

MCMC estimation of SED model parameters using multi-wavelength data of the blazar Mrk 421

Yurika Yamada¹, Yasushi Fukazawa, Yasuyuki Tanaka

Department of physics, Hiroshima University

1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan

E-mail: yurika@astro.hiroshima-u.ac.jp

Makoto Uemura

Hiroshima Astrophysical Science Center, Hiroshima University

1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan

Ryosuke Itoh

Tokyo Institute of Technology

2-12-1 Ookayama, Meguro, Tokyo 152-8552, Japan

The spectral energy distributions (SEDs) of blazars are dominated by synchrotron and inverse Compton radiation. The origin of blazar variability can be investigated from the time variation of the SEDs. In the present study, we used Markov chain Monte Carlo (MCMC) method to estimate the optimal model parameters of SEDs and their uncertainties. Our experiments using artificial data demonstrate that at least one prior probability is required to uniquely determine the solution. We used simultaneous observations of Mrk 421 with Fermi-LAT, Swift-XRT, and the 1.5-m optical telescope, Kanata from 2009 to 2011 reported in Itoh et al. (2015). We succeeded in estimating the optimal parameters and their uncertainties by using the prior probability of the variation time-scale and Doppler factor. We found that the power-law index of the electron energy distribution shows a clear positive correlation with the X-ray flux even if the uncertainties are considered.

*7th Fermi Symposium 2017
15-20 October 2017
Garmisch-Partenkirchen, Germany*

¹Speaker

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

<https://pos.sissa.it/>

1. Introduction

Blazars are a type of active galactic nuclei (AGN) whose jets are directed toward us. They are observed in a wide range of wavelengths from radio to gamma-rays. Therefore, Multi-wavelength studies have been actively performed (e.g. Abdo et al. 2011). We can estimate the physical quantity in the jets, such as the magnetic field and electron energy distribution by modeling the observed spectral energy distributions (SEDs). The number of parameters of SED models is usually $\lesssim 10$. It is difficult to estimate the optimal parameters and their uncertainties because some parameters are strongly correlated and degenerate. Several parameters are usually fixed to estimate the other parameters (e.g. M.Böttcher et al. 2013, Itoh et al. 2015). However, in this case, we cannot estimate the uncertainties of the parameters because they depend on the fixed parameters, and their errors are not considered. In addition, the standard method, like the maximum likelihood method, may be trapped by local solutions. Hence, we need a method to find a global solution of the model parameters and their uncertainties with minimum assumptions. In this study, we use the Markov chain Monte Carlo (MCMC) method to estimate the posterior probability of the parameters. It enables us to obtain the optimal parameters and their uncertainties from the probability distributions.

2. Model and Method

2.1 The Synchrotron Self–Compton (SSC) model

In this article, we use the synchrotron self-Compton model for the observed SEDs. This model assumes the situation that high energy electrons emitting synchrotron photons upscatter the photons by the inverse Compton scattering. We used the equations of synchrotron radiation, $f_{\epsilon}^{\text{syn}}$, and inverse Compton scattering, $f_{\epsilon_s}^{\text{SSC}}$, given by Finke et al. (2008), expressed as:

$$f_{\epsilon}^{\text{syn}} = \frac{\sqrt{3}\delta_D^4 \epsilon' e^3 B}{4\pi h d_L^2} \int_1^{\infty} d\gamma' N_e'(\gamma') R(x), \quad (1)$$

$$f_{\epsilon_s}^{\text{SSC}} = \frac{9}{16} \frac{(1+z)^2 \sigma_T \epsilon_s'^2}{\pi \delta_D^2 c^2 T^2} \int_0^{\infty} d\epsilon' \frac{f_{\epsilon}^{\text{syn}}}{\epsilon'^3} \times \int_{\gamma'_{\text{max}}}^{\gamma'_{\text{min}}} d\gamma' \frac{N_e'(\gamma')}{\gamma'^2} F_C(q, \Gamma). \quad (2)$$

In those equations, the parameters which can be estimated from the SED data are the magnetic field, B , the Doppler factor, δ_D , the variation time scale, T , the minimum Lorentz factor, γ_{min} , the maximum Lorentz factor, γ_{max} , and the electron energy distribution, N_e . We used N_e as a broken power law, as follows:

$$N_e(\gamma) = K_e \times \begin{cases} \left(\frac{\gamma}{\gamma_b}\right)^{-p_0}, & \gamma < \gamma_b, \\ \left(\frac{\gamma}{\gamma_b}\right)^{-p_1}, & \text{otherwise,} \end{cases} \quad (3)$$

where K_e is the electron normalization factor, γ_b is the break energy and p_0, p_1 are the electron spectral indexes. For more details about equations (1) and (2), see Finke et al. (2008).

2.2 Markov chain Monte Carlo method (MCMC)

MCMC is a method of estimating the probability distribution of parameters using random numbers. The probability distribution gives us the optimal values of parameters, as the mean or median of the distribution, and their uncertainties. According to the Bayes' theorem, the posterior

distributions of the parameters $P(\mathbf{x}|\mathbf{y})$ are proportional to the likelihood function $L(\mathbf{y}|\mathbf{x})$ times prior distribution $p(\mathbf{x})$, as follows:

$$P(\mathbf{x}|\mathbf{y}) = \frac{L(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{Z}, \quad Z = \sum_{\mathbf{x}} L(\mathbf{y}|\mathbf{x})p(\mathbf{x}). \quad (4)$$

Here, the vectors \mathbf{x} and \mathbf{y} represent the model parameters and the data, respectively. The prior distributions of a part of parameters may be given by past studies. For example, δ_D can be estimated from VLBI observations with uncertainty. In this case, we can use the probability distribution obtained from such estimations as a prior, instead of fixing the value.

We used the adaptive Metropolis algorithm to sample the posterior distribution (Robbins et al., 1951). This method learns the variance-covariance matrix of the proposed distribution from MCMC samples to perform efficient sampling. The variance-covariance matrix is updated from step $n-1$ to n , as follows:

$$\begin{aligned} \boldsymbol{\mu}_n &\leftarrow \boldsymbol{\mu}_{n-1} + h_n(\mathbf{x}_n - \boldsymbol{\mu}_{n-1}), \\ \boldsymbol{\Sigma}_n &\leftarrow \boldsymbol{\Sigma}_{n-1} + u_n \left((\mathbf{x}_n - \boldsymbol{\mu}_{n-1})(\mathbf{x}_n - \boldsymbol{\mu}_{n-1})^T - \boldsymbol{\Sigma}_{n-1} \right), \\ \sigma_n^2 &\leftarrow \sigma_{n-1}^2 + s_n(\text{FA}_n - \alpha), \end{aligned} \quad (5)$$

where, $\boldsymbol{\mu}_n$ is the mean value, $\boldsymbol{\Sigma}$ is the variance-covariance matrix, and σ^2 is the scale parameter. FA is 1 when the candidate of the next state is accepted, and 0 when it is not. α is the acceptance rate. h_n, u_n and s_n are learning coefficients, and we set $h_n = u_n = s_n = \frac{10}{n+100.0}$ in this study.

3. Results

3.1 Experiments with artificial data

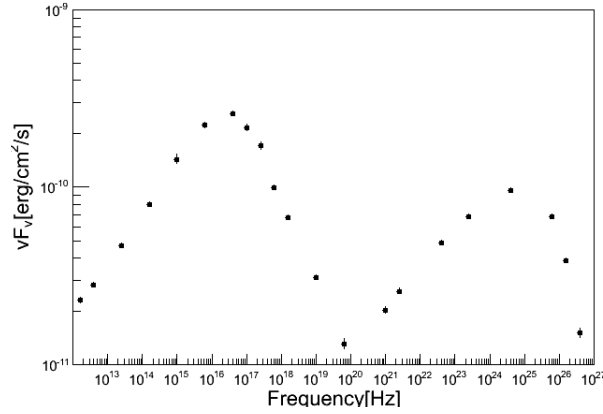


Fig 1. The SED of the artificial data.

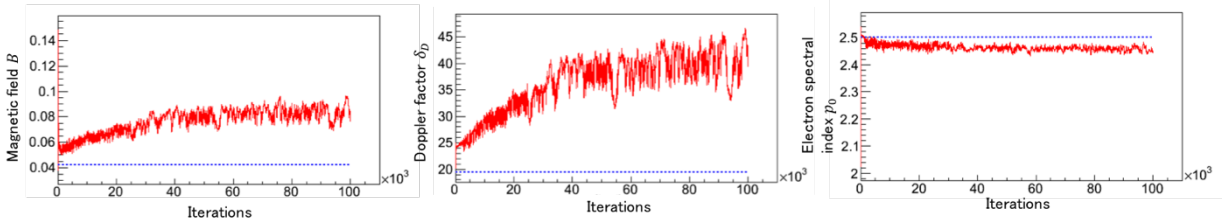


Fig 2. Trace plots of B (left), δ_D (middle), p_0 (right) obtained with the non-informative priors. The dotted line indicates each true value.

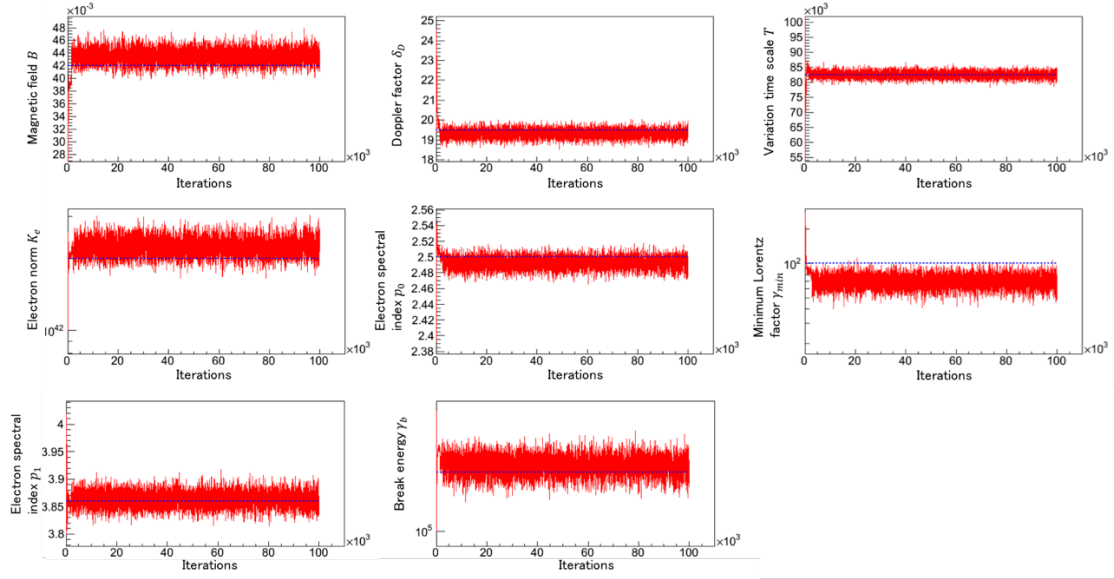


Fig 3. Trace plots all 8 parameters of the SED model obtained with the Gaussian prior for T . The dotted line indicates each true value.

We analyzed artificial SED data whose true SSC parameters were known using MCMC in order to investigate the correlation between the parameters and the necessity of prior distributions. The SSC parameters used in this study were referred from Itoh et al. (2015): $B=0.042$ G, $\delta_D=19.5$, $T=82540$ s, $K_e=2.0 \times 10^{42}$, $\gamma_B=1.2 \times 10^5$, $p_0 = 2.5$, $p_1 = 3.86$, $\gamma_{\min}=100$ and $\gamma_{\max} = 1.0 \times 10^8$. Figure 1 shows the artificial SED data. We generated the data by adding a Gaussian noise to the SED values calculated by the SSC parameters. The standard deviation of the Gaussian noise was 50% of the model SED values. The frequency coverage of the data is an ideal one. We can actually obtain only a part of those data in real observations. We confirmed that γ_{\max} was not constrained by the data. Hence, we gave its true value in MCMC iterations, and did not estimate it.

We define the likelihood function as the following normal distribution, as follows:

$$L = \begin{cases} \prod_i \frac{1}{\sqrt{2\pi\sigma_D^2}} \exp\left(-\frac{(d_i(\nu)-f_{\text{sync}})^2}{2\sigma_D^2}\right) & (\nu < 10^{21} \text{ Hz}), \\ \prod_i \frac{1}{\sqrt{2\pi\sigma_D^2}} \exp\left(-\frac{(d_i(\nu)-f_{\text{SSC}})^2}{2\sigma_D^2}\right) & (\nu > 10^{21} \text{ Hz}), \end{cases} \quad (6)$$

where, σ_D is the standard deviation and d_i is the SED data. First, we set non-informative priors to all parameters. Figure 2 shows the trace plots of B , δ_D and p_0 . The blue dotted lines indicate the true values of each parameter. As can be seen in the figure, the MCMC samples cluster with ranges of parameters significantly deviate from the true values throughout the iterations. In addition, the MCMC samples do not converge to the stationary distribution. We confirmed that B , δ_D and T are strongly correlated in the MCMC samples. All these three parameters are related to the normalization factor of the SED, as shown in equations (1) and (2). Hence, we consider that these three parameters are degenerated, and cannot be determined independently. Therefore, more informative priors are needed for them.

We used a Gaussian prior for T .

$$p(T) = \frac{1}{\sqrt{2\pi\sigma_T^2}} \exp\left(-\frac{(T-T_{\text{mean}})^2}{2\sigma_T^2}\right). \quad (7)$$

T can be estimated from the light curve. We set $T_{\text{mean}}=82540$ s, a true value of the artificial data. The standard deviation of the prior was set to be $\sigma_T=1000.0$ s. This is the maximum value to obtain the stationary distribution from MCMC samples. The trace plots obtained with the prior are shown in Figure 3. We can see that the MCMC samples of all parameters converge to the stationary distributions. This result demonstrates that the prior distribution of T is effective to estimate the SED parameters.

The actual observations provide SEDs with less data points than those in Figure 1. As a result, the MCMC convergence becomes worse than that in Figure 3. In such a case, we found that it is effective to use a prior distribution for δ_D in addition to T .

3.2 Estimation of the SED model parameters of Mrk 421

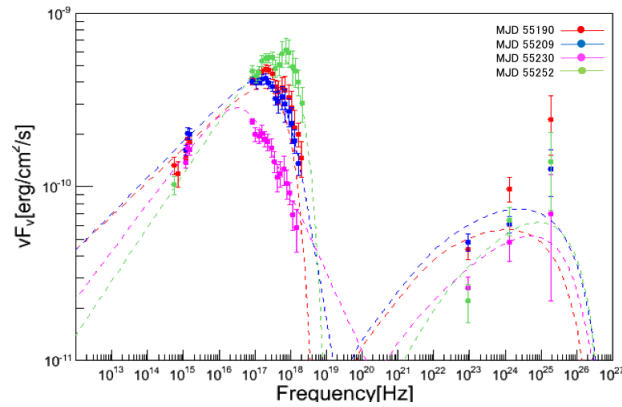


Fig. 4. SEDs of Mrk421 with the SSC model optimized by MCMC. The model SEDs are indicated as the dashed lines with the same colors as those of the data (the filled circles).

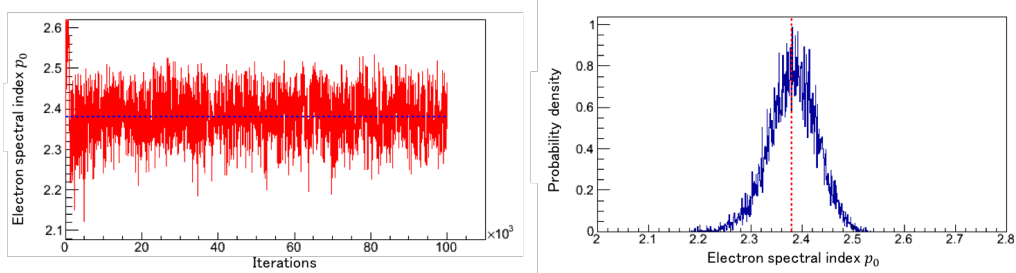


Fig. 5. The left panel shows the trace plots of parameter p_0 on MJD 55252. The right panel shows its probability distribution.

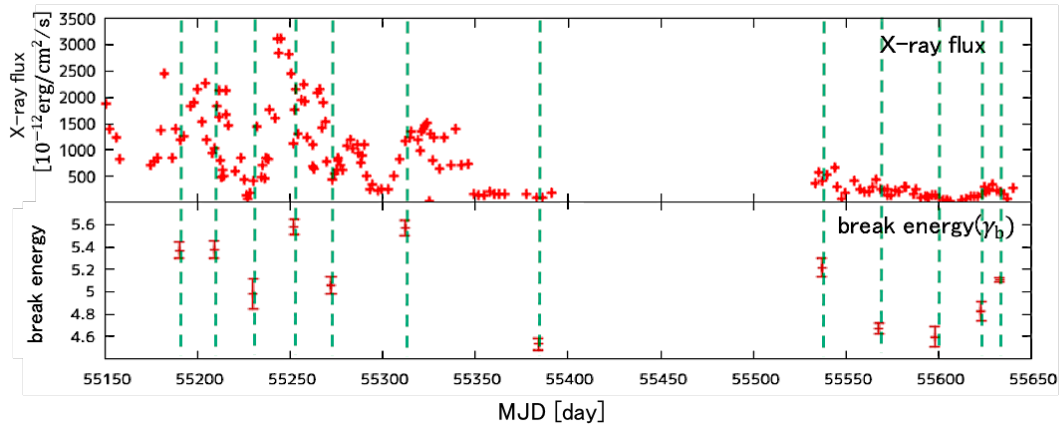


Fig. 6. The upper panel shows the 0.3-10 keV flux measured by XRT/Swift of Mrk 421 between 2009 November to 2011 March. The lower panel shows the time variations of γ_b .

As mentioned in section 3.1, we found that it was effective to use prior distributions for T and δ_D to estimate the SED model parameters. In this section, we report on the estimations of the SED parameters of Mrk 421 using this method. We used simultaneous observations of Mrk 421 with Fermi-LAT, Swift-XRT, and the 1.5-m optical telescope, Kanata from 2009 to 2011 reported in Itoh et al. (2015). We estimated the model parameters of SEDs for 12 epochs. We set a Gaussian prior of δ_D in addition to the prior of T . Abdo et al. (2011) report that δ_D of Mrk 421 is ~ 21 based on the analysis of rich SED data. We used it as the mean of the Gaussian prior. The standard deviation of the δ_D prior was set to be a large value, 10.0 to allow deviations from the mean. Four examples of the observed and optimized model SEDs are shown in Figure 4. The SSC models well explain the observations. The SED data of Mrk 421 cannot constrain not only γ_{\max} , but also γ_{\min} . Hence, we fixed γ_{\max} and γ_{\min} as 100 and 1.0×10^8 in our analysis.

Figure 5 shows the trace plot and the posterior distribution of p_0 on MJD 55252. We can confirm that the MCMC samples converges to a stationary distribution in the trace plot. The posterior distribution is symmetric without any major local peaks. The optimal value of the parameter and its uncertainty are naturally defined as the mean and 68% confidence interval of the distribution.

As the case in Figure 5, we obtained the optimal values and their uncertainties of 7 parameters for 12 epochs. We found that six parameters (B , δ_D , K_e , γ_b , p_0 , p_1) significantly changed with time even when the uncertainties were considered. Among them, we focus attention on the break energy, γ_b . Figure 6 shows the X-ray light curve and the time variation of γ_b . As can be seen in Figure 6, γ_b shows a clear positive correlation with the X-ray flux. This result suggests that the variations of the X-ray flux was mainly caused by the fluctuations of the break energy. Similar results on Mrk 421 were also reported in Itoh et al. (2015) and Bartolli et al. (2016). Our results confirmed them with uncertainties of estimations. More detailed reports on the MCMC estimated SED parameters will appear in a forthcoming paper.

4. Summary and Conclusion

We developed a method to estimate the physical quantity in jets using MCMC and succeeded in finding the optimal parameters and uncertainties. In the case of the SEDs provided by actual observations, we confirmed that B , δ_D and T are strongly correlated. In order to achieve MCMC convergence, the prior distributions are required for at least one parameter among them. The MCMC analysis of the data of Mrk 421 suggests that the variations of the X-ray flux was mainly caused by the fluctuations of the break energy.

References

- [1] Itoh, R., et al. PASJ 67,45 (2015).
- [2] Finke, J. D., Dermer, C. D., & Böttcher, M., ApJ 686,181 (2008)
- [3] Böttcher, M., et al. ApJ, 768,54 (2013).
- [4] Robbins, H. and Sutton M. The annals of mathematical statistics, 22, 400-407 (1951).
- [5] Abdo, A., et al. ApJ, 736,131 (2011).
- [6] Bartoli, B., et al. ApJSS, 222,6 (2016).