

Impact of model parameter degeneracy on leptonic radiation models

The case of blazar multiwavelength spectra

F. Apel^{1,*}, A. Omeliukh^{1,*}, A. Franckowiak^{1,2}, and J. Lederer³

¹ Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute (AIRUB), Universitätsstraße 150, 44801 Bochum, Germany

² Ruhr Astroparticle and Plasma Physics Center (RAPP Center), Germany

³ University of Hamburg, MIN Department, Bundesstraße 55, 20146 Hamburg, Germany

Received 6 June 2025 / Accepted 10 August 2025

ABSTRACT

Context. Leptonic one-zone radiation models are commonly used to describe multiwavelength data and explore the physical properties of high-energy sources, such as active galactic nuclei. However, these models often require a large number of free parameters.

Aims. In the context of a possible parameter degeneracy and the complex landscape of the parameter space, we study how the choice of the fitting procedure impacts the characterization of the source properties. Furthermore, we examine how the data coverage and the uncertainties associated with the data influence the model parameter degeneracy.

Methods. We generated simulated spectral energy distribution datasets with different properties and fit them with a numerical model, using seven free parameters to describe the relevant radiation processes. We compared different optimization algorithms and studied the parameter degeneracy using t-distributed stochastic neighbor embedding. In addition, we applied the same fitting procedures to the observational data of two sources: Mrk 501 and PKS 0735+178.

Results. We demonstrate significant degeneracies in the seven-dimensional parameter space of the one-zone leptonic models caused by the incomplete wavelength coverage of the data. Given the same goodness-of-fit function, we find that the best-fit result depends on the choice of the minimization algorithm.

Conclusions. Source properties extracted from the best-fit solution and applied to realistic datasets cannot be interpreted as the only solution given the significant degeneracies in the model parameters. We find that adding new energy ranges (e.g., MeV) and regular source monitoring would allow for the gaps in the data to be reduced, significantly decreasing the parameter degeneracy.

Key words. radiation mechanisms: non-thermal – methods: numerical – BL Lacertae objects: general

1. Introduction

Blazars are a rare subclass of active galactic nuclei (AGNs) that shoot a relativistic jet close to the observer’s line of sight (Urry & Padovani 1995). Small viewing angles and the relativistic speed of the plasma in the jet make the Doppler boosting especially efficient, leading to the high luminosity seen for blazar jets. The emission of blazars has a non-thermal nature and spans from radio frequencies to gamma rays.

Blazars are the dominant sources of extragalactic GeV and TeV gamma rays (Di Mauro et al. 2014; Ajello et al. 2015; Abdollahi et al. 2020; Bose et al. 2022). They are also suggested as neutrino source candidates (see Giommi & Padovani 2021, for a recent review) and candidate neutrino blazar associations have been identified (IceCube Collaboration 2018b,a; Kadler et al. 2016; Gao et al. 2017; Garrappa et al. 2019; Franckowiak et al. 2020; Rodrigues et al. 2021; Sahakyan et al. 2023). A good understanding of the blazar emission mechanisms is crucial for studying the nature of these sources and is essential for both gamma-ray and neutrino astronomy.

The spectral energy distribution (SED) of blazars exhibits a typical two-bump structure. It is commonly accepted that the low-energy emission (from radio to UV or X-rays) is the synchrotron emission of relativistic electrons in the jet. The nature of the high-energy emission (X-ray to gamma rays) remains unclear. The simplest way to explain the blazar spectral energy distribution is to assume that all radiation originates from one emission region called a “blob.” Depending on the particle species that produce the high-energy emission, the radiation models can be divided into leptonic, hadronic, or leptohadronic. In one-zone leptonic models, only relativistic electrons accelerated in the jet are responsible for the observed SED. The high-energy emission is explained by Compton scattering of low-energy photons by the same electrons that produce the synchrotron emission at lower energies. Those models are called synchrotron self-Compton (SSC) models. Purely hadronic models assume that gamma rays originate exclusively from proton synchrotron emission of ultra-relativistic protons, while the high-energy electron population is responsible for the low-energy emission. However, due to high proton energy requirements, the hadronic component can only be subdominant in blazars (Liodakis & Petropoulou 2020). Purely hadronic models would also require a different accretion paradigm,

* Corresponding authors: apel@astro.rub.de; omeliukh@astro.rub.de

** The work was split by the two first authors in equal parts.

which makes them highly disfavored (Zdziarski & Böttcher 2015). In leptohadronic models, both protons and electrons contribute to the high-energy part of the SED. By interacting with synchrotron photons, these protons produce hadronic cascades that contribute to the high-energy emission alongside the leptonic inverse Compton effect.

Numerical modeling of SEDs is applied to explain the properties of gamma-ray-emitting blazars (Böttcher et al. 2013; Cerruti et al. 2013; Petropoulou et al. 2014; Dermer et al. 2014; Petropoulou & Mastichiadis 2015; Cerruti et al. 2015; Paliya 2015, and many others) or for studying their neutrino emission (Petropoulou et al. 2015; Keivani et al. 2018; Gao et al. 2017; Petropoulou et al. 2020; Rodrigues et al. 2021; Gasparyan et al. 2022; Sahakyan et al. 2023, and many others). As noted in many of these works, complex radiation models typically have a high number of free parameters and require a lot of computational efforts (especially for time-dependent models) to properly fit the data. Often, models are fitted to the data based on a small number of probed parameter values or with several parameters fixed.

With the current advancements in numerical modeling, a new generation of codes was developed. Software packages such as AM³ (Klinger et al. 2024), LeHaMoC (Stathopoulos et al. 2024), or SOPRANO (Gasparyan et al. 2022) allow us to perform fast computations of blazar SEDs. This, in turn, opens up the possibility of performing data fitting and goodness-of-fit estimations with a high number of free model parameters. In fact, a high number of simulated models was already used for fitting observed blazar SEDs in Rodrigues et al. (2021); Rodrigues et al. (2024b,a); Omeliukh et al. (2025) and for training neural networks in Tzavellas et al. (2024); Bégué et al. (2024); Sahakyan et al. (2024).

This paper addresses the challenges of high-level data fitting using radiation models. We highlight that the parameters of one-zone leptonic models are degenerate (see Section 4). In a thorough comparison of different optimization algorithms, we show that the parameter degeneracy leads to ambiguity in the data fitting and physical interpretation of the models. While this work considers only blazar one-zone leptonic radiation models, similar problems arise in the radiation models applied to other classes of sources, including Seyfert galaxies, gamma-ray bursts (GRBs), tidal disruption events (TDEs), and others.

This paper is structured as follows. Section 2 describes the physical setup of one-zone leptonic models in detail. Section 3 introduces the chosen methods for visualization of the multi-dimensional data. In Section 4, we demonstrate the irregularity of the parameter space due to the nature of the radiative models and discuss its implications. Section 5 provides a comparison of the performance of different optimization algorithms tested on three sets of simulated data which are then applied to observational data in Section 6. We discuss our results in the context of current state-of-the-art numerical modeling and its applications in Section 7 and summarize our findings in Section 8. For the calculations in this paper, we adopted a flat Λ CDM cosmological model with parameters $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $\Omega_m = 0.3$.

2. Leptonic models

The leptonic SSC model is the way to explain the blazar multiwavelength emission with the smallest number of free parameters. For the numerical modeling of the SSC scenario, we utilized the open-source time-dependent code AM³ (Klinger et al. 2024). It allows us to numerically solve the system of coupled

Table 1. Leptonic model parameters.

Parameter	Description
R'_{blob} , [cm]	Radius of the spherical emission zone region
B' , [G]	Strength of the homogeneous magnetic field
Γ_b	Blob Lorentz factor
γ_e^{min}	Minimal electron Lorentz factor
γ_e^{max}	Maximal electron Lorentz factor
α_e	Power-law index of the electron energy distribution
L'_e , [erg s ⁻¹]	Total electron luminosity

integro-differential equations that describe the evolution of the particle spectra in a fully self-consistent manner.

We assumed that electrons are accelerated to a simple power-law spectrum¹ $dN/d\gamma'_e \propto \gamma'^{-\alpha_e}$ with a spectral index denoted as α_e , spanning a range of Lorentz factors from γ_e^{min} to γ_e^{max} . While more complex electron spectra are possible, we chose the simplest physically motivated spectrum for a minimalistic setup. The energy spectrum of the electrons is normalized to the total electron luminosity parameter, L'_e . These particles are injected into a single spherical blob of size of R'_b (in the comoving frame of the jet) moving along the jet with Lorentz factor Γ_b , where they encounter a homogeneous and isotropic magnetic field of strength B' . We assumed the jet is observed at an angle of $\theta_{\text{obs}} = 1/\Gamma_b$ relative to its axis, resulting in a Doppler factor of $\delta_D = \Gamma_b$. We adopted a steady-state approximation to obtain the particle spectra. The characteristic escape time is set equal to the light-crossing time for all particles. We evolved the kinetic equations over several escape timescales to ensure that the steady state is reached. During the photon propagation from the source to the observer, a part of the high-energy gamma rays is attenuated due to interactions with the extragalactic background light (EBL). Unless a different model is explicitly stated, this effect is accounted for in all models based on Domínguez et al. (2011). The parameter space of the leptonic models is represented by the seven free parameters summarized in Table 1.

3. Visualization of parameter spaces

A typical procedure of data fitting is the optimization of a goodness of fit. As a goodness-of-fit function, we chose a reduced χ^2 -function (i.e., divided per number of degrees of freedom) defined as

$$\chi^2(\theta)/\text{n.d.f.} = \frac{1}{N - N_{\text{par}} + 1} \sum_i \frac{(F_i^{\text{data}} - F_i^{\text{model}}(\theta))^2}{\sigma_i^2}, \quad (1)$$

where N is the number of data points, $N_{\text{par}} = 7$ the number of free parameters in the model, F^{data} are the observed fluxes, F^{model} are the predicted fluxes by the model, and i is the summation index that corresponds to the observed frequency values, and σ_i are the flux measurement uncertainties. The function input $\theta \in \mathbb{R}^7$ are the model parameters from Table 1. Even though, in the general case, different goodness-of-fit functions are possible, we selected the reduced χ^2 function commonly used in physics and astronomy, as its values directly indicate a poor fit ($\chi^2(\theta)/\text{n.d.f.} \gg 1$), an overfitting ($\chi^2(\theta)/\text{n.d.f.} \ll 1$) or a good fit ($\chi^2(\theta)/\text{n.d.f.} \approx 1$). Since overfitting rarely happens in the

¹ Parameters with or without prime refer to the values in the jet or observer's frame, respectively.

context of our SED modeling, we defined the model that produces the minimal value of the reduced χ^2 as the best fit within the scope of this work. If the SED contains upper limits on the flux values, they do not contribute to the reduced χ^2 value as long as the model predictions are below their values. In the opposite case, a large value (10^5) is added to the reduced χ^2 , so that a model overshooting the upper limits would be considered highly unfavorable.

To compare the results of the selected minimization procedures (Section 5), it is essential to define how close or how far multidimensional vectors of the model parameters are located with respect to each other in the parameter space depending on their goodness-of-fit value. A visualization of the parameter space reveals multiple regions with comparable χ^2 values. For this purpose, we used t-distributed stochastic neighbor embedding (t-SNE), a machine learning technique for mapping high-dimensional data into low-dimensional representations (van der Maaten & Hinton 2008). The dimensionality reduction in t-SNE is performed by matching the distribution of point similarities (defined as probabilities non-linearly proportional to the metric distances) in high-dimensional space and low-dimensional space of two abstract coordinates. As a result, points with higher similarities are arranged closer to each other in the two-dimensional map, but the ratio between the distances of points in high-dimensional space and in low-dimensional space is not conserved. An important concern of t-SNE usage is the discrimination of truly close points and artificial clustering. A hyperparameter that can affect artificial clustering is referred to as perplexity. It reflects the amount of nearest neighbors and balances the analysis of local and global data features. The details on hyperparameter tuning to ensure the adequate reflection of point similarities can be found in Appendix B.

In this paper, we use the t-SNE tool from the `scikit-learn` open-source Python library (Pedregosa et al. 2011). Before normalizing the values, we transformed those parameters with values spanning over several magnitudes (R'_{blob} , γ'_{min} , γ'_{max} , and L'_c) to logarithmic scale. This step should be performed in addition to the normalization to ensure a more balanced contribution from all parameters by avoiding skewed distributions when normalizing to a [0,1] scale. Without this step, extreme values (e.g., 10^{40} vs. 10^{47}) were dominant in the analysis, leading to smaller, but meaningful variations being neglected.

Then the low-dimensional map of the parameter space was built using t-SNE. The goodness-of-fit value that corresponds to each parameter vector is used as a color scale to highlight the regions in the parameter space with the best solutions. Fig. 1 shows an example of such a t-SNE output which represents the parameter space of the leptonic models applied to the simulated dataset 2 (see Sect. 5.1).

The darker regions in the plot correspond to the regions where good-fit solutions can be found. The plot shows 50 000 models with different combinations of model parameters. Figure 1 highlights five regions where the solutions with low reduced $\chi^2 < 7$ are located. Those solutions are shown with black markers bigger than the rest of the points. Due to the systematic and statistical uncertainties of the data, slight variations in the model parameters can still produce an SED with a small goodness-of-fit value. Therefore, the solutions are expected to form island-like structures with close (but slightly deviating) values of model parameters. Apart from this, there might be a completely different solution from the distant region in the parameter space that produces a similarly good fit. This can be seen in Fig. 1 where several distant regions contain good solutions which, in turn, are surrounded by candidates with

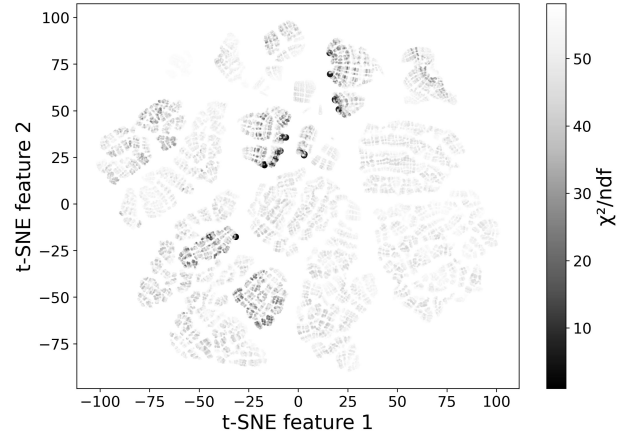


Fig. 1. Example of the seven-dimensional parameter space mapping with t-SNE (grid scan results for simulated dataset 2, see Sect. 5.1). Each point represents one model consisting of seven parameters. Darker colors correspond to lower $\chi^2/\text{n.d.f.}$ values. The plot shows the first 50 000 points with the smallest $\chi^2/\text{n.d.f.}$ value, and the perplexity is set to 30.

slightly worse goodness of fit. For the following comparisons, the close location of two points in the t-SNE plots should be interpreted as close parameter values for the corresponding models. On the contrary, points scattered over the t-SNE output are located in distant regions and correspond to physically different solutions.

4. Challenges for the parameter fitting

Minimizing the goodness-of-fit function is challenging because the function is expected to be neither convex² nor smooth³. The first reason lies in the nature of the radiation processes. Processes that have thresholds (like $\gamma\gamma \rightarrow e^+e^-$) or show very different features at different energies (e.g., the Thomson and Klein–Nishina regimes in the case of the inverse Compton effect) impact the smoothness of the goodness-of-fit function since even a small change in some of the parameters leads to a drastically different shape, meaning that the goodness-of-fit function cannot be differentiated numerically. The second reason for expecting a non-smooth and non-convex goodness-of-fit function is the limited amount of available photon flux measurements. In the ideal case, for data covering many orders of magnitudes without gaps, the shape of the SED is unambiguously defined by the spectrum of the electron population and the conditions in the emission zone. In reality, however, the SED almost always has gaps in some energy ranges due to the limited number of instruments covering only selected energy ranges. Different electron populations can produce similar photon-flux levels with differences that could be spotted only in the missing data regions. By varying, for example, the electron maximum Lorentz factor, spectral index, and electron luminosity (which normalizes the spectrum), similar levels of photon fluxes can be achieved for different combinations of these parameters. Additionally, the observed luminosity follows $L \simeq \delta_D^4 L'$, meaning that for each value of electron luminosity in the source, there is a specific bulk Lorentz factor for which the observed luminosity would remain unchanged. However, different blob Lorentz factors would

² A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if for all $0 \leq t \leq 1$ and all $x, y \in \mathbb{R}^n$ $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$.

³ All first derivatives are continuous.

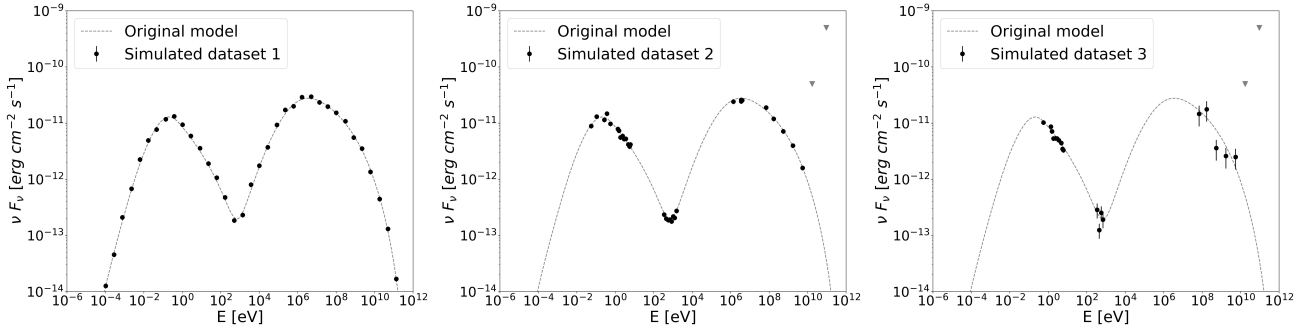


Fig. 2. Simulated datasets. Black dots show the data points in the simulated datasets. Dashed curve shows the original model that was used to generate the data. The gray triangles in the last two plots are the upper limits.

produce different SED shapes. This could be corrected by variations in the electron spectral parameters, thereby resulting in very similar SEDs. The existence of multiple good-fitting solutions suggests that any goodness-of-fit function is expected to be non-convex. Hence, formally defining optimal fits is straightforward, but computing them is challenging. We address this topic further in Section 5.

5. Comparison of minimization procedures using simulated data

In a non-smooth and non-convex parameter space, the choice of the minimization algorithm can impact the result of the fitting. This also leads to implications on the physical interpretation of the models. We selected five minimization techniques most commonly used in the literature: Minuit, grid scan, two evolutionary algorithms, and Markov chain Monte Carlo (MCMC). In this section, we describe how we tested the performance of the selected algorithm on simulated pseudo-data, with knowledge of the true values of the model parameters. To study the effects of the wavelength coverage of the data and of the size of the uncertainties associated with the data, we created three simulated datasets. We then compared the best-fit results of each algorithm with the true values.

5.1. Simulated data

Unlike the case of real blazars, where the true underlying physical parameters are unknown, fitting pseudo-data with known parameters allows us to make a direct comparison of the varying accuracy of different minimization algorithms. The simulated data were generated with AM³ using the parameters from the first row of Table 4 (true). To create a SED in a realistic flux and energy range, we chose parameter values near the best-fit leptonic model parameters for the quiescent state of the blazar PKS 0735+178 provided in Omeliukh et al. (2025). Based on this SED, we created three datasets (Fig. 2), each resembling a different wavelength coverage of the data and different uncertainties associated with the flux measurements.

For dataset 1, we equally spaced the data points on a logarithmic scale across all energies. To mimic the uncertainty of the flux measurements, we added a random value distributed normally around zero and with a standard deviation corresponding to 10% of the flux value. After the new, shifted fluxes were calculated, we assumed a 10% uncertainty for all data points. Dataset 1 represents the case of a “perfect” SED.

Dataset 2 resembles a more realistic case with good data coverage. We selected energies that correspond to K, z, I, r, and g filters in the optical range and U, B, W1, M2, and

W2 corresponding to the ultraviolet (UV) bands of the UVOT instrument onboard the *Neil Gehrels Swift* Observatory (Roming et al. 2005). This represents the typical energy range for X-ray spectra from *Swift*-XRT (Burrows et al. 2004) and NuSTAR (Harrison et al. 2013), as well as the centers of the four COMPTEL (Schoenfelder et al. 1993) energy bands (0.75–1, 1–3, 3–10 and 10–30 MeV), and typical energy bands of the gamma-ray telescope *Fermi*-LAT (Atwood et al. 2009). The two highest energy data points represent sensitivity-based upper limits. The uncertainty of the flux measurements is added in the same way as was done for dataset 1.

Dataset 3 represents a more pessimistic case with a limited amount of available data. The data covers only the optical/UV, the typical *Swift*-XRT bands in X-rays, and the *Fermi*-LAT gamma-ray range. Unlike the two previous datasets, we assumed different flux uncertainties for different energy bands: 10% in optical/UV, 30% in X-rays, and 40% in gamma rays. The procedure for accounting for the uncertainties remained the same as for dataset 1. Thus, taken together, the three datasets provide a panorama of possible situations.

5.2. Tested optimization algorithms

All selected algorithms aim to minimize the same reduced χ^2 function defined in Eq. (1). For the simulated data, they searched for the best solutions within the same parameter space limited by the boundaries shown in Table A.1.

5.2.1. Grid scan

A grid scan is an approach where values of the goodness-of-fit function are evaluated for each combination of the parameters equally spaced between defined boundaries (Table A.1). This simple setup allowed us to cover the whole parameter space homogeneously, while highlighting the regions of interest corresponding to low reduced χ^2 values. We selected ten points per parameter, probing in total 10^7 points in the parameter space. Usually, the discretization of the grid is not good enough to find a precise best-fit solution. Therefore, as a second step, we locally minimized the best result from the grid scan using Minuit (using simplex followed by migrad, see Sect. 5.2.2). Since all neighboring points of the current best-fit grid scan solution have a higher reduced χ^2 value, we adopted their values as new boundaries for searching for the new best fit. The local minimization step required an additional probing of 300 points.

5.2.2. Minuit

Minuit is a software library for numerical minimization developed at CERN (James 1998). For this work, we used the Python

Table 2. Parameters of the genetic algorithm.

Parameter	Values
Number of generations	70–80
Number of individuals	10^4
Mutation probability	0.33
Independent probability	1/7
Crossover probability	0.20

interface for Minuit called `iMinuit` (Dembinski et al. 2020). We started the minimization with the Nelder-Mead `simplex` method (Nelder & Mead 1965) due to its robustness. The independence of the gradient makes it especially suitable for the large parameter space of the leptonic models. Due to the nature of the radiation processes, small changes in the model parameters could induce such a change in the reduced χ^2 that it ends up becoming numerically non-differentiable at that point. Following the `simplex` method, we used `migrad`, which is based on gradient descent, to refine the best-fit solution locally.

We initialized the `simplex` algorithm based on some physically motivated assumptions on the model parameters derived from the SED features. We set the maximum number of function calls to 1200. Since the first run never yielded an acceptable value of reduced χ^2 , we repeated the procedure while refining the boundaries of the parameters and reducing the parameter space around the previously found region. Our new initial point in the second run was set as the best-fit solution from the previous run. The second run required around 300 function calls.

5.2.3. Genetic algorithm

Another minimization technique is the genetic algorithm (Kramer & Kramer 2017). Inspired by biological evolution, a set of candidate solutions is evaluated and evolved toward an optimal solution. We used the evolutionary functions from the DEAP module (Fortin et al. 2012) in Python. The algorithm starts with initializing a population of random parameter vectors. The fitness of each individual one is evaluated by calculating each $\chi^2/\text{n.d.f.}$ value for a given dataset. Parents with better fitness values are selected to generate the offsprings via crossover and mutation. The crossover imitates the combination of DNA sequences during reproduction. The crossover operators mix genetic material of parent solutions, and there are various approaches to achieve this. In this work, the method involves splitting the two parent solutions and reassembling them alternately. For the mutation, certain properties of the selected solutions are randomly altered to generate new offspring genes. Another parameter is the independent probability for each parameter to mutate. The crossover, mutation, and independent probability are controllable parameters of the genetic algorithm. Their values used for this work are given in Table 2.

After the variation with crossover and mutation, the previous parent population was replaced by the new offspring and the process was repeated over several generations. We took the standard deviations of the parameter values in the last generation as the uncertainty in the parameter estimation.

5.2.4. CMA-ES

Covariance Matrix Adaption Evolution Strategy (CMA-ES, Hansen 2016) is the second evolutionary algorithm used in this work. The procedure is similar to the genetic algorithm, with

the main difference linked to the mutation treatment. The mutation strength indicates how strongly the offspring differs from the parent population. CMA-ES adjusts the mutation step size based on the success of the previous mutations, which allows the algorithm to learn the mutation strength during the evolution. Compared to the genetic algorithm, this strategy requires less manual parameter tuning. While the multivariate Gaussian distribution is the generalization of the Gaussian distribution to higher dimensions, the covariance matrix can be interpreted as the generalization of the variance. From a multivariate Gaussian distribution, the candidate solutions are generated and evaluated and the sampling distribution is updated for the next iteration to increase the likelihood of selecting better solutions. Further information about CMA-ES can be found, for instance, in Hansen & Ostermeier (2001). We used the python package `pycma` by Hansen et al. (2019). The initial step size for the first generation has to be set in the beginning. After the first generation, this quantity is adjusted in each generation. In this work, the initial step size was set to $\sigma = 2$. The population size is fixed to 5000 individuals. Similarly to the genetic algorithm, the uncertainty in the parameter estimation was assumed to be the standard deviations of the corresponding parameter values in the last generation.

5.2.5. MCMC

The MCMC method allows us to draw samples from probability distributions. The posterior probability distribution allows for the optimal values of the model parameters to be found. Despite the popularity of MCMC in data fitting, Hogg & Foreman-Mackey (2018) noted that merely searching the parameter space is not a good motivation for MCMC usage since its primary function is sampling. Motivated by the common usage of MCMC in blazar modeling (see e.g., Yamada et al. 2020; Tramacere 2020; Tzavellas et al. 2024; Sciacaluga et al. 2024; Hervet et al. 2024), we investigated this method in the context of data fitting as well.

In this work, we utilized the affine invariant MCMC ensemble sampler (Goodman & Weare 2010) and the `emcee` package as its Python implementation (Foreman-Mackey et al. 2013). The `emcee` algorithm requires the user to define a log-likelihood function that evaluates the quality of the fit. Based on this function, it autonomously constructs the posterior probability density function, employing a specified number of steps, walkers, and burn-in samples.

Hogg & Foreman-Mackey (2018) recommended selecting a set of sensible parameter values as a starting point, close to the optimum, but not exactly the optimum. Following this approach, we initialized the walkers in the vicinity of the best-fit Minuit result by adding random values that follow the standard normal distribution scaled by 10% of the parameter value to the actual parameter values.

We selected a “flat” (improper) prior for the model parameters with the boundaries of the distributions fixed to the values from Table A.1. As for the log-likelihood function, we first defined it as $\ln(P) = -1/2 \log_{10}(\chi^2)$. The logarithmic scale of χ^2 was chosen to overcome regions of very low probability that may separate the regions with high probability. However, the optimal parameters found in this MCMC run did not yield a satisfactory reduced χ^2 value. We assume that the logarithmic reduced χ^2 had the effect of losing precision near the best-fit solution. To exclude this effect, we repeated the MCMC (by performing run 2) with the new initial point as the previously determined optimal solution (from run 1). To find a precise solution, we completed a second MCMC run with $\ln(P) = -1/2\chi^2$ while limiting the

Table 3. Parameters of the MCMC runs.

Parameter	MCMC run 1	MCMC run 2
Log-likelihood function	$\ln(P) = -\frac{1}{2} \log_{10}(\chi^2)$	$\ln(P) = -\frac{1}{2}\chi^2$
Log-prior	Flat	Flat
Number of walkers	50	50
Number of steps	1000	200
Burn-in samples	100	20
Initial point	From Minuit	From run 1

parameter space to the high-likelihood region identified during the first run. The settings for both MCMC runs are represented in Table 3.

5.3. Algorithm comparison

With regard to the performance of the algorithms, the grid scan offers a decisive advantage as it can effectively show multiple regions with possible good-fit solutions. Since the other algorithms converge to a single best-fit model, the grid scan is useful for considering multiple solutions. It has no controllable algorithm parameters other than the boundaries of the parameter space and the choice of the discretization. Nonetheless, it is also the most computationally expensive of all methods. In the case of more complex models, the number of required generated models (i.e., probed points in the parameter space) grows faster than the exponential result. Additionally, the results depend on the choice of the step size and the boundaries. Some of the regions of interest may be lost between the grid points if the step between the grid points is too large. On the other hand, the grid scan results can be reused for other sources by just recalculating the $\chi^2/\text{n.d.f.}$ for another dataset and followed by a computationally inexpensive local minimization.

The genetic algorithm and CMA-ES, both from the family of evolutionary algorithms, display different levels of performance. While the initial population is generated randomly (not user-defined) for both algorithms, the genetic algorithm has more controllable parameters. The CMA-ES is characterized by self-adaptivity and converges dramatically faster than the genetic algorithm. This can be seen in Figure C.1.

The Minuit approach (utilizing `simplex` and `migrad`) is the least computationally expensive method. It also requires only the parameter intervals and the initial steps to be defined by the user. Despite the need to run Minuit a couple of times in a “nested” way, each time refining the boundaries and increasing the precision, the final best-fit result has a goodness of fit similar to those from more computationally expensive algorithms. The outcome of the Minuit optimization in the parameter space of the one-zone radiation models depends on the choice of the initial point. While an “educated” guess based on the SED features may quickly lead to a good fit, random initialization may lead to local minima or even a failed fit. MCMC provides the posterior distributions of the model parameters, which enables the characterization of the parameter space near the best fit. It requires assumptions on the parameter distribution in the absence of any data (log-prior), number of walkers, and number of steps. Similarly to Minuit, the performance of MCMC in the parameter space of the one-zone leptonic models depends on the choice of the initial point. If it is initialized near the known optimum, MCMC converges fast with the required number of generated models being similar to those from the evolutionary algorithms.

If the selected initial point is far from the optimum, it may require the same or a larger number of generated models as the grid scan for convergence. The comparison of all selected algorithms is also summarized in Table C.1.

5.4. Results

The methods described above were applied to the simulated datasets described above to evaluate and compare their performance. Additionally, t-SNE was used to visualize the location of the found best-fit parameters from all optimization algorithms in the global parameter space. For this purpose, their location, together with the location of the true parameters, was mapped together with the 50 000 lowest reduced χ^2 points from the grid scan. In the t-SNE plots, points that are assigned to $\chi^2/\text{n.d.f.}$ values in the interval of interest $[(\chi^2/\text{n.d.f.})_{\min}, (\chi^2/\text{n.d.f.})_{\min} + 2]$ are highlighted as darker, larger points. The best-fit SEDs and their corresponding t-SNE maps are shown in Figs. 3–5. The best-fit model parameters found by each optimization method are listed in Table 4. In the first row of Table 4, the true parameters are presented for comparison with the results obtained from the five algorithms.

In Fig. 3, we display the results of the SED fitting for dataset 1. As shown in the left panel of Fig. 3, all the selected minimization algorithms found solutions that, overall, describe the data well. The solutions found with CMA-ES and the genetic algorithm have reduced χ^2 values of 1.6 and 1.8, respectively. As can be seen in Table 4, their best-fit parameters agree with the true values within the parameter uncertainty ranges. The grid scan, as well as MCMC and Minuit provide solutions with $\chi^2/\text{n.d.f.}$ of 2.2, 3.5, and 4.1, respectively. As the lower part of the left panel in Fig. 3 shows, they deviate from the true model mostly in the X-ray – MeV region and in the radio band. The right panel of Fig. 3 shows the location of the found best-fit model parameters in the global parameter space. As expected, the solutions from the grid scan, the genetic algorithm, and CMA-ES lie close to the true parameters. The other results from the grid scan (before the local optimization) with low values of reduced χ^2 (marked as larger black dots) highlight only one region. Both the MCMC and Minuit solutions lie outside of this region. This performance issue could be caused by our setup and would likely lead to the same close results provided by the algorithms if the algorithms were run from another initial point or with a smaller step size. Overall, in this overly optimistic case of an almost perfect SED with small uncertainties of the photon flux values, the different minimization algorithms led to similar results (see Table 4). As shown with the grid scan, the only region in the parameter space with low $\chi^2/\text{n.d.f.}$ was the one containing the true solution.

Similarly, Fig. 4 shows the results of the SED fitting for dataset 2. Despite the overall good characterization of the data points with the best-fit model, the random deviations in the IR and optical ranges made the location of the synchrotron peak ambiguous. As shown in the left panel of Fig. 4, all found best-fit solutions predict different photon fluxes in the low-energy range. Additionally, the solutions from the grid scan ($\chi^2/\text{n.d.f.} = 1.6$) and CMA-ES ($\chi^2/\text{n.d.f.} = 1.0$) deviate from the Minuit ($\chi^2/\text{n.d.f.} = 1.4$), MCMC ($\chi^2/\text{n.d.f.} = 1.3$) and genetic algorithm ($\chi^2/\text{n.d.f.} = 1.1$) solutions in the keV–MeV range where the data coverage is poor. As shown in the right panel of Fig. 4, the closest solutions to the true model were found by MCMC, Minuit, and the genetic algorithm. The values of the model parameters in the best-fit solution from the genetic algorithm and MCMC

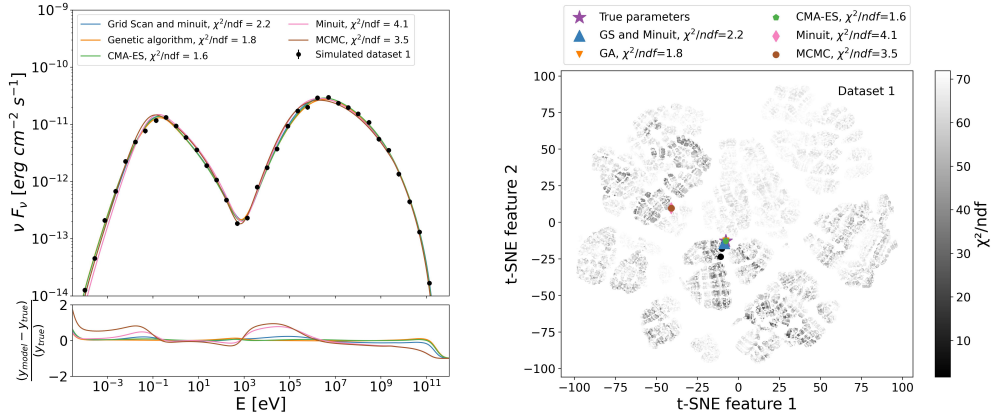


Fig. 3. Results of the SED fitting for dataset 1. Top-left: best-fit results from all selected optimization algorithms. Bottom-left: relative deviation between the true values of νF_ν and those of the best-fit solutions. Right: location of the best-fit solutions in the global parameter space shown in a t-SNE map.

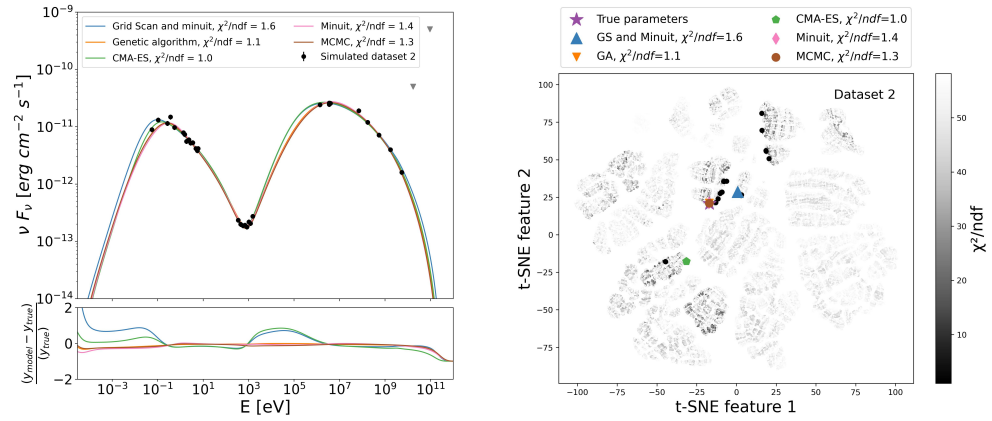


Fig. 4. Same as Fig. 3, but for dataset 2.

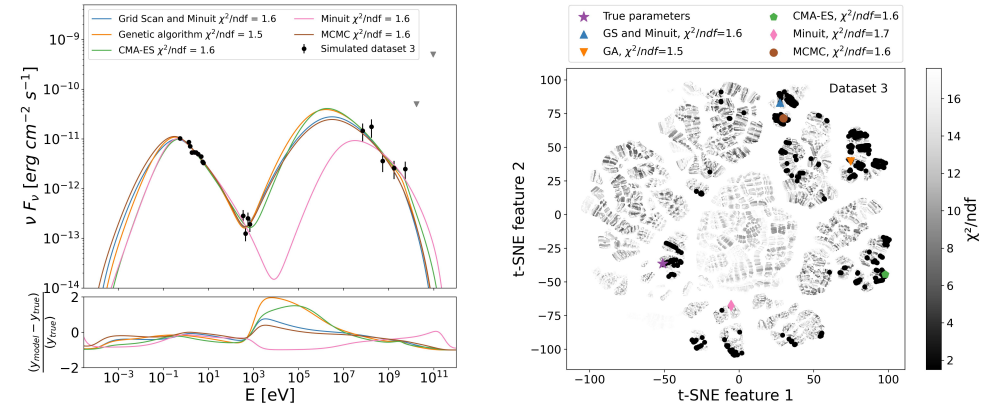


Fig. 5. Same as Fig. 3, but for dataset 3.

are also the closest to the true values when comparing their values with the obtained uncertainties to the true values. Both the CMA-ES and grid scan result in lower values of the electron luminosity compared to the true solution, which is compensated by a higher Lorentz factor and varying values of the blob radius and the magnetic field strength. Interestingly, CMA-ES, which yielded the best goodness of fit among all algorithms, is located the furthest from the true model, with none of the parameters agreeing with the true values within their uncertainties. Apart from three regions in the global parameter space where the algorithms have found their best-fit solutions, the grid scan suggests

additional regions with decent solutions. These are marked by the large black points in the right panel of Fig. 4.

The results of the dataset 3 fitting are presented in Fig. 5. As expected, with fewer data points and a worse characterization of the SED features, the degeneracy of the model parameters increases. The best-fit models, shown in the left panel of Fig. 5, all yielded almost the same values of reduced χ^2 , yet the shapes of the best-fit SEDs vary greatly. Similar to dataset 2, the biggest difference is observed in the low-energy (<0.1 eV) and MeV ranges. Additionally, due to the larger uncertainties and resulting larger scattering of the data points in GeV gamma rays,

Table 4. Best-fit model parameters for simulated data.

Dataset	Method	$\lg(R'_{\text{blob}})$ [cm]	B' [G]	$\lg(\gamma'_{\text{min}})$	$\lg(\gamma'_{\text{max}})$	$\lg(L'_e)$ [erg s ⁻¹]	Γ_b	index α	$\chi^2/\text{n.d.f.}$
True		16.58	0.41	3.30	4.88	43.39	7.8	2.82	–
Dataset 1	Grid scan	16.477±0.001	0.42±0.03	3.234±0.001	4.98±0.03	43.202±0.004	9.2±0.1	2.88±0.06	2.2
	GA	16.60 ^{+0.07} _{-0.06}	0.41±0.08	3.30±0.03	4.98±0.04	43.39 ^{+0.14} _{-0.22}	7.8±0.5	2.85±0.07	1.8
	CMA-ES	16.59±0.01	0.39±0.03	3.34±0.01	5.00±0.01	43.365±0.003	8.00±0.02	2.890±0.003	1.6
	Minuit	16.13±0.05	0.51±0.08	3.05±0.53	4.71±0.10	42.79±0.04	12.3±2.83	2.80±0.04	4.1
	MCMC	16.29 ^{+0.07} _{-0.08}	0.36 ^{+0.07} _{-0.06}	3.04 ^{+0.04} _{-0.03}	4.67 ± 0.03	42.84 ± 0.04	12.37 ^{+0.13} _{-0.17}	2.77 ± 0.06	3.5
Dataset 2	Grid scan	16.50±0.67	0.24±0.09	3.07±0.04	4.88±0.50	43.00±0.70	11.5±0.80	2.82±0.45	1.6
	GA	16.34 ^{+0.10} _{-0.13}	0.54±0.12	3.28 ^{+0.07} _{-0.08}	4.83 ^{+0.11} _{-0.15}	43.20 ^{+0.15} _{-0.22}	8.6±0.85	2.82±0.13	1.1
	CMA-ES	16.11±0.01	0.45±0.01	3.00±0.01	4.66±0.01	42.79±0.01	12.40±0.06	2.74±0.01	1.0
	Minuit	16.310±0.003	0.62±0.02	3.3016±0.0004	4.84±0.02	43.21±0.01	8.474±0.001	2.89±0.01	1.4
	MCMC	16.41 ^{+0.12} _{-0.14}	0.51 ^{+0.17} _{-0.08}	3.31 ^{+0.03} _{-0.05}	4.83 ^{+0.08} _{-0.09}	43.26 ^{+0.09} _{-0.06}	8.21 ^{+0.29} _{-0.32}	2.84 ^{+0.10} _{-0.13}	1.3
Dataset 3	Grid scan	15.88±0.06	1.20±0.01	3.21±0.02	4.39±0.31	43.06±0.11	9.0±0.17	2.83±0.25	1.6
	GA	15.47±1.15	1.32±0.33	3.04 ^{+0.20} _{-0.38}	4.34 ^{+0.25} _{-0.63}	42.63±0.92	12.9±2.3	2.91±0.09	1.5
	CMA-ES	15.000±0.006	1.74±0.08	3.000±0.006	4.99±0.02	42.14±0.04	17.8±0.4	3.34±0.15	1.6
	Minuit	16.81±0.13	0.14±0.02	3.58±0.26	5.50±0.80	42.59±0.33	11.3±0.44	3.25±0.07	1.7
	MCMC	16.21 ^{+0.22} _{-0.20}	0.84 ^{+0.23} _{-0.21}	3.29 ^{+0.09} _{-0.10}	4.55 ^{+0.25} _{-0.27}	43.23 ^{+0.15} _{-0.17}	8.05 ^{+0.92} _{-1.07}	2.84 ^{+0.18} _{-0.30}	1.6

Notes. An explanation of the parameter notations is given in Table 1.

the characterization of the high-energy peak becomes challenging for this dataset. As shown in the right panel of Fig. 5, the best-fit solutions are scattered across the parameter space and none of them are close to the true model. The grid scan results suggest that apart from the found solutions, there are multiple other regions in the parameter space that could produce similarly good solutions. Apart from the worsened characterization of the best-fit location in the global parameter space, larger flux uncertainties and fewer data have led to a worse determination of the minimum. As shown in Table 4, the uncertainties for the model parameters in dataset 3 are (on average) larger than for the previous datasets for all algorithms.

As shown with dataset 2 and dataset 3, the realistic SEDs are explained with degenerate models that result in a similar goodness of fit. With larger flux uncertainties and fewer data points, the degeneracy drastically increases. The models with the lowest $\chi^2/\text{n.d.f.}$ values are not necessarily the closest to the true parameters. It is only in the case of an almost perfect SED, the choice of the optimization algorithm does not affect the outcome.

6. Application to observational data

6.1. Selected sources

As a next step, we studied the performance of the same selected fitting procedures when applied to observational data of two blazars, PKS 0735+178 and Mrk 501. We selected those two sources, because the SED of PKS 0735+178 represents a frequent case of a source with limited available data, while the SED of Mrk 501, one of the most monitored blazars, represents a case of excellent data coverage. Despite good data coverage, fitting this source with a simple one-zone leptonic model is challenging and more complex assumption may be needed as shown in Abdo et al. (2011). However, for the context of this work, we attempted to fit it with the models given in Section 2.

PKS 0735+178 is an intermediate-frequency-synchrotron-peaked (ISP) BL Lac object with an estimated redshift of $z=0.45$

(Nilsson et al. 2012). For this work, we used the gamma-ray quiescent state data from Omeliukh et al. (2025). All the observational data were measured between January 23 and February 2, 2010. Since this source and its leptonic model from Omeliukh et al. (2025) was a prototype for our simulated data, the boundaries for the parameter space remain the same as those presented in Table A.1.

Mrk 501 is a well-known high-frequency synchrotron-peaked (HSP) BL Lac object located at a redshift of $z = 0.034$ (Ulrich et al. 1975). For the modeling, we considered the data taken during the quiescent period of Mrk 501 in 2017–2020 from Abe et al. (2023). As argued in Abe et al. (2023), the measurements in radio and optical indicate low variability in these frequencies and are interpreted to originate in the outer regions of the jet (as opposed to the compact region we intended to model in this work). Therefore, we treated these points as upper limits. Since Mrk 501 is a HSP source, its model parameters are expected to lie in a different region of the parameter space compared to the ISP source PKS 0735+178. As a reference for the typical values of one-zone leptonic model parameters of this source, we used the modeling results of Abe et al. (2023) and limited our parameter space according to them. The defined parameter space boundaries for Mrk 501 are shown in Table A.2. For consistency with Abe et al. (2023), the EBL model based on Franceschini et al. (2008) was assumed.

6.2. Results

6.2.1. PKS 0735+178

The results of the SED fitting for PKS 0735+178 along with the corresponding t-SNE map of the best-fit parameter location are shown in Fig. 6. The values of the best-fit model parameters for each minimization procedure are listed in Table 5. With the exception of the solution found by the MCMC algorithm, all methods yielded fits with $\chi^2/\text{n.d.f.}$ values close to 1, indicating that they provide equally good explanations for the dataset. In

Table 5. Parameters resulting from different parameter search methods for PKS 0735+178 data from 2010.

Method	$\lg(R'_{\text{blob}})$ [cm]	B' [G]	$\lg(\gamma'_{\text{min}})$	$\lg(\gamma'_{\text{max}})$	$\lg(L'_e [\text{erg s}^{-1}])$	Γ_b	index α	$\chi^2/\text{n.d.f.}$
Grid scan	15.27±0.04	1.26±0.37	3.221±0.005	5.17±0.27	42.47±0.42	14.9±2.48	3.27±0.54	0.9
GA	15.54 ^{+0.06} _{-0.07}	1.54±0.10	3.41 ^{+0.04} _{-0.05}	4.96 ^{+0.05} _{-0.06}	42.96 ^{+0.16} _{-0.27}	9.65±0.60	3.30±0.08	1.0
CMA-ES	15.18 ^{+0.03} _{-0.04}	1.24±0.29	3.00±0.02	4.85 ^{+0.08} _{-0.10}	42.21 ^{+0.03} _{-0.04}	18.2±0.46	3.17±0.07	1.1
Minuit	15.082±0.004	1.71±0.01	3.09±0.03	5.15±0.36	42.23±0.03	16.6±0.26	3.26±0.06	1.3
MCMC	15.71 ^{+0.16} _{-0.12}	1.22 ± 0.14	3.36 ± 0.05	4.84 ^{+0.18} _{-0.15}	42.92 ± 0.07	9.65 ^{+0.47} _{-0.35}	3.08 ^{+0.16} _{-0.13}	2.3

Notes. Explanation of the parameter notation is given in Table 1.

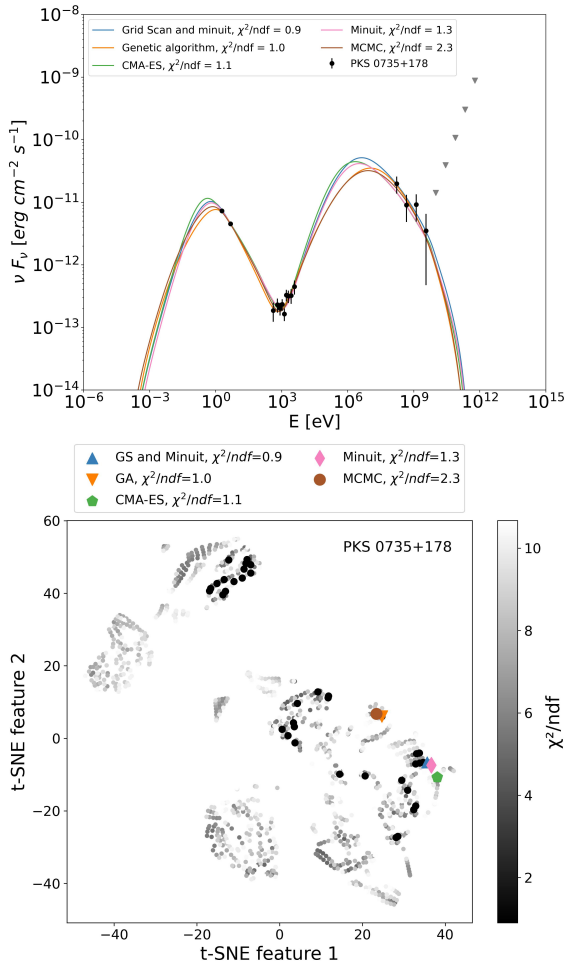


Fig. 6. Results of the SED fitting for PKS 0735+178. Top: best-fit results from all selected optimization algorithms. Bottom: location of best-fit solutions in the global parameter space shown as a t-SNE map.

contrast, MCMC resulted in a solution with $\chi^2/\text{n.d.f.}=2.3$. The higher value is primarily driven by the second data point, which has a small uncertainty. Even small deviations from this point can lead to a large contribution to the $\chi^2/\text{n.d.f.}$ value. Regardless of the goodness-of-fit values, we observe in Fig. 6 that the resulting curves differ in regions where no data are available.

As shown in the lower panel of Fig. 6, the determined model parameters lie in different regions of the parameter space. While the best results from the grid scan ($\chi^2/\text{n.d.f.} = 0.9$), Minuit ($\chi^2/\text{n.d.f.} = 1.3$) and CMA-ES ($\chi^2/\text{n.d.f.} = 1.1$) are closer to each other, results from the MCMC ($\chi^2/\text{n.d.f.} = 2.3$) and genetic algorithm ($\chi^2/\text{n.d.f.} = 1.0$) runs are outside of this cluster. This

can also be seen in the upper panel of Fig. 6, where we observe that the solutions provided by MCMC and the genetic algorithm deviate from the other three models, particularly in the radio and MeV domains, where no data are available to constrain the fits.

The t-SNE map shows that the grid scan, the CMA-ES and the Minuit results lie in those clusters where the grid scan also finds regions of favorable solutions. Interestingly, apart from the found solutions, the grid scan also suggests multiple regions of parameter space that could yield good solutions, highlighting how degenerate the parameter space is for this SED.

The SED data of PKS 0735+178 during its quiescent state is a good example of a degenerate case since the features of the SED, such as the location of the synchrotron or the high-energy peaks as well as the peak flux values, cannot be constrained. This, in turn, leads to great uncertainties in most of the model parameters. Similarly to the simulated dataset 3, almost all solutions have close $\chi^2/\text{n.d.f.}$ values despite their different best-fit model parameter values. The genetic algorithm and MCMC results deviate strongly from the other solutions. They suggest a scenario with a larger blob and higher electron luminosity. Moreover, while the other models infer bulk Lorentz factors up to $\Gamma_b > 14$, the genetic algorithm and MCMC both yield the same lower value of $\Gamma_b = 9.65$.

6.2.2. Mrk 501

The results of the SED fitting for Mrk 501 are shown in the upper panel of Fig. 7, while the lower panel of Fig. 7 shows the t-SNE map with the best-fit locations in the parameter space. The resulting model parameters are listed in Table 6. The SED of Mrk 501 built for the quiescent state has more data points than the SED of PKS 0725+178. Still, interpreting optical and radio data for Mrk 501 as upper limits makes it challenging to constrain the location of the synchrotron peak. In addition, the gap between the low-energy and high-energy peaks falls in the MeV range with no available data.

All of the solutions found in this study fit the X-ray data well. While they also match the high-energy flux levels, this type of model fails to correctly reproduce the shape of the gamma-ray spectrum, which also leads to high absolute values of the reduced χ^2 (> 26). This may be an indication that more complex model assumptions are required for this source. We assumed a simple power-law energy spectrum of the pre-accelerated electrons. As electrons cool due to synchrotron emission, the cooling break appears self-consistently in our modeling. However, [Abdo et al. \(2011\)](#) were able to successfully reproduce the Mrk 501 SED only with the introduction of multiple breaks in the electron spectrum. Alternatively, the parameters that could produce a better fit could lie outside of the considered region in the parameter space (see Table A.2).

Table 6. Parameters resulting from different parameter search methods for the data of Mrk 501.

Method	$\lg(R'_{\text{blob}})$ [cm]	B' [G]	$\lg(\gamma'_{\text{min}})$	$\lg(\gamma'_{\text{max}})$	$\lg(L'_e)$ [erg s $^{-1}$]	Γ_b	index α	$\chi^2/\text{n.d.f.}$
Grid scan	15.522±0.007	0.32±0.03	3.22±0.34	5.256±0.0004	40.63±0.02	12.5±0.02	1.97±0.05	29.7
GA	15.61 $^{+0.07}_{-0.08}$	0.22±0.03	3.14 $^{+0.11}_{-0.15}$	5.38 $^{+0.04}_{-0.05}$	40.65 $^{+0.10}_{-0.12}$	15.1±0.56	2.26±0.06	28.9
CMA-ES	16.08 $^{+0.03}_{-0.04}$	0.15±0.01	2.79 $^{+0.15}_{-0.22}$	5.51±0.02	41.13±0.04	10.2±0.3	2.04 $^{+0.03}_{-0.12}$	28.9
Minuit	15.485±0.005	0.286±0.004	2.378±0.005	5.231±0.005	40.699±0.003	13.44±0.05	1.959±0.002	33.9
MCMC	15.63 ± 0.05	0.22 $^{+0.03}_{-0.02}$	1.97 $^{+0.21}_{-0.13}$	5.33 ± 0.03	40.91 $^{+0.05}_{0.06}$	13.21 $^{+0.20}_{-0.14}$	2.05 $^{+0.03}_{-0.02}$	27.7

Notes. Explanation for parameter notation is given in Table 1.

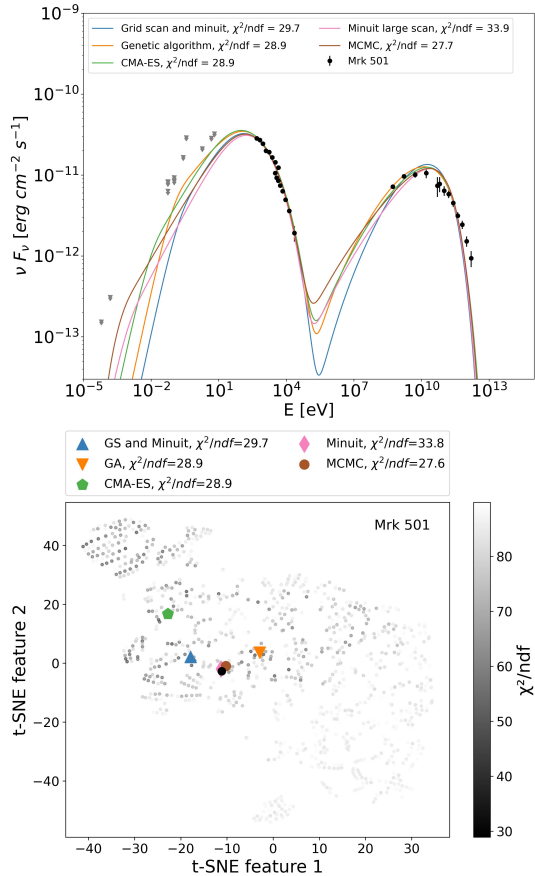


Fig. 7. Results of the SED fitting for Mrk 501. Top: best-fit results from all selected optimization algorithms. Data used for the modeling is represented as black points while the gray triangles are treated as upper limits. The discontinuity visible within the X-ray energy range could be due to the fact that these data were measured by two different instruments. Bottom: location of the best-fit solutions in the global parameter space shown as a t-SNE map.

Even though the obtained values of the reduced χ^2 cannot be considered as a good fit, the upper panel of Fig. 7 shows how different minimization algorithms produced independent solutions. All the best-fit solutions predict a very similar shape of the X-ray fluxes, but diverge in their predictions on the behavior in the dip between the two SED peaks.

The t-SNE plot (lower panel of Fig. 7) reveals that the best-fit parameters found by MCMC ($\chi^2/\text{n.d.f.} = 27.7$) and Minuit ($\chi^2/\text{n.d.f.} = 33.9$) are similar, while the solutions of CMA-ES and the grid scan belong to different families of solutions. Unlike the grid scan results from PKS 0735+178 with multiple regions of low $\chi^2/\text{n.d.f.}$ values, the grid scan results for Mrk 501 display one

prospective solution in the interval of interest (large black marker in the lower panel of Fig. 7, overlapping with the Minuit and MCMC solutions). The corresponding best-fit SEDs, as shown in the upper panel of Fig. 7, agree only in the X-ray and high-energy gamma-ray ranges, significantly deviating in their predictions in all other energy ranges.

Despite the variations in parameters that could not be contained (minimal electron Lorentz factors, bulk Lorentz factor, and electron spectral index), most of the models require a blob radius of $\sim 3 \times 10^{15}$ cm, a magnetic field strength around 0.2–0.3 G, and an electron luminosity of $\sim 5 \times 10^{40}$ erg/s. The most discrepant solution was found by CMA-ES where the blob has a three times larger radius, the required magnetic field strength is lower by a factor of 2, while the bulk Lorentz factor is lower and compensated for by a higher electron luminosity.

Overall, this case is in some respects similar to dataset 2, as there are relatively few models that provide a good fit to the data. The higher number of data points constrains the range of possible solutions, preventing the situation we encountered with dataset 3 or for PKS 0735+178; in that case, for instance, up to 20 distinct solutions were found within the parameter space of interest. However, despite this constraint, the parameter space remains degenerate. Even among the models that successfully reproduce the X-ray data, there are significant differences between the models, particularly in regions where observational data are absent.

7. Discussion

By modeling the simulated and observed SEDs of blazars, we demonstrated the expected issue of parameter degeneracy in the simplest one-zone leptonic radiation models, which was described in Sect. 4. This issue was mentioned in many previous works dedicated to blazar modeling. For instance, Ahnen et al. (2017) discussed the advantages of using a grid scan approach in the context of finding multiple solutions.

Lucchini et al. (2019) addressed the parameter degeneracy of multi-zone models and proposed breaking it by jointly fitting six consecutive blazar states. Many works, especially those utilizing MCMC, show the pairwise distributions of model parameters around the best-fit solutions. We note that the plots with such two-dimensional projections (using marginal distributions or some parameters fixed to their best-fit values) cannot capture any other significantly different solutions. In this work, we tackled this problem and showed how physically different solutions and their proximity in the parameter space can be compared using t-SNE. However, as already discussed in Section 3, a careful choice of hyperparameters such as perplexity is required.

In Section 5, we describe how we applied the grid scan, the genetic algorithm, CMA-ES, MCMC, and Minuit to simulated datasets to estimate how close the best-fit parameters found by

each method are to the true parameters. Dataset 1 represents the case of perfect data with no gaps and small uncertainties. Both evolutionary algorithms accurately identified the best-fit parameters that were consistent with the true parameter values within the uncertainty interval. The grid scan followed by a local minimization delivered best-fit parameters that were slightly off from the true parameters compared to the results of the evolutionary algorithms. We find it is likely that the true parameters could have been achieved with an improved grid discretization or local minimization within slightly broader boundaries and more function calls. Both Minuit and MCMC stopped around the local minima. Given the complexity of finding only one solution in an extremely vast parameter space, these algorithms are likely to require a larger number of probed points in the parameter space than we allocated in the present study. On the contrary, in more degenerative cases (dataset 2 and dataset 3), these methods displayed the ability to find good solutions thanks to the high probability of finding such a solution, even when probing a small number of points, covering a limited parameter space region. We note that nested sampling, such as the MultiNest algorithm used in Bégué et al. (2024) and Sahakyan et al. (2024), might be a more efficient sampling approach.

In our setup, with ten probed points per parameter, the grid scan turned out to be the most computationally expensive method. However, it is the only known approach that enables the exploration of multiple physically different solutions. It also allowed us to visualize the parameter space with t-SNE due to comprehensive parameter space coverage. While our SSC one-zone model is the simplest model for explaining a blazar SED (requiring only seven free parameters) more complex models with a higher number of free parameters would require drastically higher computational resources. However, the possibility of reusing its results for other sources or datasets offers an advantage. A known limitation of the grid scan is that solutions might fall between predefined grid points. Some studies (e.g., Ahnen et al. 2017) address this by performing additional, finer grid scans on top of the initial large grid. While this approach can improve the accuracy, it comes at the cost of increased computational time. In our approach, the second step was changed to a local minimization with Minuit.

The results showed how accurately the evolutionary algorithms captured the true solution. Both algorithms performed similarly, with CMA-ES converging much faster and offering the additional advantage of self-adapting its internal parameters. Both algorithms converge to only one solution. While it is possible to extract all the solutions from each generation to potentially identify multiple solutions, this does not offer the same coverage as the grid scan. Evolutionary algorithms explore the parameter space based on evaluations, which can lead to certain regions being excluded from further consideration as the algorithms progressively focus on specific regions with each generation. This behavior is also reflected in the parameter uncertainties. We calculated the standard deviation of the parameters from the final generation, which yielded relatively small uncertainties, as the solutions in the final generation had already converged closely towards each other. We note, however, that there is no guarantee that the results of the minimization with the evolutionary algorithms would be the same if the minimization was to be repeated with the same conditions. This is due to the random initialization of the initial population and the highly irregular structure of the parameter space.

Apart from the case of a perfect dataset, the results from our analysis of the parameter space of leptonic radiation models suggest that regardless of the method used, the complex

nature of the parameter space with multiple optima can prevent any algorithm from capturing the true values. In particular, for the models found with algorithms converging to a single solution, it can be challenging to accurately explain the emission of blazars. Ultimately, the best approach is to remain aware of the existence of multiple solutions and use methods such as the grid scan (either a nested grid scan or combined with a local minimization), as it can represent the parameter space in the most comprehensive way, rather than providing only a single solution.

For the case of the observational data, we modeled the radiation of PKS 0735+178 during a quiescent state with each of the five minimization algorithms, leading to different solutions. This state was modeled before in Bharathan et al. (2024) and Omeliukh et al. (2025). Omeliukh et al. (2025) addressed the issue of multiple possible solutions by performing a grid scan and selecting two physically different solutions (i.e., slow and fast solutions based on the bulk Lorentz factor). Our best-fit models from all minimization algorithms have a two to three times higher magnetic field strength and a much lower emission zone radius compared to both the slow and fast models for this state in Omeliukh et al. (2025). Similarly, in our solutions, the magnetic field strength is twice as high and the blob radius is one order of magnitude lower than that in Bharathan et al. (2024). This indicates that all five best-fit solutions provide new models in addition to those previously reported in the literature.

For Mrk 501, our results reproduce the overall flux levels but fail to match the gamma-ray spectral shape. The parameter boundaries were based on Abe et al. (2023) and the same dataset was used. While Abe et al. (2023) fixed the blob size to $R'_{\text{blob}} = 10^{17.06}$ cm, our solutions suggest significantly smaller emission regions. The magnetic field strengths obtained from our searches were around $B' = 0.15\text{--}0.32$ G, whereas they suggested a field strength much lower with $B' = 0.025$ G. The bulk Lorentz factor was fixed to 11 in their work, which is comparable to our range of $\Gamma_b = 10.2\text{--}15.1$. The required electron luminosity was also 10^3 times lower in our case. Despite adopting the parameter space boundaries from Abe et al. (2023), our results differ considerably. Abdo et al. (2011) used a double broken power law to successfully reproduce the SED of Mrk 501 in the quiescent state, which may be the reason why our simple power law (with naturally occurring cooling break) failed to reproduce the gamma-ray spectral shape.

Our findings, along with comparisons to other models, highlight the highly degenerate parameter space of the leptonic radiation models. Even the simplest SSC one-zone framework that has seven free parameters presents significant challenges in explaining the blazar emission. More complex models add additional parameters, further complicating the parameter space. This emphasizes the importance of careful interpretation of modeling results. To break this degeneracy, high-quality quasi-simultaneous multiwavelength data are essential. In this sense, blazar surveys and monitoring programs are crucial to constrain the possible parameter values of the one-zone models. Additionally, exploring new energy ranges, such as MeV or TeV gamma rays, would significantly constrain the parameter space. For example, a single MeV data point added to our simulated dataset 3 would reduce the number of equally well-fitting models from five to one or two. Therefore, future MeV missions such as e-ASTROGAM (De Angelis et al. 2017), COSI (Tomsick et al. 2019) or AMEGO-X (Caputo et al. 2022) would play an important role in breaking the model degeneracy. A similar effect is expected by adding data points in the TeV gamma-ray range, especially for HSP sources. The future Cherenkov

Telescope Array Observatory (CTAO, [Cherenkov Telescope Array Consortium 2019](#)) is expected to provide high-energy gamma-ray data, which will be essential for constraining the models. In the case of more complex leptohadronic models, multiwavelength polarization could potentially constrain the hadronic component ([Zhang et al. 2024](#)).

8. Summary and conclusions

In this work, we studied the fitting procedures of blazar SEDs in the context of one-zone leptonic models. We found convincing evidence that the goodness of fit is not a smooth and convex function of the model parameters, due to the nature of the underlying radiation processes and parameter degeneracies. Using the simulated pseudo-data, we observed that the degeneracies arise due to the missing data in certain energy ranges. An increase in the flux measurement uncertainty seems to further enhance these degeneracies.

We show that for a typical blazar SED with data covering the optical, X-ray, and GeV gamma-ray ranges, the choice of the fitting procedure, in particular the choice of the minimizing algorithm, leads to considerably different results. In most cases, the five tested minimizing algorithms found physically different best-fit solutions.

We applied the same fitting procedures to the observed SEDs of two blazars, PKS 0735+178 and Mrk 501. The model parameters could not be constrained for PKS 0735+178 due to unconstrained SED features such as the location of the synchrotron and the high-energy peak and the corresponding peak flux values. While the results for Mrk 501 were more consistent due to the better-characterized high-energy peak, the results of the different minimization algorithms were still degenerate, yet none of them yielded a satisfactory fit.

The degeneracy is expected to worsen further when adding extra parameters to the models as in the case of external radiation field models, for instance, or when adding hadrons to the emission zone. This creates a challenge for gamma-ray and neutrino astronomy as retrieving source properties from the SED modeling becomes ambiguous.

To reduce the parameter degeneracy it is crucial to assure a complete and simultaneous wavelength coverage of the data. Blazar monitoring programs as well as measurements in new energy ranges (e.g., MeV or very-high-energy gamma rays) could significantly improve the constraints on the model parameter space. Alternatively, taking into account multiwavelength polarization provides another approach to break the degeneracy in lepto-hadronic models.

Acknowledgements. F.A., A.F., and A.O. acknowledge the support from the DFG via the Collaborative Research Center SFB1491 *Cosmic Interacting Matters – From Source to Signal*. A.O. was supported by DAAD funding program 57552340. F.A. acknowledges support by the Helmholtz Weizmann Research School on Multimessenger Astronomy.

References

- Abdo, A. A., Ackermann, M., Ajello, M., et al. 2011, *ApJ*, **727**, 129
- Abdollahi, S., Acero, F., Ackermann, M., et al. 2020, *ApJS*, **247**, 33
- Abe, H., Abe, S., Acciari, V. A., et al. 2023, *ApJS*, **266**, 37
- Ahnen, M. L., Ansoldi, S., Antonelli, L. A., et al. 2017, *A&A*, **603**, A31
- Ajello, M., Gasparrini, D., Sánchez-Conde, M., et al. 2015, *ApJ*, **800**, L27
- Atwood, W. B., Abdo, A. A., Ackermann, M., et al. 2009, *ApJ*, **697**, 1071
- Bégué, D., Sahakyan, N., Dereli-Bégué, H., et al. 2024, *ApJ*, **963**, 71
- Bharathan, A. M., Stalin, C. S., Sahayanathan, S., Bhattacharyya, S., & Mathew, B. 2024, *MNRAS*, **529**, 3503
- Bose, D., Chitnis, V. R., Majumdar, P., & Shukla, A. 2022, *EPJ ST*, **231**, 27
- Böttcher, M., Reimer, A., Sweeney, K., & Prakash, A. 2013, *ApJ*, **768**, 54
- Burrows, D. N., Hill, J. E., Nousek, J. A., et al. 2004, *Proc. SPIE*, **5165**, 201
- Caputo, R., Ajello, M., Kierans, C. A., et al. 2022, *J. Astron. Telesc. Instrum. Syst.*, **8**, 044003
- Cerruti, M., Dermer, C. D., Lott, B., Boisson, C., & Zech, A. 2013, *ApJ*, **771**, L4
- Cerruti, M., Zech, A., Boisson, C., & Inoue, S. 2015, *MNRAS*, **448**, 910
- Cherenkov Telescope Array Consortium (Acharya, B. S., et al.) 2019, *Science with the Cherenkov Telescope Array* (World Scientific)
- De Angelis, A., Tatischeff, V., Tavani, M., et al. 2017, *Exp. Astron.*, **44**, 25
- Dembinski, H., Ongmongkolkul, P., et al. 2020, <https://doi.org/10.5281/zenodo.3949207>
- Dermer, C. D., Cerruti, M., Lott, B., Boisson, C., & Zech, A. 2014, *ApJ*, **782**, 82
- Di Mauro, M., Cuoco, A., Donato, F., & Siegal-Gaskins, J. M. 2014, *J. Cosmology Astropart. Phys.*, **2014**, 021
- Domínguez, A., Primack, J. R., Rosario, D. J., et al. 2011, *MNRAS*, **410**, 2556
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, **125**, 306
- Fortin, F.-A., De Rainville, F.-M., Gardner, M.-A., Parizeau, M., & Gagné, C. 2012, *JMLR*, **13**, 2171
- Franceschini, A., Rodighiero, G., & Vaccari, M. 2008, *A&A*, **487**, 837
- Frankowiak, A., Garrappa, S., Paliya, V., et al. 2020, *ApJ*, **893**, 162
- Gao, S., Pohl, M., & Winter, W. 2017, *ApJ*, **843**, 109
- Garrappa, S., Buson, S., Frankowiak, A., et al. 2019, *ApJ*, **880**, 103
- Gasparyan, S., Bégué, D., & Sahakyan, N. 2022, *MNRAS*, **509**, 2102
- Giommi, P., & Padovani, P. 2021, *Universe*, **7**, 492
- Goodman, J., & Weare, J. 2010, *CAMCoS*, **5**, 65
- Hansen, N. 2016, arXiv e-prints [arXiv:1604.00772]
- Hansen, N., & Ostermeier, A. 2001, *Evol. Comput.*, **9**, 159
- Hansen, N., Akimoto, Y., & Baudis, P. 2019, <https://doi.org/10.5281/zenodo.2559634>
- Harrison, F. A., Craig, W. W., Christensen, F. E., et al. 2013, *ApJ*, **770**, 103
- Hervet, O., Johnson, C. A., & Youngquist, A. 2024, *ApJ*, **962**, 140
- Hogg, D. W., & Foreman-Mackey, D. 2018, *ApJS*, **236**, 11
- IceCube Collaboration (Aartsen, M. G., et al.) 2018a, *Science*, **361**, eaat1378
- IceCube Collaboration (Aartsen, M. G., et al.) 2018b, *Science*, **361**, 147
- James, F. 1998, *MINUIT: function minimization and error analysis reference Manual*, Tech. rep., CERN
- Kadler, M., Krauß, F., Mannheim, K., et al. 2016, *Nat. Phys.*, **12**, 807
- Keivani, A., Murase, K., Petropoulou, M., et al. 2018, *ApJ*, **864**, 84
- Klinger, M., Rudolph, A., Rodrigues, X., et al. 2024, *ApJS*, **275**, 4
- Kramer, O., & Kramer, O. 2017, *Genetic Algorithms* (Springer)
- Lioudakis, I., & Petropoulou, M. 2020, *ApJ*, **893**, L20
- Lucchini, M., Markoff, S., Crumley, P., Krauß, F., & Connors, R. 2019, *MNRAS*, **482**, 4798
- Nelder, J. A., & Mead, R. 1965, *Comput. J.*, **7**, 308
- Nilsson, K., Pursimo, T., Villforth, C., et al. 2012, *A&A*, **547**, A1
- Omelikh, A., Garrappa, S., Fallah Ramazani, V., et al. 2025, *A&A*, **695**, A266
- Paliya, V. S. 2015, *ApJ*, **804**, 74
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *JMLR*, **12**, 2825
- Petropoulou, M., & Mastichiadis, A. 2015, *MNRAS*, **447**, 36
- Petropoulou, M., Lefa, E., Dimitrakoudis, S., & Mastichiadis, A. 2014, *A&A*, **562**, A12
- Petropoulou, M., Dimitrakoudis, S., Padovani, P., Mastichiadis, A., & Resconi, E. 2015, *MNRAS*, **448**, 2412
- Petropoulou, M., Murase, K., Santander, M., et al. 2020, *ApJ*, **891**, 115
- Rodrigues, X., Garrappa, S., Gao, S., et al. 2021, *ApJ*, **912**, 54
- Rodrigues, X., Karl, M., Padovani, P., et al. 2024a, *A&A*, **689**, A147
- Rodrigues, X., Paliya, V. S., Garrappa, S., et al. 2024b, *A&A*, **681**, A119
- Roming, P. W. A., Kennedy, T. E., Mason, K. O., et al. 2005, *Space Sci. Rev.*, **120**, 95
- Sahakyan, N., Giommi, P., Padovani, P., et al. 2023, *MNRAS*, **519**, 1396
- Sahakyan, N., Bégué, D., Casotto, A., et al. 2024, *ApJ*, **971**, 70
- Schoenfelder, V., Aarts, H., Bennett, K., et al. 1993, *ApJS*, **86**, 657
- Sciaccaluga, A., Tavecchio, F., Landoni, M., & Costa, A. 2024, *A&A*, **687**, A247
- Stathopoulos, S. I., Petropoulou, M., Vasilopoulos, G., & Mastichiadis, A. 2024, *A&A*, **683**, A225
- Tomsick, J., Zoglauer, A., Sleator, C., et al. 2019, in *Bulletin of the American Astronomical Society*, **51**, 98
- Tramacere, A. 2020, JetSeT: Numerical modeling and SED fitting tool for relativistic jets, Astrophysics Source Code Library [record ascl:2009.001]
- Tzavellas, A., Vasilopoulos, G., Petropoulou, M., Mastichiadis, A., & Stathopoulos, S. I. 2024, *A&A*, **683**, A185
- Ulrich, M. H., Kinman, T. D., Lynds, C. R., Rieke, G. H., & Ekers, R. D. 1975, *ApJ*, **198**, 261
- Urry, C. M., & Padovani, P. 1995, *PASP*, **107**, 803
- van der Maaten, L., & Hinton, G. 2008, *JMLR*, **9**
- Yamada, Y., Uemura, M., Itoh, R., et al. 2020, *PASJ*, **72**, 42
- Zdziarski, A. A., & Böttcher, M. 2015, *MNRAS*, **450**, L21
- Zhang, H., Böttcher, M., & Lioudakis, I. 2024, *ApJ*, **967**, 93

Appendix A: Boundaries of the parameter space

Table A.1: Boundaries of the leptonic model parameters used in fitting simulated data.

Parameter	Value range
$\log_{10}(R'_{\text{blob}})$, [cm]	[15.0, 17.5]
B' , [G]	[0.1, 5]
Γ_b	[3.0, 30.0]
$\log_{10}(\gamma_e^{\text{min}})$	[3.0, 4.0]
$\log_{10}(\gamma_e^{\text{max}})$	[4.0, 5.0]
α_e	[0.5, 3.5]
$\log_{10}(L'_e)$, [erg s ⁻¹]	[42.0, 47.0]

Notes. Parameter description: R'_b is the radius of the blob, B' is the magnetic field strength in the emission region, Γ_b is blob Lorentz factor; γ_e^{min} and γ_e^{max} are the minimum and maximum Lorentz factor of the electrons respectively, α_e is the spectral index of electrons, and L'_e is the electron luminosity.

Table A.2: Leptonic model parameters for Mrk 501.

Parameter	Value range
R'_{blob} , [cm]	[10 ^{15.5} , 10 ¹⁸]
B' , [G]	[10 ⁻³ , 0.5]
Γ_b	[5.0, 30.0]
γ_e^{min}	[10 ² , 10 ⁵]
γ_e^{max}	[10 ⁵ , 10 ⁷]
α_e	[1.0, 3.0]
L'_e / [erg s ⁻¹]	[10 ³⁹ , 10 ⁴⁴]

Notes. See Table A.1 for parameter description.

Appendix B: Perplexity and learning rate for t-SNE

As mentioned in Section 3, the hyperparameters for the application of t-SNE have to be selected carefully. The perplexity can be interpreted as the number of nearest neighbours for each point and has a significant impact on the result presentation. Another hyperparameter is the learning rate, which defines the step size for the gradient descent method in the Kullback-Leibler divergence minimization.

We selected four different leptonic models (model parameters are shown in Table B.1) and studied which hyperparameters lead to an adequate representation of the model parameter proximity in the parameter space. In the t-SNE plot, the first best solutions with a $\chi^2/\text{n.d.f.} < 9$ are marked as black points. These solutions are similar to the true solution. One of them is the best solution resulting from the grid scan and is represented in the second row of Table B.1. For this reason, t-SNE should locate the true solution in the same cluster as the marked points since we searched for a representation that would classify these points as similar. The third parameter set in the table has different parameters and should be placed in another cluster. The last dataset also has different parameters and belongs to the sets with the highest $\chi^2/\text{n.d.f.}$. Fig. B.1 and Fig. B.2 show different cases of perplexity and learning rate. While the recommendation for the perplexity is between five and 50, the learning rate is supposed to be in the range [10.0, 1000.0]. We first tested different cases of perplexity with a fixed learning rate of 300. Figure B.1 shows the t-SNE plot for perplexity values $p= 10, 30$ and 80. The left plot shows the

true solution in the same cluster as the best solutions and the different models are located in other clusters. However, it is rather difficult to distinguish different clusters and everything seems to be merged together. The perplexity of 30 seems to be perfect since the different parameter sets are arranged as we intended. Clear clusters are visible which makes it possible to differentiate between different regions. The third plot shows the case of a perplexity $p = 80$. The true solution is located in another cluster than the best solutions which gives the impression that these models could be very different. Therefore, we decided to select a perplexity of 30. If the learning rate is too low, the algorithm might get stuck in a local minimum instead of finding the global minimum of the Kullback-Leibler divergence. With a learning rate that is too high, the oscillations might be too high, and the algorithm could miss the global minimum. With a low learning rate of 10, we can notice a lack of structure and poorly separated points and clusters. In addition, the true solution is outside the cluster with the best solutions. The last two plots with a learning rate of 300 and 700 look almost similar. With a learning rate of 700, the true solution does not lie exactly on the best solutions and it indicates that they are similar but not identical. Due to that, we decided to select the learning rate of 700. Figure B.2 show three plots with different learning rates and a fixed perplexity of 30.

Appendix C: Comparison of optimization algorithms

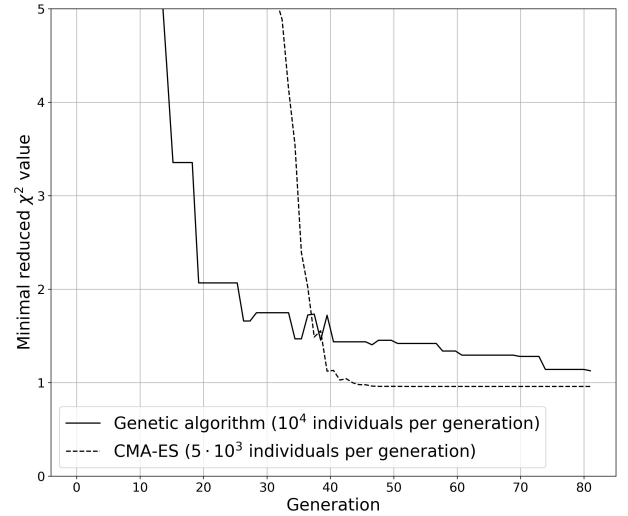


Fig. C.1: Comparison of the convergence between the genetic algorithm and the CMA-ES. The plot shows the smallest reduced χ^2 of each generation depending on the number of the generation.

Table B.1: Parameters of true solution, best grid scan result, Model 1 and Model 2 in this order.

$\lg(R'_{\text{blob}})$ [cm]	B' [G]	$\lg(\gamma'_{\text{min}})$	$\lg(\gamma'_{\text{max}})$	$\lg(L'_c [\text{erg s}^{-1}])$	Γ_b	index α
16.58	0.41	3.30	4.88	43.39	7.79	2.82
16.39	0.64	3.21	4.78	43.11	9.00	2.83
16.67	3.37	3.63	4.00	44.22	3.00	1.83
16.11	0.10	3.00	4.67	42.56	27.00	3.50

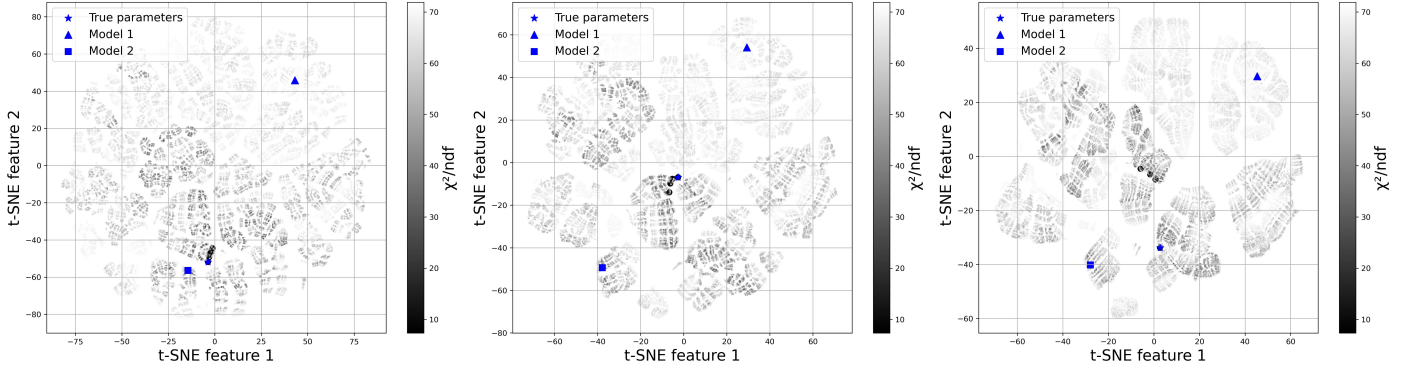


Fig. B.1: Three t-SNE plots with different perplexity values. Model 1 and Model 2 represent models that are significantly different from the true solution. The black dots are solutions with small $\chi^2/\text{n.d.f.}$ values that are close to the true parameters. The plots show the cases of perplexity values $p=10, 30$ and 80 in this order.

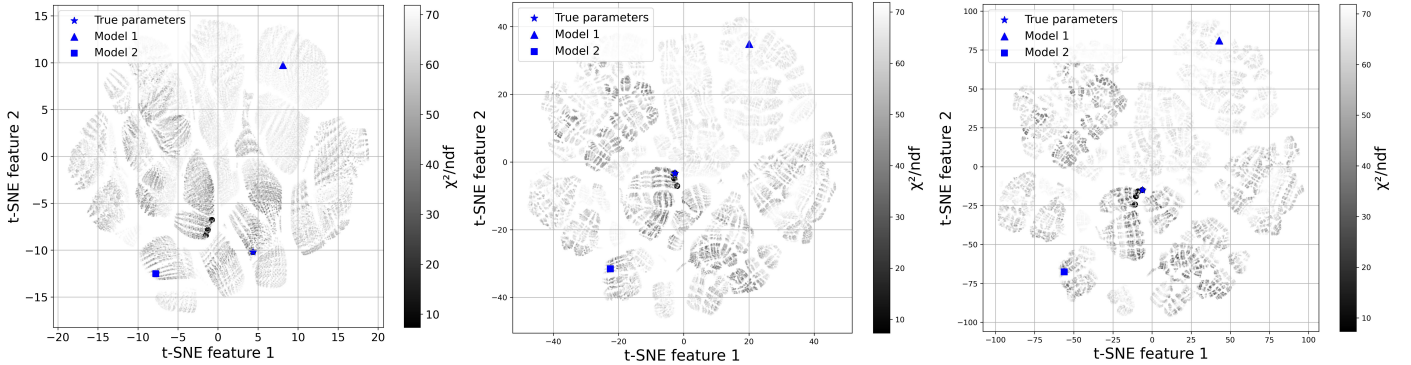


Fig. B.2: Plots of the impact of different learning rates. The perplexity is $p = 30$ for all cases. The learning rate is $\eta = 10, 100,$ and 700 consecutively.

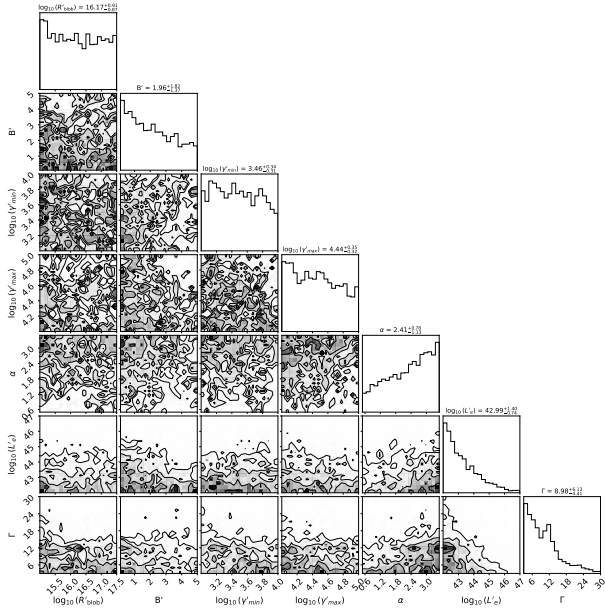
Table C.1: Overview and comparison of different algorithms

Algorithm	N_{models}	Advantage	Disadvantage
Grid scan with local minimization	10^7	Considers more than only one solution, can be reused for other sources, few controllable parameters	Computationally expensive, drastic increase in computing cost with the increase of model parameters
Genetic algorithm	10^5 – 10^6	Able to cover large part of parameter space, overcomes local minima	Only one solution, sensitive to setting of algorithm parameters
CMA-ES	10^5	Able to cover large part of parameter space, overcomes local minima, self-adaption of algorithm parameters	Only one solution
Minuit (simplex + migrad)	10^3	Computationally inexpensive, few controllable parameters	May be sensitive to local minima, only one solution, sensitive to the choice of initial point
MCMC (ensemble sampler)	10^5	Provides parameter distributions	Sensitive to the choice of initial point and algorithm parameters

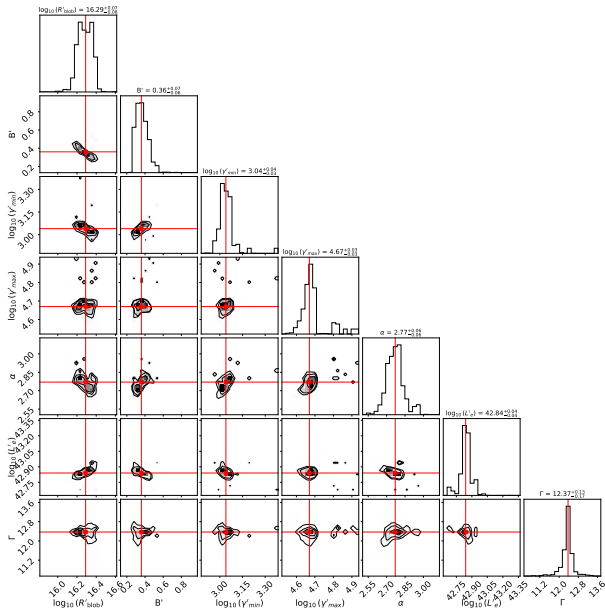
Notes. N_{models} is the total number of generated models in the selected algorithm.

Appendix D: MCMC corner plots

Figures B.1–B.5 show the two-dimensional projections of the posterior probability distributions of the model parameters. The MCMC parameter search was completed in two steps, as described in Sec. 5. The first step (run 1) corresponded to the global sampling within the boundaries defined for each SED. These boundaries were refined based on the results of the first step and the sampling was repeated (run 2). As shown in the plots for run 1 for all datasets, the selected number of walkers and steps was not enough for the algorithm to converge to one solution in the global parameter space. However, in the limited region, the convergence was achieved and the solution was found as demonstrated with the plots for run 2.

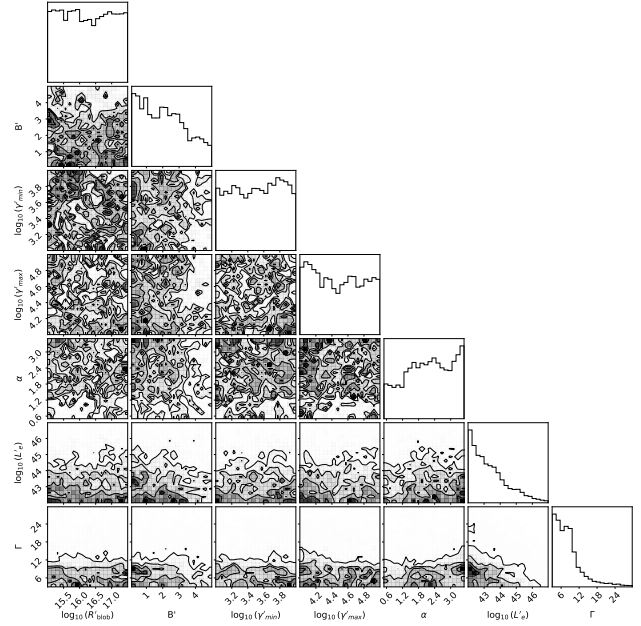


(a) MCMC run 1.

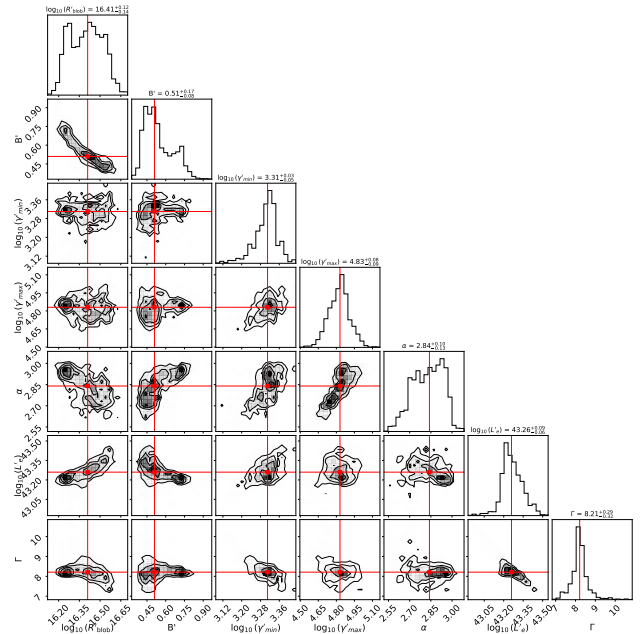


(b) MCMC run 2.

Fig. B.1: Corner plot for simulated dataset 1.

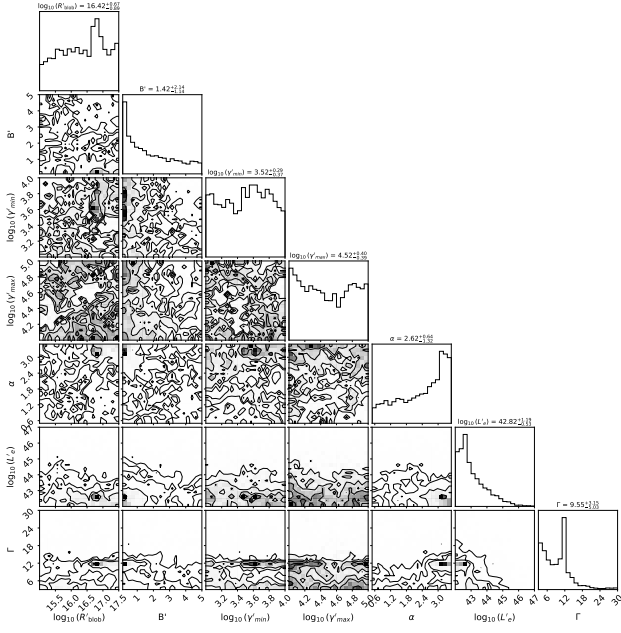


(a) MCMC run 1.

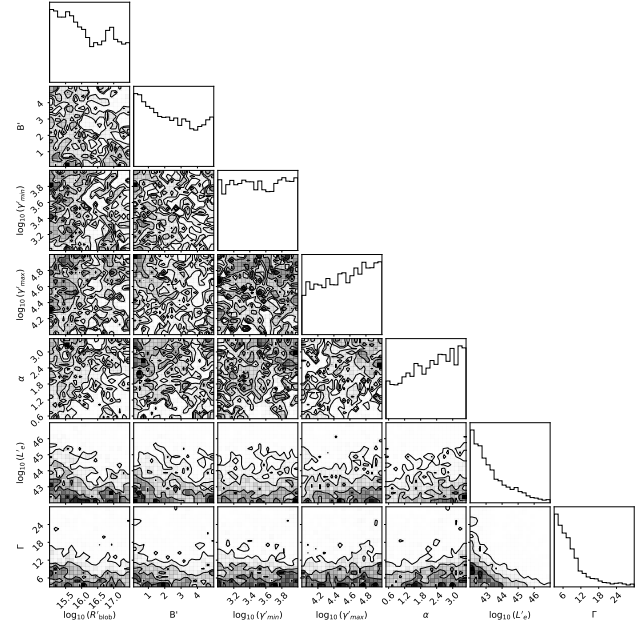


(b) MCMC run 2.

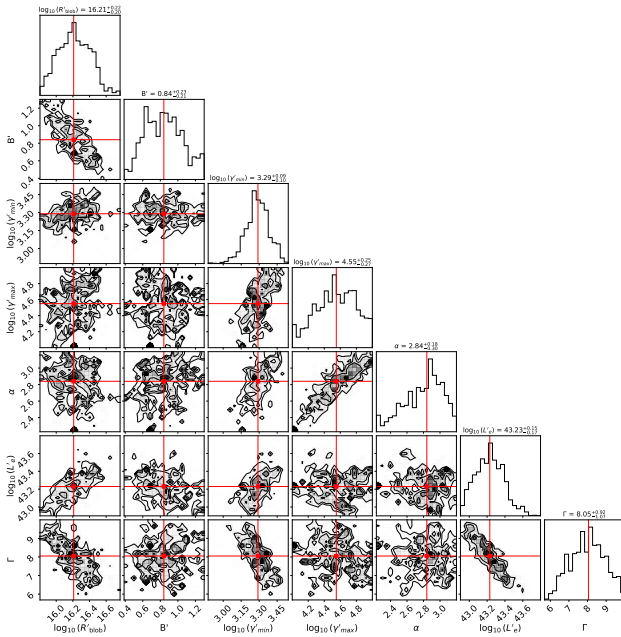
Fig. B.2: Corner plot for simulated dataset 2.



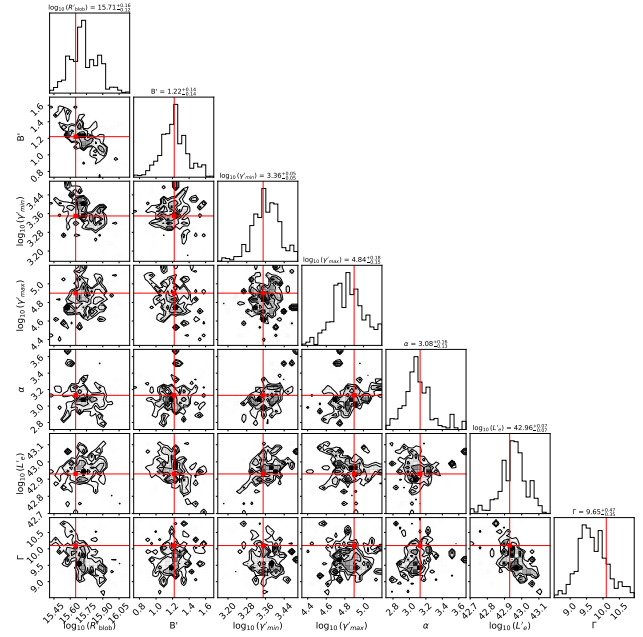
(a) MCMC run 1.



(a) MCMC run 1.



(b) MCMC run 2.



(b) MCMC run 2.

Fig. B.3: Corner plot for simulated dataset 3.

Fig. B.4: Corner plots for PKS 0735+178.

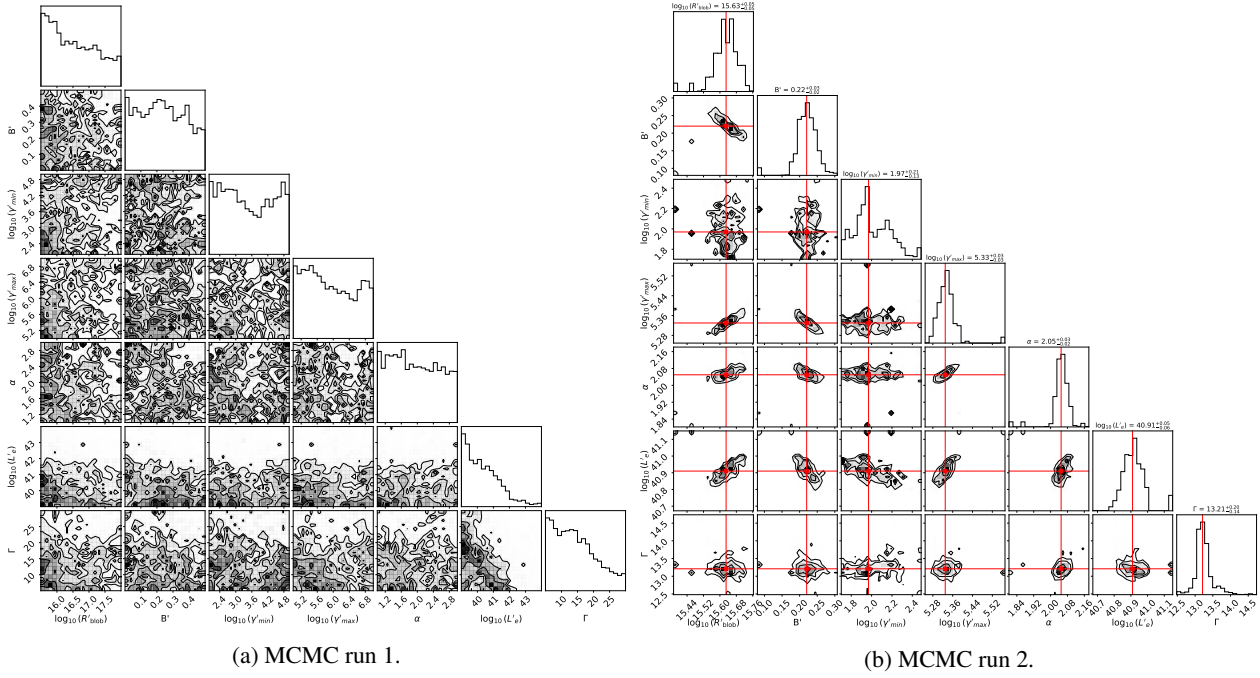


Fig. B.5: Corner plot for Mrk 501.