The Astrophysical Journal, 971:70 (13pp), 2024 August 10 https://doi.org/10.3847/1538-4357/ad5351

Modeling Blazar Broadband Emission with Convolutional Neural Networks. II. External Compton Model

N. Sahakyan[1], D. Bégué[2], A. Casotto[3], H. Dereli-Bégué[2], P. Giommi[4,5,6], S. Gasparyan[1], V. Vardanyan[1], M. Khachatryan[1], and A. Pe'er[2]
[1] ICRANet-Armenia, Marshall Baghramian Avenue 24a, Yerevan 0019, Armenia; narsahakyan@gmail.com
[2] Bar Ilan University, Ramat Gan, Israel
[3] Chief Scientist, Altair, 100 Mathilda Place, Suite 650 Sunnyvale, CA 94086, USA
[4] INAF, Osservatorio Astronomico di Brera, via Brera, 28, I-20121 Milano, Italy
[5] Center for Astrophysics and Space Science (CASS), New York University Abu Dhabi, PO Box 129188 Abu Dhabi, UAE
[6] Institute for Advanced Study, Technische Universität München, Lichtenbergstrasse 2a, D-85748 Garching bei München, Germany
Received 2024 February 12; revised 2024 May 22; accepted 2024 May 25; published 2024 August 6

Abstract

In the context of modeling spectral energy distributions (SEDs) for blazars, we extend the method that uses a convolutional neural network (CNN) to include external inverse Compton processes. The model assumes that relativistic electrons within the emitting region can interact with and up-scatter external photons originating from the accretion disk, the broad-line region, and the torus, to produce the observed high-energy emission. We trained the CNN on a numerical model that accounts for the injection of electrons, their self-consistent cooling, and pair creation-annihilation processes, considering both internal and all external photon fields. Despite the larger number of parameters compared to the synchrotron self-Compton model and the greater diversity in spectral shapes, the CNN enables an accurate computation of the SED for a specified set of parameters. The performance of the CNN is demonstrated by fitting the SED of two flat-spectrum radio quasars, namely 3C 454.3 and CTA 102, and obtaining their parameter posterior distributions. For the first source, the available data in the low-energy band allowed us to constrain the minimum Lorentz factor of the electrons, $\gamma_{\rm min}$, while for the second source, due to the lack of these data, $\gamma_{\rm min} = 10^2$ was set. We used the obtained parameters to investigate the energetics of the system. The model developed here, along with one from Bégué et al., enables self-consistent, in-depth modeling of blazar broadband emissions within a leptonic scenario.

Unified Astronomy Thesaurus concepts: Blazars (164); Flat-spectrum radio quasars (2163); Active galaxies (17)

1. Introduction

Blazars are a subtype of active galactic nuclei (AGN) characterized by jets that are closely aligned with the observer's line of sight (Urry & Padovani 1995). Their emissions, which are Doppler enhanced and highly variable, span an extensive range of frequencies from the radio to the high-energy (HE; $E > 100$ MeV) and very high-energy (VHE; $E > 100$ GeV) γ-ray bands (Padovani et al. 2017). This wide spectral range makes blazars prime candidates for multiwavelength studies. Blazars are typically classified into two major categories based on the prominence of emission lines in their optical spectra: BL Lacertae objects (BL Lacs), which exhibit weak or absent emission lines, and Flat Spectrum Radio Quasars (FSRQs), known for their prominent optical emission lines. However, this classification extends beyond the mere presence or absence of emission lines: it also reflects differences in their physical properties, including the origin of their emissions.

The spectral energy distribution (SED) of blazars is typically characterized by a double-humped structure. The first hump, peaking in the infrared (IR) through X-ray bands, is attributed to synchrotron emission from relativistic electrons within the jet. The second hump, peaking in the γ-ray band, can be explained by two different mechanisms based on the particle at the origin of the emission. First, in the leptonic scenario considered here, the HE emission arises from inverse Compton scattering of either low-energy synchrotron photons (Synchrotron Self-Compton model (SSC); see, e.g., Ghisellini et al. 1985; Maraschi et al. 1992; Bloom & Marscher 1996), or external thermal photons, resulting in the so-called External Inverse Compton (EIC) model. These photons can originate from the accretion disk, the broad-line region (BLR), or the dusty torus (see, e.g., Dermer et al. 1992; Dermer & Schlickeiser 1994; Sikora et al. 1994; Błażejowski et al. 2000). SSC models are generally effective in explaining the broadband data from BL Lacs, while EIC models are usually favored for FSRQs. Indeed for the latter class, the peak luminosity of the HE hump strongly dominates over that of the low-energy hump. In the leptonic model considered in this paper, this is referred to as Compton dominance.

The second scenario is the hadronic model, where HE/VHE emission is either from direct proton synchrotron emission (Mücke & Protheroe 2001) or from emissions by secondaries produced in photo-pion and photo-pair interactions (Mannheim & Biermann 1989; Mannheim 1993; Mücke & Protheroe 2001; Mücke et al. 2003; Böttcher et al. 2013; Petropoulou & Mastichiadis 2015; Gasparyan et al. 2022). The potential detection of VHE neutrinos from the direction of blazars, such as TXS 0506+056 (IceCube Collaboration et al. 2018a, 2018b; Padovani et al. 2018) and PKS 0735+178 (Acharyya et al. 2023; Sahakyan et al. 2023), has increasingly brought attention to hadronic models of blazars, particularly lepto-hadronic models (Ansoldi et al. 2018; Keivani et al. 2018; Murase et al. 2018; Padovani et al. 2018; Sahakyan 2018, 2019; Cerruti et al.

2019; Gao et al. 2019; Righi et al. 2019; Gasparyan et al. 2022; Sahakyan et al. 2023).

The modeling of the broadband emission from blazars has always been of high interest, as it provides crucial insights into the physics of jets. This includes understanding their structure, the mechanisms of particle acceleration within them, and the nature of their emissions, all crucial for advancing our knowledge of relativistic jets. In recent years, progress in astrophysical detectors and continuous multiband monitoring of the sky has resulted in an accumulation of extensive multiwavelength data from blazars. This wealth of data not only enables the modeling of a single snapshot SED but also allows for systematic studies of blazar emission across different periods (see, e.g., Sahakyan 2021; Sahakyan & Giommi 2022; Sahakyan et al. 2022).

To interpret this wealth of data accurately and capture the many competitive mechanisms at work in blazar jets, numerical modeling is required. Consequently, various codes have been developed to model the lepto-hadronic processes in blazar jets, such as naima (Zabalza 2015), JetSeT (Tramacere et al. 2009, 2011; Tramacere 2020), agnpy (Nigro et al. 2022), AM3 (Gao et al. 2017), ATHEνA (Mastichiadis & Kirk 1995), Böttcher13 (Böttcher et al. 2013), LeHa-Paris (Cerruti et al. 2015), LeHaMoC (Stathopoulos et al. 2024), and Simulator of Processes in Relativistic AstroNomical Objects (SOPRANO, Gasparyan et al. 2022). However, all these codes encounter significant challenges in self-consistently fitting blazar SEDs. These codes either rely on an ad hoc distribution of emitting electrons or face prohibitively long computation times that do not permit thorough exploration of the parameter space of the models. In this era of data-rich astrophysics, an accurate and comprehensive modeling of blazar SEDs, which includes the detailed treatment of electron interaction and formation of the electron spectrum, is essential for a deeper understanding of these complex astronomical objects.

To address the significant computational challenges posed by traditional radiative models, we introduced an innovative approach for the self-consistent modeling of blazar SEDs (Bégué et al. 2024). We employed a Convolutional Neural Network (CNN) trained on a leptonic SSC model that accounts for processes such as electron cooling and pair creation-annihilation. Once trained, the network effectively reproduces the radiative signature of the SSC model in a significantly reduced time and enables efficient modeling of the broadband spectrum of BL Lacs, thereby allowing to explore the parameter space through fitting procedures. Shortly after, a similar study using a different network approach was presented by Tzavellas et al. (2024).

Bégué et al. (2024) is the first publication of the project titled "Modeling blazar broadband emission with convolutional neural networks." The primary goal of the project is to transform complex and computationally demanding lepto-hadronic radiative models into efficient neural network frameworks. This transformation is aimed at enabling statistical analysis and exploration of the parameter space governing blazar emissions. The simplicity of use and quick computation time also enable us to share the stand-in model directly with the blazar community through the Markarian Multiwavelength Data Center (MMDC).[7] The current paper seeks to build upon the work we initiated in Bégué et al. (2024) by developing a CNN dedicated to an EIC model. Any EIC model presents additional challenges due to its extra parameters, which are essential for characterizing the external photon fields. Our approach involves training the CNN on a data set of spectra that are numerically derived from an EIC model of blazars, using SOPRANO (Gasparyan et al. 2022). By considering a broad range of parameters for both the emitting electrons and the external photon fields, this new CNN can be effectively applied to study the emission from any FSRQ.

The paper is organized as follows. Section 2 describes the EIC model, including the considered sources of the external photons. The parameter ranges, simulation methodology, and the implementation of the CNN are detailed in Section 3. The application of the developed CNN to model the SEDs of 3C 454.3 and CTA 102 is presented in Section 4. Finally, the conclusions are given in Section 5.

2. The Model: External Inverse Compton

2.1. The Kinetic Model

In this paper, we further develop the analysis technique presented in Bégué et al. (2024) and consider the modeling of broadband SEDs of FSRQs. These SEDs are typically characterized by a pronounced Compton dominance, suggesting that, alongside synchrotron photons, an external photon field significantly contributes to the formation of the HE and VHE components. The source and nature of these external photons vary depending on the emission region distance from the central black hole. Photons originating directly from the black hole accretion disk, those reflected from the BLR, or those emitted from the dusty torus, can all act as potential targets for external Compton scattering by relativistic electrons accelerated in the jet. This external inverse Compton mechanism has been effectively used to model the observed multiwavelength data, as well as to explain various features observed in narrow spectral bands, such as the blue bump in the optical band, which is attributed to the accretion disk, etc. (see, e.g., Ghisellini et al. 2009).

In the model used in this paper, we assume that energy is dissipated in the relativistic jet at a distance $R_{\rm diss}$ from the central black hole and the emission is assumed to originate from a spherical blob with a comoving radius R, threaded by a uniform magnetic field. In general, R and $R_{\rm diss}$ should be considered as independent parameters. Considering the already large number of parameters in our model (see below), we further assumed the relation $R = R_{\rm diss}$. Alternative assumptions, e.g., relying on the jet opening angle, could also have been considered; see Section 5.

This blob moves with a Lorentz factor Γ. As the emission region is assumed to be oriented at a small angle relative to the observer's frame, the emission is amplified by the relativistic Doppler factor, denoted as δ, which we assume to be such that $\delta = \Gamma$. The electrons are injected into the emitting region as a power law with an exponential cutoff above the minimum Lorentz factor $\gamma_{\rm min}$, which is the same distribution function we used for the SSC model (Bégué et al. 2024). This injection function is given by:

$$Q_e = \begin{cases} Q_{e,0}\gamma^{-p}\exp\left(-\dfrac{\gamma}{\gamma_{\rm max}}\right) & \gamma \geqslant \gamma_{\rm min}, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

[7] See https://mmdc.am under theoretical modeling.

Here, p represents the electron index, and $\gamma_{\rm max}$ is the electron Lorentz factor beyond which electrons are not effectively accelerated. The normalization $Q_{e,0}$ is used to scale the luminosity of the electrons:

$$L_{\rm e} = \pi R^2 \delta^2 m_e c^3 \int_1^\infty \gamma Q_e d\gamma, \quad (2)$$

where m_e denotes the electron mass.

The injected electrons interact with both internal (synchrotron) and external photons, generating the HE component. To define the spectrum of the external photons, we adopt the approach outlined by Ghisellini & Tavecchio (2009). For clarity, the relevant definitions are provided below, but we refer to Ghisellini & Tavecchio (2009) for details. We consider three distinct external radiation fields, produced by (1) the accretion disk, (2) the BLR, and (3) the dusty torus. The contribution of each field scales with the distance of the emitting region from the central source. The accretion disk is modeled following the classical Shakura & Sunyaev (1973) thin disk framework, where the disk temperature scales as:

$$T^4(R_d) = \frac{3R_{\rm s}L_{\rm d}}{16\pi\eta\sigma_{\rm SB}R_{\rm d}^3}\left[1 - \left(\frac{3R_{\rm s}}{R_{\rm d}}\right)^{\frac{1}{2}}\right]. \quad (3)$$

In this equation, $R_{\rm s}$ denotes the Schwarzschild radius, η the accretion efficiency, $\sigma_{\rm SB}$ the Stefan–Boltzmann constant, and $L_{\rm d} = \eta\dot{M}$ represents the disk luminosity, with $\dot{M}$ being the black hole mass accretion rate and $R_{\rm d}$ is the cylindrical radius between the disk fluid element and the central black hole (not to be confused with the radius of dissipation $R_{\rm diss}$). The model includes several free parameters describing the external radiation field from the disk, namely:

1. The disk luminosity $L_{\rm d}$
2. The black hole mass M, inferred through the value of $R_{\rm s}$
3. The radiative efficiency η, set to 0.1 following Ghisellini & Tavecchio (2009)
4. The inner and outer disk radii, set to $R_{\rm in} = 3R_{\rm s}$ and $R_{\rm out} = 500R_{\rm s}$. The outer radius has minimal impact as it does not significantly alter the external disk energy density, while the inner radius influences the maximum temperature.

Thus, incorporating the accretion disk's contribution into the radiative model necessitates only two additional parameters, namely the disk luminosity $L_{\rm d}$ and the black hole mass M.

The BLR is modeled as a spherical shell located at a distance $R_{\rm BLR} = 10^{17}\,{\rm cm}(L_{\rm d}/10^{45}\,{\rm erg\ s^{-1}})^{1/2}$ (Ghisellini & Tavecchio 2009). This region reflects a fraction $f_{\rm BLR} = 0.1$ of the disk luminosity. The spectrum of the BLR is assumed to be a blackbody, peaking at a frequency of $\nu_{\rm BLR} = 2.47 \times 10^{15}$ Hz (Ghisellini & Tavecchio 2009). Similarly, the dusty torus is assumed to be a spherical shell located at a distance $R_{\rm DT} = 2.5 \times 10^{18}\,{\rm cm}(L_{\rm d}/10^{45}\,{\rm erg\ s^{-1}})^{1/2}$ that emits a blackbody spectrum peaking at $\nu_{\rm IR} = 3 \times 10^{13}$ Hz and reprocesses a fraction $f_{\rm DT} = 0.5$ of the disk luminosity. To leave additional freedom to the model, we elected to allow the temperatures of the BLR and of the dusty torus to vary within limited ranges around the proposed values. It is important to note that these temperatures are measured in the frame of the central supermassive black hole, and are Lorentz transformed (meaning boosted) to the frame comoving with the blob when used in the computation.

The external fields contribute to the Compton scattering and pair production kernels. To analogously treat the external and synchrotron photon fields, we assume that the total external field spectrum is isotropic in the comoving frame. This is a strong assumption since the external field is not isotropic in the laboratory frame and the relativistic motion of the blob results in angular beaming. However, since (1) electrons are assumed to be isotropic and ultrarelativistic ($\gamma_e \gg 10^2$) in the comoving frame and (2) the observer's line of sight is aligned with the edge of the jet, the effects of anisotropies of the external radiation fields are not substantial in the resulting spectrum. For a detailed discussion on the effects of anisotropy in external radiation fields, see, for example, Finke (2016).

The temporal evolution of the electron distribution is determined by considering the interaction of electrons with the magnetic field and all photon fields, namely of external (from the accretion disk, the dusty torus, and the BLR) and internal (all leptonic processes taking place in the jet) origins. We denote the distribution functions of electrons and photons as N_e and N_γ, respectively, and the distribution function of the external photon fields as $N_{\gamma,\rm ext}$, such that $N_{\gamma,\rm ext}$ is the sum of the boosted spectra from the three external sources considered here. The photon energy and the Lorentz factor of the electrons are represented by x and γ, respectively. The kinetic equation that describes the time evolution of the electrons is approximated by the Fokker–Planck diffusion equation, while the equation for the photons is an integro-differential equation. These are expressed as follows:

$$\begin{cases} \dfrac{\partial N_e}{\partial t}(\gamma) = Q_e + \dfrac{N_e}{t_{\rm esc}} + \dfrac{\partial}{\partial\gamma}\{N_e \times [C_{\rm IC}(N_\gamma + N_{\gamma,\rm ext}) + C_{\rm sync}]\} + Q_{\gamma\gamma\to e^+e^-}(N_\gamma + N_{\gamma,\rm ext}) \\ \dfrac{\partial N_\gamma}{\partial t}(x) = \dfrac{N_\gamma}{t_{\rm esc}} + Q_{\rm\ sync}(N_e) + R_{\rm IC}(N_e, N_{\gamma,\rm ext}) - S_{\gamma\gamma\to e^+e^-}(N_\gamma + N_{\gamma,\rm ext}). \end{cases} \quad (4)$$

Here, $t_{\rm esc} \equiv t_{\rm dyn} = R/c$ represents the escape time, such that the first term on the right-hand side of each equation represents the escape of particles from the radiation zone. $C_{\rm IC}$ and $C_{\rm sync}$ denote the inverse Compton and synchrotron cooling, respectively. Terms labeled Q and S are source and sink terms, and $R_{\rm IC}$ is the redistribution kernel of Compton scattering. Distribution functions in parenthesis represent the functional dependence. Further details on the kinetic equations and their numerical solutions, along with the expressions for all the rates that appear in this set of equations, can be found in Gasparyan et al. (2022).

The set of coupled kinetic equations given by Equation (4) is solved using SOPRANO (Gasparyan et al. 2022), which is

Table 1
Properties of the Data Set Used in Training the CNN

Parameters	Units	Symbol	Minimum	Maximum	Type of Distribution
Doppler boost	⋯	δ	3	50	Linear
Blob radius	cm	R	10^{15}	10^{18}	Logarithmic
Minimum electron Lorentz factor	⋯	$\gamma_{\rm min}$	$10^{1.5}$	10^{5}	Logarithmic
Maximum electron Lorentz factor	⋯	$\gamma_{\rm max}$	10^{2}	10^{6}	Logarithmic
Injection index	⋯	p	1.8	5	Linear
Electron luminosity	erg s^{-1}	$L_{\rm e}$	10^{42}	10^{48}	Logarithmic
Magnetic field	G	B	10^{-3}	$10^{2.5}$	Logarithmic
Black hole mass	$M_\odot$	$M_{\rm BH}$	10^{7}	10^{10}	Logarithmic
Disk luminosity	erg s^{-1}	$L_{\rm d}$	$10^{43.5}$	$10^{47.5}$	Logarithmic
BLR frequency	Hz	$\nu_{\rm BLR}$	$10^{14.5}$	10^{16}	Logarithmic
DT frequency	Hz	$\nu_{\rm DT}$	$10^{12.5}$	10^{14}	Logarithmic

Note. This includes the unit (where applicable) and symbol for each parameter, along with its range and the distribution across discrete parameter values. The data set contains a total of 10^6 spectra.

based on the finite volume method for the discretization in energy and uses a semi-implicit numerical scheme, in the sense that leptonic interactions are treated in a fully implicit framework, while the targets of hadronic interactions are treated explicitly. Since there is no hadronic component in the model used in this paper, the time integration is fully implicit. The time-dependent nature of SOPRANO permits in-depth modeling of the observed multimessenger SEDs; for application of SOPRANO, see, e.g., Gasparyan et al. (2022), Abe et al. (2023), and Sahakyan et al. (2023).

In the original SOPRANO implementation, external photon fields were not included. Their contributions are now incorporated as follows. First, we assume that the external radiation fields are isotropic in the comoving frame. This is a strong assumption that is required by the approach used in SOPRANO, which assumes isotropy of all plasma components. Second, the external radiation fields are assumed to be known at any computational time. Although these fields may evolve over time based on model parameters, they must be fully specified up-front; their temporal evolution is not governed by a kinetic equation. Finally, for any interaction involving a comoving photon field, the rates are computed by considering the sum of the locally produced photon field (which evolves according to the kinetic equation) and the prescribed external radiation fields.

For the need of the current model, we assume that the system can reach equilibrium, and its evolution becomes independent of time. In order to find the equilibrium solution to the kinetic equations, we evolve the solution until time $t = 4t_{\rm dyn}$, after which we find that the variations of the spectrum are too small to be of any interest. We assume that the blob does not expand and remains the same size during the period under consideration, and the adiabatic cooling is not taken into account. As a result, the outputs of SOPRANO are used to create the model database to be used as the training set.

3. Numerical Model: Computation, Validation, and CNN

To effectively train the CNN, it is essential to have a substantial number of SEDs for a wide range of parameters. Unlike the SSC model discussed in Bégué et al. (2024), the EC model exhibits significant spectral variation, so a larger data set of training SEDs is needed for the CNN to achieve adequate accuracy. To address this, 10^6 SEDs were generated using SOPRANO, to be compared to the 2×10^5 spectra for the SSC model of Bégué et al. (2024) and the 10^4 spectra of Tzavellas et al. (2024).

3.1. Parameter Ranges and Sampling

In the EC model considered in this study, there are 11 free parameters, namely the comoving blob radius R, the Doppler factor of the emission region δ, the comoving magnetic field strength B within the emission zone, the electron luminosity $L_{\rm e}$, the minimum Lorentz factor $\gamma_{\rm min}$, the cutoff Lorentz factor $\gamma_{\rm max}$, the power-law index p, the accretion disk luminosity $L_{\rm d}$, the mass of the central supermassive black hole $M_{\rm BH}$, and the temperature/frequency of the broad-line region and dusty torus. These parameters, along with their respective units and ranges, are given in Table 1. The parameters characterizing the emission region and the electron distribution functions are consistent with those used for the SSC modeling in Bégué et al. (2024). The Doppler boost factor linearly spans a range from 3 to 50, and the power-law index p is linearly sampled between 1.8 and 5. It is important to note, however, that steep values of $p > 3$ are generally not anticipated from the theory of shock acceleration or magnetic reconnection, as discussed in, e.g., Kirk et al. (2000), Sironi & Spitkovsky (2011), and Uzdensky (2022). Despite these theoretical considerations, we have chosen to explore a broader range of values of p that allow us to comprehensively investigate the parameter space without being constrained by its boundaries. In the fits, setting the index to a given value is straightforward, and priors can be used to reduce its range.

The remaining model parameters, i.e., the emission radius R, the minimum and maximum Lorentz factors $\gamma_{\rm min}$ and $\gamma_{\rm max}$, the electron luminosity $L_{\rm e}$, the strength of the comoving magnetic field B, the mass of the central black hole $M_{\rm BH}$, the disk luminosity $L_{\rm d}$, and the temperature of the broad-line region and of the dusty torus are all sampled logarithmically within their respective ranges. Specifically, these ranges are given by $15 < \log(R) < 18$, $1.5 < \log(\gamma_{\rm min}) < 5$, $2 < \log(\gamma_{\rm max}) < 8$, $42 < \log(L_{\rm e}) < 48$, $-3 < \log(B) < 2$, $7 < \log(M_{\rm BH}) < 10$, $43.5 < \log(L_{\rm d}) < 47.5$, $14.5 < \log(\nu_{\rm BLR}) < 16$, and $12.5 < \log(\nu_{\rm DT}) < 14$.

This large range of the parameters ensures that the CNN we developed will be applicable for the modeling of any FSRQ SED. In addition, the low achievable values for the disk luminosity guaranties that the trained model can also be used to model BL Lacs, leaving external fields unconstrained. In other

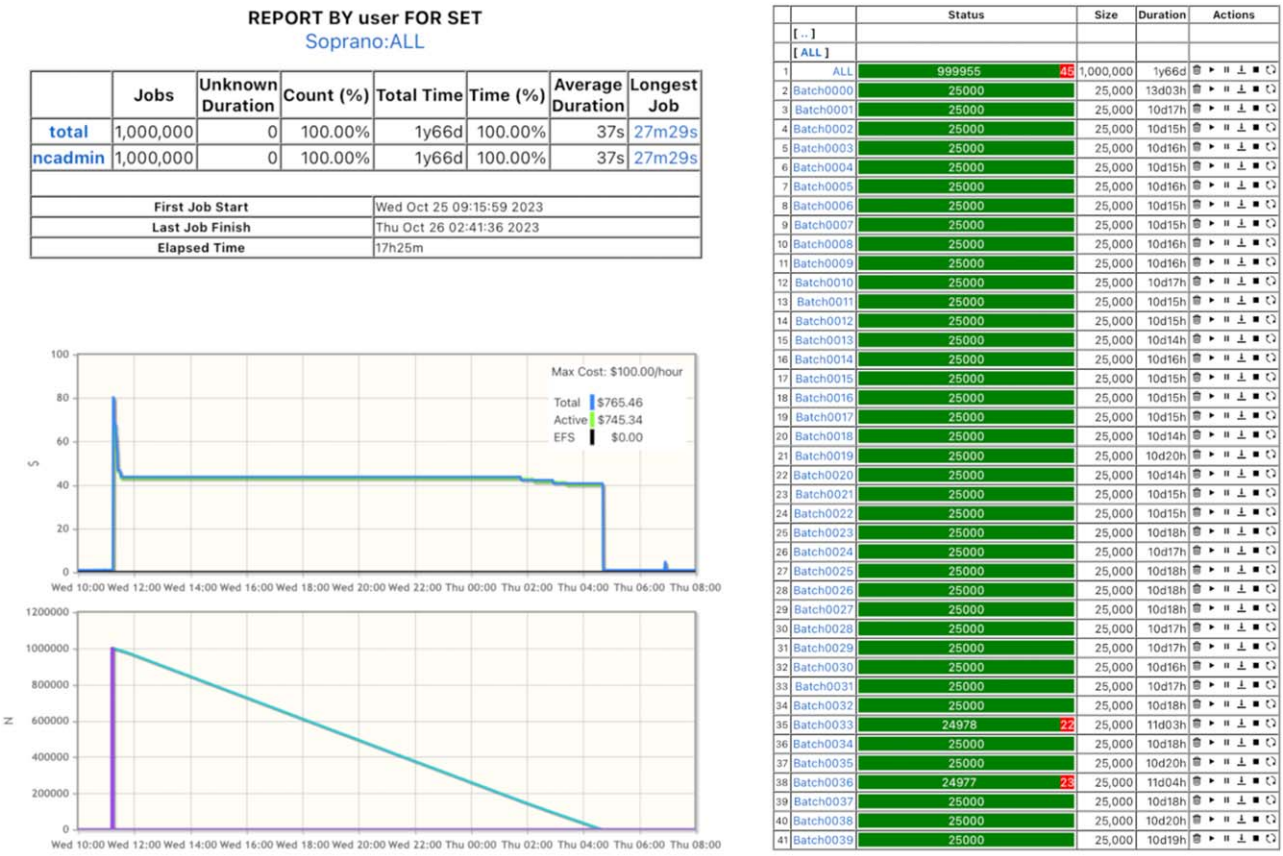

Figure 1. FNC data handling. Top left: a summary view of the workload, generated by FNC. FNC handled 1 million jobs corresponding to the computation of each individual spectrum for a total compute time of 1 yr and 66 days, performed in only 17 hr and a half. Bottom left: the top graph shows the total cost incurred for this computation as a function of time. The bottom graph shows the jobs queued waiting to be executed. Right: report showing the workload, for convenience broken down into 40 sets of 25,000 jobs each. Notice that 45 jobs have been labeled as failed after the first computation, due to the interruption of spot instances. FNC allows effortless rescheduling of those 45 failed jobs, and within a few minutes all jobs were done.

words, our approach allows for immediate model comparison between the SSC and EC model. We employ ronswanson (Burgess 2023) for generating the set of parameters to be simulated with SOPRANO. The parameters are distributed following the Latin hypercube sampling technique, which provides an efficient way to cover a multidimensional parameter space with uniform parameter distributions (see McKay et al. 2000; Viana 2016). This approach ensures a thorough, uniform, and representative sampling of the parameter space.

Finally, in the generation of the SEDs with SOPRANO, the accuracy threshold for the Newton–Raphson method was set to 10^{-10}, which is 5 orders of magnitude larger than the threshold used for the SSC model in Bégué et al. (2024). Provided the large number of spectra that needed to be computed, this adjustment was made to achieve a significant reduction of the compute time per spectra.

3.2. Computation of the Spectra on Amazon Elastic Computing Cloud

The core numerical workload consists of the generation of 10^6 spectra, where each spectrum is computed independently. This results in 1 million independent jobs, with each job executing a Python script. Every job requires precisely eight cores and runs on average for 37 s, with a maximum of over 27 minutes (see top-left panel of Figure 1). This duration differs from the run time of SOPRANO, which does not account for the loading of the cross section and for the initial setup. The peak RAM usage is slightly above 10 GB. Executing this workload on a single eight-core machine would take approximately 14 months, so it is imperative to run the workload on multiple machines. We note that we estimated the total compute time beforehand by computing two smaller databases with 500 and 5000 spectra, respectively. This step also allowed us to estimate the relevance of the parameter ranges and foresee any computational bottleneck in this large parameter space.

Although we have access to a small server with shared resources among researchers, managing a large number of independent jobs efficiently presents many challenges. It would require us to perform the computation of the spectra in large bunches via the addition of a management layer, which would also be able to keep track of computation status and failure. All of this at the price of monopolizing the compute resources for extended periods of time: even with this approach, we would still face several months of computation.

We have consequently transferred the workload to Amazon Elastic Computing Cloud (AWS EC2). For the workload manager, we used FNC, developed by Altair Labs; for more details, see A. Casotto (2024, in preparation).[8] FNC was chosen for the following reasons:

[8] Additional information can also be found on the website of Altair: https://altair.com/newsroom/articles/phoenix-rising-next-generation-job-scheduling-with-fenice.

1. Its capability to handle several million jobs, contrasting with common schedulers. FNC maximizes the resource usage while minimizing the cost of the numerical computation.
2. It keeps the state of every independent job even after its completion.
3. It includes a "Rapid Scaling" patented module that can grow and shrink the farm depending on the workload.
4. It has a browser-based interface that facilitates easy monitoring of the workload and real-time costs associated with the rental of Amazon servers.

As head node, we deployed a four-core, 32 GB machine of type t3.xlarge. For computing, AWS EC2 offers two types of instances:

1. On-demand instances, which remain available for as long as needed but which are substantially expensive;
2. Spot instances, which are offered at a discount relative to on-demand instances, with the stipulation that the instance could be revoked at any time by EC2 if it decides to give such resources to a customer willing to pay the full price.

We selected two instance families, namely c7a and r7i, based mainly on the spot cost at the time of the computation, which was about $\$0.06\,\mathrm{hr}^{-1}$ for an eight-core instance (such as c7a.2xlarge). Given that the run time of each job is short (on average 37 s), the computation was performed on spot instances, as the cost penalty for resuming an interrupted job is small.

Upon submission of the workload, the Rapid Scaling module grabs available spot instances from AWS EC2 from the preferred families. We set a limit of 600 concurrent jobs, for a total max size of the farm of $8 \times 600 = 4800$ cores. Toward the end of the computation, the Rapid Scaling module automatically returns an instance to the farm if this instance has been idle for more than 1 minute, thereby saving money by releasing unused compute instances.

In bottom-left panel of Figure 1, we show the time evolution of the total price rate (in $\$\,\mathrm{hr}^{-1}$) incurred for this computation, resulting in a total expense of about \$767. The peak cost of $\$80\,\mathrm{hr}^{-1}$ at the beginning of the computation was caused by the fact that Rapid Scaling initially estimated that 1 million queued jobs were in need of more hardware, so it tried to get as many instances as it could. Shortly after, that is, when 600 concurrent jobs were running, the workload was limited by the self-imposed policy on the number of concurrent jobs; idle instances were rapidly terminated to reduce the incurred cost.

The entire workload was completed in just over 17 hr. The FNC workload manager detected that 45 out of 10^6 jobs failed (see right panel of Figure 1), because of the sudden termination of a couple of spot instances. This is an inherent risk of using spot instances, which in our case is fully acceptable since the computation of individual spectra is short. Given that FNC maintains the state of the jobs even after completion, it was a matter of a couple of clicks in the browser-based interface to re-execute the 45 failed jobs.

3.3. Convolutional Neural Network

Once the computation is finished, the metadata are analyzed. The distribution of computation time per spectra and of the maximum error of the Newton–Raphson scheme are displayed in the upper panels of Figure 2. First, lowering the accuracy of the Newton–Raphson error from 10^{-15} to 10^{-10} resulted in a significant reduction in computation time for the spectra, averaging 25.3 s compared to 43.7 s for the SSC model, despite the increased complexity of the EC model. Additionally, this approach led to a lower fraction of failed spectra, $\sim$0.15%, defined as those not meeting the required accuracy at least once during computation. Based on these results, we conclude that the majority of the spectra computed with SOPRANO are sufficiently accurate to be utilized as a training sample for our CNN.

We employ the same CNN and optimization methodology as used in Bégué et al. (2024). Specifically, we incorporate numerical derivatives into the training output, to smooth out the spectra computed with the CNN. As highlighted in Bégué et al. (2024), this technique also improves the convergence rate. In order to obtain the most accurate CNN, the model was trained 16 times, with varying batch sizes, training fractions, and random distribution between training and validation sets. The optimal CNN was then selected based on its performances evaluated from its R2 score and mean squared error. The contribution of the external field to the observer is computed semi-analytically for each parameter set. Consequently, the total spectrum used in the fit consists of the sum of the spectrum computed by the CNN and the contribution from the external field.

In the lower panel of Figure 2, we present examples of spectra generated by SOPRANO alongside the corresponding results from the CNN for both the training set (on the two leftmost columns) and the validation set (on the two rightmost columns). The CNN demonstrates high performances with notable agreement across all spectra between the original SOPRANO outputs and the CNN results. This performance is further evidenced by the following metrics: the model achieved an averaged R^2 score of 0.87, a root mean squared error (RMSE) of 6.01×10^{-5}, a mean-absolute error (MAE) of 1.67×10^{-3}, and a criterion of 6.01×10^{-5}.

We also developed a CNN for the electron distribution function, aiming at displaying the uncertainty on the comoving distribution function of electrons. The main difference with the training of the photon spectrum is that the training is performed on $\gamma^2 N_e$ rather than on N_e. We find that this produces more accurate results, particularly at the transition between electron cooling states, namely between the slow and fast cooling regimes. While it might have been possible to further optimize the depth and the size of the network to produce acceptable results with N_e instead of $\gamma^2 N_e$, this would necessitate a large investment in work and in computation time. This is not essential for the current purpose: to depict the steady-state electron distribution function, for which a rather low accuracy is sufficient.

4. Application: Modeling Multiwavelength Emission of 3C 454 and CTA 102

The newly developed CNN presented in this study is used to model the multiwavelength SEDs of two FSRQs, namely 3C 454.3 and CTA 102. This approach follows the methodology of Bégué et al. (2024) in determining the parameters for the emitting electrons, specifically their power-law index p, luminosity L_e, $\gamma_{\rm min}$, and $\gamma_{\rm max}$, as well as the parameters of the emitting region, such as δ, B, and R. We adopt a Gaussian likelihood to sample the posterior distributions using the MultiNest algorithm (Feroz et al. 2009). For effective sampling and convergence, we utilize 1000 active points and set the

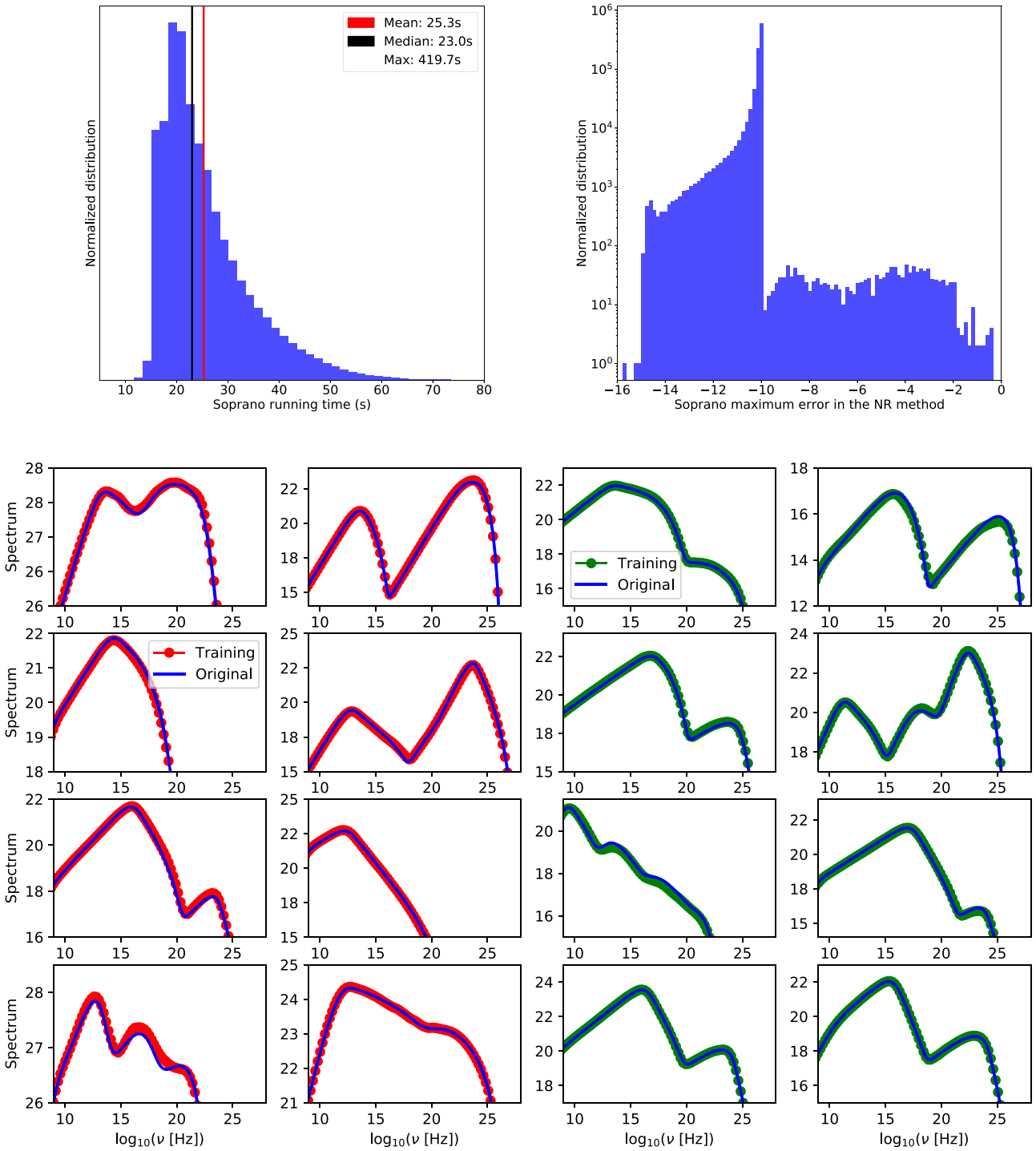

Figure 2. Upper left: distribution of the run time for each simulation, each using eight cores. Upper right: maximum error over the course of each simulation. Most simulations have errors below the targeted accuracy. Lower panel: comparison between the SED (ν vs. νF_ν) as calculated by the CNN—depicted with dots—against those generated by SOPRANO, represented by a solid line, prior to conversion into the observer's frame and unit normalization. The left panels illustrate spectra derived from the training data set, while the right panels display spectra from the validation data set. The diversity in spectra that the CNN can and does accurately reproduce is highlighted, along with the broad range of typical emitted powers predicted by the EC model.

tolerance at 0.5. We note that the extragalactic background light (EBL) absorption is taken into account via the model of Domínguez et al. (2011).

4.1. 3C 454.3

Blazar 3C 454.3, a classical FSRQ located at a redshift of $z = 0.859$, hosts a central black hole with an estimated mass of $1.5 \times 10^9\, M_\odot$ (Woo & Urry 2002; Liu et al. 2006). This source is known for its frequent flaring across various spectral bands, making it a frequent target for multiwavelength studies. The broadband SED of 3C 454.3 for the period MJD 55,519.59–55,520.19, retrieved from Sahakyan (2021), is depicted on the upper-left panel of Figure 3. This selected period corresponds to an active state of 3C 454.3, characterized by increased emission across all considered bands. In the SED

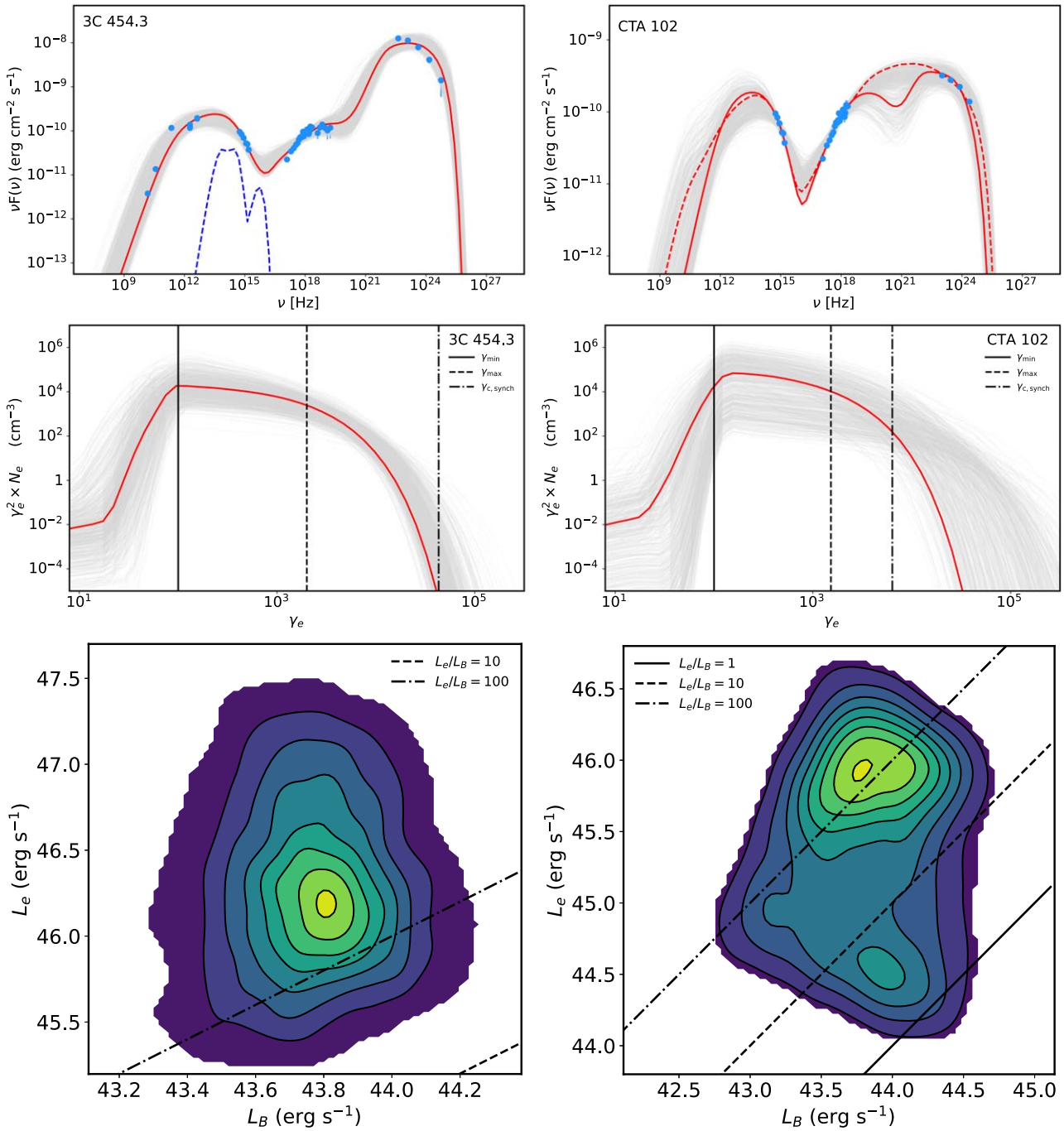

Figure 3. Upper panels: multiwavelength SEDs of 3C 454.3 (left) and CTA 102 (right), presented in blue. On each panel, the model corresponding to the maximum likelihood is shown by the solid red line, while the uncertainty associated with the model is depicted in gray. In the upper-left panel, the dashed blue line represents photons from the disk (low-energy peak) and the BLR. In the upper-right panel, the red dashed line represents the scenario for which the radius is larger and the emission is due to SSC. For both blazars, the models account for EBL absorption; in other words, the data are not EBL corrected. Middle panels: comoving electron distribution function at equilibrium for 3C 454 (left) and CTA 102 (right) corresponding to the SED shown above. The red line corresponds to the best-fit parameters, while the gray lines represent the uncertainty. The vertical black lines refer to the characteristic electron Lorentz factor $\gamma_{\rm min}$, $\gamma_{\rm max}$, and $\gamma_{\rm c,\ synch}$ for the best model parameters. Lower panels: comparison between the electron luminosity $L_{\rm e}$ and the magnetic luminosity $L_{\rm B}$ for 3C 454.3 (left) and CTA 102 (right) shown as the two-dimensional projections of the posterior probability distributions. The emission of both sources originates in a plasma far from equipartition with $L_{\rm B} \ll L_{\rm e}$. Only CTA 102 seems to have a marginal region such that $L_{\rm B} \lesssim L_{\rm e}$ in the bottom right. In fact, this region corresponds to the SSC scenario.

modeling, it is possible to vary all model parameters, but for our analysis we have fixed the black hole mass at $M_{\rm BH} = 1.5 \times 10^9\, M_\odot$, and the BLR and the dusty torus temperatures at $\nu_{\rm BLR} = 2.47 \times 10^{15}$ Hz and $\nu_{\rm IR} = 3 \times 10^{13}$ Hz, respectively. Allowing $\gamma_{\rm min}$ to vary freely over the entire range considered during model training (see Table 1) results in physically unrealistic values for other parameters, such as a very steep electron injection spectrum. Therefore, we constrained $\log(\gamma_{\rm min})$ to a narrow range between 1.5 and 2.0. During the fitting process, only the data above 10^{12} Hz were considered, as emission in the radio band is expected to be self-absorbed and should originate from a different (extended) region of the jet. The modeling results are presented on the left panels of Figure 3, where the best-fit model is indicated in red,

Table 2
Parameters Describing the SEDs in Figure 3

Parameters	3C 454.3	CTA 102
p	1.87 ± 0.18	1.97 ± 0.27
$\log_{10}(\gamma_{max})$	3.30 ± 0.16	3.18 ± 0.44
$\log_{10}(\gamma_{min})$	1.90 ± 0.10	2
δ	49.14 ± 2.53	40.74 ± 8.32
$\log_{10}(B/[\mathrm{G}])$	-0.70 ± 0.35	-0.06 ± 0.85
$\log_{10}(R/[\mathrm{cm}])$	16.13 ± 0.37	15.67 ± 0.93
$\log_{10}(L_e/[\mathrm{erg\ s^{-1}}])$	46.75 ± 0.44	45.95 ± 0.64
$\log_{10}(L_d/[\mathrm{erg\ s^{-1}}])$	47.11 ± 0.16	45.70 ± 0.96
$\log_{10}(L_B/[\mathrm{erg\ s^{-1}}])$	43.82	44.02

Note. The second column corresponds to the modeling of 3C 454.3, where all parameters of the emitting electrons, emission region, and L_d were considered free. In the modeling of CTA 102 (third column), $\gamma_{min} = 100$ is assumed. We set $\nu_{BLR} = 2.47 \times 10^{15}$ Hz and $\nu_{IR} = 3 \times 10^{13}$ Hz, and use $1.5 \times 10^9\, M_\odot$ and $8.5 \times 10^8\, M_{BH}$ for the black hole masses of 3C 454.3 and CTA 102, respectively.

while the associated uncertainty in the model is depicted in gray.

The parameters derived from the fit are summarized in Table 2, and the posterior distributions of these parameters are depicted in the Appendix, in the left panel of Figure 5. The parameters we obtained are in agreement with those typically estimated in FSRQ modeling. For instance, the power-law index of the emitting electrons is 1.87, the maximum energy of the injected electrons is 2.0×10^3, and the magnetic field strength is 0.20 G. The Doppler boosting factor, at 49.14, is relatively high and not well constrained, reaching the upper limit of our considered range. Sahakyan (2021) modeled the same period assuming an ad hoc electron distribution without considering cooling, and estimated a Doppler factor of 50.8. A high Doppler factor is necessary to account for the pronounced Compton dominance observed in this SED: the peak of the HE component surpasses the synchrotron/SSC components by nearly 2 orders of magnitude.

We use the second neural network trained on the electron distribution function to display the corresponding distribution on the middle-left panel of Figure 3. Clearly there is a large uncertainty on the electrons^ properties, in particular concerning their normalization. In addition, it is clear that electrons do not cool below γ_{min}, even though the external radiation field is quite intense. Since the maximum value of $\log \gamma_{min}$ is bounded to 2 for this fit, this could potentially inhibit electron cooling at such a small Lorentz factor, as it would require a strong magnetic field or a large density of the external photon field. In fact, the middle-left panel of Figure 3 shows the characteristic electron Lorentz factors γ_{min} and γ_{max}, alongside $\gamma_{c,synch} = 6\pi m_e c^2/(B^2 R \sigma_T)$, representing the electron Lorentz factor at which electrons would cool in a dynamical timescale by emitting synchrotron radiation. In this definition, m_e, c, and σ_T are the electron mass, the speed of light, and the Thompson cross section, respectively. It is clear that this cooling effect is subdominant, as expected due to the dominance of the external Compton component.

The fitting parameters also enable access to the energetic properties of the system. The disk luminosity, estimated based on the requirements to explain the flaring GeV data, is $L_d = 1.28 \times 10^{47}\,\mathrm{erg\ s^{-1}}$ (shown as the dashed blue line in Figure 3 upper-left panel), which is on the same order as the Eddington luminosity for a black hole mass of $M_{BH} = 1.5 \times 10^9\, M_\odot$. However, it should be noted that higher black hole masses for 3C 454.3 have also been estimated; for example, Gu et al. (2001) estimated a black hole mass of $M_{BH} = 4 \times 10^9\, M_\odot$. This luminosity is only slightly higher, by a factor 2 or 3, than the disk luminosity estimated based on the blue bump observed in the UV band (see, e.g., Bonnoli et al. 2011; Sahakyan 2021). The jet electron luminosity is $L_e = 5.61 \times 10^{46}\,\mathrm{erg\ s^{-1}}$, which is $\sim 8 \times 10^2$ times higher than the magnetic field luminosity, $L_B = \pi c R^2 \delta^2 B^2/8\pi = 6.57 \times 10^{43}\,\mathrm{erg\ s^{-1}}$. This shows that the system is far from equipartition. This is evident from the bottom-left panel of Figure 3, where we show the two-dimensional projections of the posterior probability distributions of L_e and L_b. The dashed and dotted–dashed lines represent $L_e/L_B = 10$ and $L_e/L_B = 100$, respectively. We note that the limits set on γ_{min} could influence the actual value of L_e determined by the fit and eventually change the equipartition ratio L_e/L_B, although we believe not substantially. We leave this detailed analysis for a future work.

4.2. CTA 102

CTA 102, at a redshift of $z = 1.037$ (Schmidt 1965), is another prominent FSRQ that hosts a black hole with an estimated mass of $8.5 \times 10^8\, M_{BH}$ (Zamaninasab et al. 2014). The source periodically enters a state of active emission. The SED of CTA 102 in the active state between MJD 56,196.7–56,202.3 was retrieved from Sahakyan et al. (2022). It is displayed alongside with the results of the modeling on the upper-right panel of Figure 3. Contrary to the previous model, we set the minimum Lorentz factor to $\gamma_{min} = 100$ since no low-energy data are available to help constrain its value. In addition, we set the temperatures of the external photon fields to $\nu_{BLR} = 2.47 \times 10^{15}$ Hz and $\nu_{IR} = 3 \times 10^{13}$ Hz. Moreover, we impose $M_{BH} = 8.5 \times 10^8\, M_\odot$. The results of the modeling are provided in Table 2, and the model with the best parameters, i.e., where the likelihood is maximum, is highlighted in red on the upper-right panel of Figure 3. From the red curve, it is clear that the combined contribution of all components effectively explains the observed data. The power-law index of injected electrons, $p = 1.97$, is mostly determined from the SSC contribution to the X-ray band, while the maximum electron Lorentz factor $\gamma_{max} = 1.52 \times 10^3$ is constrained by data in the optical/UV and GeV band. In this case, the estimated size of the emitting region, and consequently its distance from the central object, is 4.68×10^{15} cm. In addition, the disk luminosity is $6.25 \times 10^{45}\,\mathrm{erg\ s^{-1}}$, which is slightly lower than $L_d = 4.14 \times 10^{46}\,\mathrm{erg\ s^{-1}}$, inferred from the luminosity of the Lyα line (Pian et al. 2005). The comparison of jet luminosity in electrons and in the magnetic field is shown in the bottom-right panel of Figure 3, which demonstrates that, similar to 3C 454.3, the emitting region is far from equipartition.

As the jet propagates through the external photon fields characterized by a high comoving energy density, the injected electrons are interacting with these photons, producing the HE component. We display the uncertainty on the electron distribution function at equilibrium in the middle-right panel of Figure 3, alongside the results from the best-fit model in red. From this figure, it is seen that electrons are in the slow cooling regime for our choice of γ_{min}, which could not be constrained

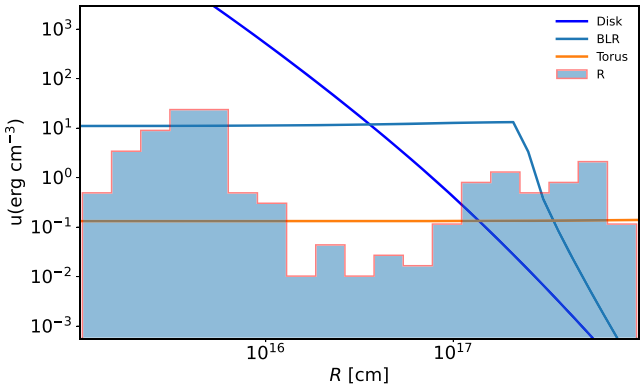

Figure 4. Variation of the external photon field comoving energy density with the radius of the emission region. Contributions from the disk, the BLR, and the torus are represented in blue, orange, and green, respectively. The posterior distribution of the radius R is shown as the light blue histogram, displaying the bimodality of the posterior. The emission can be produced either at small radius where the disk internal energy density dominates, or at large radius such that all external fields components are small and the emission is purely from an SSC origin.

independently because of the lack of data between the optical band and the radio (not associated to the jet emission).

The posterior distributions of the model parameters shown in the Appendix in the right panel of Figure 5 reveal a bimodal distribution for the radius R, at $R \sim 4 \times 10^{15}$ cm and $R \sim 5 \times 10^{17}$ cm. The location of the emission region, and consequently the main field contributing to the formation of the HE emission in FSRQs, remains an open question. In our framework, the fit drives the selection of the relevant photon field to explain the provided data by finding suitable parameter sets. This bimodality suggests the existence of two sets of conditions compatible with modeling the observed SED of CTA 102. Specifically, this indicates two distinct physical scenarios with different geometrical locations within the blazar jet that can produce the emission. In Figure 4, the comoving energy densities of all external photon fields considered in our model are displayed as a function of radius for a set Lorentz factor, alongside the posterior distribution of the radius obtained from our fit. We note that these calculations were performed for $\delta = 40.74$. Changes in δ affects all external fields nearly equally, meaning that the curves in Figure 4 will only scale up or down, while the relationships among all fields will remain unchanged. For the first maximum of the radius posterior distribution at $R \sim 4 \times 10^{15}$ cm, for which the corresponding model is shown in red in the upper-right panel of Figure 3, the photon comoving energy density is dominated by photons from the disk, indicating that the HE is due to external Compton. However, the other maximum indicates a significantly larger radius ($\sim 5 \times 10^{17}$ cm). This parameter set is also characterized by a higher L_e and γ_{max} but a lower Doppler boost δ. For this scenario, the contributions of external fields are reduced due to the much larger radius, and the electron energy density is higher, while the magnetic energy density is lower. Therefore, the emission in the HE band is attributed to the SSC emission for this parameter set. We show the most probable spectrum from this parameter mode by the red dashed line in the upper-right panel of Figure 3.

The modeling method for the SED of FSRQs presented here allows the emission location to remain unspecified, instead of presupposing the contributing fields to the formation of HE emission. Consequently, all photon fields are considered, enabling the identification of the most suitable option (in terms of model-free parameters and external photon fields) for explaining the data. This method also presents the flexibility of setting the disk luminosity L_d and the radius R such as to select beforehand the dominant external field component to be used by the model and obtain the relevant parameter distributions under these assumptions.

5. Conclusion

In this data-rich era of blazar research, the modeling of their SEDs using numerically intensive models is becoming essential. This approach is the principal way to thoroughly investigate the processes occurring within their jets. In Bégué et al. (2024), we introduced a novel framework for fitting the SED of blazars. This framework uses a CNN trained on a large set of SSC spectra generated by SOPRANO that accurately reproduces the radiative signatures of the model and is used to fit observational data and infer model parameters. Here, we extend this framework by training the CNN on a more complex model that also includes external photon fields for inverse Compton scattering and pair creation.

In this study, the CNN is trained on spectra generated by SOPRANO, incorporating the most relevant external photon fields (direct disk photons, photons from the BLR, and the dusty torus) and internal photon fields (synchrotron photons), as well as all relevant cooling processes and the pair creation process. We have shown that, despite increased complexity (7 versus 11 free parameters for the SSC and EIC models, respectively), slightly modifying the layers and dimensions of the CNN enables accurate reproduction of the model spectral features. The CNN execution time is significantly shorter compared to computations with SOPRANO (or any other code that self-consistently treats particle acceleration and injection), allowing it to be effectively coupled with fitting techniques to derive parameters that best describe the data. The application of the CNN has been demonstrated through fitting the broadband SEDs of the well-known FSRQs 3C 454.3 and CTA 102 during flaring periods, thereby constraining the parameters of the EIC model and obtaining their posterior distributions.

Although the model satisfactorily reproduces the radiative signature of particle emissions when the emission region is within an intense radiation field, several limitations had to be considered for the fitting procedure. First, the model has been trained for a wide range of parameters, which are close to physically realistic scenarios. Specifically, the disk temperature, and consequently the energy of disk photons, is determined from the black hole mass and accretion disk luminosity, both entailed with large uncertainties. Additionally, the temperatures of the BLR and the torus are free parameters that can vary within a certain range. Treating these parameters as free variables in the fitting process allows for a wider selection range, which aids in converging toward the optimal fit parameters. Conversely, fixing these parameters, particularly the black hole mass, to their estimated observational values could lead to more stringent constraints on the remaining free parameters. This last approach was used to perform the fit presented in Section 4.

Our model also suffers from limitations: linking the disk temperature to the black hole mass and accretion disk luminosity imposes limitations on accurately reproducing emissions from a hot accretion disk. Ghisellini et al. (2009) suggested that the blue bump observed in the UV band could

be attributed to thermal disk emission. To achieve a disk component peaking at $\sim 10^{15}$ Hz under the disk model considered here, a luminosity $L_{\rm d} \geqslant 10^{48}\,{\rm erg\,s^{-1}}$ would be required. This would result in a very high flux of the disk component with respect to the jet emission. This limitation boils down to assuming that the radiative efficiency of the disk η is set to $\eta = 0.1$. However, keeping this limitation in mind, the approach can be successfully applied to high-redshift blazars (see, e.g., Ghisellini et al. 2009; Sahakyan et al. 2020, 2024).

Other limitations arise from the assumption that $R = R_{\rm diss}$. (1) This assumption imposes a constraint that limits extending the emission distance to larger radii. In the current model setup, the emission region extends only a few times beyond $R_{\rm BLR}$ and does not reach, for example, 10^{19} cm or beyond. This is important, as Costamante et al. (2018) demonstrated that HE and VHE emissions from blazars can be produced outside the BLR. (2) This assumption further leads to broad jets, resulting in low compactness. Alternative assumptions could have been made. For instance, one could assume a specific jet opening angle and link the size of the emitting region R to the radius of dissipation $R_{\rm diss}$, remembering that the length in the direction orthogonal to the motion is a relativistic invariant; see, e.g., Ghisellini & Tavecchio (2009) for such a model.

To overcome and analyze the impacts of these constraints, we are training new CNNs based on models devoid of such limitations. Namely, in one of these models the external emission is modeled as a simple blackbody with independent luminosity and temperature to accurately represent cases where an excess in the UV band is observed, at the expense of losing a straightforward physical interpretation. Alternative models being considered involve different relationship between R and $R_{\rm diss}$. Results from these models and a detailed comparison between their performances in fitting the data will be presented elsewhere.

Moreover, a model could be developed where both the jet opening angle and the observer's direction are variable parameters. This would require the inclusion of anisotropic radiation fields in SOPRANO, resulting in large code modifications. However, this approach would offer the advantage of allowing independent constraints on the jet's geometrical parameters and the viewing angle, enabling the study of different classes of AGN, such as Seyfert galaxies associated with VHE neutrino emission (IceCube Collaboration et al. 2022). We do not expect that substantial changes to the CNN would be required for such a model. Indeed, the CNN method we introduced in Bégué et al. (2024) appears to be rather robust to model change.

The CNN presented here, together with the one from Bégué et al. (2024), constitutes a comprehensive and flexible methodology for self-consistent blazar modeling within leptonic scenarios. These models collectively enable in-depth investigations of the emission from both BL Lacs and FSRQs, and hold the potential to advance our understanding of the physics of relativistic jets. Similar to the SSC model, the EIC model will be made publicly available through the MMDC. This platform permits researchers to upload their data and perform self-consistent modeling, with both the SSC and EIC models, through a dedicated web interface, providing them with the best-fitting parameters and their posterior distributions.

In summary, in this study, we demonstrated that the methodology presented in Bégué et al. (2024) can be extended and applied to more complex physical scenarios, which involve more free parameters and require taking into account the contribution of several photon fields. As a natural next step, this methodology will be expanded to include the even more sophisticated hadronic and lepto-hadronic models. In these models, along with the contribution of electrons, the contributions of protons and secondaries produced in photo-pion and photo-pair interactions need to be considered. Training a CNN on these new models requires careful treatment of the interactions of all primary and secondary particles. This is not a trivial task but is essential for interpreting multimessenger data from blazar observations, including both photons and neutrinos.

Acknowledgments

We thank the anonymous referee for constructive comments, which helped to improve the clarity of the manuscript. N.S., S. G., and M.K. acknowledge the support by the Higher Education and Science Committee of the Republic of Armenia, in the frames of the research project No. 23LCG-1C004. D.B., H.D.B., and A.P. acknowledge support from the European Research Council via the ERC consolidating grant No. 773062 (acronym O.M.J.). We acknowledge with gratitude that Altair has agreed to pay for the generation of the training sample on AWS EC2.

Data Availability

All the observational data used in this paper are public. The convolutional neural network used to fit the SEDs can be shared on a reasonable request to the corresponding author. In addition, it is publicly available through the Markarian Multiwavelength Data Center (http://www.mmdc.am).

Appendix
Parameter Posteriors for 3C 454.3 and CTA 102

We show in this Appendix the parameter posterior distributions of 3C 454.3 and CTA 102 in the left and right panels of Figure 5, respectively.

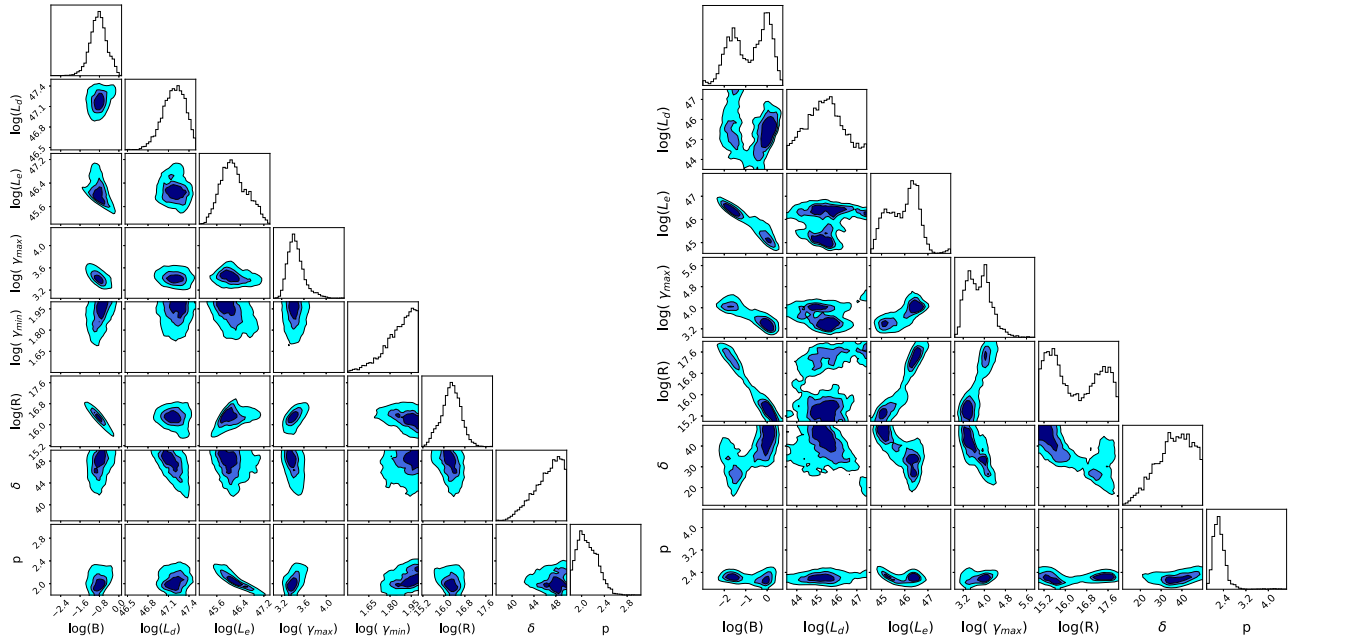

Figure 5. Left panel: parameter posterior distributions from the SED modeling of 3C 454.3 for the period MJD 55,519.59–55,520.19. Right panel: parameter posterior distributions for CTA 102, showing a bimodal distribution for the radius R alongside other parameters.

ORCID iDs

N. Sahakyan https://orcid.org/0000-0003-2011-2731
D. Bégué https://orcid.org/0000-0003-4477-1846
A. Casotto https://orcid.org/0009-0007-4522-5501
H. Dereli-Bégué https://orcid.org/0000-0002-8852-7530
P. Giommi https://orcid.org/0000-0002-2265-5003
S. Gasparyan https://orcid.org/0000-0002-0031-7759
V. Vardanyan https://orcid.org/0000-0002-3777-7580
M. Khachatryan https://orcid.org/0009-0007-7798-2072
A. Pe'er https://orcid.org/0000-0001-8667-0889

References

Abe, H., Abe, S., Acciari, V., et al. 2023, ApJS, 266, 37
Acharyya, A., Adams, C., Archer, A., et al. 2023, ApJ, 954, 70
Ansoldi, S., Antonelli, A., Arcaro, C., et al. 2018, ApJL, 863, L10
Bégué, D., Sahakyan, N., Dereli Bégué, H., et al. 2024, ApJ, 963, 71
Błażejowski, M., Sikora, M., Moderski, R., & Madejski, G. M. 2000, ApJ, 545, 107
Bloom, S. D., & Marscher, A. P. 1996, ApJ, 461, 657
Bonnoli, G., Ghisellini, G., Foschini, L., Tavecchio, F., & Ghirlanda, G. 2011, MNRAS, 410, 368
Böttcher, M., Reimer, A., Sweeney, K., & Prakash, A. 2013, ApJ, 768, 54
Burgess, J. M. 2023, JOSS, 8, 4969
Cerruti, M., Zech, A., & Boisson, C. 2019, MNRAS, 483, L12
Cerruti, M., Zech, A., Boisson, C., & Inoue, S. 2015, MNRAS, 448, 910
Costamante, L., Cutini, S., Tosti, G., Antolini, E., & Tramacere, A. 2018, MNRAS, 477, 4749
Dermer, C. D., & Schlickeiser, R. 1994, ApJS, 90, 945
Dermer, C. D., Schlickeiser, R., & Mastichiadis, A. 1992, A&A, 256, L27
Domínguez, A., Primack, J., Rosario, J., et al. 2011, MNRAS, 410, 2556
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Finke, J. D. 2016, ApJ, 830, 94
Gao, S., Fedynitch, A., Winter, W., & Pohl, M. 2019, NatAs, 3, 88
Gao, S., Pohl, M., & Winter, W. 2017, ApJ, 843, 109
Gasparyan, S., Bégué, D., & Sahakyan, N. 2022, MNRAS, 509, 2102
Ghisellini, G., Maraschi, L., & Treves, A. 1985, A&A, 146, 204
Ghisellini, G., & Tavecchio, F. 2009, MNRAS, 397, 985
Ghisellini, G., Tavecchio, F., & Ghirlanda, G. 2009, MNRAS, 399, 2041
Gu, M., Cao, X., & Jiang, D. R. 2001, MNRAS, 327, 1111
IceCube Collaboration, Aartsen, M. G., Ackermann, M., et al. 2018a, Sci, 361, 147
IceCube Collaboration, Aartsen, M. G., Ackermann, M., et al. 2018b, Sci, 361, eaat1378
IceCube Collaboration, Abbasi, R., Ackermann, M., et al. 2022, Sci, 378, 538
Keivani, A., Murase, K., Petropoulou, M., et al. 2018, ApJ, 864, 84
Kirk, J. G., Guthmann, A. W., Gallant, Y. A., & Achterberg, A. 2000, ApJ, 542, 235
Liu, Y., Jiang, D. R., & Gu, M. F. 2006, ApJ, 637, 669
Mannheim, K. 1993, A&A, 269, 67
Mannheim, K., & Biermann, P. L. 1989, A&A, 221, 211
Maraschi, L., Ghisellini, G., & Celotti, A. 1992, ApJL, 397, L5
Mastichiadis, A., & Kirk, J. G. 1995, A&A, 295, 613
McKay, M. D., Beckman, R. J., & Conover, W. J. 2000, Technometrics, 42, 55
Mücke, A., & Protheroe, R. J. 2001, APh, 15, 121
Mücke, A., Protheroe, R. J., Engel, R., Rachen, J. P., & Stanev, T. 2003, APh, 18, 593
Murase, K., Oikonomou, F., & Petropoulou, M. 2018, ApJ, 865, 124
Nigro, C., Sitarek, J., Gliwny, P., et al. 2022, A&A, 660, A18
Padovani, P., Giommi, P., & Resconi, E. 2018, MNRAS, 480, 192
Padovani, P., Alexander, D., Assef, R., et al. 2017, A&ARv, 25, 2
Petropoulou, M., & Mastichiadis, A. 2015, MNRAS, 447, 36
Pian, E., Falomo, R., & Treves, A. 2005, MNRAS, 361, 919
Righi, C., Tavecchio, F., & Pacciani, L. 2019, MNRAS, 484, 2067
Sahakyan, N. 2018, ApJ, 866, 109
Sahakyan, N. 2019, A&A, 622, A144
Sahakyan, N. 2021, MNRAS, 504, 5074
Sahakyan, N., & Giommi, P. 2022, MNRAS, 513, 4645
Sahakyan, N., Giommi, P., & Padovani, P. 2023, MNRAS, 519, 1396
Sahakyan, N., Harutyunyan, G., Gasparyan, S., & Israyelyan, D. 2024, MNRAS, 528, 5990
Sahakyan, N., Israyelyan, D., & Harutyunyan, G. 2022, MNRAS, 517, 2757
Sahakyan, N., Israyelyan, D., Harutyunyan, G., Khachatryan, M., & Gasparyan, S. 2020, MNRAS, 498, 2594
Schmidt, M. 1965, ApJ, 141, 1295
Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 500, 33
Sikora, M., Begelman, M. C., & Rees, M. J. 1994, ApJ, 421, 153
Sironi, L., & Spitkovsky, A. 2011, ApJ, 726, 75
Stathopoulos, S. I., Petropoulou, M., Vasilopoulos, G., & Mastichiadis, A. 2024, A&A, 683, A225
Tramacere, A., 2020 JetSeT: Numerical modeling and SED fitting tool for relativistic jets, Astrophysics Source Code Library, ascl:2009.001
Tramacere, A., Giommi, P., Perri, M., Verrecchia, F., & Tosti, G. 2009, A&A, 501, 879
Tramacere, A., Massaro, E., & Taylor, A. M. 2011, ApJ, 739, 66
Tzavellas, A., Vasilopoulos, G., Petropoulou, M., Mastichiadis, A., & Stathopoulos, S. I. 2024, A&A, 683, A185

Urry, C. M., & Padovani, P. 1995, PASP, 107, 803
Uzdensky, D. A. 2022, JPlPh, 88, 905880114
Viana, F. A. 2016, Quality and Reliability Engineering International, Vol. 32 (Wiley Online Library), 1975
Woo, J.-H., & Urry, C. M. 2002, ApJ, 579, 530
Zabalza, V. 2015, ICRC (The Hague), 34, 922
Zamaninasab, M., Clausen-Brown, E., Savolainen, T., & Tchekhovskoy, A. 2014, Natur, 510, 126