Modeling blazar broadband emission with convolutional neural network

N. Sahakyan & D. Bégué



Content

- What are blazars ?
- Origin of emission: problems
- Developing convolutional neural network
- Results and applications



Jetted vs. non-jetted AGN

Oľ primarily radiation emitted from the Non-jetted AGNs energy output is accretion disk, as well as from the corona and potentially from a wind outflow.



highly collimated, relativistic active galaxies that exhibit Jetted AGNs are a class of jets of plasma.





Jetted vs. non-jetted AGN

Oľ primarily radiation emitted from the Non-jetted AGNs energy output is accretion disk, as well as from the corona and potentially from a wind outflow.



highly collimated, relativistic active galaxies that exhibit Jetted AGNs are a class of jets of plasma.





Jetted vs. non-jetted AGN

Non-jetted AGNs energy output is primarily radiation emitted from the accretion disk, as well as from the corona and potentially from a wind or outflow.



Jetted AGNs are a class of active galaxies that exhibit highly collimated, relativistic jets of plasma.























Evolution of the particle distribution

Gasparyan, Bégué, Sahakyan 2022

$$egin{aligned} & \left(rac{\partial N_e}{\partial t}(\gamma) = rac{N_e}{t_{
m esc}} + rac{\partial}{\partial \gamma} [(C_{
m IC}N_\gamma + C_{
m sync})N_e imes] + Q_{\gamma\gamma o e^+e^-}, \ & \left(rac{\partial N_\gamma}{\partial t}(x) = rac{N_\gamma}{t_{
m esc}} + Q_{
m sync} + R_{
m IC}N_\gamma - S_{\gamma\gamma o e^+e^-}, \end{aligned}
ight.$$

Evolution of the particle distribution

Gasparyan, Bégué, Sahakyan 2022

$$egin{aligned} & \left(rac{\partial N_e}{\partial t}(\gamma) = rac{N_e}{t_{
m esc}} + rac{\partial}{\partial \gamma} [(C_{
m IC}N_\gamma + C_{
m sync})N_e imes] + Q_{\gamma\gamma o e^+e^-}, \ & \left(rac{\partial N_\gamma}{\partial t}(x) = rac{N_\gamma}{t_{
m esc}} + Q_{
m sync} + R_{
m IC}N_\gamma - S_{\gamma\gamma o e^+e^-}, \end{aligned}
ight.$$





Evolution of the particle distribution

Gasparyan, Bégué, Sahakyan 2022

$$egin{aligned} & \left(rac{\partial N_e}{\partial t}(\gamma) = rac{N_e}{t_{
m esc}} + rac{\partial}{\partial \gamma} [(C_{
m IC}N_\gamma + C_{
m sync})N_e imes] + Q_{\gamma\gamma o e^+e^-}, \ & \left(rac{\partial N_\gamma}{\partial t}(x) = rac{N_\gamma}{t_{
m esc}} + Q_{
m sync} + R_{
m IC}N_\gamma - S_{\gamma\gamma o e^+e^-}, \end{aligned}
ight.$$





Put the model in a fitting engine.

- After waiting for a very long time, get an answer. If it works good, but if not, it is too expensive to redo it.
- Even though, limited to a couple of dataset.



Put the model in a fitting engine.

- After waiting for a very long time, get an answer. If it works good, but if not, it is too expensive to redo it.
- Even though, limited to a couple of dataset.

Table model

- Compute the model in wide parameter space
- Interpolate between sampled parameter sets



Put the model in a fitting engine.

- After waiting for a very long time, get an answer. If it works good, but if not, it is too expensive to redo it.
- Even though, limited to a couple of dataset.

For Bayesian fit 10⁴ - 10⁵ times.
Average time for single spectrum calculations
SSC and EIC model: ~ 30 sec (using 8 core)
✓ model computation time ~ 3.5 - 35 days
hadronic model: ~ 90 sec (using 8 core)
✓ model computation time ~ 10.4 - 100 days

Table model

- Compute the model in wide parameter space
- Interpolate between sampled parameter sets



Put the model in a fitting engine.

- After waiting for a very long time, get an answer. If it works good, but if not, it is too expensive to redo it.
- Even though, limited to a couple of dataset.

For Bayesian fit 10⁴ - 10⁵ times.
Average time for single spectrum calculations
SSC and EIC model: ~ 30 sec (using 8 core)
✓ model computation time ~ 3.5 - 35 days
hadronic model: ~ 90 sec (using 8 core)
✓ model computation time ~ 10.4 - 100 days

Table model

- Compute the model in wide parameter space
- Interpolate between sampled parameter sets

Square parameter space, depending on the number of model parameters SSC (7 free parameters): $10^6 - 10^7$ times. EIC (11 free parameters): > 10^7 times.

files with size > 250 Gb



approximations

Make the model faster

numerical code

semi-analytical + numerical code
factor of ~ 2 gain but the model
needs to be computed EVERY time

approximations

Make the model faster

Machine learning:

✓ Model will be computed once and can be used with different datasets.

numerical code

 \checkmark It does not require a square parameter space, which makes it less computationally expensive. It allows for model failure which can be removed from the training

semi-analytical + numerical code factor of ~ 2 gain but the model needs to be computed EVERY time

approximations

Make the model faster

Machine learning:

✓ Model will be computed once and can be used with different datasets.

numerical code

 ✓ It does not require a square parameter space, which makes it less computationally expensive. It allows for model failure which can be removed from the training

Latin hypercube sampling method is a widely popular technique in the creation of surrogate models as it presents several advantages. First, it allows to specify the number of simulations to be computed. As a byproduct, this method does not require to specify parameter spacing. Second, it ensures uniform sampling across all parameters. Lastly, it avoids the regular sampling of parameters, which is typical in grid scan techniques. semi-analytical + numerical code
 factor of ~ 2 gain but the model
 needs to be computed EVERY time





Workflow of the method

Parameter Sampling

This large range of parameters guarantees that any blazar SED can be reproduced.

synchrotron self Compton

Parameter	Units	Symbol	Minimum	Maximum	Type of
Doppler boost	_	δ	3	50	L
Blob radius	cm	R	10^{15}	10^{18}	Loga
Minimum electron injection Lorentz factor	-	$\gamma_{ m min}$	$10^{1.5}$	10^5	Loga
Maximum electron injection Lorentz factor	-	$\gamma_{ m max}$	10^{2}	10^{8}	Loga
Injection index	-	p	1.8	5	L
Electron luminosity	erg.s ⁻¹	L_e	10^{42}	10^{48}	Loga
Magnetic field	G		10^{-3}	10^{2}	Loga

Bégué, Sahakyan, et al., 2024, ApJ, 963, id.71

External inverse Compton

Parameters	Units	Symbol	Minimum	Maximum	Type of
Doppler boost	_	δ	3	50	Ι
Blob radius	cm	R	10^{15}	10^{18}	Log
Minimum electron Lorentz factor	-	$\gamma_{ m min}$	$10^{1.5}$	10^5	Log
Maximum electron Lorentz factor	-	$\gamma_{ m max}$	10^{2}	10^{6}	Log
Injection index	-	p	1.8	5	1
Electron luminosity	erg.s ⁻¹	L_e	10^{42}	10^{48}	Log
Magnetic field	G	B	10^{-3}	$10^{2.5}$	Log
Black hole mass	M_{\odot}	$M_{ m BH}$	10 ⁷	10^{10}	Log
Disk luminosity	erg.s ⁻¹	L_d	$10^{43.5}$	$10^{47.5}$	Log
BLR frequency	Hz	$ u_{ m BLR}$	$10^{14.5}$	10^{16}	Log
DT frequency	Hz		$10^{12.5}$	10^{14}	Log

Sahakyan, Bégué, et al., 2024, arXiv:2402.07495

- distribution
- linear
- arithmic
- arithmic
- arithmic
- Linear arithmic arithmic



garithmic







Spectrum generation: Time dependent approach

Monthly Notices

of the Royal Astronomical Society

lssues 🔻	Advance articles	Submit 🔻	Purchase	Ale

Monthly Notices of the Roy 🔻



Volume 509, Issue 2 January 2022

< Previous Next >

JOURNAL ARTICLE

Time-dependent lepto-hadronic modelling of the emission from blazar jets with SOPRANO: the case of TXS 0506 + 056, 3HSP J095507.9 + 355101, and 3C

About 🔻

rts

279 Get access >

S Gasparyan 🖾, D Bégué, N Sahakyan

Monthly Notices of the Royal Astronomical Society, Volume 509, Issue 2, January 2022, Pages 2102–2121, https://doi.org/10.1093/mnras/stab2688 Published: 29 September 2021 Article history ▼





$$\frac{\partial N_{p}}{\partial t} = C_{p\gamma \to p\pi} + C_{p\gamma \to e^{+}e^{-}} + C_{synch} - S_{\gamma p \to n\pi} + Q_{\gamma n \to p\pi} \qquad \qquad \frac{\partial N_{\mu}}{\partial t} = Q_{\pi_{\pm}} - S_{\mu} + C_{synch}$$
$$\frac{\partial N_{n}}{\partial t} = -S_{n\gamma \to p\pi} + Q_{p\gamma \to n\pi} + C_{n\gamma \to n\pi} \qquad \qquad \frac{\partial N_{\nu,\zeta}}{\partial t} = Q_{\pi_{\pm}} + Q_{\mu}$$
$$\frac{\partial N_{\pi_{\pm}}}{\partial t} = Q_{p\gamma \to \pi} + Q_{n\gamma \to \pi} - S_{\pi} + C_{synch} \qquad \qquad \frac{\partial N_{e^{\pm}}}{\partial t} = Q_{\mu} + Q_{p\gamma \to e^{+}e^{-}} + Q_{\gamma\gamma \to e^{+}e^{-}}C_{IC} + C_{synch}$$

Q: sínk term S: source term C: coolíng term



$$\frac{\partial N_{\mu}}{\partial t} = Q_{\pi_{\pm}} - S_{\mu} + Q_{\pi_{\pm}} - S_{\mu} + Q_{\mu}$$

















FNC is the synthesis of decades of experience with jobs schedulers

PBS + Grid Engine + Accelerator



Scalable

• From 1 job to 20 million jobs in queue

Small, Quick

- Small memory footprint
- Speed: clocked up to 70k+ tasks/second

Feature Rich

- Full-cycle scheduler
- Cost-driven job placement
- Workload analysis (via simulation)
- Prediction of job duration + job size
- PMIx support (evolution of MPI)
- SAGA (storage-aware scheduling)
- **Rapid Scaling in cloud**

Contact: casotto@altair.com



REPORT BY user FOR SET

Soprano:ALL

	Jobs	Unknown Duration	Count (%)	Total Time	Time (%)	Ave Dur			
total	1,000,000	0	100.00%	1y66d	100.00%				
ncadmin	1,000,000	0	100.00%	1y66d	100.00%				
First Job Start			Wed Oct 25 09:15:59 2023						
Last Job Finish			Thu Oct 26 02:41:36 2023						
Elapsed Time			17h25m						



2024, arXiv:2402.07495 al., et égué, р Sahakyan,

		Status	Size	Duration	Actions					
	[]									
	[ALL]									
1	ALL	999955 <mark>45</mark>	1,000,000	1y66d	莭	۲	Ш	Ŧ	•	0
2	Batch0000	25000	25,000	13d03h	莭	•	II	1	•	0
3	Batch0001	25000	25,000	10d17h	茴	•	П	1	•	0
4	Batch0002	25000	25,000	10d15h	茴	•	II	1		0
5	Batch0003	25000	25,000	10d16h	Ŵ	•	п	1	•	0
6	Batch0004	25000	25,000	10d15h	⑪	•	П	Ŧ	•	0
7	Batch0005	25000	25,000	10d16h	⑪	Þ	II	1	•	0
8	Batch0006	25000	25,000	10d15h	莭	•	п	1	•	0
9	Batch0007	25000	25,000	10d15h	⑪	۲	II	ł	•	0
10	Batch0008	25000	25,000	10d16h	莭	•	II	Ţ	•	0
11	Batch0009	25000	25,000	10d16h	Ü	•	Ш	Ţ	•	0
12	Batch0010	25000	25,000	10d17h	尙	Þ	II	ł		0
13	Batch0011	25000	25,000	10d15h	Ö	•	п	Ţ	•	0
14	Batch0012	25000	25,000	10d15h	莭	•	П	ţ	•	0
15	Batch0013	25000	25,000	10d14h	莭	Þ	п	ł		0
16	Batch0014	25000	25,000	10d16h	莭	►	п	Ţ	•	0
17	Batch0015	25000	25,000	10d15h	莭	•	Ш	Ŧ	•	0
18	Batch0016	25000	25,000	10d15h	面	Þ	11	ł		0
19	Batch0017	25000	25,000	10d15h	世	•	п	Ţ	•	0
20	Batch0018	25000	25,000	10d14h	前	•	II	Ŧ		0
21	Batch0019	25000	25,000	10d20h	莭	Þ	11	Ţ	•	0
22	Batch0020	25000	25,000	10d14h	莭	•	П	Ŧ	•	0
23	Batch0021	25000	25,000	10d15h	茴	•	II	ŧ	•	0
24	Batch0022	25000	25,000	10d15h	ů	•	п	Ţ	•	0
25	Batch0023	25000	25,000	10d18h	尙	•	Ш	ł	•	0
26	Batch0024	25000	25,000	10d17h	前	•	п	ł	•	0
27	Batch0025	25000	25,000	10d18h	<u>ٿ</u>	•	II	Ţ	•	0
28	Batch0026	25000	25,000	10d18h	莭	Þ	Ш	ł	•	0
29	Batch0027	25000	25,000	10d18h	莭	•	п	ł	•	0
30	Batch0028	25000	25,000	10d17h	莭	•	п	ļ	•	0
31	Batch0029	25000	25,000	10d17h	莭	Þ	II	4		0
32	Batch0030	25000	25,000	10d16h	莭	•	п	Ţ	•	0
33	Batch0031	25000	25,000	10d17h	莭	•	П	ţ		0
34	Batch0032	25000	25,000	10d18h	莭	•	п	ł	•	0
35	Batch0033	24978 22	25,000	11d03h	莭	•	п	Ţ	•	0
36	Batch0034	25000	25,000	10d18h	Û	•	П	ł		0
37	Batch0035	25000	25,000	10d20h	ŵ	•	II	Ţ	•	0
38	Batch0036	24977 23	25,000	11d04h	₫	•	п	Ŧ	•	0
39	Batch0037	25000	25,000	10d18h	莭	•	п	ł		0
40	Batch0038	25000	25,000	10d20h	前	•	п	Ţ	•	0
44	Batch0039	25000	25.000	10d19h	莭	•	п	ł	•	0



Monitoring the data



long tail extending beyond 700 s. These extended durations correspond to spectra characterized by a high compactness with small radius R, large electron luminosity Le, and small injection Lorentz factor gmin.

Bégué, Sahakyan, et al., 2024, ApJ, 963, id.71

The computation of the spectra by SOPRANO can fail. The solution is obtained with the Newton-Raphson root finding algorithm, which can, in some instances, not converge toward the solution with the required accuracy (10^{-15}) , close to machine accuracy. The total number of spectra with at least one failed time iteration is 3693, constituting fewer than 2% of all calculated spectra. From the distribution of the maximum error across a full simulation is evident that only a small fraction of the spectra are unreliable, with most spectra having a maximal error below 10^{-10} .







Modeling the relationship between the input parameters and their corresponding spectra

Convolutional Neural Networks (CNNs)

CNNs, initially designed for image analysis, have evolved to become versatile tools for processing sequential data, including time series. Their ability to automatically extract hierarchical features makes them well-suited for capturing complex temporal dependencies present in time series data.











Independent outputs











Hidden Layers



Independent outputs







Hidden Layers



Independent outputs





A deep network is not necessary to produce an accurate representation of the numerical model. The CNN contains only eight layers in this order: a first dense layer transforms the seven inputs to a high dimensional vector, five 1D convolutional layers with different kernel sizes and strides, one maxpooling layer followed by a 1D convolutional layer, and a final dense layer, mapping to the 150 outputs.



Bégué, Sahakyan, et al., 2024, ApJ, 963, id.71

Final results

Ъ validation split into 10%spectra is set. Ъ a 10% test set, of training sample and 80% Our set,



et al., 2024, ApJ, 963, id.71 ahakyan, S Bégué,



Bégué, Sahakyan, et al., 2024, ApJ, 963, id.71

Workflow of the method

Bégué, Sahakyan, et al., 2024, ApJ, 963, id.71

Workflow of the method

Application

The trained CNN is used with multinest to fit the broadband SEDs of 4 blazars

Synchrotron self-Compton model

Mrk 421 (z = 0.031, data from MW campaign in 2009) 1ES 1959+650

The broadband SEDs of Mrk 421 during the 4.5 month long multiwavelength campaign in 2009. The data and the errors are in blue, the red line is the model corresponding to the best parameters, i.e., maximizing the likelihood, and the gray spectra represent one in 10 randomly selected samples from the MCMC sampling, representing the model uncertainty.

Bégué, Sahakyan, et al., 2024, ApJ, 963, id.71

Parameter posterior distributions for Mrk 421 during the multiwavelength campaign of 2009. The contours give, from outward to inward, the 20%, 40%, and 75% confidence regions. Apart from the radius R, all parameters are well constrained.

Synchrotron self-Compton model

Mrk 421

1ES 1959+650 (z = 0.031, data from MW campaign in 2009)

The broadband SEDs of 1ES 1959+650 on the 2016 June 14. The data and the errors are in blue, the red line is the model corresponding to the best parameters, i.e., maximizing the likelihood, and the gray spectra represent one in 10 randomly selected samples from the MCMC sampling, representing the model uncertainty.

Bégué, Sahakyan, et al., 2024, ApJ, 963, id.71

Parameter posterior distributions for 1ES 1959+650. The magnetic field, the electron luminosity, and the electron index are well constrained. In contrast, the other parameters remain somewhat unconstrained due to the high uncertainty in the position of the peak energy of the synchrotron bump.

External inverse Compton

→ 3C 454.3 (z = **0.859**)

CTA 102

Multiwavelength SED of 3C 454.3, presented in blue. The model corresponding to the maximum likelihood is shown by the solid red line, while the uncertainty associated with the model is depicted in gray.

Sahakyan, Bégué, et al., 2024, arXiv:2402.07495

Parameter posterior distributions from the SED modeling of 3C 454.3 for the period MJD 55519.59-55520.19

External inverse Compton

3C 454.3

CTA 102 (z = 1.037)

Multiwavelength SED of CTA 102, presented in blue. The model corresponding to the maximum likelihood is shown by the solid red line, while the uncertainty associated with the model is depicted in gray. the red dashed line represents the scenario for which the radius is larger and the emission is due to SSC

Sahakyan, Bégué, et al., 2024, arXiv:2402.07495

Parameter posterior distributions for CTA 102, showing a bimodal distribution for the radius R alongside other parameters.

MMDC

MMDC

MMDC

Home

About

Data access

Theoretical Modeling

Articles

Team

Contact

Conclusions

- Astrophysical data in the 1970s were sent by post on magnetic tape.
- If There is now exponential growth in astronomical data volumes, driven by advancements in observational technologies and an increasing number of telescopes observing in different bands.
- i AI and ML have revolutionized our approach to studying the universe, aiming to uncover hidden characteristics within data, enabling faster simulations, improved observations, and deeper understandings of cosmic phenomena.
- Matter Astrophysical research will largely benefit from new developments in AI/ML.

