our hyperparameters. The GridSearch is a hyperparameter tuning method (Bergstra & Bengio 2012) that consists of train the NN with several combinations of hyperparameters. In our case, our hyperparameter is the number of neurons in each layer. We created a shell script that runs all the combinations possible and saves the chosen metric. The trainings with the better metrics are the one we use to further analyze. Our best model is composed of two architectures, one is to analyze the RIAF and the other to analyze the jet region. The RIAF model is composed of 4 layers with 56, 60, 99, and 99 neurons, respectively. The jet model is composed of 5 layers with 10, 44, 66, 99, and 130 neurons. Each layer has a ReLU activation except for the last layer which is a linear function. The mean absolute error as $L$ and the optimizer is Adam (Kingma & Ba 2014).

To summarize, we build our neural network from scratch using the GridSearch technique. Our architecture is composed of hidden layers meaning that it is a deep neural network. We tested other architectures with different hyperparameters, but this is the best model.

2.3 AGNNES’s Performance

SED calculations for LLAGN are time expensive. The fiducial code which we generated our training RIAF SED’s sample took a minute per spectrum$^2$. The trained NN, AGNNES, calculates the same component in approximately 0.3 ms. It is $\sim 400000$ times faster than the original RIAF code. The time improvement allowed us to implement more robust statistic methods to LLAGN SED fitting, which is absent in previous works. The original jet calculation is much faster than the RIAF component, took around $\sim 0.2$ seconds each jet SED. AGNNES calculates the same component in approximately 0.4 ms. The improvement, for the jet, is close to $\sim 500$ times.

AGNNES results are an approximation from the original calculation. We had an extremely high speed-up in calculations, but the NN introduces inherently small errors when compared to the fiducial code. AGNNES, on average, can reproduce very well the validation data – see appendix C. Comparing the AGNNES predictions with the validation data sample, we defined $\Delta SED$ as the “distance” between the model and AGNNES prediction:

$$\Delta SED(\nu) = \log_{10}(\nu L_{\nu})_{\text{original}} - \log_{10}(\nu L_{\nu})_{\text{AGNNES}}$$ (2)

Considering the validation sample and comparing original data against AGNNES predicted one, for every point (i.e. frequency). We found the averaged value of $\Delta SED = 0.02 \pm 0.05$ (see 2), we plotted in figure 1. For the RIAF component the NN predicted slightly smaller emission values. The jet component presented $\Delta SED = 0.00 \pm 0.01$. These $\Delta SED$ values are AGNNES’s uncertainty.

The original data is not perfect. Sometimes the original code randomly fails while calculates the inverse Compton emission, returning a lower value for the emission. We automatically generated the data, and we identified the code failures and discarded the wrong data automatically, but it was not 100% efficient. The data set have few “bad” SEDs mixed in the whole set; this can affect the NN training and be responsible for the slightly systematic deviation of 0.02 dex in $\log_{10}(\nu L_{\nu})$. However, LLAGNs are very difficult to observe, and the observations typically have considerable uncertainties, higher than 0.02 dex. The theoretical models for the LLAGN emission have their uncertainties. The NN averaged deviation from the original model is negligible if compared with the usual error bars in measurements and model uncertainties.

2.4 Fitting Method

The current approach for LLAGN SED fitting in literature is an iterative method in which one changes all the parameters

$^2$ Depends on the hardware and parameters set. For our available computers, it took 0.5-2min
these values as a histogram. We obtained Figure 1. From the validation set, we calculated the difference between log \( \nu L_\nu \) bottleneck of LLAGN SED calculation is in the Comptonization. The SED calculation is computationally expensive. The fiducial code takes minto generate a single SED. It is tough to explore a vast range of parameters; there is insufficient time for this. For example, to run \( 10^5 \) SEDs, we would need approximately \( 5 \) days with no pause. With a faster method to calculate SEDs, we could explore a broader range of parameters and apply more robust fitting methods.

Time performance is the strongest point of AGNNES when we compare it with the iterative method. A well trained NN can do the IC step smoothly without high time spending. The NN is not a perfect copy of the original model. During the training, there is error propagation, and the NN final result is an approximation of the real value; for simplicity, we assumed the original code gives the perfect SED value, and we compared AGNNES with it. We discussed the NN performance in section 2.3.

We implemented the fitting procedure using Markov-chain Monte Carlo method. We performed the chain calculation with the Python package emcee (Foreman-Mackey et al. 2013). We used emcee to estimate the posterior distributions for the parameters – \( \delta, s, M_0, p, M_j, \epsilon_e, \epsilon_B \) – that better describe the data. We defined the likelihood function as a pure Gaussian likelihood. The final likelihood is the sum of all seven parameters’ likelihoods. The priors of parameters have initially been the limits of the data set presented in sections 2.1.1 and 2.1.2.

We started the MCMC with a flat distribution ball around random initial values with small radius, considering all parameters as adimensional. Our MCMC chain ran with 300 walkers for the number of steps \( N \): \( N > 30000 \). We neglected the first 20\% steps as burn-in. The value of \( N \) depends strongly on the used dataset. We estimated the MCMC convergence using the integrated autocorrelation time \( \tau \) implemented in emcee following Goodman & Weare (2010) method; it should be \( N/50 \geq \tau \).

3 RESULTS

We chose some objects with available SED data in the literature to fit. We aimed to find the best constraints for the accretion flow parameters that reproduce sources’ SED. Our chosen objects were: M87, NGC 4261, and NGC 315. The data points of all SEDs are available in appendix A.

To obtain the fit, we used existing independent measurements and theoretical models as priors. In our work, we considered all our priors as flat distributions, for simplicity. For the SEDs plots, we adopted the following convention: The red dashed line is the RIAF contribution, the blue dash-dotted line is the jet contribution, and the solid black line is the sum of them. The grey shaded area is a set of one hundred curves generated by the MCMC method and represents the uncertainties in the total sum. We showed the posterior distribution of the free parameters in appendix B.

We showed the fitting results in table 1. The columns show the seven free parameters for RIAF and jet models, and the last one is the reduced \( \chi^2 \) calculated between the best fit and the observational data points. For observational points without uncertainty, we assume an uncertainty of 0.05 dex.

In this work, we favored the scenario in which the RIAF dominates the high energy emission. We did not consider IC emission from jet photons. The jet contribution was considered relevant mainly for low frequencies – up to near-infrared – but neglected in higher frequencies for the galaxies NGC 4261 and NGC 315. In our jet fitting, we considered data points above infrared as upper limits for the jet emission. For M87 we modeled the whole emission favoring the jet model due to known observational priors (Doeleman et al. 2012; Kuo et al. 2014).

3.1 M87

The elliptical galaxy M87 harbors a supermassive black hole of \( 6.5 \times 10^9 M_\odot \) at 16 Mpc of distance from us (Event Horizon Telescope Collaboration et al. 2019). The source presents strong radio emission due to a prominent relativistic jet. We showed the SED data points and the best fit in figure 2.

For this source, we used as data points the high-resolution observations presented in Prieto et al. (2016) (c.f. Table 4). We treated observations with the lower spatial resolution – presented in the same paper for different wavelengths – as upper limits.

The jet emission is a known feature of M87 (Doeleman et al. 2012). We tested models considering RIAF and Jet emission, we found the best result considering only the jet emission. Our fit
3.2 NGC 4261 and NGC 315

NGC 4261 is an elliptical galaxy at ~ 30 Mpc of distance, located at the Virgo cluster. It harbors a SMBH with $5.2 \times 10^8 M_\odot$ (Tremaine et al. 2002). The SED data is available in table A2. de Menezes et al. (2020) presented data points and similar modeling for the same galaxy. We showed the SED data points and our best fit in figure 3.

NGC 315 is another elliptical galaxy at ~ 64 Mpc of distance, located at Pisces constellation. It contains a SMBH of $7.8 \times 10^8 M_\odot$ (Woo & Urry 2002). The SED data is available in table A3. de Menezes et al. (2020) presented data points and similar modeling for the same galaxy. NGC 315’s SED data points and the best fit are in figure 4.

For these sources, we considered the optical data as upper limits. In the optical frequency range, there is a high possibility of stellar contamination. Stellar clusters typically inhabit the center of galaxies, and these populations emit a considerable amount of light in the optical band (Lauer et al. 2005). Since our primary interest here is the AGN emission, we neglected the potentially contaminated points by stars.

For NGC 4261 and NGC 315, our models were more focused on the RIAF component, especially for higher frequencies. In our model, the innermost regions of the RIAF produce most of the X-ray emission. We used the jet component to explain mostly the radio region of the SED. We did not consider the jet emission to explain the bump around 100GHz and the most energetic data point is consistent with the best fit with $3\sigma$.

A thermal population of electrons in the accretion flow cannot reproduce the SED of M87 with such accretion rate. It is necessary to change the electronic distribution or assume the jet dominates the observed energy output entirely. The results for M87 fitting parameters are in table 1, and we only considered the synchrotron jet in our model.

| Object | $\delta$ | $M_0$ ($M_\odot$) | $s$ | $M_f$ ($M_\odot$) | $p$ | $\epsilon_e$ | $\epsilon_0$ | $\chi^2_{red}$ |
|--------|---------|------------------|----|------------------|----|-------------|-------------|--------------|
| M87    | –       | –                | –  | –                | 6.6$^{+0.7}_{-0.5}$ x 10$^{-7}$ | 2.54$^{+0.04}_{-0.07}$ | 4.7$^{+0.4}_{-0.3}$ x 10$^{-2}$ | 1.2$^{+0.3}_{-0.2}$ x 10$^{-5}$ | 10.0 |
| NGC 4261 | 0.3$^{+0.2}_{-0.3}$ x 10$^{-4}$ | 9.9$^{+0.02}_{-0.05}$ x 10$^{-3}$ | 0.613$^{+0.003}_{-0.003}$ | 3.4$^{+0.0}_{-0.1}$ x 10$^{-5}$ | 2.4$^{+0.1}_{-0.3}$ | 2.4$^{+1.3}_{-3.9}$ x 10$^{-4}$ | 1.2$^{+4.3}_{-6.6}$ x 10$^{-4}$ | 57.8 |
| NGC 315 | 0.3$^{+0.1}_{-0.2}$ x 10$^{-4}$ | 9.9$^{+0.01}_{-0.03}$ x 10$^{-3}$ | 0.507$^{+0.003}_{-0.003}$ | 4.7$^{+0.0}_{-1.0}$ x 10$^{-3}$ | 2.6$^{+0.2}_{-0.1}$ | 3.7$^{+1.1}_{-1.1}$ x 10$^{-3}$ | 1.2$^{+0.3}_{-0.2}$ x 10$^{-5}$ | 6.7 |

Table 1. The final results of AGNNES’s fit to our galaxy sample.
For both NGC 4261 and NGC 315, assuming the RIAF dominates high energy emission – there are no strong constraints in the high-frequency jet emission ($\nu > 10^{10}$Hz) –, this assumption makes the value of $p$ poorly constrained. The jet model fits the radio points, and this gives us information about the accretion rate $M_{\text{j}}$ and the distribution of energy, represented by $\epsilon_e$ and $\epsilon_B$. In our model, we restrained $M_0 < 10^{-2}M_{\text{Edd}}$. We trained AGNNES to calculate SEDs with accretion rates up to $10^{-1}M_{\text{Edd}}$ but we expect even lower accretion rates for LLAGNs. Our assumption of the maximum value for $M_0$ resulted in the truncated distributions for $M_0$ in NGC 315 and NGC 4261, respectively figures B3 and B2.

If the sources have higher accretion rates, the RIAF assumption is not valid. This truncation can be a hint that the emission could not come from an inner hot accretion disc. The X-ray emission for NGC 4261 could be better constrained if we allow an accretion rate of $M_0 = 3.2 \times 10^{-2}M_{\text{Edd}}$.

For NGC 4261 and NGC 315, the fitting was not able to fully explain the UV emission. Both galaxies presented much lower observational value in UV than our RIAF model prediction.

## 4 Discussion

AGNNES’s accuracy is an essential feature of this work. Using the parameters set calculated by AGNNES as the best fit, we plotted figure 5. In the top panel of this figure, the solid black line is AGNNES SED fit, and the dashed green line is the SED calculated with the fiducial code – for the same parameters. For all sources, both curves are very similar. In the bottom panel, there are the residuals between the two curves. The grey zones represent the error of AGNNES for the uncertainty of 1$\sigma$ and 3$\sigma$ (see figure 1). Considering all frequencies, from radio to X-ray, the fiducial code and AGNNES are equivalent.

### 4.1 Comparison with previous models

Astronomers have studied M87 extensively over the past sixty years (Feltén 1968; Schreier et al. 1982; Biretta et al. 1999; Doeleman et al. 2012; Event Horizon Telescope Collaboration et al. 2019). Recently, Bandyopadhyay et al. (2019) modeled this source with a combination of RIAF and synchrotron jet. They used a population of nonthermal electrons, which is absent in our work, and they manage to fit the bump at ~100GHz combining RIAF and jet emission. If they considered only thermal electron RIAF, the disc emission is not strong enough compared to the jet emission.

For NGC 4261 and NGC 315, de Menezes et al. (2020) did the previous SED modeling for the same data set with our original code with iterative procedures. In this work, we reported an improved version of these fits with AGNNES. Our results suggested that $M_0 > 10^{-2}M_{\text{Edd}}$, which is not consistent with the RIAF scenario – notwithstanding our model fits well the data. de Menezes et al. (2020) Gamma-ray observations indicate the emission comes from a jet emitting synchrotron (low-frequency) and synchrotron self-Compton (high-frequency).

### 4.2 Model limitations

The original model used to calculate the SEDs has its limitations, and AGNNES inherited all of them. The RIAF SED followed a straightforward model, with some strong approximations. First of all, it is a 1-dimensional calculation. It was assumed an axisymmetric accretion flow, and we integrated over the $z$-component (considering a cylindrical geometry). Furthermore, the code works on stationary accretion flow, hence it cannot account time variability.

Our jet SED took account of synchrotron emission only. We did not calculate the Inverse Compton effect in the jet photons. Some observations support a scenario with synchrotron self-Compton emission in the jet dominating the LLAGN emission (Nagar et al. 2005; Finke et al. 2008; Takami 2011).

The SED data of every object was not simultaneously measured. Many authors observed each data point at a different time, with years of separation. Moreover, we did not take into account any variability. We assumed all data comes from the quiescent state of the source. Such assumptions can be a problem for galaxies with few data points and not known variability patterns. M87 has more reliable data because it is a source continuously observed.

AGNNES works properly inside its training parameters range (see section 2.1). For fainter sources with very low luminosity, like SgrA*, AGNNES can not fit the SED reliably. For the SgrA* case, the accretion is too small to be inside our training specifications. The fiducial code can calculate SEDs for objects with such low accretion rates. However, numerical errors increase considerably as the accretion rate decreases. To generate data for low accretion rates are harder than for mildly higher accretion rates. The difficulty in creating SEDs with $M_0 \geq 10^{-5}$ impacted on AGNNES results. Below this $M_0$ value, our NN can not calculate a correct SED. AGNNES best reproduces LLAGN SEDs with $(\nu L_{\nu})_{\text{peak}} > 10^{36}$ erg/s.

## 5 Summary

The primary objective of this work was to optimize the calculations of LLAGN SEDs, and the modelling of observed broadband spectra. Several previous works used the iterative method to calculate the radiative emission from RIAF and jet for LLAGNs. The fiducial code spent few minutes to generate one SED with two components. To implement a more robust statistical analysis, we need a faster approach to calculate a single SED. AGNNES performs this calculation much faster than the original code: it is ~400000x faster for the RIAF and ~500x faster for the jet. We achieved a substantial speed-up allowing us to fit some sources (M87, NGC 4261, NGC 315). We constrained the better distribution of the parameters (see section 2) to fit the SED for each object using the MCMC method. AGNNES is very accurate with respect to the fiducial method.

We built AGNNES for physical systems with fitting parameters inside its training range (see section 2.1). A possible future improvement is to generate more data for smaller BH masses and lower accretion rates to enhance the AGNNES training dataset. Our original code did not take into account a nonthermal leptonic population, and this can be an improvement in the future.

SED calculations are very time expensive if you need to take into account some iterative procedures as calculating Inverse Compton emission. Our work achieved a vast speed-up of some thousand times in the calculation of a single SED. This is a demonstration of the power of deep learning algorithms for astrophysical problems.

## Acknowledgements

We used Python (Oliphant 2007; Millman & Aivazis 2011) to organise all SED data and to make all figures. In this work we used
Deep Learning model for LLAGNs emission

Figure 5. Top panel: Comparison between SED calculated with AGNNES (solid black line) and following the original code (dashed green line). Bottom panel: The residuals between the original code SED and AGNNES’s prediction. The darker grey region is $1\sigma$ uncertainty on the AGNNES result; the lighter grey region is $3\sigma$ region.

several packages as pandas (McKinney), NumPy (Van Der Walt et al. 2011), SciPy (Virtanen et al. 2019) and Matplotlib (Hunter 2007).

Figure C1 was built using the software In-n-avg (LeNail 2019).

We acknowledge useful discussions with João Paulo Peçanha Navarro, Felipe Lucas Gewers, Raniere de Menezes, Amanda Rubio and Stephane V. Werner. This work was supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) under grants 2016/24857-6, 2017/01461-2 and 2019/10054-7. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Quadro P6000 GPU used for this research.

SUPPORTING INFORMATION

The LLAGN SED fitting code will be made publicly available on github upon publication of this manuscript.

REFERENCES

Abramowicz M. A., Fragile P. C., 2013, Living Reviews in Relativity, 16, 1

Agudo I., Thum C., Gómez J., Wiesemeyer H., 2014, A&A, 566, A59

Almeida I., Nemmen R., 2020, Monthly Notices of the Royal Astronomical Society, 492, 2553

Almeida I., Nemmen R., Wong K.-W., Wu Q., Irwin J. A., 2018, MNRAS, 475, 5398

Asensio Ramos A., Ramos Almeida C., 2009, ApJ

Asmus D., Höing S., Gandhi P., Smette A., Duschl W., 2014, MNRAS, 439, 1648

Ball N. M., Brunner R. J., Myers A. D., Strand N. E., Alberts S. L., Tcheng D., Llorà X., 2007, The Astrophysical Journal, 663, 774

Bandyopadhyay B., Xie F.-G., Nagar N. M., Schleicher D. R., Ramakrishnan V., Arévalo P., López E., Diaz Y., 2019, Monthly Notices of the Royal Astronomical Society, 490, 4606

Banerji M., et al., 2010, Monthly Notices of the Royal Astronomical Society, 406, 342

Bell E. F., et al., 2004, ApJ, 608, 752

Bergstra J., Bengio Y., 2012, J. Mach. Learn. Res., 13, 281–305

Biretta J. A., Sparks W. B., Macchetto F., 1999, ApJ, 520, 621

Blandford R. D., Begelman M. C., 1999, MNRAS, 303, L1

Breu P. G., Foley C. N., Boekholt T., Zwart S. P., 2020, Monthly Notices of the Royal Astronomical Society, 494, 2465–2470

Bundy K., et al., 2006, ApJ, 651, 120

Capetti A., Kleijn G. V., Chiaberge M., 2005, A&A, 439, 935

Comrie A. C., 1997, Journal of the Air & Waste Management Association
APPENDIX A: OBSERVATIONAL DATA

The observational data of the modelled galaxies are available in tables A1-A3. Data for M87 were extracted from Prieto et al. (2016) and data for NGC315 and NGC4261 were the same as presented by de Menezes et al. (2021).

REFERENCES

Doelmann S. S., et al., 2012, Science, 338, 355
Event Horizon Telescope Collaboration et al., 2019, ApJ, 875, L1
Faber S. M., et al., 2007, ApJ, 665, 265
Fathiavasiri H., 2020, Deep Learning Prediction of Quasars Broad Ly-alpha Emission Line (arXiv:2006.05324)
Feltzen J. E., 1968, The Astrophysical Journal, 151, 861
Ferrarese L., Ford H. C., Jaffe W., 1996, ApJ, 470, 444
Finke J. D., Dermer C. D., Böttcher M., 2008, ApJ, 686, 181
Firth A. E., Lahav O., Somerville R. S., 2003, Monthly Notices of the Royal Astronomical Society, 339, 1195
Foreman-Mackey D., 2016, The Journal of Open Source Software, 24
Foreman-Mackey D., Hogg D. W., Lang D., Goodwin J., 2013, PASP, 125, 806
George D., Huerta E., 2018, Physics Letters B, 778, 64–70
Giovannini O., Masegosa J., Márquez I., Guerrero M. A., Dultzin-Hacyan D., 2006, A&A, 460, 45
Goodman J., Weare J., 2010, Communications in computational science, 5, 65
Gu Q.-S., Huang J.-S., Wilson G., Fazio G., 2007, ApJ, 671, L105
Ho L. C., 1999, ApJ, 516, 672
Ho L. C., 2002, ARA&A, 46, 475
Hunter J. D., 2007, Computing in science & engineering, 9, 90
Jones D. L., 1997, ApJ, 484, 186
Junor W., Biretta J. A., 1995, The Astronomical Journal, 109, 500
Kelley H. J., 1960, Ars Journal, 30, 947
Kingma D., Ba J., 2014, International Conference on Learning Representations
Kuo C. Y., et al., 2014, ApJ, 783, L33
Lauer T. R., et al., 2005, The Astronomical Journal, 129, 2138
Liu H., Wu Q., 2013, ApJ, 764, 17
Lonsdale C. J., Doelmann S. S., Phillips R. B., 1998, The Astronomical Journal, 116, 8
MarulandaJ. P., SantaC., Romano A. E., 2020, Deepend Learning Gravitational Wave Detection in the Frequency Domain (arXiv:2004.01050)
McCulloch W. S., Pitts W., 1943, The bulletin of mathematical biophysics, 5, 115
McKinney W., 2012, Computing in Science & Engineering, 13, 9
Morabito D., Niell A., Prestwich R., Mo这种e K., 2007, The Astronomical Journal, 91, 1038
Nagar N. M., Falcke H., Wilson A. S., 2005, A&A, 435, 521
Narayan R., Yi I., 1995, ApJ, 452, 710
Narayan R., Yi I., 1995, A&A, 356, 711
Narayan R., Mahadevan R., 1995, Nature, 374, 623
Narayan R., Mahadevan R., Grindlay J. E., Phopham R. G., Gammie C., 1998, ApJ, 492, 554
Nemmen R. S., Storchi-Bergmann T., Yuan F., Eracleous M., Terashima Y., Wilson A. S., 2006, ApJ, 643, 652
Nemmen R. S., Storchi-Bergmann T., Eracleous M., 2014, MNRAS, 438, 2804
Odevahn S., 1995, Publications of the Astronomical Society of the Pacific Oliphant T. E., 2007, Computing in Science & Engineering, 9, 10
Pacheco-Sanchez J., Alejo R., Cruz-Reyes H., Alvarez Ramírez F., 2019, Fuel
Paczynski B., Wiita P. J., 1980, A&A, 88, 23
Perlman E. S., Sparks W. B., Radomski J., Packham C., Fisher R. S., Piña R., Biretta J. A., 2001, The Astrophysical Journal Letters, 561, L51
Prieto M. A., Fernández-Ontiveros J. A., Markoff S., Espada D., González-Martín O., 2016, MNRAS, 457, 3801
Quataert E., Gruzinov A., 1999, ApJ, 520, 248
Rodríguez A. C., Kacprzak T., Lucey A., Amara A., Siguer R., Fluri J., Hofmann T., Refregier A., 2018, Computational Astrophysics and Cosmology, 5
Rohde D. J., Drinkwater M. J., Gallagher M. R., Downs T., Doyle M. T., 2005, Monthly Notices of the Royal Astronomical Society, 360, 69
Ruder S., 2016, arXiv preprint arXiv:1609.04747
Schreier E. J., Gorenstein P., Feigelson E. D., 1982, The Astrophysical Journal, 261, 42
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Sharma P., Quataert E., Hammett G. W., Stone J. M., 2007, ApJ, 667, 714
Sikora M., Stawarz L., Lasota J.-P., 2007, ApJ, 658, 815
Spada M., Ghisellini G., Lazzati D., Celotti A., 2001, MNRAS, 325, 1559
Specht D. F., 1991, IEEE transactions on neural networks
Storrie-Lombardi L., Mahadevan R., 2016, A&A, 804, 101
Takami H., 2011, MNRAS, 413, 1845
Terasima Y., Iyomoto N., Ho L. C., Ptak A. F., 2002, ApJS, 139, 1
Tremaine S., et al., 2002, ApJ, 574, 740
Ulvestad J. S., Ho L. C., 2001, ApJ, 562, L133
Valdés J. J., Bonham-Carter G., 2006, Neural Networks, 19, 196
Van Der Walt S., Colbert S. C., Varoquaux G., 2011, Computing in Science & Engineering, 13, 22
Verdoes Kleijn G. A., Baum S. A., de Zeeuw P. T., O'Dea C. P., 2002, AJ, 123, 1334
Virtanen P., et al., 2019, arXiv e-prints, p. arXiv:1907.10121
Wan E. A., 1990, Trans. Neur. Netw.
Wang J.-H., Urry C. M., 2002, ApJ, 579, 530
Woo J.-H., Urry C. M., 2002, ApJ, 579, 530
Wu Q., Yuan F., Cao X., 2007, ApJ, 669, 96
Yu Z., Yuan F., Ho L. C., 2011, ApJ, 726, 87
Yuan F., Narayan R., 2014, ARA&A, 52, 529
Yuan F., Quataert E., Narayan R., 2003, ApJ, 598, 301
Yuan F., Wu M., Bu D., 2012, ApJ, 761, 129
Yuan F., Gan Z., Narayan R., Sadowski A., Bai X.-N., 2015, ApJ, 804, 101
Zedzias A., Birkinshaw M., Worrall D., Peters A., Fabbiano G., 2005, ApJ, 627, 711
de Menezes R., Nemmen R., Finke J. D., Almeida I., Rani B., 2020, Monthly Notices of the Royal Astronomical Society
parameters in the following order: $\delta$, $M_0$, $s$, $M_J$, $p$, $\epsilon_E$, $\epsilon_B$ – the first three from the RIAF modeling, and the last four from the jet modeling.

**APPENDIX C: NEURAL NETWORK DIAGNOSTICS**

We built AGNNES with two different NNs: one for the RIAF component, another for the jet. Figures C2-C3 show how close is AGNNES prediction to the original calculation. C2 is the RIAF component and C3 is the jet component. In these plots, we are using some SEDs from the validation sample.

For C4 and C5 we plotted the original value of SED against AGNNES’s prediction for several frequencies, respectively the RIAF and jet components. The perfect result should be a straight line $y = x$, our results were very close to this. The further the point is from the line $x = y$, the higher is the error in the NN prediction. Some of the simulations are not physically correct, so these are underrepresented since the model learns mostly from physically accurate simulations. Another possibility to improve the results of the $x = y$ is to feed more data to the model with every emission line shape equally represented. Finally, statistical learning is the base of NN models, so they learn based on an error function and its derivatives. Even NNs with high accuracies have an error associated. NNs with accuracy equals to 1 are often overfitted and cannot generalize well.

This paper has been typeset from a TeX/LaTeX file prepared by the author.

**APPENDIX B: PARAMETER DISTRIBUTION**

In figures B1-B3 we present the posterior distributions of our free parameters from the SED fits in figures 2-4. The plot was done using the Python package corner (Foreman-Mackey 2016). The values presented in table 1 are the mean of the distribution and the uncertainties reported are correspondent to 1$\sigma$. We showed the

\begin{table}
\centering
\begin{tabular}{lll}
\hline $\nu$ (Hz) & $\nu L_\nu$ (erg s$^{-1}$) & Reference \\
\hline 1.60E+09 & 1.87E+038 & Giovannini et al. (1990) \\
8.40E+09 & 1.41E+039 & Morabito et al. (1986) \\
2.20E+10 & 2.48E+039 & Junor & Biretta (1995) \\
8.60E+10 & 4.43E+039 & Lee et al. (2008) \\
1.00E+11 & 1.61E+040 & Lonsdale et al. (1998) \\
2.30E+11 & 7.25E+040 & Doeleman et al. (2012) \\
2.60E+11 & 1.09E+041 & Whysong & Antonucci (2004) \\
2.80E+11 & 1.51E+041 & Perlman et al. (2001) \\
2.47E+12 & 1.28E+041 & Prieto et al. (2016) \\
3.32E+12 & 1.39E+041 & Prieto et al. (2016) \\
3.70E+12 & 7.55E+040 & Prieto et al. (2016) \\
4.99E+12 & 6.81E+040 & Prieto et al. (2016) \\
6.32E+12 & 8.40E+040 & Prieto et al. (2016) \\
8.93E+12 & 5.15E+040 & Prieto et al. (2016) \\
1.10E+13 & 4.53E+040 & Prieto et al. (2016) \\
1.27E+13 & 7.40E+040 & Prieto et al. (2016) \\
1.36E+13 & 4.73E+040 & Prieto et al. (2016) \\
2.06E+13 & 2.75E+040 & Prieto et al. (2016) \\
2.42E+13 & 1.95E+040 & Prieto et al. (2016) \\
2.42E+13 & 7.90E+040 & Prieto et al. (2016) \\
\hline
\end{tabular}
\caption{SED data for M87.}
\end{table}

\begin{table}
\centering
\begin{tabular}{lll}
\hline $\nu$ (Hz) & $\nu L_\nu$ (erg s$^{-1}$) & Reference \\
\hline 1.63E+09 & 2.40E+38 & Jones & Wehrle (1997) \\
5.00E+09 & 5.88E+38 & Nagar et al. (2005) \\
8.39E+09 & 1.24E+39 & Jones & Wehrle (1997) \\
1.50E+10 & 6.71E+39 & Nagar et al. (2005) \\
1.66E+13 & (5.39 \pm 0.24)E+41 & Asmus et al. (2014) \\
2.50E+13 & (4.79 \pm 1.10)E+41 & Asmus et al. (2014) \\
3.64E+14 & > 6.03E+39 & Ferrarese et al. (1996) \\
4.41E+14 & > 4.60E+39 & Ferrarese et al. (1996) \\
5.41E+14 & > 3.96E+39 & Ferrarese et al. (1996) \\
(4.8–24)E+17 & 1.03E+41 & Zezas et al. (2005) \\
\hline
\end{tabular}
\caption{SED data for NGC 4261.}
\end{table}

\begin{table}
\centering
\begin{tabular}{lll}
\hline $\nu$ (Hz) & $\nu L_\nu$ (erg s$^{-1}$) & Reference \\
\hline 1.40E+09 & 2.86E+39 & Capetti et al. (2005) \\
2.50E+09 & 5.28E+39 & Lazio et al. (2001) \\
5.00E+09 & 9.11E+39 & Nagar et al. (2005) \\
1.50E+10 & 6.71E+39 & Nagar et al. (2005) \\
8.62E+10 & (2.28 \pm 0.13)E+41 & Agudo et al. (2014) \\
2.29E+11 & (3.08 \pm 0.23)E+41 & Agudo et al. (2014) \\
3.75E+13 & (1.47 \pm 0.08)E+42 & Gu et al. (2007) \\
5.17E+13 & (7.72 \pm 0.59)E+41 & Gu et al. (2007) \\
6.67E+13 & (4.63 \pm 0.45)E+41 & Gu et al. (2007) \\
8.33E+13 & (2.95 \pm 0.35)E+41 & Gu et al. (2007) \\
3.68E+14 & 2.00E+41 & Verdoes Kleijn et al. (2002) \\
5.40E+14 & 9.48E+40 & Verdoes Kleijn et al. (2002) \\
(4.8–24)E+17 & 4.36E+41 & Gonzalez-Martin et al. (2006) \\
\hline
\end{tabular}
\caption{SED data for NGC 315.}
\end{table}

MNRAS 000, 1–9 (2021)
Figure B1. Corner plot for the fit parameters of M 87. In the main diagonal were plotted the posteriori distribution of the parameters. The other plots show the correlation between each pair of variables.
Figure B2. Corner plot for the fit parameters of NGC 4261. In the main diagonal were plotted the *posteriori* distribution of the parameters. The other plots show the correlation between each pair of variables.
Figure B3. Corner plot for the fit parameters of NGC 315. In the main diagonal were plotted the posteriori distribution of the parameters. The other plots show the correlation between each pair of variables.
Figure C1. The architecture (a) is the one used to predict the emission lines from RIAF mode and (b) to predict the emission line from jets.
Figure C2. Plots comparing AGNNE output with the expected value for the RIAF component.
Figure C3. Plots comparing AGNNeS output with the expected value for the jet component.
Figure C4. From the validation data set, we selected frequencies and plotted the SED value generated by the original code and the SED predicted by AGNNES for the RIAF component, respectively y-axis and x-axis. The perfect fit is the thin curve $y = x$ plotted in red.
Figure C5. From the validation data set, we selected frequencies and plotted the SED value generated by the original code and the SED predicted by AGNNES for the jet component, respectively y-axis and x-axis. The perfect fit is the thin curve $y = x$ plotted in red.