

RELATIVISTIC JETS
IN THE ERA OF X-RAY POLARIMETRY

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Abel Lawrence Peirson V
April 2023

© Copyright by Abel Lawrence Peirson V 2023
All Rights Reserved

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Roger W. Romani) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Roger D. Blandford)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Peter F. Michelson)

Approved for the Stanford University Committee on Graduate Studies

Abstract

Measuring X-ray polarization has been a major goal in astrophysics for the last 40 years. With the launch of NASA’s Imaging X-ray Polarimetry Explorer (IXPE) in December 2021, soft X-ray polarimetry (1-10 keV) is now possible. In this thesis, I introduce an optimal signal extraction method for X-ray polarimeters based on gas pixel detectors. Using IXPE’s thus improved sensitivity, I make progress towards identifying the high-energy emission processes and magnetic field geometries responsible for radiation in blazar jets. First, I develop a simulation-based inference framework for measuring X-ray polarization as a function of time, energy, and position in gas pixel detectors, proving that for a fixed model budget the framework is optimal. Second, I propose a simple geometrical jet model to study the polarization properties of leptonic blazar jets, including their observed EVPA rotations and their detection prospects with IXPE. Finally, I interpret IXPE’s measurements of BL Lacertae across the synchro-Compton transition region to suggest leptonic-dominated emission processes in blazars, and of Mrk 421 to provide evidence for helical magnetic fields as the source of EVPA rotations.

Preface

This thesis is based in part on the following eight peer-reviewed publications, invited textbook chapter, one publication in submission, and one publication in preparation.

- Peirson and Romani [1], “The Polarization Behavior of Relativistic Synchrotron Jets”
- Peirson and Romani [2], “The Polarization Behavior of Relativistic Synchrotron Self-Compton Jets”
- Peirson et al. [3], “Deep Ensemble Analysis for Imaging X-ray Polarimetry”
- Peirson and Romani [4], “Towards Optimal Signal Extraction for Imaging X-ray Polarimetry”
- Liodakis, Peirson, and Romani [5], “Prospects for Detecting X-ray Polarization in Blazar Jets”
- Peirson, Liodakis, and Romani [6], “Testing High-energy Emission Models for Blazars with X-Ray Polarimetry”
- Peirson and Romani [7], *A Deep Ensemble Approach to X-ray Polarimetry*
- Peirson [8], *Neural Network Analysis of X-ray Polarimeter Data*
- Peirson and the IXPE team (2023), ”X-ray Polarization of BL Lacertae in Outburst”
- Peirson and Romani (2023), in prep., ”Geometric Model for X-ray Polarization Rotation in Mrk 421”

Acknowledgments

First and foremost I would like to thank my advisor Roger Romani for encouraging my interests, tirelessly fighting my corner, and generally being my biggest supporter throughout the PhD. I have never met anyone with greater clarity of thought and I am very proud to have been his student. I would also like to thank my close group members and collaborators, Yannis Liodakis and Dinesh Kandel for both treating me as part of their families and sharing the journey here at Stanford together.

Thank you to my friends in the bay many of whom have profoundly shaped my interests and the way I think. Especially: Taylor Howell for incessant discussions about optimization, get-rich-quick schemes, and long surfs. Dylan Freedman for instilling in me the great endeavor of side projects and for how to work ergonomically. We shared many a night programming on the sofa together. Christopher Fifty and Ehsan Amid for taking my interests in AI to the next level. John Steidley for opening my eyes to the world of rationalism. Leo Mehr for making triathlon bearable and sharing the most beautiful house in SF together. And thank you to all of my friends back in the UK, who no matter how anxious I get about losing touch always welcome me back with open arms.

Lastly, I would like to thank my family in Spain, the US and the UK for always making a base for me no matter how deeply I was entrenched in this PhD.

I acknowledge and appreciate support from the NASA FINESST and Stanford Data Science programs during my tenure at Stanford.

Contents

Abstract	iv
Preface	v
Acknowledgments	vi
1 Introduction	1
1.1 Preliminaries	3
1.1.1 Imaging X-ray Polarimetry	3
1.1.2 Polarization estimation	11
1.1.3 Neural Networks for Machine Learning	15
1.1.4 Blazars	22
2 Neural Networks for Track Reconstruction	26
2.1 Introduction	26
2.2 Dataset	27
2.3 Geometric bias	28
2.3.1 Hexagonal to square conversion	29
2.4 Deep ensemble setup	30
2.4.1 Emission angles	30
2.4.2 Absorption points	31
2.4.3 Photon energy	31
2.5 Removing tail tracks	32
2.6 Training and ensemble selection	34
2.7 Performance	35
2.7.1 Emission angles	35
2.7.2 Tail vs. peak classification	37
2.7.3 Absorption points	38
2.7.4 Photon energy	38

3	Polarization Estimation	43
3.1	Introduction	43
3.2	Modulation factor	43
3.3	Weighted maximum likelihood estimator	45
3.3.1	Deep ensembles	46
3.3.2	Simulation-Based inference	47
3.4	Performance	48
3.4.1	Weights	48
3.4.2	Comparison	49
4	Synchrotron Polarization of Relativistic Jets	51
4.1	Introduction	51
4.1.1	Blazar EVPA Variability	51
4.2	Relativistic PA Rotation	53
4.3	Geometrical Jet Model	56
4.4	Radiation Model	59
4.5	Application to Mrk 501	63
4.6	Conclusion	64
5	Synchrotron Self-Compton Polarization of Relativistic Jets	66
5.1	Introduction	66
5.2	ICS Basics & Seed Photon Polarization	67
5.3	Multi-zone Effects	71
5.4	Light Travel Effects	74
5.4.1	Finite Bulk Lorentz factor	75
5.4.2	Non-zero Viewing Angle of a Conical Jet	76
5.4.3	Finite Travel Time at Large Jet Radius: Seed Photon Build-up	76
5.5	SSC with All Effects	78
5.5.1	Effective number of emission zones	79
5.6	Conclusion	84
6	X-ray Polarization Prospects for Blazars	87
6.1	Introduction	87
6.2	Detecting X-ray polarization in blazar jets	87
6.2.1	X-ray variability	88
6.2.2	Expected X-ray polarization	90
6.2.3	Blazar Detectability	91
6.2.4	Summary	95

6.3	Detecting X-ray polarization in ISP blazars	97
6.3.1	Multiwavelength observations and modeling	100
6.3.2	<i>IXPE</i> Measurement Simulations	103
6.3.3	Discussion	105
7	X-ray Polarization of BL Lacertae in Outburst	109
7.1	Introduction	109
7.2	Data analysis	110
7.2.1	Time variability	112
7.3	Discussion & Conclusions	114
8	Geometric Model for X-ray Polarization Rotation in Mrk 421	118
8.1	Introduction	118
8.2	Rotation Models	119
8.2.1	Helical field in a relativistic jet	121
8.2.2	Rotation with underlying constant polarization	122
8.3	Data Analysis	122
8.3.1	Polarization reconstruction	122
8.3.2	Model fitting	124
8.4	Discussion	127
8.5	Conclusion	128
9	Conclusions	131
A	Computational Jet Model	133
B	Observations of BL Lacertae	135
B.1	X-ray observations	135
B.2	Time-resolved spectro-polarimetric analysis.	137
B.2.1	<i>XSPEC</i> analysis	137
B.3	Multiwavelength observations	138

List of Tables

3.1	Sensitivity analysis for two power law spectra ($dN/dE \sim E^{-N}$; PL2 for $N = 2$, PL1 for $N = 1$) each normalized to produce 10^5 2-8 keV photons when folded through IXPE’s energy response. MDP ₉₉ gives the sensitivities for the various reconstruction methods; smaller MDP ₉₉ is better. Mom. denotes moments analysis. DNN denotes deep ensemble. Percentages denote tail cut thresholds, 95% means events with DNN predicted tail probability greater than 95% are removed.	50
4.1	Radio and X-ray polarization fractions for Figures 9 (top panel) and 10. For the fig. 4.9 row $\bar{\sigma}$ describes the fluctuations about the mean $\bar{\Pi}$, while for fig. 4.10 it represents the uncertainty in the amplitude \bar{A} of the best-fit sinusoid.	64
4.2	Parameters used for the Mrk 501 fits of section 4.5. The top and bottom rows are for the solid and dashed fits respectively. Both fits use $W_j = 2 * 10^{37}W$, $L_j = 5 * 10^{20}m$ and $\gamma_{\min} = 10$	64
5.1	Parameters used for the plotted jet models. All models additionally assume jet power $W_j = 1.3 \times 10^{37}W$, electron spectral index $\alpha = 1.85$, and electron energy range $\gamma_{\min} = 10$ to $\gamma_{\max} = 3.3 \times 10^4$	85
6.1	X-ray flux modeling results. The X-ray fluxes are all in erg/cm ² /s (log).	97
6.2	X-ray flux and polarization. The X-ray fluxes are all in erg/cm ² /s (log). Polarization degree is in %. The table lists sources with $> 0.5\%$ X-ray polarization and X-ray flux $> 5 \times 10^{-13}$ erg/cm ² /s. The table lists only the first 10 sources. The table is published in its entirety in the machine-readable format. A portion is shown here for guidance regarding its form and content.	98
6.3	Detectability duty cycle. The table is sorted according to detection probability and lists only sources with DP > 1%.	98
6.4	Sources without optical polarization. The X-ray fluxes are all in erg/cm ² /s (log). Column $\Pi_{O,\min}$ lists the minimum optical polarization degree (%) required for an IXPE detection at 100ksec.	98

6.5	X-ray spectral parameters. The N_{H} values are given in units of $\times 10^{22} \text{ cm}^{-2}$	101
6.6	Polarized jet model best fit parameters. Jet power W_j , electron high energy cutoff before exponential decay E_{max} , electron population power law index α , full conical jet opening angle in lab frame θ_{open} , bulk Lorentz factor Γ_{bulk} , initial magnetic field strength B_0 , jet observation angle in lab frame θ_{obs} , and initial equipartition fraction A_{eq} . (f) denotes a flaring state. The number of magnetic field zones N_{zones} is selected from [1, 7, 19, 37, 64, 128]. All blazar models shown here use 37 magnetic field zones except CGRaBS J0211+1051, which uses 19.	103
7.1	Spectro-Polarimetric XSPEC model fits to joint <i>NuSTAR</i> , <i>XMM-Newton</i> , and <i>IXPE</i> spectra, fig. 7.1. Γ – photon index, Π – polarization fraction, ψ – EVPA. For the sum of two power-laws, the polarization fraction of the second power-law component is fixed to zero. Since the single power-law fit is inadmissible, uncertainties are likely underestimated for this model’s parameters. Upper limits are at 99% confidence. . .	112
7.2	Time-binned polarization results presented in fig. 7.3. MDP99 is the minimum detectable polarization, i.e., the 99% confidence level for non-zero polarization fraction. All errors are 68.3% (1σ) confidence intervals. Error on polarization fraction Π_X is non-symmetric because $0 \leq \Pi_X \leq 1$. PL1 fraction is the fraction of flux density in the bin from the lower energy (synchrotron) power-law component. For polarization measurements less than the MDP99 we display upper limits at 99% confidence. . .	114
8.1	MDP99 values for different polarization reconstruction methods for the joint Mrk 421 observations.	124
8.2	Natural log marginal likelihood differences from the constant rotation for weighted NNs. RC: constant rotation with constant polarization component, BH: boosted helix. More strongly preferred models have a higher log marginal likelihood difference. Negative log marginal likelihood means the constant rotation is preferred.	125
8.3	Log marginal likelihood differences from the constant rotation for weighted moment analysis. RC: constant rotation with constant polarization component, BH: boosted helix. More strongly preferred models have a higher log marginal likelihood difference. Negative log marginal likelihood means the constant rotation is preferred.	130
8.4	Log marginal likelihood differences from the constant rotation for the standard moment analysis. RC: constant rotation with constant polarization component, BH: boosted helix. More strongly preferred models have a higher log marginal likelihood difference. Negative log marginal likelihood means the constant rotation is preferred.	130
B.1	Major quasi-simultaneous X-ray observations related to the 2022 November 27 <i>IXPE</i> pointing of BL Lac.	136

B.2	<i>Swift</i> -XRT exposures in the vicinity of the <i>IXPE</i> pointing of BL Lac. Figure B.1 displays the exposures superimposed on the <i>IXPE</i> light curve. To measure flux density and Γ , single absorbed power-laws were fit to the full XRT 0.3–10 keV range.	137
B.3	Sum of two power-law spectro-polarimetric XSPEC model fits to time binned <i>IXPE</i> and <i>Swift</i> -XRT spectra, fig. 7.3. Photon indices Γ_1, Γ_2 , apec temperature kT , and N_H are fixed to their full-spectrum fit values; see table 7.1. High-energy PL polarization is fixed to zero. Π – polarization fraction, ψ – EVPA. Relative power-law normalizations in each time bin dictate the low-energy hump fractions in table 7.2.	138
B.4	Multi-wavelength polarization observations of BL Lac during the IXPE pointing. The uncertainties in Π and ψ are either the uncertainty of the measurement or, in the case of multiple measurements, the median uncertainty. σ_Π and σ_ψ show the standard deviation of the observations.	139

List of Figures

1.1	The IXPE observatory fully deployed for science observations. Three cylindrical mirror units direct X-rays to three corresponding detector units (gas pixel detectors). Figure from Ramsey et al. [9].	3
1.2	Design of an IXPE GPD. An incoming X-ray photon interacts with a dimethyl ether gas molecule to produce a photoelectron between the entrance window and the gas electron multiplier (GEM). Secondary charges created along the photoelectron's path are multiplied by the GEM and read out at the anode plane. This forms a 2D photoelectron track image on hexagonal pixel grid; fig. 1.3 gives some examples. The volume of the active gas region between the top Be window and the GEM is 15mm \times 15mm \times 10mm. Figure from Baldini et al. [10].	4
1.3	A selection of simulated photoelectron track events from IXPE GPDs. Pixel color density represents charge deposited, blue crosses show photon x-y absorption point and blue lines the initial photoelectron direction. Pixels are thresholded at 10 counts. Track morphologies vary widely and depend strongly on energy.	5
1.4	Photoelectron angles for a 100% polarized ($p_0 = 1$) 6.4keV simulated line source. Black gives the true photoelectron angles, blue gives the recovered photoelectron angles for the standard moment analysis.	6
1.5	Layout of IXPE GPD readout pixel array. There are approximately 90000 pixels. More pixels allow for better track reconstruction. Figure from Bellazzini et al. [11]. . .	8
1.6	<i>Left:</i> Recovered energy histogram for a 6.4keV line source. A simple linear function of total charge deposited is used here for recovered energy, as in the standard moment analysis. The long low energy tail is produced by events converting in the window or GEM. <i>Right:</i> Fraction of events that are ‘tails’ as a function of energy. Red and blue traces show the Be window and GEM conversion respectively. The jump in GEM conversions at 8.9 keV Cu edge is prominent.	9

1.7	Polarization sensitivity, measured by the modulation factor μ (§1.1.2), as a function of absorption point height in IXPE’s GPD. Dotted lines denote the edges of the GPD, the top at $z=10.83\text{mm}$ is the Beryllium window, and the bottom at $z=0.83\text{mm}$ is the Cu-coated GEM.	10
1.8	Confidence intervals from full posterior distribution (eq.1.18) over polarization parameters (p_0, ϕ_0) given estimates $(\hat{p}_0, \hat{\phi}_0)$ using N events and instrument modulation factor μ . Confidence intervals for 68.2%, 95.5% and 99.7% (1σ , 2σ , 3σ) are displayed. The full marginal posterior distributions are plotted on the wings with 68.2% (1σ) confidence intervals drawn.	14
1.9	Feed-forward neural network with one hidden layer. The neural network takes vector \mathbf{x} as input and produces vector y as output.	16
1.10	Toy regression problem using a deep ensemble. Right panel shows individual DNN prediction functions. Left panel gives the estimated epistemic uncertainty of the deep ensemble. Figure from Antorán, Allingham, and Hernández-Lobato [12].	21
1.11	The three different blazar subclasses based on the frequency of the peak of the synchrotron component in their SED. Low synchrotron peaked blazars (LSP) have peak frequencies lower than 10^{14}Hz , intermediate synchrotron peaked blazars (ISP), $10^{14}\text{Hz} < \nu_{peak_S} < 10^{15}\text{ Hz}$, and high synchrotron peaked blazars (HSP), $> 10^{15}$. Plot reproduced from Abdo et al. [13].	23
1.12	(a) Pure leptonic emission model for example blazar TXS 0506+056. This model cannot produce observed neutrino emission. (b) Hadronic dominated emission model. Ions accelerated in the jet produce secondary leptons via the Bethe-Heitler process or photo-pion production [14]. This model overproduces the X-ray emission. (c) Hybrid emission model; a mostly leptonic emission model with subdominant contribution from hadronic processes. Plots reproduced from Gao et al. [15].	25
2.1	Example square conversions of a 6.4 keV hexagonal track (left panel). The six panels to the right show shifts along the 120° GPD axes; shifting odd rows (upper) or even rows (lower). For each hexagonal track, CNNs are fed column-wise pairs of square conversions, along with the energy, absorption point (blue dot) and initial photoelectron direction (blue line) as labels.	28
2.2	Example tail events (top row) and peak events (bottom row) for three different recovered energies. All plots follow the same spatial and color scale. These tail events convert in the GEM. Color denotes charge deposited in a given detector pixel. Note that tail events are more compact, short and with high charge density at maximum.	33
2.3	Mean squared error on emission angle prediction as a function of photon energy, measured using eq. 2.10, for the classical moment analysis and deep ensemble.	34

2.4	Emission angle recovery for unpolarized, $p_0 = 0$, (left two panels) and 100% polarized, $p_0 = 1$, (right two panels) simulated IXPE data for 4.0 and 7.0 keV. The true photoelectron angle distribution is shown in black; standard moment analysis reconstruction is in blue and deep ensemble in red.	35
2.5	<i>Top:</i> Histograms of the DNN emission angle residual for three different DNN predicted aleatoric VM concentration parameters $\hat{\kappa}^a$. The DNN predicted VM error distributions are over-plotted. The DNN is able to correctly predict its emission angle aleatoric error distribution. <i>Bottom:</i> Histograms of emission angle residuals at three different energies for DNNs and the moment analysis. All methods are unbiased (centered at zero) and at higher energies the DNNs have better recovery of the true emission angle θ , as evidenced by the larger peak heights and lower tails.	36
2.6	Distribution of deep ensemble predicted concentration parameters $\hat{\kappa}$ across the IXPE energy spectrum (left). The right-hand plot shows the $\hat{\kappa}$ distribution for two specific energies 3.0 keV (red), 6.4 keV (blue). High concentration parameter κ means a low predictive uncertainty.	37
2.7	Histograms of the emission angle residual at three different DNN predicted concentration parameters $\hat{\kappa}$ for tail events only. The predicted DNN error distributions are overplotted. When tail events are treated as if they were normal events, the DNNs do not learn to predict the appropriate uncertainties.	40
2.8	<i>Left:</i> The solid curve shows the fraction of peak events retained as a function of the tail probability cut, while the dashed curve shows the fraction of the cut sample from the remaining tail events. Insets show the confusion matrices, normalized by column, for our adopted 70% cut and a 35% cut. <i>Right:</i> The top curves show the peak retention as a function of energy (black solid – 70% cut, red dot-dash – 35% cut). Below, the dotted curve shows the uncut fraction of the sample due to tail events, while the lower black and red curves show the residual tail pollution (70% and 35% cut, respectively). Depending on how harmful tail events are to the desired measurements, different cut levels could be appropriate.	40
2.9	Photon absorption point localization accuracy using the root mean squared error. The deep ensemble predictions (red) do appreciably better than the moment analysis and matches a barycenter estimate at the lowest energies. For IXPE, all methods are adequate, as localization is much better than the size of the point source image produced by IXPE’s mirrors. One quarter of IXPE’s half power diameter is illustrated (black dotted line).	41

2.10	Response for three true energies. Note that while the P21 analysis suppressed the large tail rate seen in the Moments processing, events leaked to a high energy tail for medium to low true energies. Our morphological tail cut further suppresses the GEM/window tail, avoids the high energy leakage and achieves comparable or better peak width.	41
2.11	<i>Left:</i> Mean absolute error in predicted energy as a function of true energy. <i>Right:</i> FWHM ($3.46 \times$ Median Absolute Deviation) of predicted energy distribution for a source at energy E. The grey band marks the limiting energy resolution in the purely-exponential multiplication regime. NNs with 70% tail cut performs better on every metric, more accurate per track and tighter resolution at all energies. All methods suffer from the large increase of tail events above 9 keV.	42
3.1	Modulation factor or optimal event weight as a function of von Mises concentration parameter κ . The function is monotonic, higher concentration parameters are always weighted more.	47
3.2	Probabilistic graphical model representing the hierarchical data generating process for photoelectron track images in an X-ray polarimeter. Shaded nodes are observed.	48
3.3	Measured μ as a function of deep ensemble event weight w for large test data sets of 1 – 10 keV simulated events. Each μ bin contains 20,000 individual track events. Closed circles represent a PL1 source ($dN/dE = E^{-1}$) convolved with IXPE’s effective area, open circles a flat spectrum. Spectra with tail event cuts applied are in black, and full spectra in red.	49
4.1	The polarization fraction and EVPA in the R-band against the time in days during an observed rotation by RoboPol [16] of the LBL blazar J1751+0939. Filled points mark their identification of a ‘rotating’ epoch.	52
4.2	EVPA in the observers frame on the plane of the sky as a function of jet bulk gamma Γ for a helical B-field with a 45° pitch angle. The direction of the jet is 4° off our line of sight in the \hat{x} direction on the plane of the sky, corresponding to 0° on the plot. The B-field is sampled every $\phi_B = 20^\circ$ from $0 - 360^\circ$. The black lines on the y-axis denote the observed EVPAs without RPAR (i.e. $\Gamma = 1$). The solid cyan and blue lines mark $\phi_B = 0^\circ, 120^\circ$ respectively.	53
4.3	Arrows denoting the observed EVPA on the plane of the sky for the jet described in section 4.2 at the 3 different Γ cuts. The black arrows show the initial EVPA absent RPAR (i.e. for stationary $\Gamma = 1$), the orange arrows show the RPAR-shifted EVPAs. The projected jet axis is along the green arrow. Arrows are plotted as single headed for clarity, although observations will span 180° (i.e. -90 to $+90$ in section 4.2).	54

4.4	EVPA bias along jet axis against Γ_{bulk} for a set of 127 random B-field zones in the jet frame. The bias is calculated as the difference in squared EVPA components parallel (\hat{x}) and perpendicular (\hat{y}) to the jet axis. The bias is averaged over 500 random realizations. $\theta_{\text{obs}} = 4.0^\circ$	56
4.5	Map of zones in a jet section [17]. Each zone has a B-field direction, with random initially isotropic orientation during the stochastic phase. During a rotating phase the central zones (shaded) turn helical. The computations typically use 127 zones (6 concentric rings); fewer are shown here for clarity.	57
4.6	Schematic showing the shape of the ‘conical’ jet model with some of the important parameters labelled, [18]. A ‘section’ of the jet is a dx slice.	58
4.7	EVPA (blue) and polarization fraction (red) for stochastic and rotating phases in a jet with projection on the plane of the sky along the x axis. We are viewing at $\theta_{\text{obs}} = 4.0^\circ$ with $\Gamma = 5$ and $\theta_{\text{op}} = 9.5^\circ$. RPAR has been applied. This low Γ means the vast majority of zones in the jet contribute. The black line shows the same helical rotation without RPAR or turbulence. The rotation phase spans 770° with starting phase $\phi = 3\pi/2$	59
4.8	As for section 4.3 for the rotational phases of a conical jet viewed at $\theta_{\text{obs}} = 1.5^\circ$. This represents a 770° rotation with starting phase $\phi = 160^\circ$. Stepping is less prominent.	60
4.9	Energy variation for EVPA and Π during a non-rotating (stochastic) phase. Colors denote low (Radio, red), mid-range (Optical, green) and near cut-off (e.g. X-ray, blue) synchrotron bands. This is for a jet with $\theta_{\text{obs}} = 1.5^\circ$, $\Gamma = 5$ and $\theta_{\text{op}} = 9.5^\circ$. Both panels use the random seed for the generated B-fields. The top panel shows a RPAR affected jet: substantial energy dependence is seen, with low energy (radio) PA better aligned with the parent jet. Low energy fluctuations are also somewhat smaller. In the lower panel RPAR effects are ignored and the behavior is essentially achromatic with lower overall polarization fraction.	62
4.10	Energy dependence during a rotating phase (the 770° rotation shown in section 4.3), as for fig. 4.9. Strong sinusoidal variation of the radio polarization fraction occurs as the rotating helical zone EVPA aligns parallel and perpendicular to the jet axis. Since the random zones’ EVPA are increasingly aligned with this axis as one progresses along the jet (i.e. observes at lower frequency; see top panel, fig. 4.9) the total Π has a sinusoidal modulation. The strength of the modulation depends on the ratio of helical zones to random zones.	63

4.11 A fit to Mrk 501's SED (bottom panel, approximately contemporaneous fluxes drawn from the ASDC compilation for 2010; https://tools.asdc.asi.it/SED/) using the radiative model. The top two panels show the energy dependence of Π and EVPA averaged over 200 iterations for the stochastic phase, with solid line for a low δ solution and a dashed line for high δ . The observed EVPA is expected to fluctuate about these mean values during stochastic phases. The jet projection on the plane of the sky is $\theta_{\text{sky}} = -40^\circ$, as observed for Mrk 501.	65
5.1 Sketch of the SSC scattering geometry to illustrate the definition of angles. The coordinate system is chosen such that the magnetic field B lies in the (x-z) plane and the scattered photon direction $\mathbf{v}_{\mathbf{k}'}$ is along the z axis.	68
5.2 The polarization fraction, observed PA angle and SED for a jet with 19 independent B-field zones jet for a single dx step (i.e. no cooling). The SSC emission is calculated using only local synchrotron seed photons (i.e. evaluating Eq. (17) instantaneously). Published models typically use such a single zone. Red denotes SSC, blue synchrotron and black their combination. The jet parameters are tabulated in the Appendix. . .	72
5.3 Plot of an example full jet synchrotron SED (blue solid) made up of emission from 19 B-field zones. The zones are grouped into 5 different Doppler factor sets. Dashed lines show the SEDs from zone sets with differing $\theta_{\text{obs},i}$. We can see that the total emission (blue solid) has contributions from fewer total zones at high energy. Once the SED is exponentially dominated, the zone with the highest γ_{max} dominates the observed flux.	73
5.4 Effect of combined Doppler boosting and γ_{max} spread at the synchrotron SED. Each dashed line now represents an individual zone. Again, N_{eff} drops at high energy. . .	74
5.5 <i>Bottom Panel:</i> The number of zones contributing half of the observed flux (normalized to the number contributing at the SED peak) for a typical full jet blazar synchrotron SED. The blue lines represent the dependence for the Doppler effect of fig. 5.3. The solid line shows a jet with large θ_{op} viewed near its edge (large Doppler spread) while the dashed line shows a small θ_{op} viewed near the axis (small Doppler spread). Red curves give the case where γ_{max} varies between zones. The solid red curve is the scenario proposed in Marscher [17] where γ_{max} depends on $B_{ }$, the shock-aligned field (here we assume a transverse shock). The dashed red curve represents the same γ_{max} distribution as the solid line case, but with random B orientation. <i>Top Panel:</i> The corresponding polarization fractions; colors and lines as below.	75
5.6 Simplified schematic displaying the light travel effect described in §4.2. The jet is split into horizontal slices of width dx_{ch} , the scale in the jet frame over which the magnetic field vectors decorrelate. These B -field vectors are denoted by the red arrows. Note that for jets viewed at large off-axis angles multiple slices are viewed at a given time.	77

5.7	<i>Top panel:</i> Synchrotron energy density as a function of distance along the jet x for three different energy bands (red: mm wavelength, cyan: optical, blue: 10 keV). Solid lines are for the jet center, while the dashed lines are for the edge. The black vertical dotted line represent $x = R_0$. <i>Middle panel:</i> Fraction of the integrated synchrotron density due to the local zone (solid line central zone, dashed line edge zone). Neighboring zones dominate for $x > R_0$. <i>Bottom panel:</i> Build up of the on-axis synchrotron seed photon population at three energies. The dotted blue line represents the corresponding 1-10 keV SSC flux. These are fluxes measured in the jet frame; jet parameters are as in §5.	79
5.8	SSC power in the jet frame as a function of the (lab frame) seed photon energy – i.e. the seed photon spectrum for a given Compton energy. The solid line is for keV SSC emission while the dashed line shows the seed spectrum producing MeV emission. For the overall SED and assumed jet parameters, see §5.1.	80
5.9	The geometrical effects of Doppler boosting on the number of effective emitting zones. Jet frame θ_{op} increases from top to bottom; the fraction of zones within our $1/\Gamma$ effective viewing circle decreases (blue curves). As θ_{obs} increases off-axis, the fraction of the zones in the $1/\Gamma$ viewing cone also decreases, until $\theta_{\text{obs}} > 1/\Gamma$ and more (distant) zones are boosted at a similar level, causing N_{eff} to grow again. The exact functional form is controlled by the weighting function, in this case the Doppler factor D^4 , and the jet geometry, in this case a cone. We note that treating N_{eff} as the number of zones that dominates half of the flux is not a perfect proxy, the true N_{eff} is controlled by a weighted distribution. The red lines show the same relationship for a $\theta_{\text{op}} = 40^\circ$ jet affected by LT effects (§4.2). At low energies (e.g. mm wavelengths), the emission occurs at sufficiently large radii (see fig. 5.7) that light travel effects substantially increase N_{eff} for jets viewed off-axis.	81
5.10	<i>Left Panel:</i> The optical synchrotron (red) and X-ray SSC (blue) polarization fraction plotted against N_{eff}/N . Triangle points are for an $N = 19$ zone jet, circles for $N = 37$. Solid lines show a Doppler-dominated jet. For the dotted lines zone intensities are randomly distributed. The synchrotron cyan dashed lines show $N_{\text{eff}}^{-1/2}$, while the SSC equivalents show the best fit power laws ($N=37$ gives $p = 0.29$; $N=19$, $p = 0.31$). <i>Right Panel:</i> The SSC/synchrotron polarization ratio <i>versus</i> N_{eff}/N (line types as in left panel). Note that the SSC angle-averaging is more effective for a randomly distributed jet, leading to a lower ratio. Values are averaged over many simulations with isotropic B-fields; error bars show the residual errors on the mean. The jet parameters used in the simulations are given in the Appendix Table.	82

5.11 Correlation plot of the X-ray SSC polarization fraction (top) and PA (bottom) against the corresponding optical synchrotron for two different viewing angles (N_{eff}) (red: $\theta_{\text{obs}} = 4.5^\circ$, black: $\theta_{\text{obs}} = 0.1^\circ$). Note Π_{SSC} can persist even when chance cancellation makes Π_{Sync} for the Earth line-of-sight; this leads to high $\Pi_{\text{SSC}}/\Pi_{\text{Sync}} >> 1$ and a non-zero intercept. EVPAs are strongly correlated, but show scatter due to similar differences in the zone-sampling for the Earth line-of-sight.	83
5.12 Example fully treated jet SED for a single random B-field draw, with $\theta_{\text{obs}} = 6.0^\circ$ and the remaining jet parameters as in fig. 5.10. This simulation uses 19 B-field zones and serves to display some typical polarization behavior. Note the substantial jump in EVPA between the Sync and SSC components. Note also the substantial Π increase above the synchrotron and SSC peaks due to the decreased N_{eff} and the average Π_{Sync} decrease for sub-infrared energies due to §4.2. The shaded bands show the standard deviations in individual realization over many isotropic B-field draw; EVPA average has no preferred direction for either component, with an average value 0.	84
6.1 Example of a unimodal (top panel) and a bimodal (bottom panel) flux distribution in the 2-10 keV energy range in log-space. The black dashed line shows the best-fit model. In the bimodal case, the red lines show the individual components.	88
6.2 The dashed lines show the shift from observed (optical, “x”) to predicted (X-ray, “dots”) Π for a few objects in each source class: LSP (black), ISP (red) and HSP (green). The diagonal lines show the <i>IXPE</i> sensitivity for a source with photon index ≈ 2 in a given exposure time. Inset shows the frequency dependence of the effective zone number using the median Lorentz factor and viewing angle from [119].	92
6.3 Predicted X-ray polarization degree versus the X-ray flux: LSP (black), ISP (red) and HSP (green). Lines show the 1σ variability range in each quantity. Sources lacking at least 3 observations each in S_X and Π_O are shown by open squares. Stars have > 3 measurements in one quantity, solid squares have > 3 in both. Red dashed lines show simple estimates for <i>IXPE</i> sensitivity for a source with photon index ≈ 2 in a given exposure time.	93
6.4 Required <i>IXPE</i> exposure time as a function of (presently unknown) optical polarization level. This Π_O has been corrected to an expected Π_X , using the sources SEDs. The best prospects are labeled. The sources are colored according to their SEDs: LSPs (black), ISPs (red) and HSPs (green).	96
6.5 Polarized leptonic jet model fits to all blazars and states. ‘(f)’ denotes a flaring state. Black traces show the expected total SED for the best fit jet parameters. Grey shaded regions around the black trace show 1σ model deviations due to different random magnetic field zones. Vertical grey lines denote <i>IXPE</i> ’s sensitive band, 1 – 10 keV, and insets show close-ups of this region.	102

6.6	Leptonic (SSC) jet model polarization fraction predictions. The jet models used in each panel are the same as those in Fig. 6.5. Black observations denote the average measured optical polarization over multiple epochs (MJD 56432 – 57893 for Robopol measured S5 0716+714 and J0211+1051 [19] and MJD 58019 – 58267 for TXS 0506+056). Lines and shaded regions mean the same as Fig. 6.5 with the addition of two transparent models, which represent randomly selected model realizations.	104
6.7	Violin plots of the true polarization fraction distribution (blue, left-hand-side) and the measured polarization fraction distribution (red, right-hand-side) for 2 – 4 keV and 4 – 8 keV energy bins, and 300 ks/500 ks exposures. The distributions of true polarization fractions are extracted from our jet model fits (Fig. 6.6) and are the same for both exposure times. Measured polarization fraction distributions assume a single observation with true polarization equal to the expected value. Dashed black traces represent the minimum detectable polarization (MDP_{99}) for each measurement bin.	106
7.1	Quasi-simultaneous X-ray spectra during the <i>IXPE</i> observation. The fitted model is the sum of two absorbed power-laws and an apec component. Dotted lines in the top panel show the individual power-law and apec components. The middle panel shows the polarized spectra with constant polarization for the low-energy component; the high energy component has fixed polarization degree $PD=0$. The significance of the spectral fit residuals is shown in the bottom panel.	111
7.2	Polarization fraction and EVPA confidence levels (68%, 95%, 99%) for the first time bin over the 2–4 keV energy range, fig. 7.3. Gray shaded region represents the VLBI-determined jet axis projection on the plane of the sky [20].	113
7.3	Light curves, polarization fractions, and EVPAs in the optical (top three panels) and X-ray (bottom three panels) bands. X-ray soft power flux contribution fraction is shown in the middle panel. In the X-ray panels, the data are split into two energy bins and three equal time bins. Times of observations from satellites other than IXPE are indicated in the flux panel. Errors are 68.3% (1σ) confidence intervals and upper limits are 99% confidence. Shaded horizontal regions in EVPA panels represent the jet axis projection on the plane of the sky [20]. Dotted lines in the X-ray polarization-fraction panel represent the MDP_{99} level for each bin individually. For X-ray polarization measurements, only the first time bin at 2–4 keV exceeds MDP_{99} .	117

8.1	(a) A helical magnetic field moving through a stationary acceleration/emission region. The helix advances along $+z$ with B_3 entering the active region first. (b) The same helix as in (a) viewed from 70° . (c) Observed EVPA rotation pattern from synchrotron emission in the stationary acceleration zone as viewed from (a) solid, (b) dashed. Labelled red lines denote specific phase points highlighted in (a) and (b).	120
8.2	Full Mrk 421 rotation fit for the highest marginal likelihood model, the boosted helix. Throughout, model traces are multiple samples from the posterior distribution and points are binned observation derived with weighted NNs. Left panel gives observation binned count spectra with model flux variation overlayed. Middle panel shows binned observed EVPAs with rotation model overlayed. Grey shaded region gives Mrk 421 43GHz jet axis direction [20] and horizontal lines are the fitted model jet axis predictions. Right panel is the binned observed polarization fraction, with a step plot denoting the MDP99s for each bin. Horizontal lines are the model polarization fraction prediction.	126
8.3	Same as fig. 8.2 but for the highest marginal likelihood model on the first Mrk 421 observation only. Here, a rotation with added constant polarization component is strongly preferred. This model predicts a variation in the polarization fraction shown in the right hand panel. Red horizontal lines in the right hand panel are posterior samples of the constant polarization fraction, black lines are the polarization fraction of the rotation component.	127
8.4	Same as fig. 8.2 but for the highest marginal likelihood model on the second Mrk 421 observation only, a boosted helix.	128
8.5	Posterior distribution of the preferred boosted helix model for the joint Mrk 421 dataset.	129
B.1	<i>IXPE</i> light curve. Black lines are three equal time bin separators; red shaded regions indicate XRT exposures.	136
B.2	Polarization fraction and EVPA confidence levels (68%, 95%, 99%) for the low-energy power-law in fig. 7.1. Gray shaded region represents the VLBI-determined jet axis projection on the plane of the sky [20].	137
B.3	(a) Absorbed sum of two power-laws with apec fit to first <i>IXPE</i> time bin with first <i>Swift</i> XRT observation. Photon indices and apec temperature are fixed to values in table 7.1. The high-energy power-law polarization is fixed to zero. (b) Second <i>IXPE</i> time bin with second XRT observation. (c) Third <i>IXPE</i> time bin with second XRT observation. (d) Polarization fraction and EVPA confidence levels (68%, 99%) of the low-energy PL for each time bin: black – T1, red – T2, blue – T3. Grey shaded region represents the VLBI-determined jet axis projection on the plane of the sky [20]. . .	140

B.4 Optical BVRI observations of BL Lac before and during the 2022 November 27–30 <i>IXPE</i> pointing, showing brightness in magnitudes (top panel), polarization degree (middle panel), and polarization position angle (bottom panel). The duration of the <i>IXPE</i> observation is marked by the gray shaded area.	141
B.5 <i>Fermi</i> -LAT 3-day cadence light curve, showing the outburst that triggered the <i>IXPE</i> observation. A single power-law model with a free photon index is used to determine the photon fluxes. The two gray-shaded regions mark two previous <i>IXPE</i> observations, while red shaded are indicates the observation discussed in this paper.	142

Chapter 1

Introduction

Polarization is a fundamental property of light; along with energy and direction, the state of a photon can be fully described by two additional degrees of freedom: its polarization state. In the classical limit, polarization manifests as the 2D oscillation direction of the light wave's electric field in the plane perpendicular to the direction of motion. As such, polarization is intimately connected to the oscillation geometry of the light source. Polarization can be induced by ordered oscillations of charged particles such as during scattering, birefringence, and synchrotron radiation. All of these processes are extremely relevant in high-energy astrophysics – measuring polarization can thus reveal information about astrophysical sources inaccessible using only brightness and frequency.

Optical polarization measurements have been available since the turn of the 19th century with the invention of the Wollaston prism and similar types of polarizers. In principle, the technology has been available since the observation of birefringence in calcite crystals by Erasmus Bartholinus in 1669. Such prisms split incident light into two independent polarization states measured separately. X-ray polarization measurements, on the other hand, became possible only very recently. Novick et al. [21] made the first astrophysical measurement of X-ray polarization when they detected the Crab nebula using Thomson scattering and Bragg reflection polarimeters. Since then, measuring X-ray polarization has become a major goal in astrophysics. The recent development of photoelectron tracking detectors [11] has greatly improved the prospects of doing so. The gas pixel detector (GPD) [22] has brought soft X-ray polarimetry (1-10 keV) to the PolarLight CubeSat test [23], NASA's just launched Imaging X-ray Polarimetry Explorer (IXPE) [24], and the potential Chinese mission, eXTP [25]. Indeed, IXPE is the first dedicated X-ray polarimeter since the 1970s, opening a new window on the universe.

X-ray polarization measurements offer rich opportunities to probe the magnetic field topology and emission physics of high energy astrophysical sources, such as accreting black holes and astrophysical jets [26, 27]. One especially interesting class of source is blazars. Blazars are relativistic jets pointed very close to our line of sight emanating from active galactic nuclei, bright supermassive black holes

at the centers of galaxies. They are detectable across the whole electromagnetic spectrum, are the most consistently luminous sources in the universe, and display high variability down to timescales of minutes [28]. Despite great improvements in our understanding of these objects since their discovery with the advent of radio astronomy in the 1930s, there remain many open questions. Most of the emission from blazars can be attributed to high-energy leptons in the jet emitting synchrotron and inverse-Compton radiation; however, recent observations by IceCube collaboration have connected blazar TXS 0506+056 to multiple astrophysical neutrino events [29], suggesting hadronic emission processes in blazars. In the optical and radio, blazar polarization vectors have been known to continuously rotate up to more than 720° over periods of days or weeks. These polarization rotations have so far defied simple explanation; they are thought to arise from helical magnetic fields or magnetic reconnection. Finally, particle acceleration and jet launching processes each have multiple competing mechanisms with a lack of consensus on their relative contributions. Measuring X-ray polarization can shed light on all of these fundamental questions – X-ray polarimetry probes jet geometry and magnetic field orientation closer to the jet acceleration region than ever observed before. Blazars’ high variability means a very sensitive instrument is required to extract meaningful polarization measurements.

X-ray polarimeters based on GPDs measure polarization by tracking photoelectrons excited by individual incoming source photons and estimating their emission direction. Their sensitivity is limited by the data analysis algorithm used to recover source polarization, spatial structure, and energy given a measured set of photoelectron track images. Deep neural networks have achieved state-of-the-art performance on a wide variety of machine learning tasks and are becoming increasingly popular in domains such as speech recognition [30], natural language processing [31], bioinformatics [32] and especially computer vision [33]. Going from photoelectron track images to emission angle estimates can be classified as a computer vision problem, so it is not surprising that neural networks would be well suited to track reconstruction. The Cherenkov Telescope Array (CTA) [34] team have applied deep learning methods to differentiate between cosmic rays and gamma rays. Notably they also have to deal with a hexagonal pixel grid, also found in imaging X-ray polarimeters. The IceCube collaboration has begun the use of graph neural networks to identify 3D neutrino tracks [35].

This thesis brings together recent advances in machine learning and X-ray polarimetry to improve the state-of-the-art in X-ray polarimetry and significantly increase science output of IXPE and future high-energy polarimetry missions. I use these sensitivity improvements to make progress towards answering longstanding questions about blazar jets, including the nature of polarization rotations and the emission mechanism behind their observed radiation.

1.1 Preliminaries

Throughout this thesis, I regularly use the following standard notation:

- p_0, Π – Source polarization fraction.
- ψ_0, θ_0, ϕ – Source electric vector position angle (EVPA).
- θ – Photoelectron emission angle.
- Γ – Bulk Lorentz factor.
- \sim – ‘is distributed as’.

1.1.1 Imaging X-ray Polarimetry

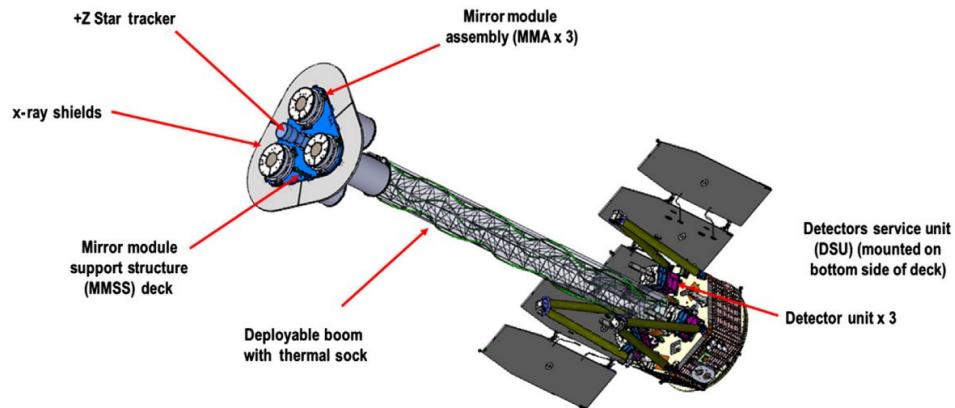


Figure 1.1: The IXPE observatory fully deployed for science observations. Three cylindrical mirror units direct X-rays to three corresponding detector units (gas pixel detectors). Figure from Ramsey et al. [9].

The IXPE observatory consists of three mirror modules and three detector units, fig. 1.1. The mirror modules are each grazing incidence, Wolter-I type optics made up of 24 600mm long shells of electroformed nickel-cobalt alloy and have a focal length of 4m [9]. These mirrors result in a 19 – 28 arcsec angular resolution (half power diameter) and a 12.9 arcmin field of view. The entire observatory has a mass of 170kg and measures 5.2m in length at full extension [36]. The three detector units at the base of the observatory house gas pixel detectors (GPDs), each rotated 120° with respect to the other two.

X-ray polarization telescopes using GPDs, like IXPE, directly image electron tracks formed from photoelectrons scattered by incoming X-ray photons. This technique has the capability to build up an image of extended sources (e.g., supernova remnants or pulsar wind nebulae). Fig. 1.2 shows a

schematic of the GPD and fig. 1.3 gives example photoelectron tracks at various photon energies, measured by IXPE's GPDs. IXPE's GPDs consist of an active gas volume 15mm × 15mm × 10mm sandwiched between a 50 μm Beryllium window and a 0.7mm gas electron multiplier (GEM) [10]. The GEM uses a 470V gain to amplify the photoelectron track signal, which is readout on a hexagonally pixelated application specific integrated circuit (ASIC). GPD sensitivity is limited by the track analysis algorithm used to recover source polarization, spatial structure, and energy, given a measured set of electron track images.

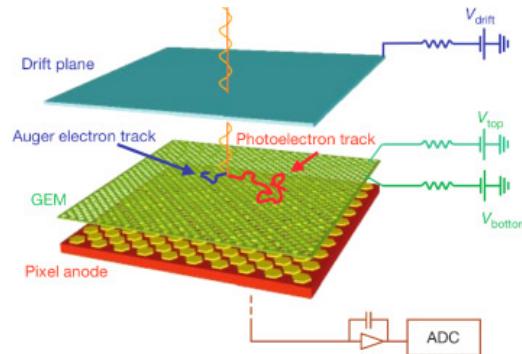


Figure 1.2: Design of an IXPE GPD. An incoming X-ray photon interacts with a dimethyl ether gas molecule to produce a photoelectron between the entrance window and the gas electron multiplier (GEM). Secondary charges created along the photoelectron's path are multiplied by the GEM and read out at the anode plane. This forms a 2D photoelectron track image on hexagonal pixel grid; fig. 1.3 gives some examples. The volume of the active gas region between the top Be window and the GEM is 15mm × 15mm × 10mm. Figure from Baldini et al. [10].

In the 1 – 15 keV range, the differential cross-section for emission of a K shell photoelectron [37],

$$\frac{d\sigma}{d\Omega} = 4\sqrt{2}r_0^2 \frac{Z^5}{137^4} \left(\frac{m_e c^2}{h\nu} \right)^{\frac{7}{2}} \frac{\sin^2 \psi \cos^2(\theta - \theta_0)}{(1 - \beta_e \cos \psi)^4}, \quad (1.1)$$

where θ_0 is the normal incidence X-ray's EVPA, θ the initial azimuthal emission direction of the photoelectron, ψ is the polar angle from the incoming photon direction, Z is the atomic number of the target atom, and β_e is the emitted photoelectron velocity over c . In this energy range, photoelectrons are emitted predominantly transverse to the incoming photon since β_e is small. IXPE's GPDs, use Dimethyl Ether (DME), CH_3OH_3C , as the active gas in the detector, fig. 1.2. K-shell electron targets are mostly provided by the Oxygen O . Given a photon source with polarization fraction p_0 and EVPA θ_0 , the azimuthal emission angles θ thus follow the distribution

$$\theta \sim \frac{1}{2\pi} (1 + p_0 \cos[2(\theta - \theta_0)]). \quad (1.2)$$

By measuring a large number of individual photoelectron emission angles $\{\hat{\theta}_i\}_{i=1}^N$, one can recover the

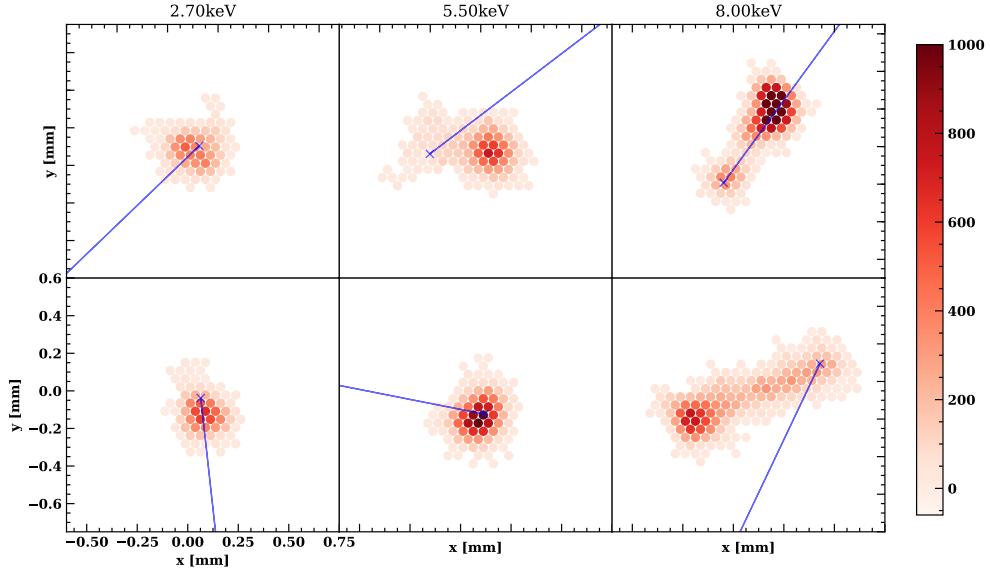


Figure 1.3: A selection of simulated photoelectron track events from IXPE GPDs. Pixel color density represents charge deposited, blue crosses show photon x-y absorption point and blue lines the initial photoelectron direction. Pixels are thresholded at 10 counts. Track morphologies vary widely and depend strongly on energy.

above distribution to extract the source polarization parameters: polarization fraction ($0 \leq p_0 \leq 1$) and EVPA ($-\pi/2 \leq \theta_0 < \pi/2$). In practice, the recovery of photoelectron emission angles from track images is imperfect. Track images are noisy due to Coulomb scattering and diffusion, and, especially for low energies, are often barely resolved. For example, in fig. 1.4 the distribution of recovered emission angles using IXPE’s classical track reconstruction is significantly blurred compared to the true distribution.

The best classical track reconstruction method for GPDs is a moment analysis described by Bellazzini et al. [11] and in the previous chapter. Impressive accuracies for the emission angle and photon absorption point are achieved from a simple weighted combination of track moments. Photon energy estimates are proportional to the total collected GPD charge for a track. The track ellipticity is quantified to provide a rough proxy for track reconstruction quality. High ellipticity tracks typically have more accurate angle estimates. However, simple moments cannot capture all image information, especially for long high energy tracks, and so a more sophisticated image analysis scheme can lead to improved track emission angle, absorption point, and energy recovery.

Once the emission angles have been reconstructed, $\{\hat{\theta}_i\}_{i=1}^N$, classical polarization estimation uses a maximum likelihood estimator (MLE) or direct curve fit to calculate $(\hat{p}_0, \hat{\theta}_0)$. This assumes individual tracks contribute equally to the final polarization estimate. In fact, photoelectron tracks

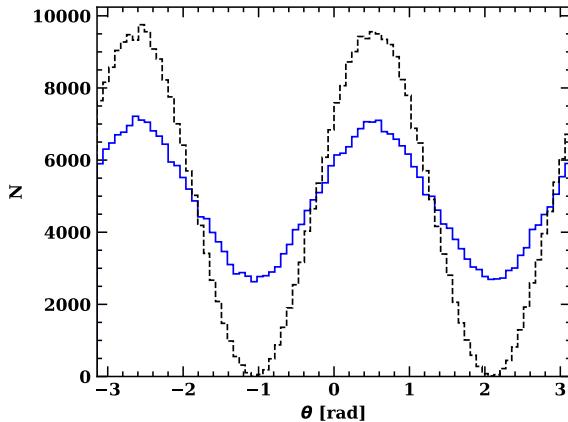


Figure 1.4: Photoelectron angles for a 100% polarized ($p_0 = 1$) 6.4keV simulated line source. Black gives the true photoelectron angles, blue gives the recovered photoelectron angles for the standard moment analysis.

are very morphologically diverse, even for the same photon energy (cf. fig. 1.3), and so emission angle estimates are highly heteroskedastic: some emission angles are much better estimated than others. This is especially true for tracks of different photon energies, important in the case of broadband polarization estimates. Assuming an equal contribution from all estimated emission angles results in sub-optimal polarization recovery. To ameliorate this, IXPE’s standard analysis performs an event cut removing tracks in the bottom 20% of estimated ellipticities; this does provide a marginal improvement in the recovered polarization signal. However, a detailed analysis with proper estimation of emission angle uncertainties and their inclusion in the likelihood function can substantially improve the signal-to-noise ratio of recovered polarization.

Measuring the X-ray polarization of a source using an imaging X-ray polarimeter is, from a data analysis perspective, a two step process.

1. *Track reconstruction.* The telescope is pointed at the source and collects photoelectron track images, see fig. 1.3. Individual track images are processed to extract all of the relevant features: emission angles, absorption points and photon energies.
2. *Polarization estimation.* The extracted emission angles from many individual photoelectron tracks are combined to estimate the source polarization parameters.

For spatially extended, time-varying or spectrally varying sources, the extracted emission angles could be grouped based on absorption points, arrival time or photon energy, i.e. in spatial, time and/or energy bins. Polarization estimation from these grouped emission angles follows the same procedure.

This chapter describes classical approaches for both steps, a prerequisite for the neural network approach. In classical polarization analysis the two steps are entirely disconnected. §3 will show how these two steps can be more closely connected using neural network uncertainty estimates.

Track reconstruction

Photoelectron track images contain all the extractable information in imaging X-ray polarimeters. More specifically, the various sensitivities of an imaging X-ray polarimeter can be attributed to a few photoelectron track features that can be extracted from individual track images:

- *Polarization sensitivity*: photoelectron emission angle θ , §1.1.1.
- *Spatial sensitivity*: photon absorption point (x, y) , §1.1.1.
- *Spectral sensitivity*: photon energy E , §1.1.1.

Tracks are imaged by charge deposition onto an array of pixels. For example, IXPE’s GPD pixel array is a $\sim 15\text{mm} \times 15\text{mm}$ square tiled with hexagonal pixels on a $50\mu\text{m}$ pitch [10]; see fig. 1.5. The charge deposited in each pixel is measured in integer counts, typically ranging from 0 to 1000+, fig. 1.3. Track images are identified on the pixel array by clustering analysis and the region of interest is cropped. Each track image comes as a list of (x, y, c) tuples: the pixel coordinates x, y and charge deposited c .

There are a number of instrument specific settings that affect track morphology. For example, a pixelwise charge threshold is usually applied to individual track images to reduce small noisy pixel clusters. The GPD gas pressure and composition can also be varied; lower pressure allows for longer tracks that are easier to reconstruct but reduces the detector quantum efficiency. The methods presented in this chapter generalize to all imaging polarimeter settings. The examples shown here assume IXPE-specific GPD settings.

Emission angle reconstruction

Source polarization is encoded in the photoelectron emission angles θ . These are the image plane projected initial directions of the photoelectron at the interaction point. Track images capture the entire track of the photoelectron in the image plane, so in principle it should be possible to recover the initial direction. In practice, track images are relatively noisy representations of the photoelectron path. Charge diffusion means the further away from the pixel array a photoelectron is emitted, the less well defined its track image will be.

Track images are typically asymmetric and follow a similar pattern: the photoelectron starts at the absorption point and moves along its initial direction. Thanks to the low atomic number of the target Oxygen atom in DME, an Auger electron is also almost always isotropically released at the absorption point [37]. As the photoelectron moves away from the absorption point, it is

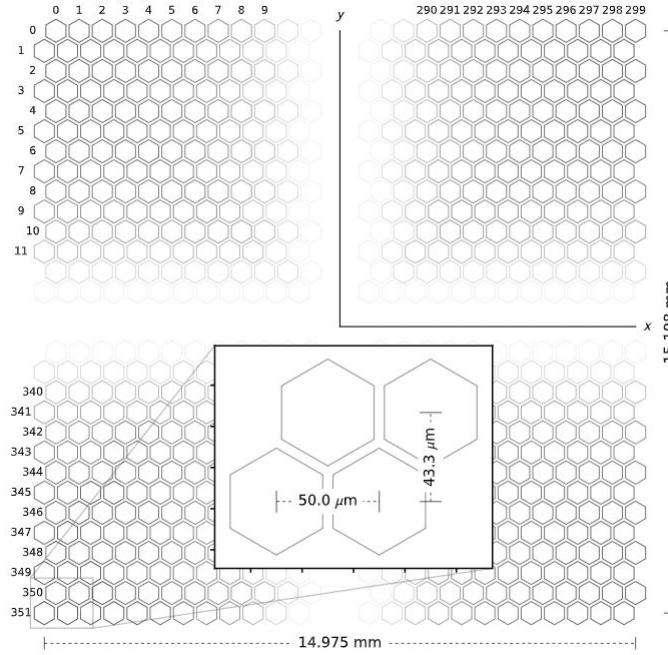


Figure 1.5: Layout of IXPE GPD readout pixel array. There are approximately 90000 pixels. More pixels allow for better track reconstruction. Figure from Bellazzini et al. [11].

Coulomb scattered away from the initial direction. As the photoelectron slows down, it ionizes more gas molecules, depositing more charge in the pixel array and culminates in a final Bragg peak. Low energy photoelectrons travel less far from the absorption point and leave nearly circularly symmetric track images.

The scattering of the photoelectron from its initial direction and short, symmetric photoelectron tracks make emission angle reconstruction challenging. The moment analysis described in Bellazzini et al. [11] uses the known structure of photoelectron tracks to improve their emission angle estimates. They initially calculate the principal axis of the track image by maximizing the second charge moment about the track barycenter. This provides a simple initial estimate of the emission angle, but is skewed by the Bragg peak – the dense final part of the track that is usually not correlated to the initial direction because of Coulomb scattering. To correct for the Bragg peak bias, the location of the Bragg peak is calculated by observing the sign of the third charge moment about the track barycenter. This identifies which side of the track contains the Bragg peak. The absorption point is estimated as being a multiple of the second moment away from the track barycenter along the principal axis, away from the Bragg peak. Finally, the emission angle is calculated by using the second moment about the absorption point, using only pixels in the initial half of the track (those not on the Bragg peak side). See the previous chapter for more details on the moment analysis.

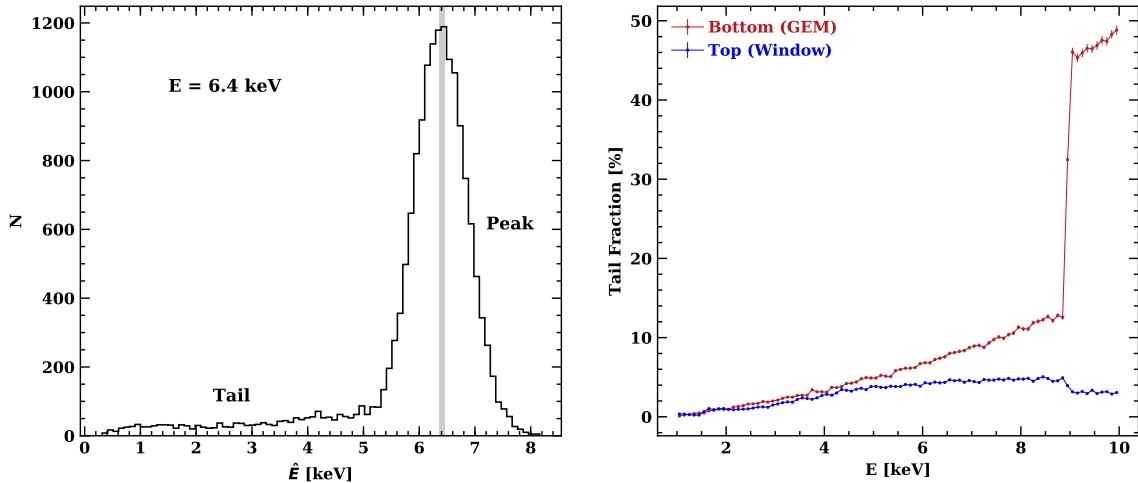


Figure 1.6: *Left:* Recovered energy histogram for a 6.4keV line source. A simple linear function of total charge deposited is used here for recovered energy, as in the standard moment analysis. The long low energy tail is produced by events converting in the window or GEM. *Right:* Fraction of events that are ‘tails’ as a function of energy. Red and blue traces show the Be window and GEM conversion respectively. The jump in GEM conversions at 8.9keV Cu edge is prominent.

Absorption point reconstruction

Photon absorption points (x, y) on the image plane along with the known point spread function (PSF) of the telescope optics allow for the spatial resolution of extended sources. Better absorption estimates yield better spatial resolution. A simple estimator for the absorption point of an individual photon from its track image is the track barycenter, but this can be strongly biased by the photoelectron Bragg peak §1.1.1. In moment analysis [11], the absorption point is estimated jointly with the emission angle by identifying and excluding the Bragg peak. For low energy symmetric tracks where it is impossible to separate the Bragg peak, the track barycenter is used as the standard absorption point estimator.

Energy reconstruction

An X-ray photon’s energy is proportional to the average charge deposited in the detector. Since the charge deposited is subject to statistical fluctuations, the energy resolution is limited and this limit depends on the effective detector Fano factor. Classical polarimetry uses a simple linear model calibrated on real detector track images to reconstruct the photon energy. Thus, the total track charge is assumed proportional to the photon energy. This method approaches but does not meet the theoretical best energy resolution.

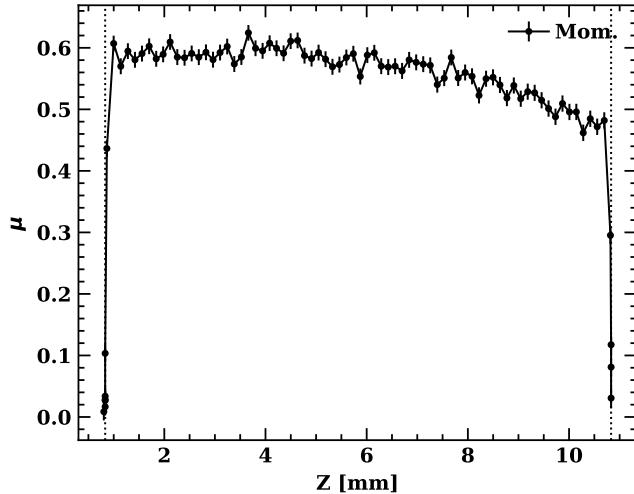


Figure 1.7: Polarization sensitivity, measured by the modulation factor μ (§1.1.2), as a function of absorption point height in IXPE’s GPD. Dotted lines denote the edges of the GPD, the top at $z=10.83\text{mm}$ is the Beryllium window, and the bottom at $z=0.83\text{mm}$ is the Cu-coated GEM.

Events converting outside of the gas volume

So far, the track images considered are formed when photons interact with the detector gas. However, incoming X-ray photons can also convert in detector components just outside of the main gas volume, with electrons penetrating the gas triggering the GPD and producing photoelectron tracks. For IXPE’s GPDs, these external interactions occur in the thin beryllium entrance window at the top of the GPD and in the gas electron multiplier copper material at the bottom, fig. 1.2. The track images formed by these external interactions have different morphological properties compared to those in the detector gas. Notably, they have less total charge deposited for the same photon energy, because some of the photoelectron energy can be lost in the solid window or GEM material. In addition they tend to have a denser ‘core’ since the photoelectrons surviving to the gas volume tend to have trajectories with a large vertical component. The left panel of fig. 1.6 shows the linearly reconstructed energy (§1.1.1) for a 6.4keV line source. External events form a low energy tail on the predicted energy histogram – external events are described here tail events, producing tail tracks, while events converting in the detector gas are called peak events. The right panel gives the population fraction of tail events as a function of photon energy. Tail tracks are mostly subdominant, but with a linearly increasing fraction as photon energy increases. At energies above 8.9 keV, the Copper K absorption edge causes a large increase in tail events in the GEM, increasing the tail track fraction.

Tail tracks also have significantly reduced polarization signal compared to peak tracks. The emitted photoelectron is scattered more strongly in the dense GEM and window materials than in the detector gas; as a result, the initial photoelectron direction is often irrecoverable. Fig. 1.7 gives

the simulated polarization sensitivity as a function of vertical direction along IXPE’s GPD, using the moment analysis to estimate photoelectron emission angles. There is a clear drop in sensitivity at both vertical edges of the detector where tail tracks dominate the absorption. Additionally, this figure serves to illustrate the decrease in polarization sensitivity with increasing vertical distance from the pixel array. This is due to charge diffusion in the detector gas blurring tracks before they reach the pixel read-out array (§1.1.1).

If included, tail tracks can cause problems during polarization estimation. Not only do tail tracks degrade the detector polarization sensitivity, the energy-dependent tail fraction can make it difficult to properly calibrate the polarization sensitivity and degrade the energy resolution for continuous spectra. In classical imaging polarimetry, tail tracks are cut from the line source calibration by excluding the low energy tail from the recovered energy histogram, fig. 1.6. For real, continuous sources this is not possible, so tail tracks remain an important problem for classical polarimetry approaches. §2.5 will describe how a NN analysis can identify tail tracks based on their morphological differences and decrease their contamination of the polarization signal.

1.1.2 Polarization estimation

Polarization estimation can begin once all relevant track features have been recovered. The basic problem is to estimate source polarization parameters (p_0, θ_0) and their uncertainties from a set of reconstructed emission angles $\{\hat{\theta}_i\}_{i=1}^N$. As described in the introduction, true emission angles θ exhibit a sinusoidal modulation with period π that depends on the source polarization

$$p(\theta|p_0, \theta_0) = \frac{1}{2\pi}(1 + p_0 \cos[2(\theta - \theta_0)]), \quad (1.3)$$

where $0 \leq p_0 \leq 1$, $-\pi/2 \leq \theta_0 < \pi/2$ and $-\pi \leq \theta < \pi$. Reconstructed emission angles $\hat{\theta}$ are imperfectly recovered, so these follow an adjusted distribution

$$p(\hat{\theta}|p_0, \theta_0) = \frac{1}{2\pi}(1 + p_0 \mu \cos[2(\hat{\theta} - \theta_0)]). \quad (1.4)$$

The modulation factor $0 \leq \mu \leq 1$ is a measure of the detector polarization sensitivity: how well the emission angles are recovered. It can be thought of as a calibration term. A higher modulation factor means better polarization sensitivity and $\mu = 1$ means perfect emission angle reconstruction, $\theta = \hat{\theta}$. The modulation factor is the recovered polarization fraction \hat{p}_0 for a 100% polarized source, $p_0 = 1$. Fig. 1.4 gives an example for a $p_0 = 1$ source, the moment analysis recovered emission angles follow the blue $\mu < 1$ distribution while the true emission angles have $\mu = 1$. The modulation factor is a function of photon energy, higher energy tracks have higher modulation factor because they have longer tracks, thus better emission angle estimates. Track reconstruction algorithms and detector hardware both affect the instrument modulation factor. The modulation factor can be considered an average of the reconstruction quality of all tracks for a particular source energy spectrum. In

§3 we will revisit the modulation factor from first principles; for now it can be treated as a known constant for a given source energy spectrum, determined by instrument calibration.

Stokes parameters

It is usually simpler to estimate the normalized linear Stokes parameters $(\mathcal{Q}, \mathcal{U})$ instead of the polarization fraction and EVPA (p_0, θ_0) . These are an alternative representation of the source linear polarization. Disregarding circular polarization, the Stokes parameters are defined as

$$Q = Ip_0 \cos 2\theta_0, \quad (1.5)$$

$$U = Ip_0 \sin 2\theta_0, \quad (1.6)$$

where I is the source intensity, and the normalized Stokes parameters are

$$\mathcal{Q} = p_0 \cos 2\theta_0, \quad (1.7)$$

$$\mathcal{U} = p_0 \sin 2\theta_0, \quad (1.8)$$

$-1 \leq \mathcal{Q} \leq 1$, $-1 \leq \mathcal{U} \leq 1$, from which the polarization fraction and EVPA can be derived:

$$p_0 = \sqrt{\mathcal{Q}^2 + \mathcal{U}^2}, \quad (1.9)$$

$$\theta_0 = \frac{1}{2} \arctan \frac{\mathcal{U}}{\mathcal{Q}}. \quad (1.10)$$

Using Stokes parameters, the probability density eq.1.4 can be rewritten

$$p(\hat{\theta} | \mathcal{Q}, \mathcal{U}) = \frac{1}{2\pi} (1 + \mathcal{Q}\mu \cos 2\hat{\theta} + \mathcal{U}\mu \sin 2\hat{\theta}). \quad (1.11)$$

This is a more convenient form because the probability density is linear in the Stokes parameters, unlike for (p_0, θ_0) . Some researchers even prefer to work solely with Stokes parameters. Since likelihoods are invariant to reparameterization, estimating the Stokes parameters and converting back to (p_0, θ_0) is equivalent to estimating (p_0, θ_0) directly, although some care must be taken in propagating the uncertainties.

Methods

With the modulation factor known, (p_0, θ_0) and/or $(\mathcal{Q}, \mathcal{U})$ can be recovered from the emission angles $\{\hat{\theta}_i\}_{i=1}^N$ in a number of ways. The simplest method is a binned curve fit, i.e. directly fitting the probability density eq.1.11 to the histogram of emission angles $\{\hat{\theta}_i\}_{i=1}^N$, fig.1.4. Since the Stokes parameters are linear in the probability density eq.1.11 they can be fit using binned least squares. The resulting estimates $(\hat{\mathcal{Q}}, \hat{\mathcal{U}})$ are unbiased and equivalent to the maximum likelihood estimator

(MLE), so long as the bin widths are small enough and each bin contains enough events such that Poisson statistics can be well approximated by Gaussian. This is not always the case, so directly applying the MLE is often a better approach.

The MLE is the estimator $(\hat{\mathcal{Q}}, \hat{\mathcal{U}})$ that maximizes the likelihood function:

$$L(\{\hat{\theta}_i\}_{i=1}^N | \mathcal{Q}, \mathcal{U}) = \prod_{i=1}^N \frac{1}{2\pi} (1 + \mathcal{Q}\mu \cos 2\hat{\theta}_i + \mathcal{U}\mu \sin 2\hat{\theta}_i). \quad (1.12)$$

Here L is a function of $(\mathcal{Q}, \mathcal{U})$ and the observations $\{\hat{\theta}_i\}_{i=1}^N$ are fixed. Computationally, it is easier to minimize the negative log-likelihood function

$$-\log L = N \log 2\pi - \sum_{i=1}^N \log (1 + \mathcal{Q}\mu \cos 2\hat{\theta}_i + \mathcal{U}\mu \sin 2\hat{\theta}_i). \quad (1.13)$$

This expression must be minimized numerically to find $(\hat{\mathcal{Q}}, \hat{\mathcal{U}})$. Confidence intervals are calculated by numerically integrating the negative log likelihood function in the vicinity of the final estimator. Stokes parameter estimators and their errors can be simply transformed into $(\hat{p}_0, \hat{\theta}_0)$ if desired.

It is often inconvenient to evaluate the MLE and its errors numerically. An analytical solution for the MLE exists if $|\mathcal{Q}\mu| \ll 1, |\mathcal{U}\mu| \ll 1$. In practical X-ray polarimetry this limit is nearly always satisfied; polarimeters typically do not achieve much better than $\mu \lesssim 0.5$ averaged over a continuous spectrum and real astrophysical sources generally have low polarization $p_0 \lesssim 0.3$. By applying the Taylor expansion $\log(1+x) \approx x - x^2/2$ to the negative log-likelihood eq. 1.13 and minimizing the resulting quadratic form one finds

$$\hat{\mathcal{Q}} = \frac{1}{\mu} \frac{(\sum_{i=1}^N \cos 2\hat{\theta}_i - \sum_{i=1}^N \cos 2\hat{\theta}_i \sin 2\hat{\theta}_i) \sum_{i=1}^N \sin^2 2\hat{\theta}_i}{\sum_{i=1}^N \cos^2 2\hat{\theta}_i \sum_{i=1}^N \sin^2 2\hat{\theta}_i - (\sum_{i=1}^N \cos 2\hat{\theta}_i \sin 2\hat{\theta}_i)^2}, \quad (1.14)$$

$$\hat{\mathcal{U}} = \frac{1}{\mu} \frac{(\sum_{i=1}^N \sin 2\hat{\theta}_i - \sum_{i=1}^N \cos 2\hat{\theta}_i \sin 2\hat{\theta}_i) \sum_{i=1}^N \cos^2 2\hat{\theta}_i}{\sum_{i=1}^N \cos^2 2\hat{\theta}_i \sum_{i=1}^N \sin^2 2\hat{\theta}_i - (\sum_{i=1}^N \cos 2\hat{\theta}_i \sin 2\hat{\theta}_i)^2}. \quad (1.15)$$

Notice that $\mathbb{E}[\sum_{i=1}^N \cos 2\hat{\theta}_i \sin 2\hat{\theta}_i] = 0$, $\mathbb{E}[\sum_{i=1}^N \cos^2 2\hat{\theta}_i] = N/2$ and $\mathbb{E}[\sum_{i=1}^N \sin^2 2\hat{\theta}_i] = N/2$. For the relatively large N in X-ray polarimetry, these terms are effectively constants, so the expressions reduce to

$$\hat{\mathcal{Q}} = \frac{2}{N\mu} \sum_{i=1}^N \cos 2\hat{\theta}_i, \quad (1.16)$$

$$\hat{\mathcal{U}} = \frac{2}{N\mu} \sum_{i=1}^N \sin 2\hat{\theta}_i. \quad (1.17)$$

Because these are unbiased and MLE, they are the minimum variance unbiased estimators for $(\mathcal{Q}, \mathcal{U})$. In other words, so long as $|\mathcal{Q}\mu| \lesssim 1, |\mathcal{U}\mu| \lesssim 1$ and $N \gtrsim 1000$, these are the best possible estimators

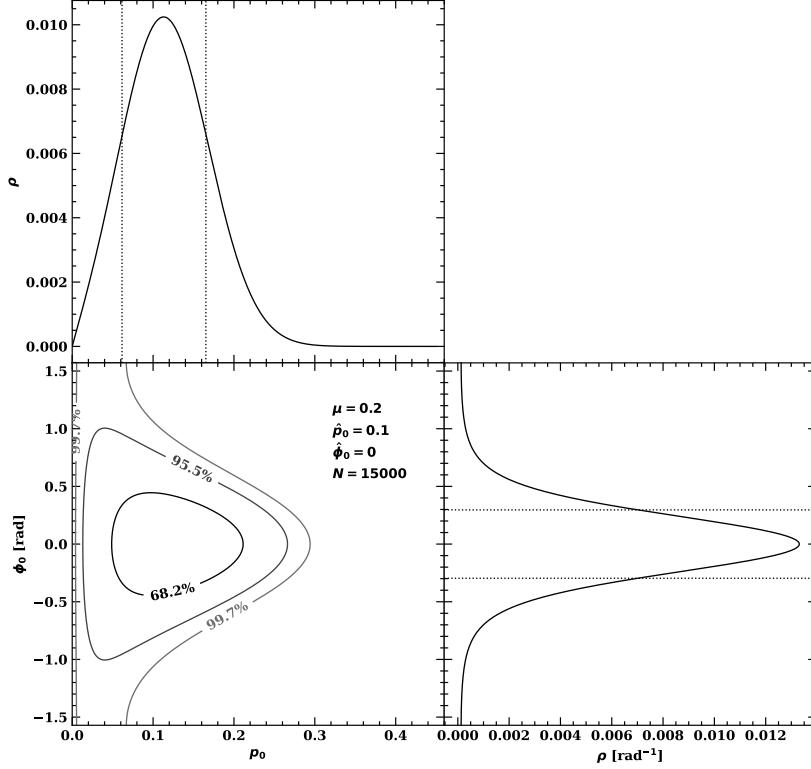


Figure 1.8: Confidence intervals from full posterior distribution (eq.1.18) over polarization parameters (p_0, ϕ_0) given estimates ($\hat{p}_0, \hat{\phi}_0$) using N events and instrument modulation factor μ . Confidence intervals for 68.2%, 95.5% and 99.7% ($1\sigma, 2\sigma, 3\sigma$) are displayed. The full marginal posterior distributions are plotted on the wings with 68.2% (1σ) confidence intervals drawn.

for $(\mathcal{Q}, \mathcal{U})$.

Kislat et al. [38] derive the full analytical posterior distribution for both $(\mathcal{Q}, \mathcal{U})$ and (p_0, θ_0) given the estimators eqs.1.16,1.17. They find the $(\mathcal{Q}, \mathcal{U})$ posterior follows a bivariate normal distribution while the (p_0, θ_0) posterior is

$$\begin{aligned} p(p_0, \theta_0 | \hat{p}_0, \hat{\theta}_0) &= \frac{\sqrt{N} \hat{p}_0 \mu^2}{2\pi\sigma} \times \\ &\exp \left[-\frac{\mu^2}{4\sigma^2} \left\{ \hat{p}_0^2 + p_0^2 - 2\hat{p}_0 p_0 \cos(2(\hat{\theta}_0 - \theta_0)) \right. \right. \\ &\quad \left. \left. - \frac{\hat{p}_0^2 p_0^2 \mu^2}{2} \sin^2(2(\hat{\theta}_0 - \theta_0)) \right\} \right], \end{aligned} \quad (1.18)$$

where

$$\sigma = \sqrt{\frac{1}{N} \left(1 - \frac{p_0^2 \mu^2}{2} \right)}. \quad (1.19)$$

The posterior assumes a uniform prior over (p_0, θ_0) . With the posterior in hand, any desired confidence interval can be computed analytically or numerically. High μ and high N reduce the width of the posterior so are both desirable to minimize the errors on recovered polarization parameters. Note that the estimator \hat{p}_0 is not unbiased, its posterior is highly asymmetric for low p_0 since $p_0 \geq 0$. For $p_0 = 0$ some polarization fraction $\hat{p}_0 > 0$ will always be measured. The amount of polarization likely to be measured when $p_0 = 0$ depends on the width of the posterior, i.e. N and μ , and is often used as a measure of detector sensitivity or signal-to-noise ratio (SNR), §1.1.2. Fig. 1.8 shows the posterior distribution for an example observation.

The recommended approach to polarization estimation is using eqs. 1.16, 1.17 for point estimates and the derived posterior eq. 1.18 for confidence intervals. This approach is expedient and optimal given the instrumental constraints of imaging X-ray polarimetry. In §3 we extend eqs. 1.16 – 1.18 to include neural network uncertainty estimates on predicted emission angles $\hat{\theta}$.

Minimum detectable polarization (MDP)

The standard figure-of-merit used in X-ray polarimetry to compare instrument sensitivity is minimum detectable polarization (MDP) [39]. MDP₉₉ is the polarization fraction that has a 1% probability of being exceeded by chance for an unpolarized ($p_0 = 0$) source. This can be found by integrating the posterior distribution eq.1.18:

$$\int_0^{\text{MDP}_{99}} \int_{-\pi/2}^{\pi/2} p(p_0, \theta_0 | \hat{p}_0, \hat{\theta}_0) d\theta_0 dp_0 = 0.99, \quad (1.20)$$

$$\text{MDP}_{99} = \frac{4.29}{\mu\sqrt{N}}. \quad (1.21)$$

On fig. 1.8, the MDP₉₉ would be the one sided 99% confidence interval for the p_0 marginal distribution (top panel).

The MDP₉₉ depends only on the modulation factor (signal) and the Poisson counting noise $1/\sqrt{N}$; it is effectively an inverse of the SNR. Polarimeters with lower MDP₉₉ can expose for shorter times and get the same confidence on their polarization measurements. Track reconstruction and polarization estimation approaches that increase the modulation factor for a fixed number of tracks N will decrease the MDP₉₉. Sections 2 and 3 will describe how a neural network approach can improve both track reconstruction and polarization estimation to minimize the MDP₉₉.

1.1.3 Neural Networks for Machine Learning

This section briefly covers deep neural network and machine learning concepts relevant to X-ray polarimetry, but for a proper introduction to this important field see Goodfellow, Bengio, and Courville [40] and Hastie, Tibshirani, and Friedman [41].

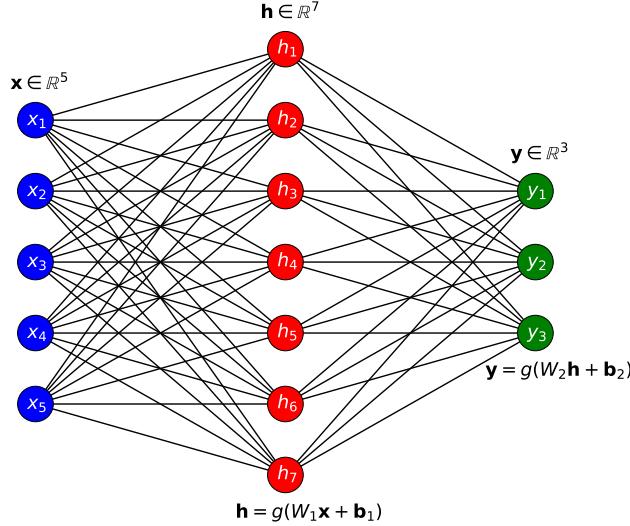


Figure 1.9: Feed-forward neural network with one hidden layer. The neural network takes vector \mathbf{x} as input and produces vector \mathbf{y} as output.

Deep neural networks (DNNs) are a type of supervised machine learning algorithm. Machine learning algorithms learn from data, as opposed to a fixed algorithm programmed by human beings. Supervised machine learning algorithms learn from labelled data: known input-output pairs $\{x_i, y_i\}_{i=1}^N$. An example from X-ray polarimetry would be track image (input), emission angle (output). Supervised machine learning algorithms can be used to solve a variety of tasks. Most tasks belong to one of two groups: regression or classification. In regression the algorithm is trained to predict a continuous numerical value given an input. In classification, the algorithm predicts the probability an input belongs to a particular class from a preselected group of finite classes. In the case of track reconstruction, we are interested mainly in regression.

In essence, DNNs approximate the unknown function f^* that maps problem inputs to outputs $y = f^*(x)$. DNNs are organized as a forward network of function layers each with their own learnable parameters or weights w , e.g. $f(x) = f_3(f_2(f_1(x; w_1); w_2); w_3)$. The initial function layer f_1 is known as the input layer, the final layer f_3 is the output layer, and all layers in between are the hidden layers. Using many hidden layers tends to improve the predictive capacity of the DNN, whence the term 'deep learning'. During training, the DNN $f(x)$ is iteratively directed towards $f^*(x)$.

If the layer functions f_1, f_2, f_3 were linear then the DNN $y = f(x)$ would be a linear model $y = w^T x$, equivalent to a simple least squares approach. DNNs use non-linear layer functions. This allows the DNN to learn non-linear relationships between the input features $y = w_2^T f_1(x; w_1)$. Here $f_1(x; w_1)$ is a non-linear function with learnable parameters w_1 . In this way DNNs are able to

learn what functions of the input features should be combined, unlike a linear model that linearly combines input features directly. A simple example of a non-linear layer function used in DNNs is $f_1(x; w_1, b) = g(w_1^T x + b)$ where w_1 is a matrix and g is a non-linear function such as tanh, the sigmoid function or a rectified linear unit (ReLU). DNNs using this kind of layer are universal function approximators in the asymptotic limit [42], so they should in principle well approximate any $f^*(x)$. In practice, memory, training and data limitations make this difficult to achieve. Many DNNs use domain specific layer functions and architectures to reduce these limitations and converge to $f^*(x)$ more quickly, see §1.1.3. Figure 1.9 shows a schematic of a simple DNN.

Training

As in supervised linear models like least squares regression, DNNs are trained to minimize a specific loss function over the observed data. In most tasks this normally means minimizing the negative log-likelihood, as in §2.2. For regression tasks, if the output distribution family $p(y|x)$ is unknown, a Gaussian should usually be the default since it is the least informative choice. For a Gaussian likelihood, the negative log likelihood is

$$L(\{y_i, x_i\}_{i=1}^N | \mathbf{w}) \propto \sum_{i=1}^N \|y_i - \hat{y}(x_i | \mathbf{w})\|^2, \quad (1.22)$$

the mean squared error (MSE) of prediction for the dataset. The loss function $L(\mathbf{w})$ is minimized over the DNN weights \mathbf{w} . MSE is a common loss function for regression with DNNs.

For classification tasks, $p(y|x)$ is usually assumed to follow a Bernoulli (in binary, 2-class classification) or categorical distribution (in \mathcal{C}_2 class classification). In the Bernoulli case, the negative log-likelihood is

$$L(\{y_i, x_i\}_{i=1}^N | \mathbf{w}) \propto \sum_{i=1}^N -y_i \log(\hat{y}(x_i | \mathbf{w})) - (1 - y_i) \log(1 - \hat{y}(x_i | \mathbf{w})). \quad (1.23)$$

In order to ensure $0 \leq \hat{y}_i \leq 1$, DNNs for classification usually end with a softmax layer [43]. These project DNN outputs onto the appropriate simplex, i.e., makes sure all predicted probabilities sum to one. For K output classes

$$\text{Softmax}(\mathbf{x})_i = \frac{e^{x_i}}{\sum_{i=1}^K e^{x_i}} \quad (1.24)$$

Since DNNs use non-linear functions, eq. 1.22 or 1.23 cannot be minimized analytically. Non-linearity makes the loss function complex with multiple local minima and saddle points in a very high dimensional weight space $\{\mathbf{w}\}$. This challenging minimization problem is usually tackled with variations on the stochastic gradient descent (SGD) algorithm [40]. At each iteration SGD approximates

the gradient of the loss function using a subset of the training data

$$g_i = \nabla_{\mathbf{w}} L(\mathbf{w}) \approx \sum_{i=1}^n \nabla_{\mathbf{w}} \|y_i - \hat{y}(x_i|\mathbf{w})\|^2, \quad n \ll N \quad (1.25)$$

and takes a step in parameter (\mathbf{w}) space along the direction of the negative gradient

$$\mathbf{w}_{i+1} = \mathbf{w}_i - \epsilon g_i. \quad (1.26)$$

The process is repeated until suitable convergence. The weights are initialized randomly and the learning rate ϵ is a tuneable scalar for the specific application. The gradient for each individual w_i can be calculated by using the chain rule through each layer of the DNN, in a process called backpropagation [40]. In order for SGD to work, the DNN has to use a differentiable loss function and layer functions f_1, f_2, f_3 . The whole dataset N is not used to calculate the gradient because it is too slow and often too large to fit into memory. Furthermore, DNNs benefit from using small batches n to approximate the gradient because the added stochasticity lets the minimization escape saddle points and local minima [40]. Usually DNNs are trained on graphical processing units (GPUs), specialized computer hardware for fast and parallelizable linear algebra operations.

Validation and model selection

DNNs typically have on the order of millions, sometimes billions [44], of trainable weights \mathbf{w} . The number of weights is often higher than the number of training examples, making overfitting the labelled dataset likely. This is the central challenge of machine learning: the algorithm must perform well on new inputs, not just those in the training dataset. To ensure model generalization to unseen examples, part of the training data is set aside for validation. The trainable weights \mathbf{w} are not trained on these examples. DNNs come with a number of additional model parameters – hyperparameters – that can be adjusted to improve model generalization; examples include the learning rate ϵ , the neural network architecture (type and number of layers) and any regularization. During training, the DNN is evaluated on the validation examples to tune the hyperparameters. This process is known as model selection.

Regularization is an important hyperparameter for improving model generalization. In exactly the same way as linear models can be regularized, for example in ridge regression, one can regularize a DNN by adding a penalty on the weights \mathbf{w} to the loss function. A common example is L2 regularization:

$$L(\mathbf{w}) \propto \lambda \|\mathbf{w}\|^2 + \sum_{i=1}^N \|y_i - \hat{y}(x_i|\mathbf{w})\|^2. \quad (1.27)$$

The additional term restricts the model space by encouraging the weights \mathbf{w} to be close to zero. The hyperparameter λ can be tuned to maximize generalization on the validation examples. This type

of regularization is equivalent to a Gaussian prior on the weights with mean zero. By restricting the model space, regularization helps prevent overfitting and improves generalization when λ is appropriately tuned.

Convolutional Neural Networks

Some DNNs have specialized layer functions that provide an inductive bias helpful for specific inputs. Convolutional neural networks (CNNs) [45] were designed for processing inputs that have neighbouring position correlations, for example 2D images, where adjacent pixel values are often correlated, or 1D time series where adjacent time points are often correlated. CNNs use layer functions that apply a convolution to the input. The convolution kernel function is learned during training. This kind of layer function is invariant under translation of the input, an important prior for working with image data. CNNs have revolutionized image processing and computer vision applications since winning the Imagenet benchmark competition in 2012 [33].

Recently, a certain class of CNNs called residual networks (ResNets) have achieved state-of-the-art results in multiple image benchmark tasks [46]. ResNets introduce 'skip' connections between certain layers that allow the CNN to ignore the layers in between. This alleviates many of the difficulties in training extremely deep neural networks.

Multitask learning

So far, only DNNs with single objectives and outputs have been considered. However, DNNs can be trained to solve multiple tasks at once, producing a vector of outputs. Solving multiple tasks at once often performs better than solving each task individually. This is partly because multiple tasks have a regularizing effect on each other, penalizing model complexity and overfitting on any single task leading to better overall model generalization [47]. Moreover, since each task shares the same input representation, e.g., an image, the features learned for a single task can help the others train faster and better.

In a multitask learning setting, the loss function includes the terms for each individual task. For example, in regression with three tasks and an MSE loss for each

$$L(\mathbf{w}) \propto \sum_{i=1}^N \|y_i^1 - \hat{y}^1(x_i|\mathbf{w})\|^2 + \alpha \sum_{i=1}^N \|y_i^2 - \hat{y}^2(x_i|\mathbf{w})\|^2 + \beta \sum_{i=1}^N \|y_i^3 - \hat{y}^3(x_i|\mathbf{w})\|^2, \quad (1.28)$$

where the DNN outputs concatenated vector $(\hat{y}^1, \hat{y}^2, \hat{y}^3)$. The relative size of each task loss term is controlled by the hyperparameters α, β . As usual, these hyperparameters should be tuned to maximize the model performance on unseen examples. More often in multitask learning, each task has a different form of loss function.

Uncertainty quantification

Once trained, DNNs can achieve state-of-the-art prediction accuracy in many domains. However, these models cannot capture the uncertainty inherent in their predictions. Predictive uncertainty quantification can be crucial, especially in applications where the real-world data distribution differs from the training data distribution. In this scenario, DNNs can extrapolate, producing overconfident but highly inaccurate predictions sometimes with disastrous consequences [48]. Predictive uncertainty is best represented by a posterior distribution on the predicted parameters given the inputs, however, simply quoting the variance of the posterior, often assumed to be Gaussian, is common practice.

There are two germane types of uncertainty one can model [49]. *Aleatoric uncertainty* captures noise inherent in the data. This is equivalent to statistical uncertainty or unexplained variance. On the other hand, *epistemic uncertainty* accounts for uncertainty in the model parameters – uncertainty which captures our ignorance about which model generated our collected data. This uncertainty can be reduced given the appropriate choice of model or enough data, and is often referred to as model uncertainty, systematic uncertainty or explained variance. More explicitly [50], training a model $\hat{y} = f(\mathbf{x})$ to approximate $y = f^*(\mathbf{x}) + \epsilon$, where the statistical noise ϵ is normally distributed $\epsilon \in \mathcal{N}(0, \sigma_a^2)$, the expected error is

$$\begin{aligned}\mathbb{E}\|y - \hat{y}\|^2 &= \mathbb{E}\|y - f^*(\mathbf{x}) + f^*(\mathbf{x}) - f(\mathbf{x})\|^2 \\ &= \mathbb{E}\|y - f^*(\mathbf{x})\|^2 + \mathbb{E}\|f^*(\mathbf{x}) - f(\mathbf{x})\|^2 \\ &= \sigma_a^2 + \sigma_e^2,\end{aligned}\tag{1.29}$$

where σ_a^2 and σ_e^2 are the aleatoric and epistemic errors, respectively. Cross terms in eq. 1.29 go to zero since ϵ is independent of y and \hat{y} . Eq. 1.29 highlights how epistemic uncertainties arise from a systematic difference between the true and learned models, while aleatoric uncertainties are irreducible statistical fluctuations about the true model.

Aleatoric uncertainties can be further broken down into two types: homoskedastic and heteroskedastic uncertainties. Homoskedastic uncertainties are the same for all inputs, whereas heteroskedastic uncertainties depend on the input \mathbf{x} , so $\epsilon \in \mathcal{N}(0, \sigma_a^2(\mathbf{x}))$. Almost all real data has heteroskedastic uncertainty; some examples are better measured than others. Photoelectron tracks are extremely heteroskedastic.

Uncertainty quantification in DNNs is a quickly developing field with many possible approaches. Modelling aleatoric uncertainties is simpler and a similar method is shared between approaches. This involves training the DNN to predict its own aleatoric uncertainty by enhancing the loss function. The next section, §1.1.3, provides a concrete example. Epistemic uncertainty is more difficult to quantify since it usually involves sampling over different models. Bayesian neural networks [49] do this by randomly subsampling the trained DNN weights \mathbf{w} and making multiple predictions,

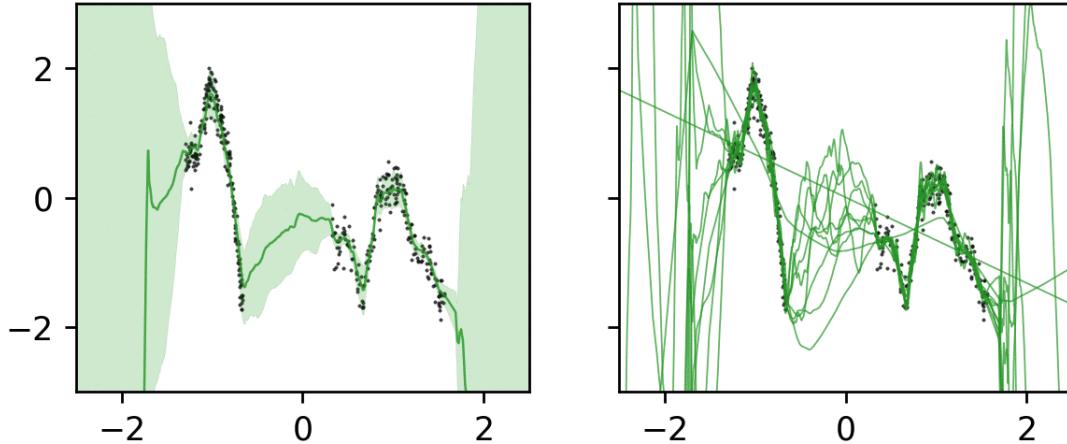


Figure 1.10: Toy regression problem using a deep ensemble. Right panel shows individual DNN prediction functions. Left panel gives the estimated epistemic uncertainty of the deep ensemble. Figure from Antorán, Allingham, and Hernández-Lobato [12].

while deep Gaussian processes [51] or mixture density networks [50] model the epistemic uncertainty intrinsically. For practical applications, which require robust and scaleable uncertainty predictions as well as high prediction accuracy, deep ensembles [52] are the state of the art [53].

Deep ensembles

Deep ensembles are made up of an ensemble of individual DNNs, each trained independently on the same dataset to predict the desired output features. Different random initializations of the same NN at the start of training leads to widely different prediction functions [54]. Deep ensembles exploit this property by incorporating the results of many differently initialized NNs, increasing the diversity of predictors. This not only yields higher prediction accuracy, but also allows for state-of-the-art aleatoric and epistemic uncertainty quantification [55, 53].

To model the aleatoric uncertainty in a regression task, each individual DNN composing the ensemble is trained to minimize the following negative log-likelihood loss function over all training examples i [52]

$$L(y_i \mid \mathbf{x}_i) = \frac{\log(\hat{\sigma}^2(\mathbf{x}_i))}{2} + \frac{\|y_i - \hat{y}(\mathbf{x}_i)\|_2^2}{2\hat{\sigma}^2(\mathbf{x}_i)}. \quad (1.30)$$

The DNN predicts the aleatoric uncertainty $\hat{\sigma}_a$ for each example prediction, as well as the example label \hat{y} as in standard DNN regression, eq. 1.22. This form of loss function assumes the aleatoric error has a Gaussian distribution, usually a reasonable approximation. In practice the DNNs are trained to predict the log variance $\hat{s}(\mathbf{x}_i) = \log(\hat{\sigma}^2(\mathbf{x}_i))$

$$L(y_i \mid \mathbf{x}_i) = \frac{\hat{s}(\mathbf{x}_i)}{2} + \frac{e^{-\hat{s}(\mathbf{x}_i)} \|y_i - \hat{y}(\mathbf{x}_i)\|_2^2}{2} \quad (1.31)$$

since this is more numerically stable [49].

During model prediction, predictions from all DNNs composing the ensemble are combined. The final predicted label is the mean over the ensemble predictions. For an ensemble with M DNNs $\hat{y}(\mathbf{x}) = (1/M) \sum_{j=1}^M \hat{y}^j(\mathbf{x})$. Similarly, the final aleatoric uncertainty is given by $\hat{\sigma}_a^2(\mathbf{x}) = (1/M) \sum_{j=1}^M \hat{\sigma}_a^2(\mathbf{x})^j$. The variance of the label predictions \hat{y}^j over the ensemble is used to model the epistemic uncertainty, resulting in total predicted uncertainty

$$\begin{aligned}\hat{\sigma}^2(\mathbf{x}) &= \hat{\sigma}_a^2 + \hat{\sigma}_e^2 \\ \hat{\sigma}^2(\mathbf{x}) &= \frac{1}{M} \sum_{j=1}^M \hat{\sigma}_a^2(\mathbf{x})^j + \frac{1}{M} \sum_{j=1}^M \left[\hat{y}^j(\mathbf{x}) - \frac{1}{M} \sum_{j=1}^M \hat{y}^j(\mathbf{x}) \right]^2.\end{aligned}\quad (1.32)$$

Very few DNNs are needed in the ensemble, $M \approx 5 - 10$, for the method to achieve state-of-the-art uncertainty quantification [52]. Figure 1.10 visualizes the epistemic uncertainty prediction for a toy regression problem. In the area far from the training examples, the DNNs composing the ensemble strongly disagree in their predictions, giving a high epistemic uncertainty.

In multitask problems, the aleatoric uncertainties can actually replace the hyperparameters that determine the relative importance of individual tasks in the loss function [56]. The DNNs automatically tune their own hyperparameters by learning the aleatoric uncertainties during training.

1.1.4 Blazars

Blazars are active galactic nuclei whose powerful relativistic jets point at small angle θ_{obs} to the Earth line-of-sight [57], so that the Doppler-boosted jet emission dominates the observed spectral energy distribution (SED). This broad multiwavelength SED, from radio to γ -rays, is characterized by a low energy peak caused by synchrotron radiation from energetic electrons, and a high energy peak generally attributed to Inverse Compton scattering of photons by these same electrons [58]; although, in some models hadronic processes may also contribute to the high energy flux [59]. The Compton scattered seed photons can either be from the synchrotron emission (SSC) or from an external source such as the accretion disk or broad line region (EC). The sources are further subdivided by the frequency of the νF_ν synchrotron peak [13], with $\log \nu_{\text{sy}} < 14$ labeled LSP (low synchrotron peaked, and most Flat spectrum Radio Quasars, FSRQ), $\log \nu_{\text{sy}} \approx 14$ labelled ISP (intermediate synchrotron peaked), and $\log \nu_{\text{sy}} > 15$ called HBL (high synchrotron peaked). Here the frequency is in Hz. Figure 1.11 visualizes blazar characteristic two-humped SED, with the three blazar subclasses labelled.

Dramatic variability, on timescales down to minutes in a few cases [60], is another hallmark of blazar emission. While radiation-zone models can reproduce the general two-humped emission pattern, many details remain to be explained and the underlying mechanisms of jet energization and collimation are still a subject of debate [61].

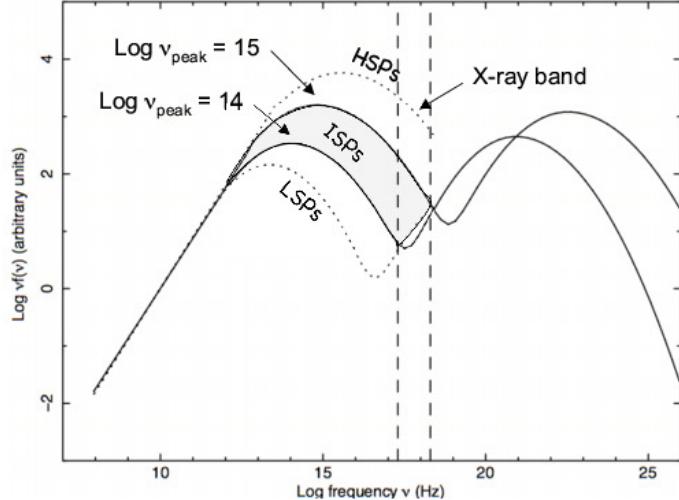


Figure 1.11: The three different blazar subclasses based on the frequency of the peak of the synchrotron component in their SED. Low synchrotron peaked blazars (LSP) have peak frequencies lower than 10^{14} Hz, intermediate synchrotron peaked blazars (ISP), 10^{14} Hz $< \nu_{peak_S} < 10^{15}$ Hz, and high synchrotron peaked blazars (HSP), $> 10^{15}$. Plot reproduced from Abdo et al. [13].

An attractive origin for how these jets are energized and launched with bulk Lorentz factor Γ is the Blandford and Znajek [62] process, so that the jet axis may be associated with the spin axis of the central black hole and the angular momentum axis of the surrounding accretion disk. The jet e^+/e^- obtain an energy distribution extending to $\gamma_{\max} \sim 10^4$ or higher, often attributed to shock acceleration or magnetic reconnection. Radiation from these particles spiraling in the embedded magnetic field B can be used to constrain the geometry and energetics of the emission zone and, by extension, the jet accelerator.

Another open question concerning blazars is jet composition and the nature of the high-energy emission component. While synchrotron self-Compton (SSC) emission from the primary lepton population in the jet is known to contribute, hadronic processes producing a secondary lepton population via the Bethe-Heitler process or a pion cascade could also make up a significant proportion of the high-energy emission. Jet composition is of particular interest as blazars have been proposed as candidate sources of TeV-energy neutrinos and ultra-high-energy cosmic-rays (UHECR), which would require a significant hadronic component in some blazar jets. The possible 3σ association of ISP blazar TXS 0506+056 with neutrino IceCube-170922A event motivates this connection [63]. Current blazar neutrino models assume either lepto-hadronic emission (e.g., [64]) or subdominant hadronic components where the hadronic-based emission only dominates the SED in the transition region (e.g., [15]). Figure 1.12 gives three example blazar composition scenarios and their resulting observed emission.

In studying jet geometry and composition, polarization can be particularly useful. Radio VLBI studies have long shown that the pc-scale jet can be substantially polarized. Recently much effort has been spent measuring the optical polarization properties of blazars, since this probes even smaller scales, closer to the acceleration zone. This polarization is often quite variable, offering new dynamical information on the jet structure [e.g. 65, 66]. In addition to typical stochastic behavior of polarization fraction (Π) and angle (PA), Blinov et al. [65] found periods of relatively steady rotation of the PA, sometimes extending many $\times\pi$, lasting weeks or months. These rotations are sometimes associated with flares in total intensity and drops in Π [16]. Various models have been proposed to explain this behavior, including a turbulent stochastic model [17], a spiraling jet [67] and a helical kink propagating along a conical jet [68]. Although these pictures can accommodate multicycle rotations, they fail to address the optical trends found in Blinov et al. [16].

In the following chapters, we show how measuring the X-ray polarization of blazars with *IXPE* [69] can distinguish between different emission processes and can probe closer to the jet acceleration zone than ever before.

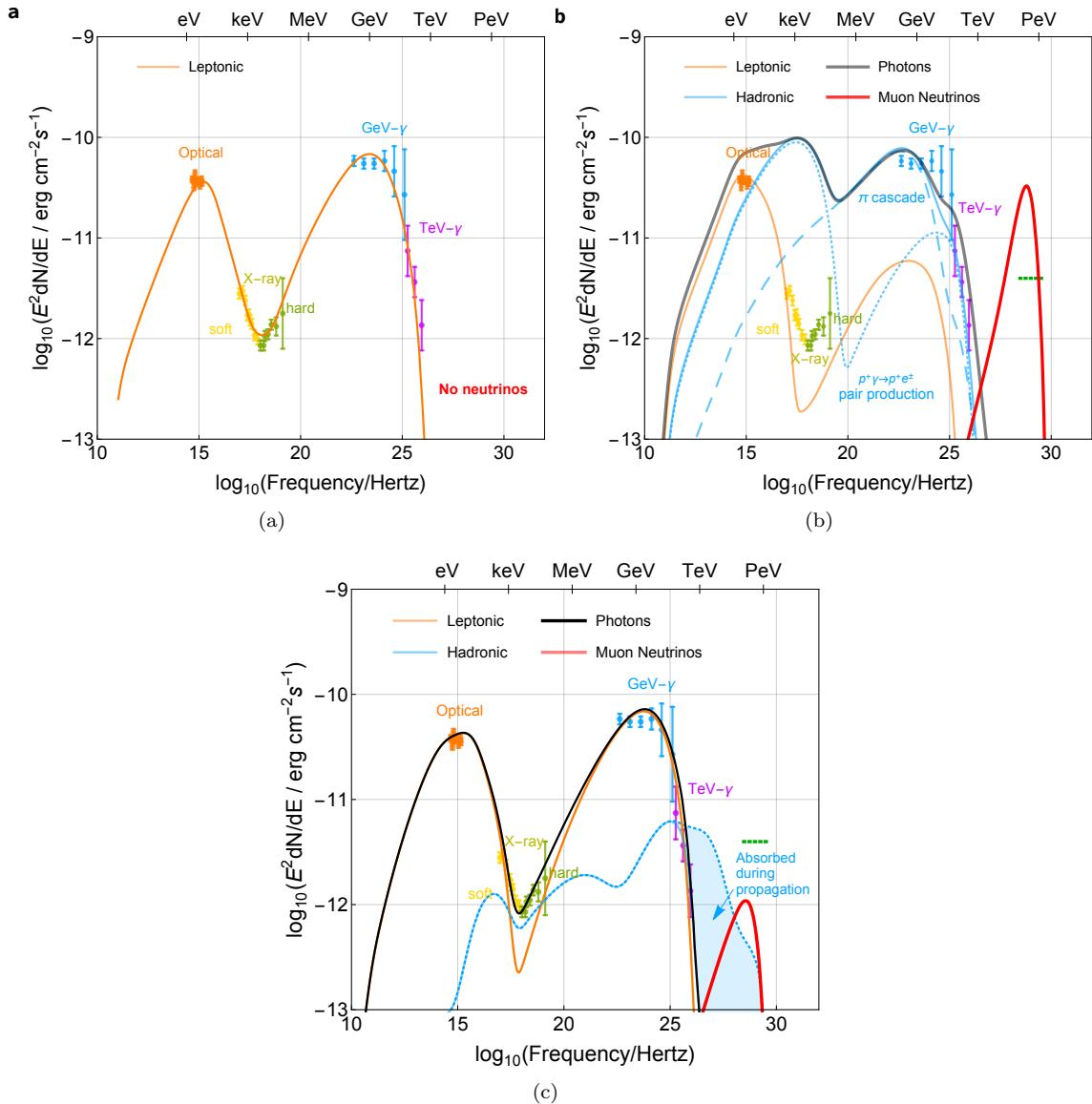


Figure 1.12: (a) Pure leptonic emission model for example blazar TXS 0506+056. This model cannot produce observed neutrino emission. (b) Hadronic dominated emission model. Ions accelerated in the jet produce secondary leptons via the Bethe-Heitler process or photo-pion production [14]. This model overproduces the X-ray emission. (c) Hybrid emission model; a mostly leptonic emission model with subdominant contribution from hadronic processes. Plots reproduced from Gao et al. [15].

Chapter 2

Neural Networks for Track Reconstruction

2.1 Introduction

Neural networks are supervised learning models that are the state of the art for classification and regression problems in many fields, especially computer vision [70]. They are highly non-linear models trained on large datasets of labelled input-output example pairs. Neural networks are often criticized as black box models with no output interpretability. However, recent work in uncertainty quantification for neural networks [71] has demonstrated these models are capable of trustworthy uncertainty estimates on their predictions. Neural networks' success in dealing with image data and their advances in uncertainty quantification make them a perfect candidate for X-ray polarimetry track analysis [72, 3].

Track reconstruction is an image input, multi-output regression problem §1.1.1. From a machine learning perspective, this problem is perfectly suited to a multitask CNN approach. CNNs learn translation invariant representations §1.1.3, an important constraint for working with image inputs, and a multitask approach is not only more computationally efficient but also likely to improve the individual task performances §1.1.3. In practice, a deep ensemble of CNNs is better than a single CNN for track reconstruction:

- An ensemble of models provides better predictive performance and generalization, §1.1.3.
- Trustworthy uncertainty quantification will be important for the next step: polarization estimation, §3.
- Deep ensembles provide automatic hyperparameter tuning of individual task importance for the multitask loss function, §1.1.3.

Some authors [72] have argued for an end-to-end deep learning approach to polarization estimation, combining track reconstruction with polarization estimation into a single large supervised learning problem to go directly from a set of tracks to source polarization. Although initially attractive because one could use source polarization directly as the neural network learning signal, avoiding the need for simulated detector events, this approach is difficult. Training would be costly (perhaps infeasible) since many individual events would be required for a single training example. Moreover, the observer may wish to adjust partitioning of the data set into e.g., time, energy, and spatial subsets, with differing polarization. The best partition will often not be obvious before the analysis starts, thus combining the properties of individual tracks allows a more efficient exploration of binning options. Finally detector-dependent artifacts can best be handled from individual tracks (with fine spatial positioning), rather than point spread function (PSF) weighted groups of tracks.

The neural network approach presented in this section and the next splits imaging X-ray polarimetry into two mostly separate data analysis steps, as described in §1. This section describes step 1, track reconstruction §1.1.1: extracting the relevant features (emission angles, absorption points, energies, and their associated uncertainties) from photoelectron track images as well as possible. This section follows the current state-of-the-art deep ensemble approach for track reconstruction taken by Peirson et al. [3] and Peirson and Romani [4].

2.2 Dataset

In any supervised learning problem, a dataset of labelled examples is an essential first step, without this, training a model is not possible. In the case of track reconstruction, one needs photoelectron track images labelled with known photoelectron 2D emission directions, 2D absorption points on the detector grid and scalar photon energies. Unfortunately, these labelled events cannot be collected from the detector in the lab so must be simulated. Most telescope detectors have event simulators to study detector performance; for example, IXPE has a GPD Monte Carlo event simulator built in GEANT4 [73] as part of its software suite. In any application where simulated data is used to train a machine learning model, care must be taken to make sure the simulated training data truly matches the real data the algorithm is likely to see. Systematic difference between the two is known as covariate shift and can cause reductions in prediction accuracy and uncertainty quantification on real data compared to the simulated data. In the case of IXPE, simulated events have been rigorously verified to match with real lab-collected detector events [10] at the expected telescope settings, including photons absorbed outside the gas volume §1.1.1. For IXPE trained neural network approaches, the performance on polarization estimation for simulated events generalizes to real detector data [3, 4], suggesting minimal covariate shift.

Once labelled photoelectron track events can be reliably simulated for the X-ray polarimeter of choice, one needs to choose the amount and type of training and testing data to be simulated. For

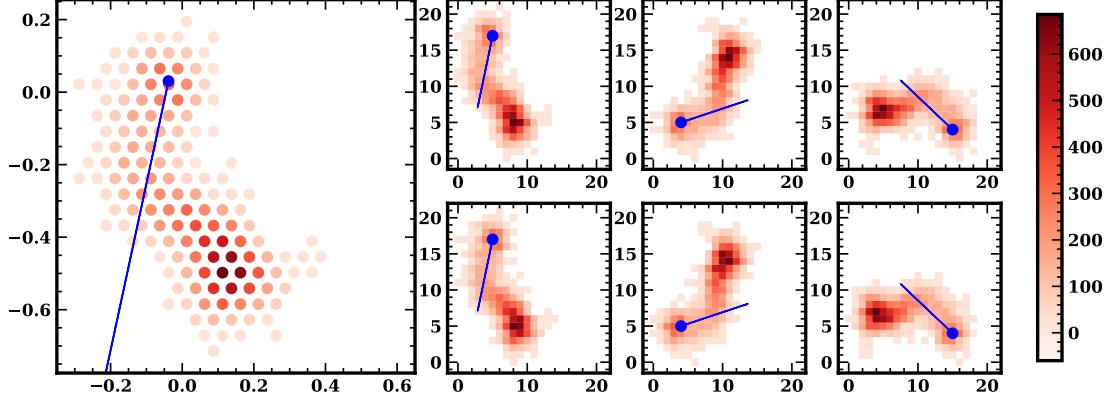


Figure 2.1: Example square conversions of a 6.4 keV hexagonal track (left panel). The six panels to the right show shifts along the 120° GPD axes; shifting odd rows (upper) or even rows (lower). For each hexagonal track, CNNs are fed column-wise pairs of square conversions, along with the energy, absorption point (blue dot) and initial photoelectron direction (blue line) as labels.

IXPE’s GPDs, [3, 4] use a training set of approximately 3.3 million tracks simulated uniformly for an unpolarized source between 1 - 11keV, for performance across a broad spectrum. They keep separate an additional 200k simulated tracks for validation and testing. The energy distribution of the training set does not make much difference to track reconstruction or polarization prediction, so long as no energies are greatly underrepresented nor labeled energy bins significantly larger than the energy resolution. An unpolarized source is simulated to make sure emission angles are uniformly distributed, so as not to cause any learning bias. The next section describes additional data pre-processing necessary for non-square grids.

In general, the larger the DNN (in terms of parameters and layers), the more data is required for training.

2.3 Geometric bias

Imaging polarimeters like IXPE use hexagonal pixels for two reasons:

1. Hexagonal arrays are the densest possible 2D packing, meaning more pixels per unit area thus improved track resolution.
2. Hexapolar grid effects are orthogonal to the quadrupolar polarization signal.

The discrete nature of the pixel grid means that track elongations and emission directions are better resolved along some axes of symmetry than others, potentially creating preferred directions for reconstruction. This bias would only be visible when the pixel size is significant compared to the track size, i.e., for small, low energy tracks. The hexagonal grid used in the IXPE GPDs

is designed to minimize any polarization systematics arising from these biases. Although not all imaging polarimeters use hexagonal pixels, this section describes some potential neural network solutions for dealing with non-square pixels.

A hexagonal grid is not natively compatible with standard CNN implementations for square matrices. It is possible to transform from a hexagonal to a square grid, however a naive transformation can lead to polarization systematics. Polarization estimation, step 2 §3, expects unbiased estimators for the emission angles. CNNs learn to prefer some emission angles over others, possibly interfering with the polarization signal. In a naive hexagonal to square conversion, emission angle bias happens partly because the CNN convolutional kernels are not spatially equivariant in hexagonal space. A possible solution is to work directly with hexagonal convolutions [74], but Peirson et al. [3] describe an expedient method.

2.3.1 Hexagonal to square conversion

It is possible to leverage existing square CNN architectures by using an appropriate polarization-systematic-free hexagonal to square conversion. From Peirson et al. [3]:

There are two main ways of converting between hexagonal and square grids: interpolation and pixel shifting. Interpolation places a fine square grid on top of the hexagonal image and interpolates. Interpolation should generally be avoided since it adds noise to the raw data and is not easily reversible. Pixel shifting rearranges pixels by shifting alternate rows and then rescaling [74] ... To avoid any polarization bias from the conversion, we pixel shift each track along each of the six hexagonal axes. Hexagonal tracks are rotated so that rows align horizontally (this can be done in three different ways, separated by 120°), then alternate rows are shifted (this can be done in two different ways, left and right) so that the track resembles a rectangular grid. We convert the rectangular grid into a square image by defining the leftmost track pixel and bottom track pixel as the left edge and the base of the image respectively. A square image size of 50x50 pixels is used to fit all track sizes for energies up to 11 keV. Since square track images are defined independently of the absolute hexagonal coordinate values, the initial hexagonal track rotation can be performed about any axis.

A single hexagonal track thus produces six square conversions, fig. 2.1, two for each 120° angle. A single training example for the CNNs is formed by stacking the corresponding square conversion pair, similarly to color channels in a *rgb* image CNN problem – in this case with only two channels. At test time all 3 pairs are evaluated by the CNNs and the predicted angles are rotated back to their original direction.

This approach artificially approximates spatial equivariance of CNN convolution kernels in the hexagonal space and removes all relevant prediction bias on emission angles θ (and later p_0, θ_0) introduced

when converting from hexagonal to square coordinates.

2.4 Deep ensemble setup

The specific CNN model that will compose the deep ensemble has to be chosen before any training begins. CNN architectures are usually compared by their performance on natural image classification datasets, like imangenet (224x224 pixels) or CIFAR (32x32 pixels) [75]. The residual network architecture (ResNet, He et al. [46]) described in §1.1.3 and related architectures like DenseNet [76] became the state of the art for image processing benchmarks in 2017. Peirson et al. [3] use the ResNet-18 CNN architecture to compose the deep ensemble; this has 18 total layers. The results shown in this section are using ResNet-18s, but any DNN appropriate for images could be used. Indeed, as of 2020, transformer architectures [77] have supplanted residual networks as the latest state of the art in image processing.

During training, each CNN minimizes a loss function. Since track reconstruction is a multitask problem, the loss function will include multiple terms for emission angle loss, absorption point loss, and energy loss. The DNNs will also predict the aleatoric uncertainty on each of its estimates. Each task loss function term is considered separately and they are all combined at the end of the section.

2.4.1 Emission angles

Since emission angles are periodic, a Gaussian aleatoric uncertainty is inappropriate. Instead of the Gaussian negative log-likelihood (NLL) used in §1.1.3, the von Mises (VM) distribution NLL is a better loss function choice. This is the maximum entropy distribution for circular data with specified expectation. For a random 2D unit vector \mathbf{x}

$$\text{VM}(\mu, \kappa) \equiv p(\mathbf{x}|\mu, \kappa) = \frac{\exp(\kappa\mu^T \mathbf{x})}{2\pi I_0(\kappa)}. \quad (2.1)$$

where I_0 is the modified Bessel function of the first kind. This can also be considered a 1D distribution over the polar angle θ of vector \mathbf{x} :

$$\text{VM}(\theta_\mu, \kappa) \equiv p(\theta|\theta_\mu, \kappa) = \frac{\exp(\kappa \cos(\theta - \theta_\mu))}{2\pi I_0(\kappa)}. \quad (2.2)$$

The VM distribution is parameterized by concentration parameter κ ; for large κ the VM converges to a Gaussian with variance $\sigma^2 = 1/\kappa$. For small κ the VM converges to a uniform distribution. This more appropriately reflects the distribution of predictions $\hat{\theta}$, which are clearly periodic.

Predicting scalar periodic values directly is tricky for standard DNNs, so the emission angle θ can be parameterized as a 2D vector $\mathbf{v} = (\cos 2\theta, \sin 2\theta)$. Now the DNNs attempt to predict the 2D vector $\hat{\mathbf{v}}$ [3]. Only $-\pi/2 \leq \hat{\theta} < \pi/2$, as opposed to $-\pi \leq \hat{\theta} < \pi$, are required for polarization

estimation since the EVPA $-\pi/2 \leq \theta_0 < \pi/2$, i.e. polarization is a quadrupolar signal. However, the full 2π emission angle can be useful when dealing with systematics and extended source analysis. If predicting the full 2π emission angle is necessary, an additional loss term for $\mathbf{v}_2 = (\cos\theta, \sin\theta)$ can be included [3]. Predicting only \mathbf{v}_2 gives very poor results for low energy tracks because of direction ambiguity, and should be avoided.

Each DNN model in the ensemble computes the emission angle loss for track image \mathbf{x} with true direction θ as the VM NLL

$$L_\theta(\mathbf{v} \mid \mathbf{x}) = -\hat{\kappa}^a(\hat{\mathbf{v}} \cdot \mathbf{v}) + \log I_0(\hat{\kappa}^a). \quad (2.3)$$

where $\hat{\kappa}^a$ are the DNN predicted aleatoric VM uncertainties. The epistemic uncertainties can also be assumed to follow $\text{VM}(0, \kappa^e)$; κ^e can be estimated from the output of a deep ensemble with M DNNs $\{\hat{\theta}_j\}_{j=1}^M$ using the appropriate maximum likelihood estimator:

$$\bar{R}^2 = \left(\frac{1}{N} \sum_{j=1}^M \cos 2\hat{\theta}_j \right)^2 + \left(\frac{1}{N} \sum_{j=1}^M \sin 2\hat{\theta}_j \right)^2 \quad (2.4)$$

$$\frac{I_1(\hat{\kappa}^e)}{I_0(\hat{\kappa}^e)} = \bar{R}, \quad (2.5)$$

The total VM predictive error on each track angle $\hat{\theta}$ is approximated by summing the aleatoric and epistemic variances, as in §1.1.3:

$$\frac{1}{\hat{\kappa}} = \frac{1}{\hat{\kappa}^e} + \frac{1}{M} \sum_{j=1}^M \frac{1}{\hat{\kappa}_j^a} \quad (2.6)$$

2.4.2 Absorption points

To predict the (x,y) photon absorption point coordinates given a track image \mathbf{x} , a simple Gaussian NLL loss function suffices:

$$L_{\text{abs}}(x, y \mid \mathbf{x}) = \frac{1}{2} \log \hat{\sigma}_{\text{abs}}^2 + \frac{1}{2\hat{\sigma}_{\text{abs}}} \left\| \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} \right\|^2. \quad (2.7)$$

As in eq.1.31, the DNN actually predicts the log variance for numerical stability. If desired, the epistemic error for the absorption point predictions can be calculated in the same way as §1.1.3.

2.4.3 Photon energy

Similarly to the absorption point loss eq.2.7, the energy loss can be stated as a simple MSE with the aleatoric error included:

$$L_E(E \mid \mathbf{x}) = \frac{1}{2} \log \hat{\sigma}_E^2 + \frac{1}{2\hat{\sigma}_E} \|E - \hat{E}\|^2, \quad (2.8)$$

and like the absorption points, the epistemic error can be calculated over the deep ensemble as in §1.1.3.

Predicting the energy, while simple at first sight, is more tricky than absorption points because of tail tracks. The presence of tail tracks, §1.1.1, can cause the DNNs to produce biased energy estimates \hat{E} with this loss function. Tail tracks look like lower energy tracks than their true energy label suggests, and this can lead to two possibilities:

- If tail tracks are difficult to distinguish from peak tracks this will cause the DNN to skew its overall predictions upward in energy to minimize eq.2.8.
- If low energy tail tracks are difficult to distinguish from high energy tail tracks, then most tail tracks will be assigned the mean energy of the training distribution as this minimizes the expected value of eq.2.8. This is usually undesirable for astrophysical spectra.

In principle these two effects could be remedied in a number of ways, for example by adjusting the loss function eq.2.8 [3] or by a careful look at the predicted energy uncertainty (tail tracks should have much higher uncertainties). In practice, since tail tracks also affect emission angle reconstruction and by extension polarization estimation, the recommended approach is to remove tail tracks from the analysis to the greatest extent possible [4]. This works well only if tail tracks are distinguishable from peak tracks. For IXPE GPD events this is fortunately the case, but a different approach may be better for other imaging X-ray polarimeters. In §2.5 we describe a DNN approach to identify and remove tail tracks.

The full DNN loss function to be minimized is now

$$L(\theta, x, y, E \mid \mathbf{x}) = L_\theta + L_{\text{abs}} + L_E + \alpha \|\mathbf{w}\|^2, \quad (2.9)$$

where the final term is an L2 regularization on the DNN weights controlled by hyperparameter α , to prevent overfitting the training data. The predicted aleatoric uncertainties act as automatically tuned hyperparameters for the remaining loss terms [56]. Given a track image input \mathbf{x} , an individual DNN outputs an 8 dimensional vector: $(\hat{\mathbf{v}}, \hat{\kappa}^a, \hat{x}, \hat{y}, \hat{\sigma}_{\text{abs}}, \hat{E}, \hat{\sigma}_E)$.

2.5 Removing tail tracks

Photoelectron tracks from events converting outside the main detector gas are known as tail tracks, §1.1.1. For both DNNs and classical methods they cause problems with photon energy prediction, discussed in §2.4. Furthermore, tail tracks have comparatively high emission angle prediction error (low polarization sensitivity §1.1.1) and §2.7 will show that this error is not properly captured by deep ensemble predictive uncertainties. One possible approach to remedy the tail track issue is to attempt to excise tail tracks from the data analysis entirely.

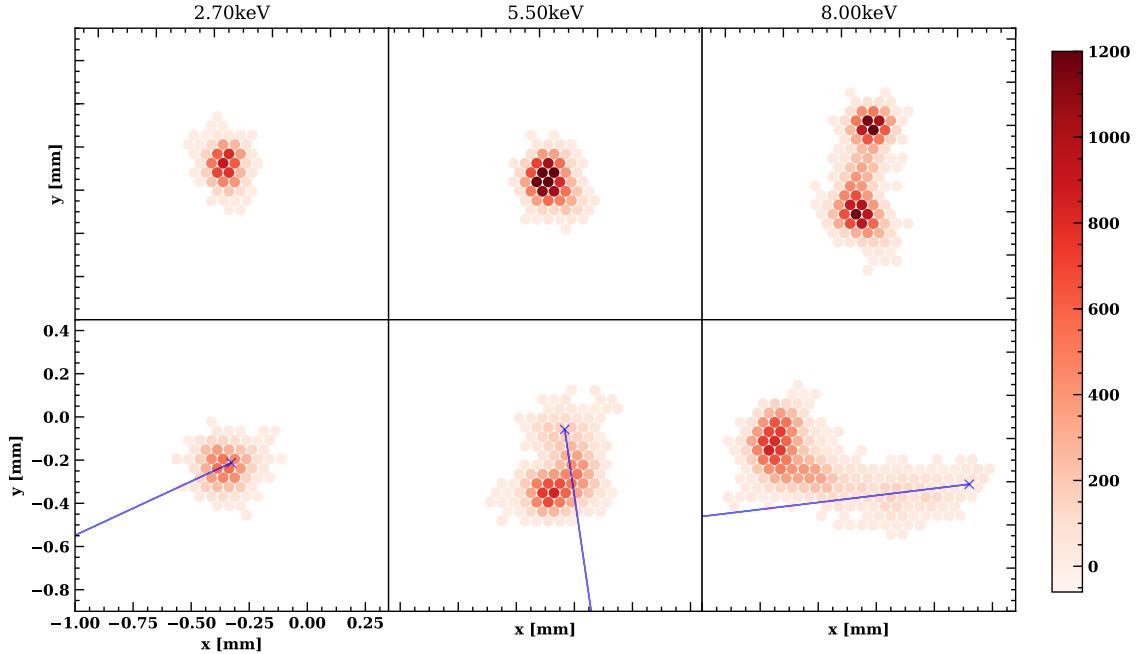


Figure 2.2: Example tail events (top row) and peak events (bottom row) for three different recovered energies. All plots follow the same spatial and color scale. These tail events convert in the GEM. Color denotes charge deposited in a given detector pixel. Note that tail events are more compact, short and with high charge density at maximum.

For IXPE’s GPDs, tail tracks differ in morphology from peak events converting in the gas. Fig. 2.2 compares events with the same linearly recovered energy (largely determined through the summed pixel counts, the summed energy deposition in the gas). Since the GEM and window material have a lower mean free path for photoelectrons, detected tail events are typically due to photo-electrons ejected close to the GEM/window normal. Their tracks are thus more compact, with higher counts/pixel for the same recovered energy. Window events have larger drift diffusion than GEM conversions. Distinguishing peak and tail tracks can be formed as a computer vision classification problem, so a DNN would be a good model choice and could be trained on simulated events to recognize the differences.

A good approach for peak vs. tail track classification would be to include the classification task into the existing track reconstruction deep ensemble, §2.4. This would likely improve performance over training a separate model and at first glance would only require an additional loss term in eq.2.9. However, this approach would not fix the problem of improper energy predictions and emission angle uncertainties unless more complex cross terms to eq.2.9 were also added. The cross terms would ensure, for example, that tracks likely to be tail events have higher energy uncertainty. Opting for

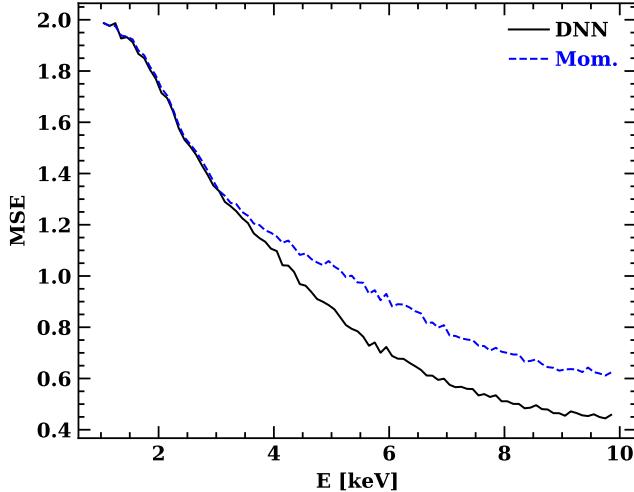


Figure 2.3: Mean squared error on emission angle prediction as a function of photon energy, measured using eq. 2.10, for the classical moment analysis and deep ensemble.

a simpler approach, Peirson and Romani [4] trained a separate peak vs. tail classifier DNN and used only peak tracks to train the track reconstruction deep ensemble. The Peirson and Romani [4] separation approach is described in this section, but a unified deep ensemble approach to tail vs peak classification would likely work better and is discussed as a future direction in §9.

Setting up a peak vs. tail DNN classification model is the same as setting up a multitask deep ensemble §2.4, only now the loss function to be minimized is different. In binary classification tasks, DNNs minimize the cross-entropy loss, eq. 1.23. Given a single input, a DNN outputs a single scalar between 0 and 1 that represents the probability an input is the first class. Peirson and Romani [4] use exactly the same ResNet-18 architecture and data preprocessing, §2.2–2.3, when training the peak vs. tail classifier DNN; the only notable differences are the loss function and data labels.

2.6 Training and ensemble selection

Individual DNN training procedures should follow the guidelines in §1.1.3, §1.1.3. Specifics will depend on the particular DNN architecture chosen. CNNs for computer vision problems train particularly well using stochastic gradient descent with momentum [78] as the optimization algorithm. Normalizing the inputs and outputs before training typically helps improve convergence. From Peirson et al. [3]:

The ResNet-18 architecture trains in a reasonable amount of time (~ 15 hours for 150 epochs on 4 Nvidia Titan GPUs, using a batch size of 2048). Before training, we normalize the track images, subtracting the pixel-wise mean from each track image and

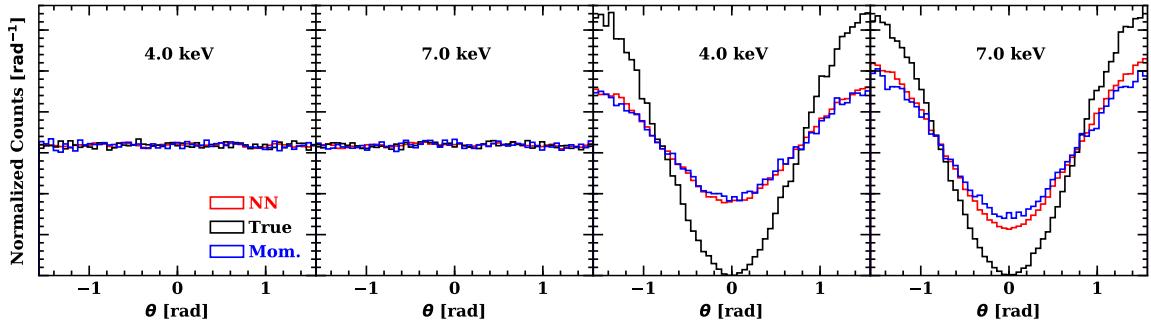


Figure 2.4: Emission angle recovery for unpolarized, $p_0 = 0$, (left two panels) and 100% polarized, $p_0 = 1$, (right two panels) simulated IXPE data for 4.0 and 7.0 keV. The true photoelectron angle distribution is shown in black; standard moment analysis reconstruction is in blue and deep ensemble in red.

dividing by the pixel-wise standard deviation (where the mean and standard deviation are calculated over the full training set). The track energy and absorption point labels are similarly processed. Normalizing the training data helps prevent vanishing and exploding gradients during the NN training procedure and lead to faster convergence. We use stochastic gradient descent with momentum as our optimizing algorithm, typical in computer vision tasks [78], with a stepped decaying learning rate starting at 0.01. We choose batch sizes of 512, 1024, 2048 tracks ... We tune the hyperparameters to minimize the MSE for the validation set.

Once multiple DNNs are trained with optimized hyperparameters, individual DNNs should be chosen at random to compose the ensemble. This ensures accurate epistemic uncertainty estimates.

2.7 Performance

The performance of deep ensembles in track reconstruction is compared to the classical moment analysis §1 on simulated IXPE data. The results shown are for peak events only unless otherwise specified.

2.7.1 Emission angles

Two important measures of performance should be evaluated for emission angle recovery: the accuracy of recovery and the quality of uncertainty estimates. The accuracy of emission angle recovery can be measured, for example, by the MSE on a test dataset

$$\frac{1}{N} \sum_{i=1}^N \|\mathbf{v}_i - \hat{\mathbf{v}}_i\|^2. \quad (2.10)$$

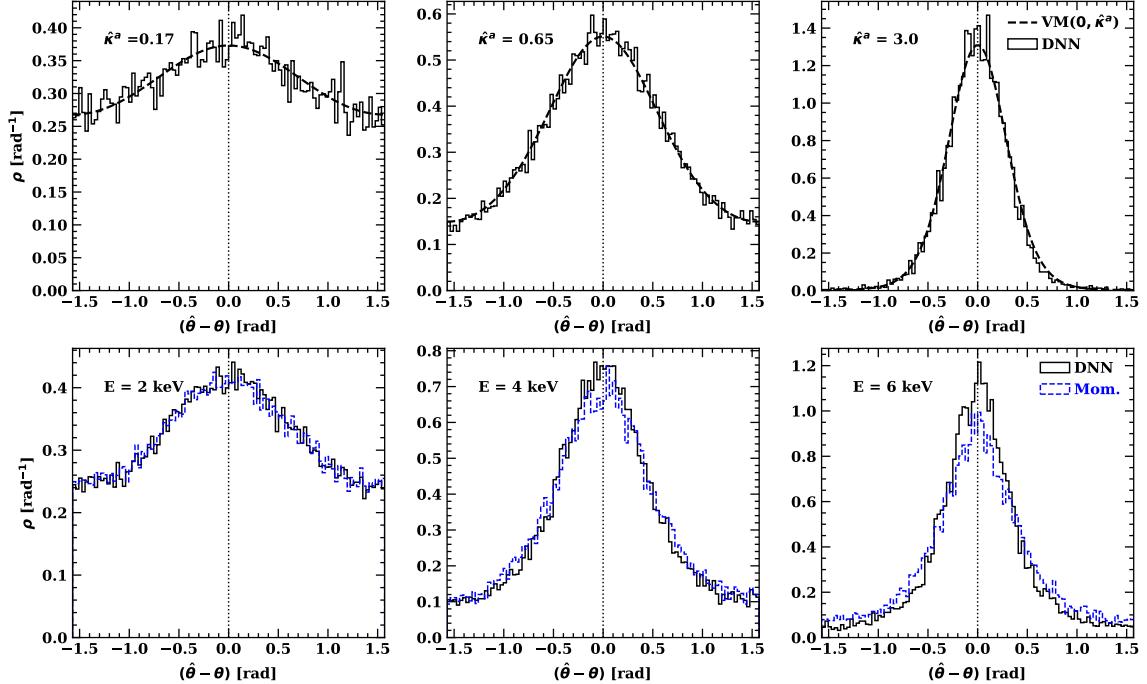


Figure 2.5: *Top:* Histograms of the DNN emission angle residual for three different DNN predicted aleatoric VM concentration parameters $\hat{\kappa}^a$. The DNN predicted VM error distributions are overplotted. The DNN is able to correctly predict its emission angle aleatoric error distribution. *Bottom:* Histograms of emission angle residuals at three different energies for DNNs and the moment analysis. All methods are unbiased (centered at zero) and at higher energies the DNNs have better recovery of the true emission angle θ , as evidenced by the larger peak heights and lower tails.

Fig. 2.3 gives the above MSE as a function of the true track energy on an unpolarized IXPE test dataset for both classical moment analysis and deep ensemble approaches. At low energies, deep ensembles fail to improve significantly over the moment analysis; when resolution is low, track image moments contain all the relevant information.

It is also essential to confirm that the estimates $\hat{\theta}$ are unbiased. Fig. 2.4 gives the distribution of emission angles $\hat{\theta}$ recovered by the deep ensemble and moment analysis for an unpolarized and polarized source at various energies. For both methods, emission angle estimates show no systematic biases, and at higher energies for the polarized source the improved signal recovery of the deep ensemble is visible. As an additional check, the bottom row of fig. 2.5 displays the distribution of $\hat{\theta} - \theta$ for the moment analysis and deep ensemble at three different true energies. Both the moment analysis and deep ensemble have symmetric error distributions centered at zero for all energies. At higher energies, the lower dispersion and thus higher accuracy of the deep ensemble is again clear.

A simple way to assess the deep ensemble predicted uncertainty estimates $\hat{\kappa}$ is to consider the distribution of $\hat{\theta} - \theta$ for fixed $\hat{\kappa}$. According to the assumptions in §2.4, $\hat{\theta} - \theta$ should follow a $VM(0, \hat{\kappa})$

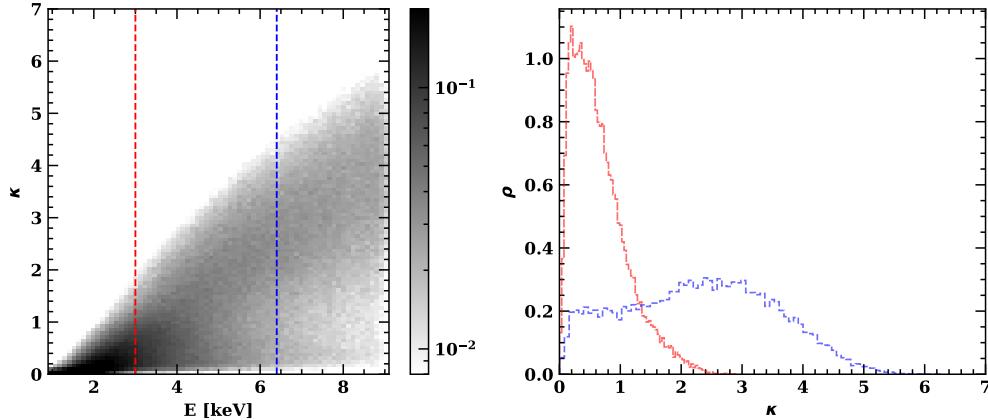


Figure 2.6: Distribution of deep ensemble predicted concentration parameters $\hat{\kappa}$ across the IXPE energy spectrum (left). The right-hand plot shows the $\hat{\kappa}$ distribution for two specific energies 3.0 keV (red), 6.4 keV (blue). High concentration parameter κ means a low predictive uncertainty.

if the DNNs have been properly trained. In the top row of fig. 2.5, the distributions of $\hat{\theta} - \theta$ for three different fixed $\hat{\kappa}$ are displayed with $\text{VM}(0, \hat{\kappa})$ overlayed. In all cases, the distributions match very closely. Epistemic errors are negligible for the vast majority of peak tracks, suggesting good model choice and training convergence.

It is also instructive to visualize the distribution of deep ensemble predicted uncertainties as a function of energy, fig. 2.6. For higher energies a wider range of concentration parameters are predicted and the $\hat{\kappa}$ peak moves progressively higher. Below 2keV, barely any polarization signal is recoverable with IXPE. Fig. 2.6 highlights the strong heteroskedasticity in imaging X-ray polarimetry.

For a deep ensemble trained on both peak and tail tracks, the predicted uncertainties for tail tracks are untrustworthy. Comparing the error distributions in fig. 2.7 to 2.5, tail track error is not captured by deep ensemble uncertainty predictions, even with the epistemic uncertainty included. Furthermore, at low energies, tail track emission angle estimates are potentially biased.

2.7.2 Tail vs. peak classification

For tail vs. peak classification, there is no equivalent classical method to compare the DNN performance, so this section lists the classification results achieved by Peirson and Romani [4]. They use the predicted tail probability as an event threshold cut. Fig. 2.8 shows the effectiveness of different tail probability cuts. If all events with a DNN predicted tail probability higher than 70% are removed, a loss of only 3% of true peak events is incurred while 66% of true tail events are successfully removed. Using a DNN classifier, only a small loss to signal events need be incurred to

remove the majority of harmful tail events. Fig. 2.8 also shows how remaining tail and peak events are distributed in terms of true energy after a tail probability threshold cut. Remaining tail events come from all energies in proportion to their original population, while misidentified peak events come from middle energies where peak and tail events overlap most.

At least for IXPE GPDs, it is certainly possible to meaningfully distinguish between tail and peak events. The upcoming section on photon energy recovery and §3 demonstrate how tail cuts based on DNN probabilities can be used to improve energy and polarization resolution.

2.7.3 Absorption points

Absorption point accuracy can be evaluated using the MSE. In fig. 2.9 the classical moment analysis and deep ensemble predictions are compared at different photon energies. For IXPE, the track barycenter provides a better prediction of the absorption point for very low energies – the deep ensemble recovers this transition while the moment analysis does not.

For current generation X-ray polarimeters like IXPE, the telescope PSF is much larger than any absorption point discrepancy, so absorption point accuracy is less essential than emission angle or photon energy accuracy. However, future missions with higher spatial resolution may require very accurate absorption point estimates for detailed imaging of extended sources. Even for IXPE the improved localization can mitigate polarization fringes about such sources.

2.7.4 Photon energy

Energy recovery can be adversely affected by tail events. While the MSE in predicted energy is lower for deep ensemble methods compared to classical linear charge recovery, these results should be treated with caution. Much of this improvement is specious, coming from a mistreatment of tail events. Fig. 2.10 plots the IXPE recovered energy histograms of the classical linear method (blue), a deep ensemble trained including tail events (black) and a deep ensemble trained on only peak events with tail events removed at a 70% peak vs. tail classifier probability (red). Around the true energy peak all methods perform similarly, with a slight advantage for the DNN based methods. As the energy increases, all methods show a low energy tail in reconstructed energy which is produced by tail events. The standard deep ensemble that does not differentiate peak and tail tracks also shows a high energy tail in reconstructed energy. This tail arises from low energy peak events ‘looking’ like higher energy tail events, so pushing some events up to higher energy tends to minimize the MSE. As discussed in §2.4, high energy tails in reconstructed energy are undesirable for astrophysical spectra. Training a deep ensemble with only peak events prevents this high-energy tail behavior.

Cutting events based on the predictions of the peak vs. tail classifier successfully reduces tails in the recovered energy histograms. A tail event cut improves the energy resolution at the cost of potentially losing some peak events and thus some polarization signal. §3.4 will discuss and quantify the polarization signal loss due to tail event cuts.

The energy prediction accuracies (mean absolute error and FWHM) for all three methods as a function of photon energy are given in fig. 2.11. Removing tail events provides a significant improvement to the energy prediction accuracy.

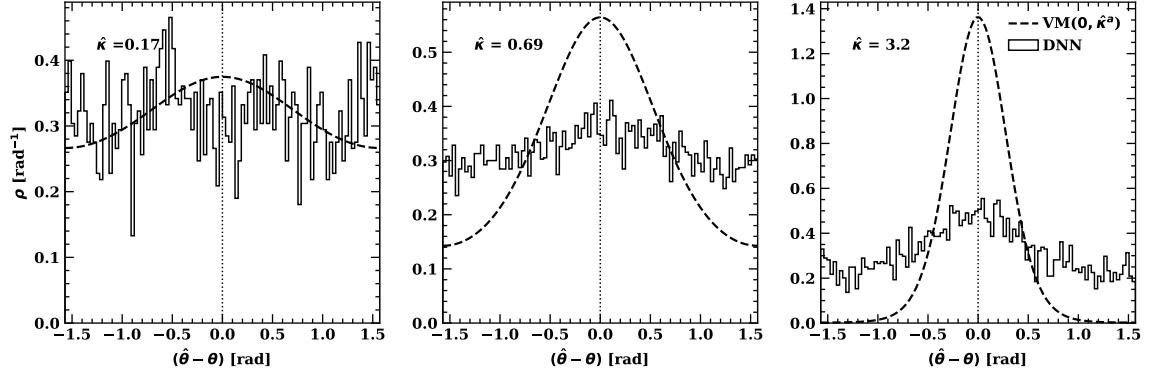


Figure 2.7: Histograms of the emission angle residual at three different DNN predicted concentration parameters $\hat{\kappa}$ for tail events only. The predicted DNN error distributions are overplotted. When tail events are treated as if they were normal events, the DNNs do not learn to predict the appropriate uncertainties.

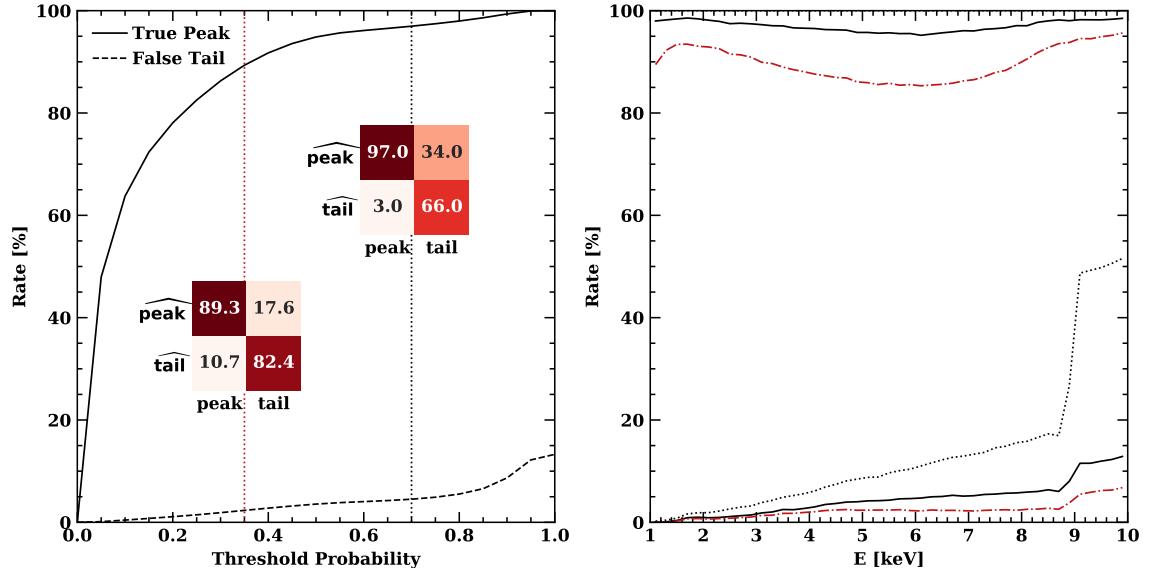


Figure 2.8: *Left:* The solid curve shows the fraction of peak events retained as a function of the tail probability cut, while the dashed curve shows the fraction of the cut sample from the remaining tail events. Insets show the confusion matrices, normalized by column, for our adopted 70% cut and a 35% cut. *Right:* The top curves show the peak retention as a function of energy (black solid – 70% cut, red dot-dash – 35% cut). Below, the dotted curve shows the uncut fraction of the sample due to tail events, while the lower black and red curves show the residual tail pollution (70% and 35% cut, respectively). Depending on how harmful tail events are to the desired measurements, different cut levels could be appropriate.

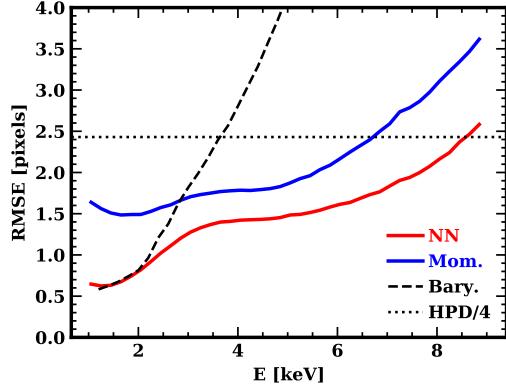


Figure 2.9: Photon absorption point localization accuracy using the root mean squared error. The deep ensemble predictions (red) do appreciably better than the moment analysis and matches a barycenter estimate at the lowest energies. For IXPE, all methods are adequate, as localization is much better than the size of the point source image produced by IXPE’s mirrors. One quarter of IXPE’s half power diameter is illustrated (black dotted line).

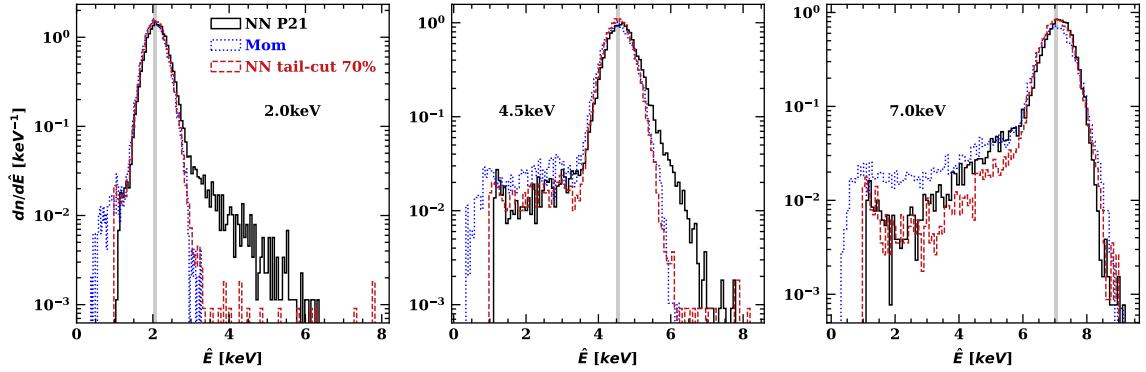


Figure 2.10: Response for three true energies. Note that while the P21 analysis suppressed the large tail rate seen in the Moments processing, events leaked to a high energy tail for medium to low true energies. Our morphological tail cut further suppresses the GEM/window tail, avoids the high energy leakage and achieves comparable or better peak width.

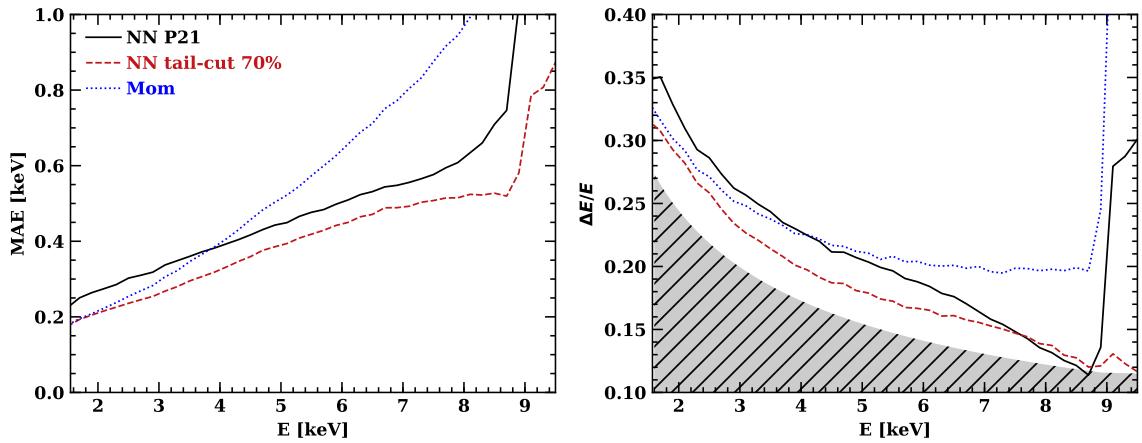


Figure 2.11: *Left:* Mean absolute error in predicted energy as a function of true energy. *Right:* FWHM ($3.46 \times$ Median Absolute Deviation) of predicted energy distribution for a source at energy E . The grey band marks the limiting energy resolution in the purely-exponential multiplication regime. NNs with 70% tail cut performs better on every metric, more accurate per track and tighter resolution at all energies. All methods suffer from the large increase of tail events above 9 keV.

Chapter 3

Polarization Estimation

3.1 Introduction

§1.1.2 explained how to best estimate polarization parameters (p_0, θ_0) given a group of reconstructed photoelectron emission angles $\{\hat{\theta}_i\}_{i=1}^N$. This section explains how to best estimate polarization parameters given a group of reconstructed photoelectron emission angles *and their associated uncertainties* $\{\hat{\theta}_i, \hat{\kappa}_i\}_{i=1}^N$. These uncertainties can vary for different individual emission angles; the measurements are heteroskedastic. Emission angles with larger uncertainties should contribute less to the final polarization parameter estimates. In theory, this section need not use DNNs and the methods described can be used for any track reconstruction method that provides (trustworthy) emission angle uncertainties; the uncertainties can be parameterized by any appropriate distribution. In practice, the predicted emission angle uncertainties come from deep ensembles and the uncertainties are assumed to follow a von Mises family of distributions.

3.2 Modulation factor

To develop a polarization estimation method that incorporates emission angle uncertainties, it is essential to specify how these uncertainties affect the recovered emission angle distribution eq. 1.3. For a source with polarization parameters (p_0, θ_0) , the true photoelectron emission angles follow the distribution eq. 1.3, while the recovered or measured emission angles follow eq. 1.4. The difference between the two is captured by the modulation factor μ . The modulation factor summarizes the effect of uncertainty in individual emission angle measurements.

Consider a single emission angle measurement $\hat{\theta}_i$. Since track reconstruction methods are imperfect and contain sources of error (even DNN based ones as was shown in §2), $\hat{\theta}_i$ can be considered a random variable:

$$\hat{\theta}_i = \theta_i + \epsilon_i \tag{3.1}$$

where θ is the true emission angle, which follows the distribution eq. 1.3, and the measurement error ϵ_i is a random variable with

$$\mathbb{E}[\epsilon_i] = 0, \text{Var}[\epsilon_i] = \sigma_i^2, p(\epsilon_i = 0) = p(\epsilon_i = 2\pi) \quad (3.2)$$

This assumes the $\hat{\theta}_i$ estimate is unbiased; reasonable given both the deep ensemble and moment analysis in §2.7 are for the most part unbiased estimators. The specific distribution for ϵ_i will depend on the track reconstruction method, but all distributions should follow the properties in eq. 3.2: be unbiased, have a finite variance, and be periodic, since $\hat{\theta}$ itself is periodic.

For any ϵ_i distribution with the above properties, it is possible to find the distribution for $\hat{\theta}_i$ as the convolution of the θ_i, ϵ_i distributions:

$$p(\hat{\theta}_i | p_0, \theta_0, \sigma_i^2) = \frac{1}{2\pi} (1 + \mu_i p_0 \cos[2(\hat{\theta}_i - \theta_0)]), \quad (3.3)$$

where $0 \leq \mu_i \leq 1$, $\mu_i(\sigma_i^2)$ and $\mu_i(\sigma_i^2 = 0) = 1$. In other words, the distribution of estimators $\hat{\theta}_i$ are the same as the distribution of the true values θ_i but with a reduced individual modulation factor $\mu_i(\sigma_i^2)$ that depends on the measurement error σ_i^2 . The measurement noise will blur the sinusoidal modulation signal by a factor μ_i for the specific event i .

The final distribution for a set of recovered emission angles $\{\hat{\theta}_i\}_{i=1}^N$ can now be simply formulated as an equal mixture model of the distributions of all individual recovered emission angles, eq.3.3:

$$p(\hat{\theta} | p_0, \theta_0) = \frac{1}{N} \sum_{i=1}^N \frac{1}{2\pi} (1 + \mu_i p_0 \cos[2(\hat{\theta} - \theta_0)]), \quad (3.4)$$

$$= \frac{1}{2\pi} (1 + (\frac{1}{N} \sum_{i=1}^N \mu_i) p_0 \cos[2(\hat{\theta} - \theta_0)]), \quad (3.5)$$

$$= \frac{1}{2\pi} (1 + \mu p_0 \cos[2(\hat{\theta} - \theta_0)]). \quad (3.6)$$

Thus, the final modulation factor of a set of emission angles with varying individual measurement errors $\{\hat{\theta}_i, \sigma_i^2\}_{i=1}^N$ is an average over all individual modulation factors

$$\mu = \frac{1}{N} \sum_{i=1}^N \mu_i(\sigma_i^2), \quad (3.7)$$

where individual modulation factors each depend on the individual measurement errors σ_i^2 .

For a specific example of an individual modulation factor, consider the case where ϵ_i is $\text{VM}(0, \kappa_i)$ between $[0, \pi]$. This is the uncertainty distribution learned by the deep ensembles in §2.

$$p(\epsilon_i) = \frac{1}{2\pi I_0(\kappa_i)} \exp(\kappa_i \cos 2\epsilon_i). \quad (3.8)$$

This distribution meets all the assumptions stipulated in eq. 3.2. Evaluating the convolution of the

θ (eq.1.3) and ϵ (eq.3.8) distributions, one finds

$$p(\hat{\theta}_i \mid p_0, \theta_0, \kappa_i) = \frac{1}{2\pi} \left(1 + \frac{I_1(\kappa_i)}{I_0(\kappa_i)} p_0 \cos[2(\hat{\theta}_i - \theta_0)] \right). \quad (3.9)$$

Comparing to eq. 3.3

$$\mu_i = \frac{I_1(\kappa_i)}{I_0(\kappa_i)}, \quad (3.10)$$

and the modulation factor for set of tracks $\{\hat{\theta}_i, \kappa_i\}_{i=1}^N$ is

$$\mu = \frac{1}{N} \sum_{i=1}^N \frac{I_1(\kappa_i)}{I_0(\kappa_i)}, \quad (3.11)$$

3.3 Weighted maximum likelihood estimator

In §1.1.2, it was assumed when estimating polarization parameters that the modulation factor is a constant, μ . The previous section showed how the modulation factor arises from the measurement errors of many individual emission angle estimates. If individual uncertainties are known, rather than assuming these lead to one constant μ , individual modulation factors can be directly incorporated into the likelihood function for the polarization parameters.

With individual emission angle measurement errors known, the likelihood eq. 1.12 can be expressed as a more informative likelihood:

$$L(\{\hat{\theta}_i\}_{i=1}^N \mid \mathcal{Q}, \mathcal{U}) = \prod_{i=1}^N \frac{1}{2\pi} \left(1 + \mathcal{Q}\mu_i \cos 2\hat{\theta}_i + \mathcal{U}\mu_i \sin 2\hat{\theta}_i \right). \quad (3.12)$$

Now each estimated emission angle $\hat{\theta}_i$ comes with its own individual modulation factor μ_i , as opposed to a global μ . This likelihood can be maximized in exactly the same way as eq. 1.12. Following §1.1.2, assuming $|\mathcal{Q}\mu_i| \ll 1$, $|\mathcal{U}\mu_i| \ll 1$, the maximum likelihood estimators for Stokes' parameters $(\mathcal{Q}, \mathcal{U})$ are now

$$\hat{\mathcal{Q}} = \frac{2}{\sum_{i=1}^N \mu_i^2} \sum_{i=1}^N \mu_i \cos 2\hat{\theta}_i, \quad (3.13)$$

$$\hat{\mathcal{U}} = \frac{2}{\sum_{i=1}^N \mu_i^2} \sum_{i=1}^N \mu_i \sin 2\hat{\theta}_i. \quad (3.14)$$

Following the notation in Kislat et al. [38] and Peirson and Romani [4], one can define the individual event weights

$$w_i = \mu_i \quad (3.15)$$

and the total modulation factor (now a weighted average)

$$\mu = \frac{\sum_{i=1}^N w_i \mu_i}{\sum_{i=1}^N w_i} = \frac{\sum_{i=1}^N w_i^2}{\sum_{i=1}^N w_i}. \quad (3.16)$$

Then eqs.3.13,3.14 are

$$\hat{Q} = \frac{2}{\mu \sum_{i=1}^N w_i} \sum_{i=1}^N w_i \cos 2\hat{\theta}_i, \quad (3.17)$$

$$\hat{U} = \frac{2}{\mu \sum_{i=1}^N w_i} \sum_{i=1}^N w_i \sin 2\hat{\theta}_i. \quad (3.18)$$

Now when estimating the polarization parameters, each event is weighted by its expected signal $w_i = \mu_i$. As in §1.1.2, Kislat et al. [38] derive the posterior distribution for estimators eqs. 3.17,3.18. The posterior distribution in the weighted case is the same as eq.1.18 but now N is replaced with

$$N_{\text{eff}} = \frac{(\sum_{i=1}^N w_i)^2}{\sum_{i=1}^N w_i^2}, \quad (3.19)$$

the effective number of events. Because the posterior distribution has changed, now

$$\text{MDP}_{99} = \frac{4.29}{\mu \sqrt{N_{\text{eff}}}}. \quad (3.20)$$

Weighting events reduces the total number of events contributing to a measurement of the polarization parameters, making $\sqrt{N_{\text{eff}}}$ smaller, but it increases the modulation factor μ by more. Overall the MDP₉₉ is reduced, improving the polarization measurement. Since $w_i = \mu_i$ is the maximum likelihood estimator, this is the optimal event weighting scheme. It is also possible to arrive at this conclusion in reverse: starting with weighted estimators eq. 3.17,3.18 and MDP₉₉, eq.3.20, Peirson and Romani [4] prove that $w_i = \mu_i$ minimizes the MDP₉₉, or equivalently maximizes the signal-to-noise ratio.

3.3.1 Deep ensembles

A DNN approach to track reconstruction using deep ensembles gives accurate estimates of individual VM uncertainties κ_i , §2.7. If measurement uncertainties are VM, individual modulation factors are given by eq. 3.10. Then, for a set of deep ensemble predicted emission angles and uncertainties, $\{\hat{\theta}_i, \hat{\kappa}_i\}_{i=1}^N$, weighted maximum likelihood estimators for the polarization parameters eq. 3.17,3.18 should use

$$w_i = \mu_i = \frac{I_1(\hat{\kappa}_i)}{I_0(\hat{\kappa}_i)}. \quad (3.21)$$

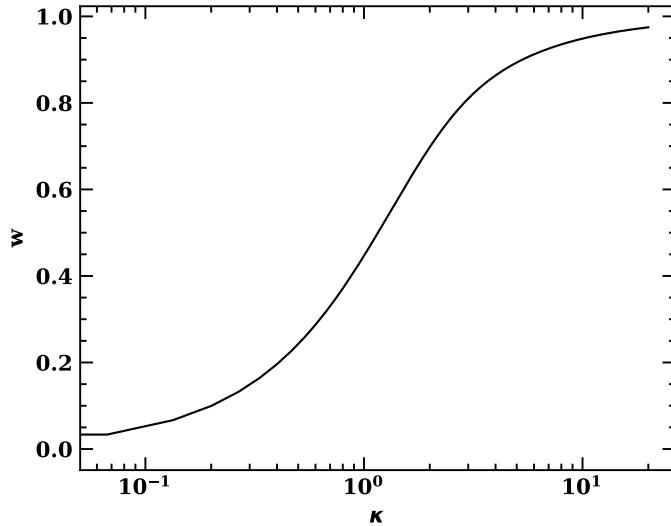


Figure 3.1: Modulation factor or optimal event weight as a function of von Mises concentration parameter κ . The function is monotonic, higher concentration parameters are always weighted more.

Fig. 3.1 plots the optimal weighting (or equivalently, the modulation factor) as a function of κ . As expected, $0 \leq \mu_i \leq 1$ and an increasing concentration parameter κ (decreasing uncertainty) always increases the event weight.

3.3.2 Simulation-Based inference

The framework for polarization estimation described in this chapter, when viewed from Bayesian perspective, can be thought of as a form of simulation-based inference [79]. The process to go from photoelectron track images to polarization parameters can be represented as a probabilistic graphical model, fig. 3.2. The source polarization parameters, polarization fraction and EVPA: p_0, ϕ , generate N photoelectron directions θ , which go on to each produce an observed photoelectron track image x in the detector. The probabilistic relationship between the polarization parameters and the photoelectron direction is known, eq. (1.2), but there is no equivalent analytical probabilistic relationship between track images and their initial photoelectron direction. With a deep ensemble trained on simulated track images, we are able to learn this missing relationship, completing the graph and allowing for hierarchical inference of the polarization parameters.

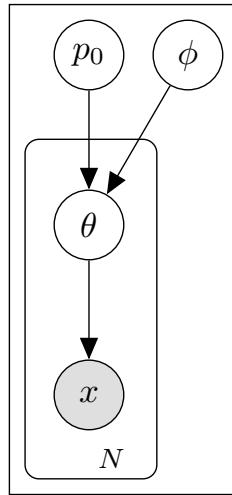


Figure 3.2: Probabilistic graphical model representing the hierarchical data generating process for photoelectron track images in an X-ray polarimeter. Shaded nodes are observed.

3.4 Performance

It is now possible to compare the full polarimetry data analysis pipeline for a DNN approach with the classical moment analysis. Track reconstruction, step 1 of the pipeline, has already been compared in §2.7; this section will reveal how both track reconstruction improvements and uncertainty estimation affect polarization recovery. The quality of polarization recovery is measured by the MDP₉₉, eq.3.20.

3.4.1 Weights

Before comparing methods, it is important to confirm that individual modulation factors μ_i (or weights w_i) predicted by the deep ensemble for each event, eq.3.21, are accurate. This can be checked by observing a 100% polarized source and grouping events with same predicted μ_i , then empirically measuring the modulation factor μ of the group with eq.1.9. The measured and predicted μ should always match.

Fig. 3.3 gives the measured modulation factor μ as a function of deep ensemble predicted weights $w_i = \mu_i$ for two polarized IXPE datasets with 1.5×10^6 events. Any deviation from $y = x$ means the deep ensemble predicted uncertainties are imperfect. When most tail events are removed using a DNN tail cut, §2.5, the deep ensemble predicted weights are close to perfect and show minimal spectral dependence. If tail events are allowed to remain in the spectrum, deep ensemble weights stray from the optimal MLE for highly weighted events. This is because it is difficult for the deep ensemble to learn the appropriate uncertainties $\hat{\kappa}$ for tail events, §2.7. The effect is reduced for the PL1 spectrum where there are fewer high energy events and thus fewer tail events.

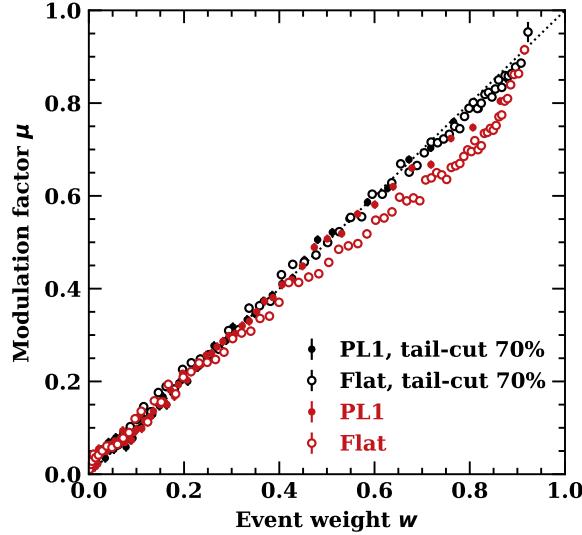


Figure 3.3: Measured μ as a function of deep ensemble event weight w for large test data sets of 1 – 10keV simulated events. Each μ bin contains 20,000 individual track events. Closed circles represent a PL1 source ($dN/dE = E^{-1}$) convolved with IXPE’s effective area, open circles a flat spectrum. Spectra with tail event cuts applied are in black, and full spectra in red.

3.4.2 Comparison

To expose which parts of the DNN approach are providing the biggest MDP_{99} improvements, table 3.1 gives the performance breakdown of the different components of the deep ensemble approach. The test datasets are inspired by realistic astrophysical spectra convolved with IXPE’s effective area, to give an idea of realized instrument improvements when using DNNs. Since the absolute MDP_{99} values are specific to simulated IXPE events, the relative improvements between methods are the important takeaways as these are likely to generalize to all imaging X-ray polarimeters.

Two baseline classical methods are provided in table 3.1, the standard moment analysis described in §1 and a weighted version. The weighted version uses event weights derived from moment analysis predicted track ellipticities. Events with higher ellipticities are weighted more in polarization estimation, eq.3.17,3.18. The specific function that transforms ellipticities to weights is optimized numerically. The ellipticities provide a rough proxy for emission angle uncertainty. Even this approximate uncertainty calculation can yield significant MDP_{99} improvements.

The base DNN approach predicts emission angles using a deep ensemble but does not use the predicted uncertainty estimates. The improved emission angle estimates yield a marginal improvement over the moment analysis and are worse than a moment analysis with ellipticity based weights. Improvements are small since most astrophysical spectra are soft, containing many low energy photons, where DNN and moment analysis emission angle accuracies are comparable. Clearly, the main

Spectrum	Method	MDP ₉₉ (%)
PL2	Mom.	5.36 ± 0.03
	Mom. w/ Ellip. weights	4.89 ± 0.03
	DNN	5.10 ± 0.03
	DNN w/ wts.	3.99 ± 0.02
	DNN w/ wts. 95%	3.98 ± 0.02 ←
	DNN w/ wts. 70%	4.08 ± 0.02
PL1	Mom.	4.80 ± 0.02
	Mom. w/ Ellip. weights	4.37 ± 0.02
	DNN	4.50 ± 0.02
	DNN w/ wts.	3.58 ± 0.01
	DNN w/ wts. 95%	3.57 ± 0.01 ←
	DNN w/ wts. 70%	3.83 ± 0.01

Table 3.1: Sensitivity analysis for two power law spectra ($dN/dE \sim E^{-N}$; PL2 for $N = 2$, PL1 for $N = 1$) each normalized to produce 10^5 2-8 keV photons when folded through IXPE’s energy response. MDP₉₉ gives the sensitivities for the various reconstruction methods; smaller MDP₉₉ is better. Mom. denotes moments analysis. DNN denotes deep ensemble. Percentages denote tail cut thresholds, 95% means events with DNN predicted tail probability greater than 95% are removed.

improvement to be had in X-ray polarimetry is properly accounting for emission angle heteroskedasticity by appropriately weighting events. This is born out in the weighted deep ensemble MDP₉₉ results, where a < 0.75 reduction in MDP₉₉ from the standard moment analysis is achieved for both spectra. Compared to the classical ellipticity weight approach, the deep ensemble achieves a ~ 0.82 MDP₉₉.

A 0.75 reduction in MDP₉₉ means a $0.75^2 = 0.563$ reduction in the number of counts (and observing time) required to reach the same signal-to-noise ratio. This makes a lot more science possible for an X-ray polarimetry mission, with no changes to the existing hardware.

Removing tail events using the tail vs peak classifier, §2.5, can provide a small additional sensitivity boost. However, improvements in energy resolution, §2.7, require substantially stronger tail exclusion with a threshold of 70% or less. There is a trade-off between achieving good spectral performance and minimizing the MDP₉₉ because the DNN tail vs peak classifier is not perfect. When tail event cuts are applied, some peak events are lost. What tail cut threshold should be chosen will depend on the specific telescope and science goals.

It is important to note again that all of the results presented in this chapter are for *simulated* IXPE events. Fortunately, tests of the DNN techniques presented here on real IXPE GPD data have yielded similar relative improvements over classical techniques, once real detector systematics are taken into account. If the simulated detector events are close enough to real events, DNN performance will generalize.

Chapter 4

Synchrotron Polarization of Relativistic Jets

4.1 Introduction

We describe a geometric model for synchrotron radiation from blazar jets, involving multiple emission zones with turbulent magnetic fields and a transient core with a helical B field. Including the effects of jet divergence, particle cooling and the Relativistic PA rotation (RPAR) to the observer frame, we find polarization behavior consistent with recent data from monitoring campaigns. We predict that under some circumstances multi- π rotation phases should exhibit relativistically-induced steps in rate $d\text{PA}/dt$ and modulation in polarization Π that can be helpful in pinning down the jet Γ and θ_{obs} . Also, RPAR enhances waveband differences that will be particularly interesting for comparing radio, optical and, soon, X-ray PA and Π variations.

4.1.1 Blazar EVPA Variability

The polarization fraction Π and electric vector position angle θ_{EVPA} of blazar emission have long been known to exhibit stochastic variability. Indeed optical polarization variability is a defining property of the BL Lac class. Recent monitoring campaigns have revealed new polarization patterns. The typical behavior is a stochastic variation about $\Pi \sim 0.05 - 0.15$ fluctuating with Π/σ_Π and $\sigma_{\theta_{\text{EVPA}}} \sim 1$. In addition, periods of relatively steady rotation of the EVPA, sometimes extending many $\times \pi$, can occur lasting weeks or months [65], after which the EVPA returns to the stochastic phase. These may be associated with flares in the total intensity [16], but this is not always the case.

A few other trends have been noted. Blinov et al. [16] indicate that Π is on average smaller in the rotating phases. However there are many examples where π increases during rotation; these

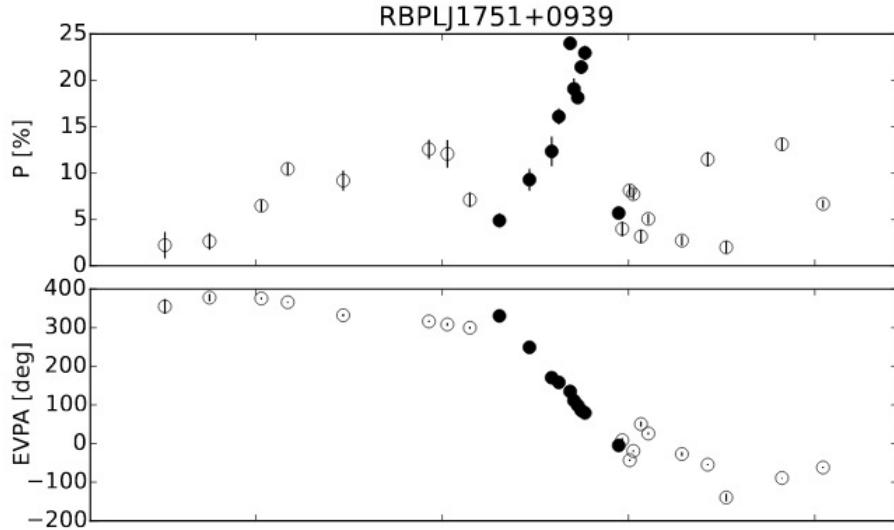


Figure 4.1: The polarization fraction and EVPA in the R-band against the time in days during an observed rotation by RoboPol [16] of the LBL blazar J1751+0939. Filled points mark their identification of a ‘rotating’ epoch.

seem more common for the long $\Delta PA > \pi$ rotations (I. Liidakis, priv. comm.). Also σ_{Π} of a given source appears to be similar in the stochastic and rotating phases. Thus while a systematic B structure should be present to drive the large angle EVPA swings an underlying stochastic process must continue. Further, there is a tendency for the mean EVPA during the stochastic phase to correlate on the sky with the projected jet axis [80]. Radio and optical polarization behavior can be similar [81], with Π often higher in the optical. There appear to be an association between GeV flares and rotation phases [82]. Also rotations appear to be recurrent in some blazars and absent in others. However none of these trends is universal. An example of a source making a transition from stochastic to rotation phase and back is shown in section 4.1.1.

Attempts to model such behavior have taken a variety of forms. In Hughes, Aller, and Aller [83] a set of multiple shocks in the jet was used to reproduce the stochastic fluctuations, while in Marscher [17] multizone turbulence in a conical standing shock was posited to generate the stochastically variable emission. Such stochastic variation can induce epochs of relatively constant PA sweep, but it was shown [16] that the incidence (and persistence over many $\times \pi$) of the observed sweeps are inconsistent with purely stochastic models. Accordingly, models for the rotating phase generally invoke helical structures in the jet. For example Zhang et al. [84] invoke a helical field energized by a standing shock. Such fields are suggested by Faraday rotation gradients transverse to the local jet directions in several nearby blazars e.g. PKS 0745+241, PKS 0820+225, Mrk 501, 3C 371 [85] and

could quite naturally be attributed to field symmetries imposed at the jet base by the Blandford-Znajek process [62]. Alternatively, Lyutikov and Kravchenko [67] assume that the jet itself takes on a helical form, e.g. due to precession, with an aligned embedded field. A third picture [68] posits a helical kink propagating along a conical jet with an embedded toroidal B field. Each of these pictures can accommodate smooth multicycle rotations.

We explore here a heuristic model that incorporates the main features above, in an attempt to reproduce the range of observed optical polarization phenomena and to predict new correlations to be tested with multiwavelength observations. We start (§2) with a description of the important, but under-appreciated, effect of relativistic boosting on the observed polarization. We then describe (§3) a toy geometry with multiple zones transitioning between random and helical magnetic patterns and propagating downstream in a conical jet. We then couple this with a radiation model that follows the cooling and synchrotron radiation of the e^+/e^- (§4) and comment on the patterns and multiwavelength correlations of the resulting polarization signal. §5 applies this picture to the bright HBL Mrk 501, and we conclude with general predictions for future extensions and comparisons with the data.

4.2 Relativistic PA Rotation

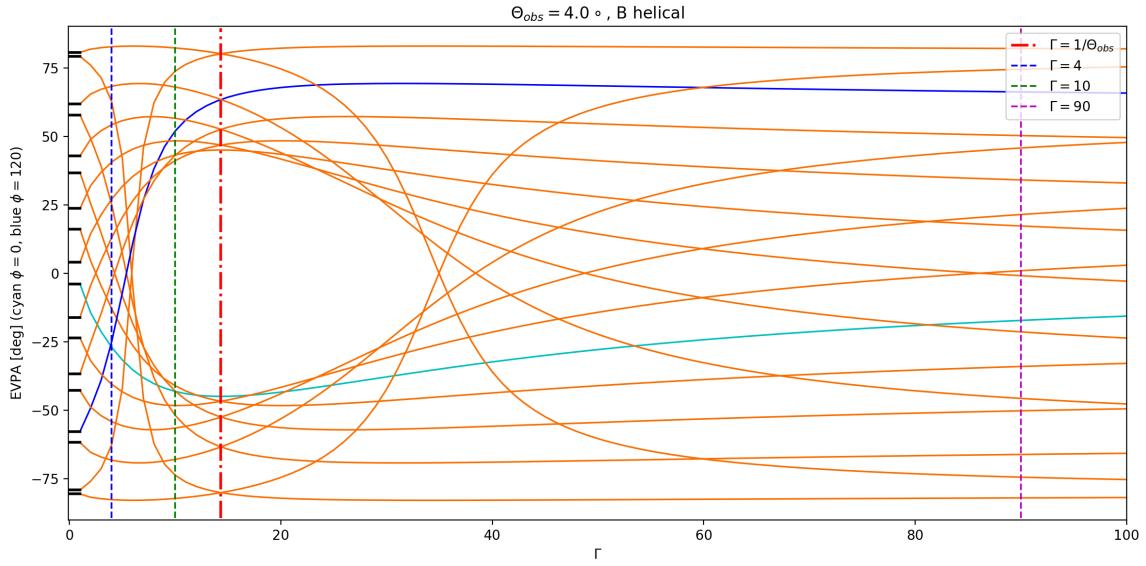


Figure 4.2: EVPA in the observers frame on the plane of the sky as a function of jet bulk gamma Γ for a helical B-field with a 45° pitch angle. The direction of the jet is 4° off our line of sight in the \hat{x} direction on the plane of the sky, corresponding to 0° on the plot. The B-field is sampled every $\phi_B = 20^\circ$ from $0 - 360^\circ$. The black lines on the y-axis denote the observed EVPAs without RPAPR (i.e. $\Gamma = 1$). The solid cyan and blue lines mark $\phi_B = 0^\circ, 120^\circ$ respectively.

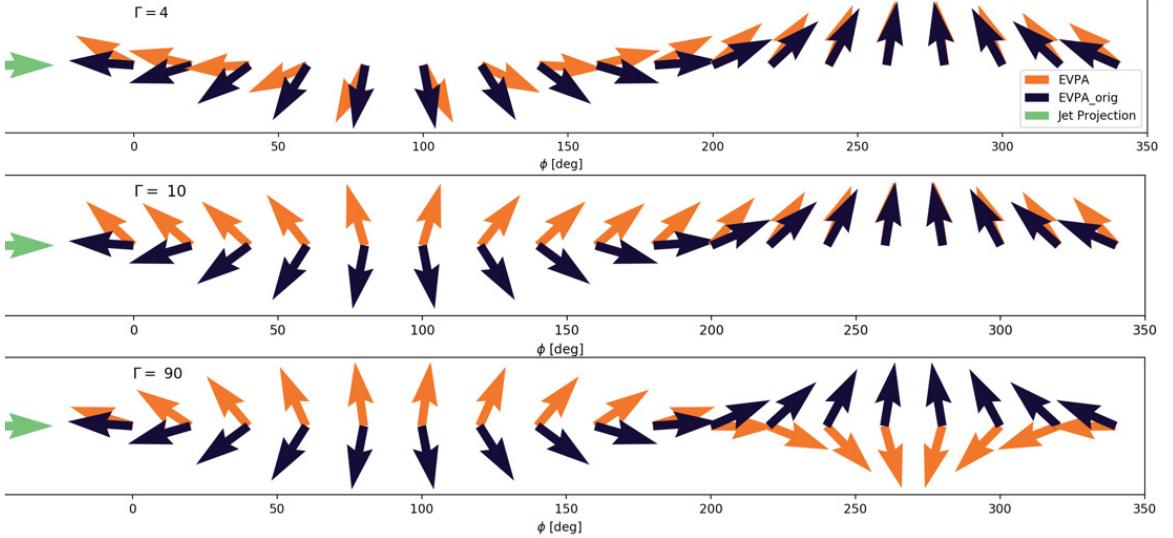


Figure 4.3: Arrows denoting the observed EVPA on the plane of the sky for the jet described in section 4.2 at the 3 different Γ cuts. The black arrows show the initial EVPA absent RPAR (i.e. for stationary $\Gamma = 1$), the orange arrows show the RPAR-shifted EVPAs. The projected jet axis is along the green arrow. Arrows are plotted as single headed for clarity, although observations will span 180° (i.e. -90 to $+90$ in section 4.2).

In most cases models assume that the observed polarization direction is that of the EVPA in the jet fluid frame θ_{jf} . However, Lyutikov, Pariev, and Blandford [86], following Blandford and Koenigl [87], show that the relativistic transformations can induce significant rotation, with the direction and magnitude of the observed angle depending on θ_{jf} , Γ_{bulk} , θ_{obs} and $\hat{\underline{v}}$ (the direction of the jet). We refer to this process as Relativistic PA rotation (RPAR) in the following discussion. Taking $\hat{\mathbf{e}}' = \hat{\mathbf{n}}' \times \hat{\mathbf{B}}'$ as the EVPA in the jet frame, they show that

$$\hat{\mathbf{e}} = \frac{\mathbf{n} \times \mathbf{q}'}{\sqrt{q'^2 - (\mathbf{n} \cdot \mathbf{q}')^2}}, \quad (4.1)$$

$$\mathbf{q}' = \hat{\mathbf{B}}' + \mathbf{n} \times (\mathbf{v} \times \hat{\mathbf{B}}') - \frac{\Gamma}{1 + \Gamma} (\hat{\mathbf{B}}' \cdot \mathbf{v}) \mathbf{v}. \quad (4.2)$$

Here prime quantities are measured in the jet frame, \mathbf{n} is the vector to the observer, $\hat{\mathbf{B}}$ is the magnetic field vector and \mathbf{v} is the jet bulk velocity vector.

For radiation emitted in a jet zone containing a helical B-field, where

$$\hat{\mathbf{B}}' = \frac{\sin(\phi_B) \hat{\mathbf{x}}' + \cos(\phi_B) \hat{\mathbf{y}}' + \tan(\Psi_B) \hat{\mathbf{z}}'}{\sqrt{1 + \tan^2(\Psi_B)}} \quad (4.3)$$

and $\hat{\mathbf{z}}'$ is parallel the jet axis. ϕ_B represents the phase angle along the helix and Ψ_B is the pitch

angle. We can see the effects of RPAR on the observed EVPAs for different Γ_{bulk} regimes in Figures 2 and 3. The jet shown here is pointed 4° off our line of sight in the \hat{x} direction on the plane of the sky. The field is pitched at $\Psi_B = 45^\circ$ to the jet axis. This pitch angle affects the magnitude of the RPAR effect [as in 88]. RPAR effects are reduced for smaller pitch angles and increased for larger ones; we thus choose $\Psi_B = 45^\circ$ for illustrative purposes. As section 4.2 shows there are three characteristic regimes for the RPAR effects. In the ‘low Γ ’ case RPAR induces a counter-clockwise shift of the EVPA for half a cycle and clockwise for the other half (with the phase controlled by the component of the helical B along the jet axis). This case is represented by the blue dotted line of section 4.2 and the first row of section 4.2. The net effect is that, for a smoothly increasing ϕ_B (an EVPA rotating smoothly in the fluid frame), the lab EVPA rotates quickly for half a cycle and slows in the second half, centered perpendicular to the projected jet axis. Thus an intrinsic smooth rotation is seen as a set of ‘stair-steps’ spanning 360° .

For somewhat faster jets, the ‘high Γ ’ case, the RPAR bias (here toward $+y$) dominates so the EVPA is driven increasingly transverse to the jet. In this regime the observed rotation can switch between clockwise and counter-clockwise rotations for the same field geometry. This is the green dotted line of section 4.2 and the second row of section 4.2. For very large jet bulk motion, the ‘extreme Γ ’ case, the observed EVPA is actually reflected across the jet axis from the initial vector. One can understand these behaviors by picturing a relativistic electron emitting synchrotron radiation in the fluid frame within a uniform B-field at arbitrary pitch angle, where the radiation is intermittently directed towards the observer. The electron will emit beamed radiation polarized mostly perpendicular to the B-field. As this emission is boosted to the lab frame, slightly off axis, the radiation initially directed to the side is boosted toward the observer, rotating the PA. For extreme Γ , the radiation initially at large angle to the line of sight, is boosted toward the observer and begins to dominate the received signal. The EVPA flips across the projected jet axis. Of course, the bulk of the radiation is boosted into an angle $1/\Gamma$ (red dot-dashed line). Thus to see the ‘extreme Γ ’ case one needs to observe so far off-axis that the observer sees only a tiny portion of the observable jet power. Thus virtually all observers will be in the first two cases, with typical jet alignment resulting in the ‘Low Γ ’ case and only jets viewed well off-axis reaching the ‘High Γ ’ regime. Typical HBL parameters inferred from VLBI observations and SED fits have $\theta_{\text{obs}} = 1^\circ - 5^\circ$, generally in the ‘Low Γ ’ case. Thus ‘stair-steps’, regular slope variations in extended rotation phases, will often be present, although the amplitude is quite sensitive to field pitch angle and jet parameters.

In section 4.2 one can see that for small Γ the observed EVPA is generally driven toward the jet axis by RPAR. Although for the particular pitch angle Ψ_B shown there is a range of ‘Large Γ ’ vectors that are driven perpendicular to B, when one averages over all magnetic inclinations the former effect dominates. At large Γ the forward boosting drives to the jet axis as well. Thus we see that in section 4.2 that a random set of initial EVPA vectors is driven to the jet axis, except for a range around $1/\theta_{\text{obs}}$, which would be seldom observed. As noted in the introduction, a statistical

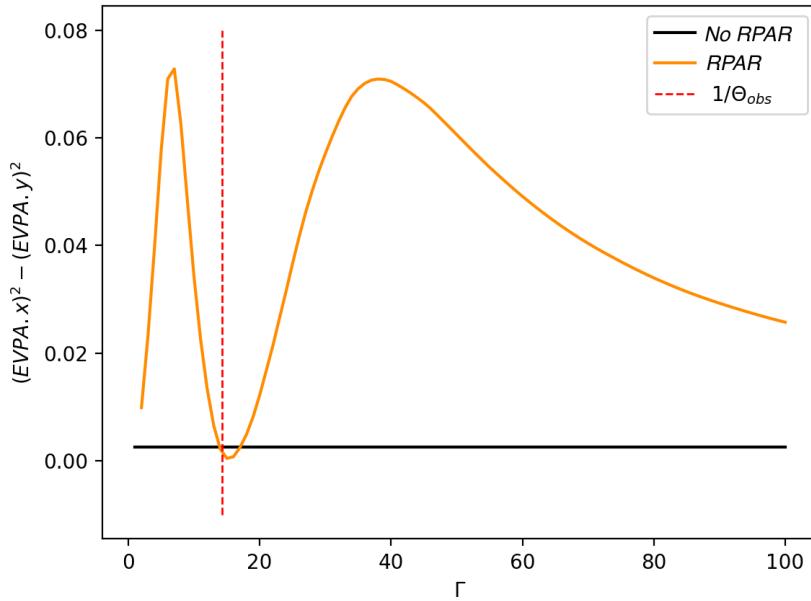


Figure 4.4: EVPA bias along jet axis against Γ_{bulk} for a set of 127 random B-field zones in the jet frame. The bias is calculated as the difference in squared EVPA components parallel (\hat{x}) and perpendicular (\hat{y}) to the jet axis. The bias is averaged over 500 random realizations. $\theta_{obs} = 4.0^\circ$.

bias for such alignment is indeed observed [80].

4.3 Geometrical Jet Model

The modest, and fluctuating, blazar polarization in the stochastic phases suggests that many zones contribute to the radiation seen with uncorrelated, and varying magnetic field orientation. We thus follow many authors (e.g. Marscher (2014)'s TEMZ model) in assuming a multizone emission region. We attribute this to shock-induced turbulence. While a typical spectral index indicates a saturated polarization level $\Pi_{max} \sim 0.7$, we more commonly observe $\Pi \approx 10\%$, indicating $N \sim (\Pi_{max}/\Pi)^2 \sim 50$ uncorrelated zones contribute to the observed emission. At any one time one receives radiation from an angle $\sim 1/\Gamma$ about the line of sight. Thus if the jet opening angle θ_{op} is $< 1/\Gamma$, then this is the full number of zones seen at a given radius; wider jets will have $\sim N$ zones in the observed subset (see section 4.3). With large Γ and small viewing angle we in general expect radiation from a single excitation radius (spanning N zones across the jet) to dominate the received radiation at any one time. The detected spectrum will represent a time averaged output of this excited radiating zone as it travels downstream in near-synchrony with its early emission.

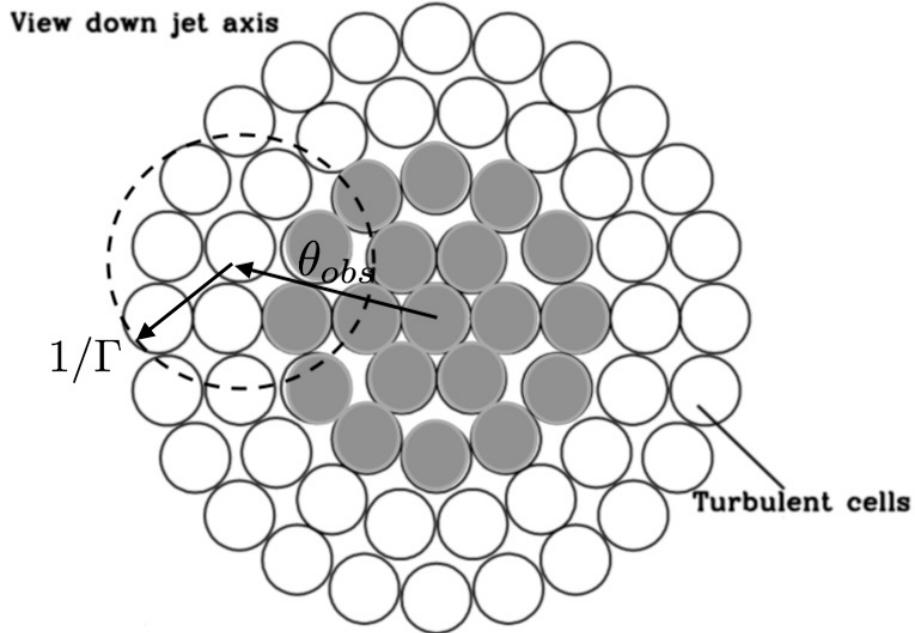


Figure 4.5: Map of zones in a jet section [17]. Each zone has a B-field direction, with random initially isotropic orientation during the stochastic phase. During a rotating phase the central zones (shaded) turn helical. The computations typically use 127 zones (6 concentric rings); fewer are shown here for clarity.

However, as noted, the prevalence of persistent rotating phases is inconsistent with a simple superposition of random polarization vectors [16] and the presence of rotations $> \pi$ demands a deterministic underlying structure. Again with high Γ and small θ_{obs} we observe only one jet radial zone at a given time. Thus if there is an underlying helical field passing through the excitation zone (e.g. a standing shock) then only one portion of the helix will be visible to the observer at any time and the polarization will mark the orientation of its field. To implement this, we assume that a fraction of the jet zones are imprinted with this coherent, helically varying, field (shaded zones in section 4.3). Given the modest Π during the rotating phase and the similarity of σ_Π to that of the stochastic phase, the helical coherent field occupies only a small fraction of the observed jet zones. To illustrate strong rotational signals we use $\sim 1/6$ here, although a smaller fraction may be coherent in actual jets.

We expect the jet to be imperfectly collimated, so that its cross section will increase as it propagates. This has two principal effects. First, each zone of the jet has a different θ_{obs} . If one views within θ_{op} then values from 0 to $\sim \min(1/\Gamma, \theta_{op})$ contribute. This averages out some of the RPAR effects. Second, the spreading jet affects the coherent field geometry. Imagine a helical field injected into such a conical jet. Assuming uniform constant Γ , the distance and time between rotations remains fixed. However the radius of the spiral grows. Hence the pitch angle flattens

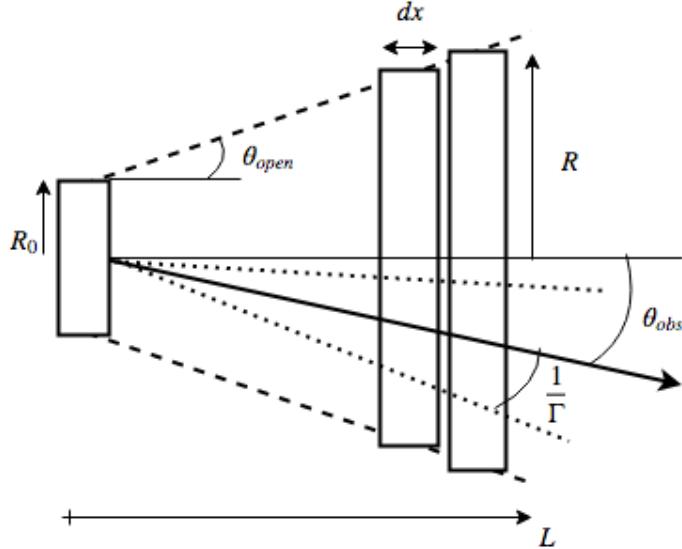


Figure 4.6: Schematic showing the shape of the ‘conical’ jet model with some of the important parameters labelled, [18]. A ‘section’ of the jet is a dx slice.

and the field become more nearly transverse. This leads to evolution in the PA behavior as the jet propagates.

We now show these geometrical effects in two simulated realizations. Both have low $\Gamma = 5$ and opening angle $\theta_{op} = 9.5^\circ < 1/\Gamma$, and we show a 770° rotating bracketed by stochastic phases. RPAR effects and beaming geometry can introduce an especially interesting behavior in extended rotating phases. Especially in the low Γ regime, even a smoothly varying helical field produces non-uniform sweep of the observed PA (section 4.2, top line). For extended rotation events we can expect to see this rate modulation as a series of rotation rate steps. section 4.3 shows a simulation of this behavior. This is most prominent when the effective θ to our line of sight of the helical zones within the $1/\Gamma$ solid angle corresponds with the appropriate Γ given by Figures 2 and 3. Possible examples of such stepped sweeps are seen in PKS 1510-089 [89] and S5 0716+71 [90].

Of course, in this picture the coherent fields in a portion of the N zones during the rotating phase induce higher polarization. This is not universally seen. In as far as rotations are associated with flare events, turbulence associated with the flare power injection could increase the effective N , lowering Π . Also increased activity during flares may drive higher energy emission closer to the standing shock, where the initially isotropic zones dominate (see below), also lowering Π . Regardless, as long as the fraction of zones participating in the rotation is not large, the polarization fluctuations σ_Π will remain substantial and closely resemble those in the non-rotating phase [16].

Next we show the behavior for a conical jet for the same Γ and θ_{op} but a lower θ_{obs} in section 4.3. Π is larger and fluctuations are smaller since the helical zones now dominate the central position

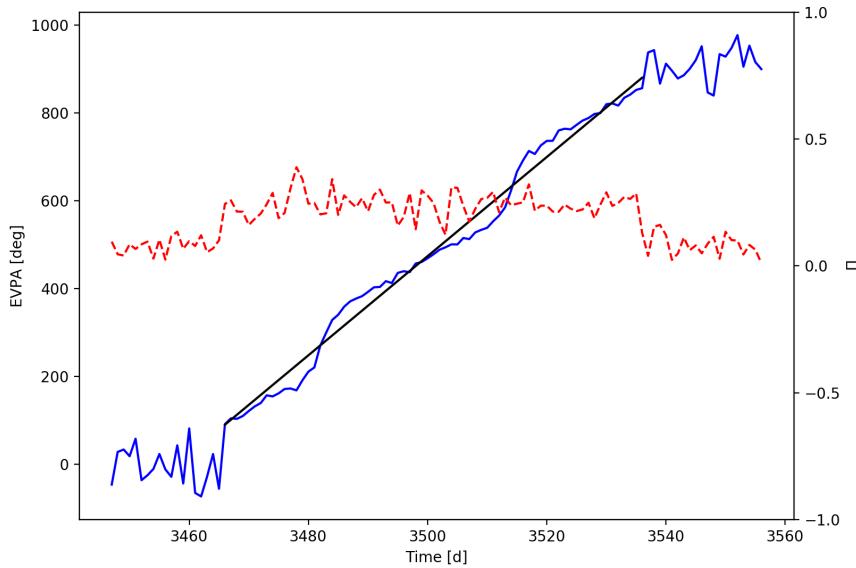


Figure 4.7: EVPA (blue) and polarization fraction (red) for stochastic and rotating phases in a jet with projection on the plane of the sky along the x axis. We are viewing at $\theta_{\text{obs}} = 4.0^\circ$ with $\Gamma = 5$ and $\theta_{\text{op}} = 9.5^\circ$. RPAR has been applied. This low Γ means the vast majority of zones in the jet contribute. The black line shows the same helical rotation without RPAR or turbulence. The rotation phase spans 770° with starting phase $\phi = 3\pi/2$.

in our line of sight. The PA sweep is steady in the rotation phase now with only small stepping behavior present. A reduction in the effective θ of the helical zones to our line of sight means that for the same Γ we move to the left on section 4.2, bringing us to lower RPAR regime. The stair-stepping effect is much weaker here. Indeed the lack of measurable steps during the rotation phase of J1751+0939 (section 4.1.1) places some limits on θ_{obs} and Γ . This source has a VLBI estimated $\beta_\perp = 7.9 \pm 0.8$ [91] and we find that $\theta_{\text{obs}} < 2^\circ$ and $\Gamma > 11$ when pitch angle Ψ_B is near 0° . Steeper pitch angles place more stringent constraints, e.g. $\Psi_B = 25^\circ$ implies $\theta_{\text{obs}} \leq 0.8^\circ$ and $\Gamma \geq 16$.

4.4 Radiation Model

Our radiation model is inspired by Potter and Cotter [18] and uses the basic jet setup shown in section 4.3. We take the jet to be composed of an electron-positron plasma (hereafter electron). Under the assumption of equipartition between electron energy and magnetic field energy, the radius and electron population are initialized at the base of the jet from the input parameters: the length of the jet L_0 , the total jet power W_j , the bulk Lorentz factor Γ , the magnetic field strength at the base B_0 , the maximum initial electron energy E_{max} , the electron power law index α , the jet observation

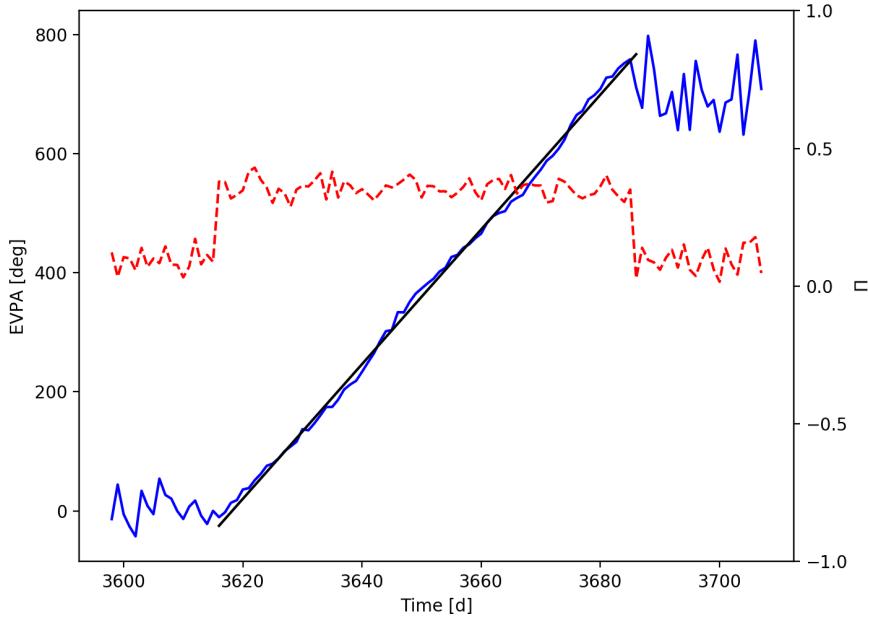


Figure 4.8: As for section 4.3 for the rotational phases of a conical jet viewed at $\theta_{\text{obs}} = 1.5^\circ$. This represents a 770° rotation with starting phase $\phi = 160^\circ$. Stepping is less prominent.

angle θ_{obs} and the jet opening angle θ_{op} . Furthermore, it is assumed that the electrons and magnetic field are homogeneously distributed throughout each jet section and that the magnetic field energy is conserved.

The initial electron population is assumed to have a power law energy distribution with an exponential cutoff [92] and is represented by a discrete set of energy bins.

$$N(E_e) = AE_e^{-\alpha} e^{\frac{-E_e}{E_{\max}}}. \quad (4.4)$$

For each jet zone we divided into sections of length dx (see section 4.3). In a particular section, the synchrotron emission from each electron energy bin is calculated using the B field in that section. The synchrotron losses then determine the evolution in the electron energy bin populations for the next section. Furthermore, since we are viewing down the jet, the self-absorption opacities for each jet section are calculated and applied to the spectrum - these depend on θ_{obs} and synchrotron photon energy. This is a quiescent jet model, without flux variability in time; only the B-field directions in the zones change for the purpose of polarization modeling.

To find the polarization in the conical jet, we calculate the electrons' synchrotron power per unit frequency perpendicular and parallel to the B-field projection on the plane of the sky for each zone

in each section. These are [93]:

$$P_{\perp}(\omega) = \frac{\sqrt{3}q^3 B \sin \alpha}{4\pi mc^2} [F(x) + G(x)], \quad (4.5)$$

$$P_{\parallel}(\omega) = \frac{\sqrt{3}q^3 B \sin \alpha}{4\pi mc^2} [F(x) - G(x)]. \quad (4.6)$$

Here $F(x)$ and $G(x)$ are the modified Bessel functions [e.g. 94]. This radiation is then subject to RPAR and the intensity is Doppler boosted by δ^4 . We next sum, using Stokes' parameters, the contribution of all zones and sections to obtain the total powers in a coordinate system aligned with the projected jet axis on the plane of the sky. These then provide the polarization fraction:

$$\Pi(\omega) = \frac{P_{\perp}(\omega) - P_{\parallel}(\omega)}{P_{\perp}(\omega) + P_{\parallel}(\omega)}. \quad (4.7)$$

The projected net EVPA (relative to the jet axis) can also be referenced to an absolute angle.

Note that in this model the electron population evolves (cools) along the jet. For a conical jet we assume that the field also evolves, becoming increasingly transverse as the jet expands. This means that we expect measurable differences in Π and PA as a function of observed frequency. Since the high energy electrons cool most quickly, for appropriate parameters their radiation may come only from the base of the jet. Lower energies come from a large range of jet radii and thus, for a conical jet, a more transverse field structure. Thus we expect an energy-dependent shift in the observed polarization properties. Note also that, for a given number of emission zones N , the conical jet model will have a higher polarization than the simple cylindrical jet case. This is both because the jet divergence means that the received radiation is dominated by zones close to the line of sight and because down-stream fields are increasingly transverse and hence, even for the stochastic case, increasingly coherent. In particular, at low (radio) energies one expects EVPA increasingly aligned with the projected jet axis (fig. 4.9). This is indeed observed.

The change in the effective B field pitch angle as one moves along the jet also introduces energy-dependent shifts in the EVPA at a given phase. From the underlying geometry, with small θ_{obs} and conical jets, the effects are relatively subtle, since with the jet viewed nearly end-on, even dramatic pitch angle changes make modest change to the projection on the sky. However, RPAR effects make the observed PA sensitive to the full polarization vector and greatly enhance the sensitivity to the magnetic field inclination, even for nearly aligned jets.

Interestingly, this can introduce a rotation-phase dependent modulation of amplitude \bar{A} in the degree of polarization Π , which is especially strong for the low energy synchrotron emission from larger distances in the spread jet (fig. 4.10). It will be interesting to see if this pattern can be recovered from monitoring observations at high (core dominated) radio frequencies.

By taking the electron population to be homogeneously distributed across the jet zones, our

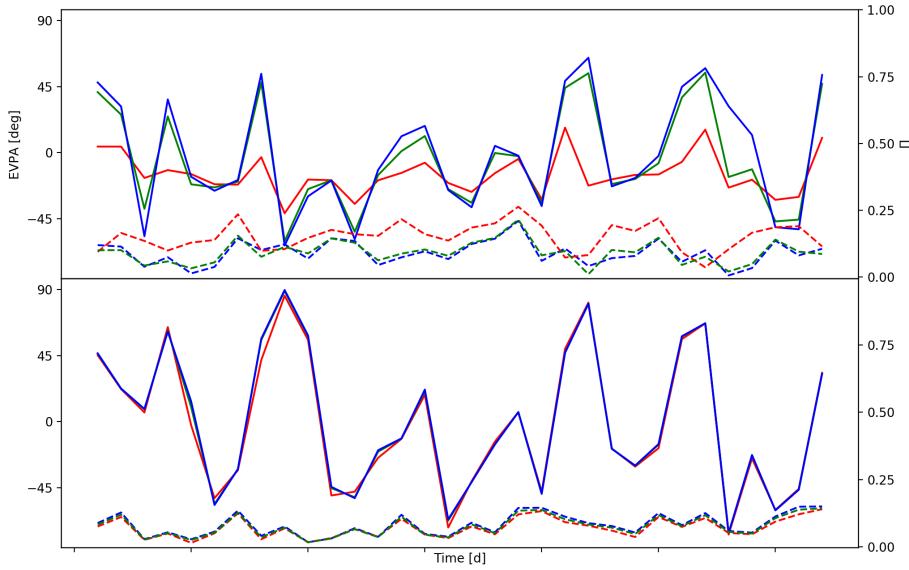


Figure 4.9: Energy variation for EVPA and Π during a non-rotating (stochastic) phase. Colors denote low (Radio, red), mid-range (Optical, green) and near cut-off (e.g. X-ray, blue) synchrotron bands. This is for a jet with $\theta_{\text{obs}} = 1.5^\circ$, $\Gamma = 5$ and $\theta_{op} = 9.5^\circ$. Both panels use the random seed for the generated B-fields. The top panel shows a RPAR affected jet: substantial energy dependence is seen, with low energy (radio) PA better aligned with the parent jet. Low energy fluctuations are also somewhat smaller. In the lower panel RPAR effects are ignored and the behavior is essentially achromatic with lower overall polarization fraction.

model assumes the main source of energy dependence in polarization to originate from the B-field change along the jet linked with RPAR. However, having higher energy electrons relegated to fewer zones as in Marscher and Jorstad [95] or Angelakis et al. [96] has been invoked to explain the blazar sequence polarization trends described in the studies of Itoh et al. [97] and Angelakis et al. [96] which suggest LBLs have higher polarization fraction (up to 40%) and variability than HBLs (up to 10%) in the optical. Indeed if the X-ray polarization fraction is observed to be much greater than the optical in HBL sources, one could plausibly extend this model by having X-ray synchrotron emission from fewer zones as above.

In this chapter we treat only polarization of the synchrotron emission, so this model describes up to the peak ν_{sy} . This is in the X-ray band for HBL, but typically in the IR/optical band for other blazars. Treatment of Inverse-Compton regime polarization is covered in Zhang et al. [98]. In general IC components should have a lower polarization fraction, including when SSC dominates. We expect this will decrease the observed Π in our model for a given number of radiation zones N . When external photon fields dominate, the observed Π will be even smaller.

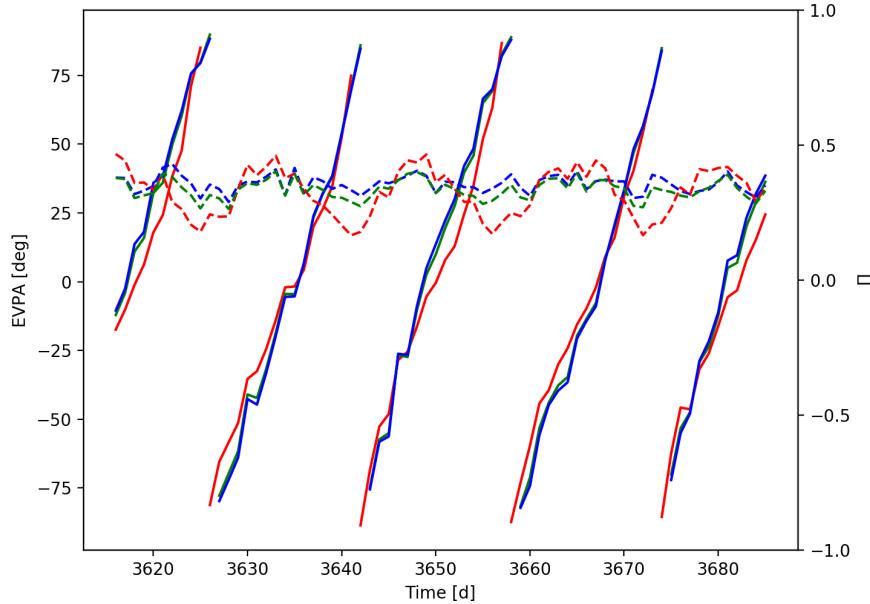


Figure 4.10: Energy dependence during a rotating phase (the 770° rotation shown in section 4.3), as for fig. 4.9. Strong sinusoidal variation of the radio polarization fraction occurs as the rotating helical zone EVPA aligns parallel and perpendicular to the jet axis. Since the random zones' EVPA are increasingly aligned with this axis as one progresses along the jet (i.e. observes at lower frequency; see top panel, fig. 4.9) the total Π has a sinusoidal modulation. The strength of the modulation depends on the ratio of helical zones to random zones.

4.5 Application to Mrk 501

One motivation for this study is the new prospect of measuring blazar X-ray polarization with *IXPE* [69] or other upcoming missions. To date only a handful of X-ray polarization measurements have been made (OSO-8, PoGo+, X-Calibur) mostly of the Crab Nebula with no blazar measurements as of yet [99]. However new facilities should provide a number of good polarization measurements, making an evaluation of energy dependence timely. For many LBLs, thermal emission is important in the optical band and synchrotron emission does not dominate. Also, the radio emission is often dominated by larger scale jet flux far from the acceleration zone. So intraband comparisons of HBLs are especially interesting since the optical and even the X-ray can come from the synchrotron peak, allowing multiband comparisons to probe the RPAR and jet geometry effects described above.

We thus illustrate the various polarization phenomena with simulations of an HBL source, Mrk 501. This blazar displays substantial optical PA variability, including EVPA rotation and can be well measured by IXPE in a few day's exposure. Mrk 501's Doppler factor δ has been estimated

Fig	$\bar{\Pi}_R$	$\bar{\sigma}_R$	\bar{A}_R	$\bar{\Pi}_X$	$\bar{\sigma}_X$	\bar{A}_X
4.9	0.15	0.051	—	0.09	0.045	—
4.10	0.32	0.031	0.11	0.36	0.029	0.016

Table 4.1: Radio and X-ray polarization fractions for Figures 9 (top panel) and 10. For the fig. 4.9 row $\bar{\sigma}$ describes the fluctuations about the mean $\bar{\Pi}$, while for fig. 4.10 it represents the uncertainty in the amplitude \bar{A} of the best-fit sinusoid.

as $\sim 6 - 22$, leaving much leeway in choosing the Γ , θ_{obs} and θ_{op} . As shown in §2, different Γ lead to different RPAR behaviors, so we have modeled for two values consistent with allowed δ range, adjusting jet parameters to fit the overall SED (bottom panel of section 4.5). Table 2 shows the selected fit parameters.

δ	$B_0[G]$	γ_{max}	α	$\theta_{\text{op}}[^\circ]$	$\theta_{\text{obs}}[^\circ]$	Γ
6.6	40	$4.9 * 10^4$	1.95	13.0	4.0	7.5
17.4	30	$6.3 * 10^4$	1.95	3.9	3.3	17.5

Table 4.2: Parameters used for the Mrk 501 fits of section 4.5. The top and bottom rows are for the solid and dashed fits respectively. Both fits use $W_j = 2 * 10^{37} W$, $L_j = 5 * 10^{20} m$ and $\gamma_{\text{min}} = 10$.

Both sets of parameters fit the SED reasonably well while producing different polarization behavior. The solid line set provides a higher average polarization fraction across all bands as slightly fewer B-field zones lie within its $1/\Gamma$ range (section 4.3). Also, for this model, rotating phases (not shown) will produce strongly ‘stepped modulation’, due to the proximity of the helical zones to the line of sight. In contrast, the dashed line set does not produce EVPA rotations since the helical zones are located on the periphery of its more restricted $1/\Gamma$ range. Note that in this picture, we can predict that low Γ jets are, in general, more likely to produce rotation events. Finally, the difference in the low energy polarization behavior provides additional observables that are sensitive to θ_{obs} and θ_{op} .

4.6 Conclusion

We have explored a simple geometrical model of a conical blazar jet with multiple emission zones across the interior. By introducing a coherently rotating helical field in a subset of these emission zones and by noting that the jet magnetic fields can become increasingly transverse as the jet expands, we have been able to mimic a variety of observed blazar PA behavior. Our study treats the often neglected effect of the jet boost on the observed EVPA (RPAR) which provides interesting effects on the observed PA behavior in some regimes. Finally, by computing with a simple emission model in this jet geometry we have seen that the energy dependence of some polarization observables can also display useful dependence on the model parameters.

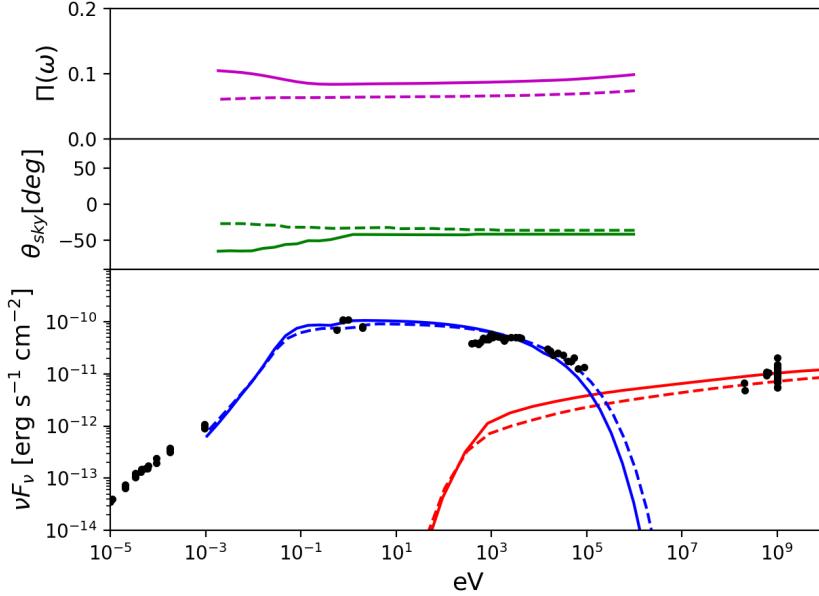


Figure 4.11: A fit to Mrk 501’s SED (bottom panel, approximately contemporaneous fluxes drawn from the ASDC compilation for 2010; <https://tools.asdc.asi.it/SED/>) using the radiative model. The top two panels show the energy dependence of Π and EVPA averaged over 200 iterations for the stochastic phase, with solid line for a low δ solution and a dashed line for high δ . The observed EVPA is expected to fluctuate about these mean values during stochastic phases. The jet projection on the plane of the sky is $\theta_{\text{sky}} = -40^\circ$, as observed for Mrk 501.

Although many of our ingredients have been considered in past studies, by combining these into a single model, we have been able to reproduce a large fraction of the observed EVPA phenomena. Moreover, the model makes interesting predictions for the (modest) energy-dependence of EVPA across the synchrotron peak of the blazar emission. With the possibility of X-ray polarization measurements in the near future, the model allows for some useful comparison with observed data sets. Novel dependence of the EVPA observables on Γ and θ_{obs} offer new ways of constraining these important parameters.

Of course extensions are needed: for most blazars IC is relevant in X-ray emission, so we should add simulation of this more weakly polarized emission to the model. But the interesting patterns in Π and PA introduced by the jet expansion and the relativistic boost provide a range of observables that can be sought in extensive polarization monitoring programs. These patterns can be useful in constraining jet geometry and, eventually, in guiding detailed RMHD modeling that will follow the shocks and acceleration giving rise to the energetic electron populations responsible for the observed (polarized) blazar emission.

Chapter 5

Synchrotron Self-Compton Polarization of Relativistic Jets

5.1 Introduction

We describe a geometric model for synchrotron and synchrotron self-Compton (SSC) radiation from blazar jets, involving multiple emission zones with turbulent magnetic fields and fully self-consistent seed photon mixing for SSC. Including the effects of jet divergence, particle cooling and the Relativistic PA rotation (RPAR) to the observer frame, we find that the multi-zone model recovers simple predictions for SSC polarization, but describes new dependencies on jet viewing geometry and zone multiplicity. Increasing the zone number decreases both synchrotron and SSC polarization, but with different scaling. A rise in synchrotron polarization fraction Π_{Sync} at high energies is guaranteed by basic relativity considerations, and strengthened by jet non-uniformity. Finite light travel time effects can suppress the synchrotron polarization at energies well below the ν_{Sync} peak. In general Π_{Sync} and Π_{SSC} are correlated with $\Pi_{\text{SSC}}/\Pi_{\text{Sync}} \approx 0.3$, but individual realizations can lie far from this trend. This study lets us estimate Π across the SED, leading to predictions in the X-ray band helpful for planning observations with *IXPE* and other upcoming X-ray polarization missions.

In chapter 4, we modeled a simple multizone conical jet model, optionally with a helical core field, and found that, when a proper treatment of relativistic PA rotation (RPAR) is included, it can explain many of the synchrotron emission trends mentioned above. In addition it makes a number new, testable predictions, which can help in interpreting future optical/IR polarization campaigns. However, we soon expect to measure X-ray polarizations with space missions such as *IXPE*; This band is often well above the ν_{Sync} synchrotron peak and, for LSP and ISP, may include significant Compton flux.

In this chapter we extend our conical jet model to include the multizone treatment of SSC

polarization in blazars. We particularly focus on the transition region between synchrotron and SSC dominated flux, as this will be *IXPE*'s range for many ISP. We start by reviewing the residual polarization after Compton scattering. We then describe how the synchrotron emission of multiple zones are combined for Compton re-processing. §4 notes how light-travel time effects additionally modify the polarization seed flux; this varies along the jet so that the effects vary with energy band. A numerical realization of this model lets us check how the averaging scales with the number of effective jet zones (which can be constrained by e.g. optical Π_O). We conclude with full Sync+SSC simulations for representative blazar parameters.

5.2 ICS Basics & Seed Photon Polarization

Bonometto, Cazzola, and Saggion [100] have developed an analytical formalism to evaluate inverse-Compton polarization for scattering in the Thomson scattering regime, with Bonometto and Saggion [101, hereafter BCS] treating the SSC case. Analytic solutions are difficult in the Klein-Nishina regime, but Krawczynski [102] provides a general Monte-Carlo based framework, verifying and extending the BCS results. To date such pseudo-analytic treatments have been applied to homogeneous single-zone jet models. The results provide useful upper limits on the plausible ICS polarization. For example Poutanen [103] explore how the magnetic field orientation in a uniform jet affects Π_{SSC} , while McNamara, Kuncic, and Wu [104] run a single zone Monte-Carlo model to show how X-ray polarization would differ between synchrotron, SSC and Externally-dominated Compton (EC) emission. Finally Zhang and Böttcher [105] argued from single zone SSC simulations that polarization measurements can distinguish between weakly polarized leptonic Compton emission and strongly polarized hadronic models.

However, with the new (especially optical) evidence for incoherent jets with multiple zones contributing to the polarized flux, these studies are inadequate to describe any but the most basic differences between jet models. In addition, many of these previous efforts do not fully incorporate the important RPAR rotation of the emitted polarization to the Compton scattering zone and on to the observer frame. Marscher [17] introduce a multizone framework; in Peirson and Romani [1] we explored synchrotron emission in this picture, including the effects above. Here we summarize the jet geometry before describing additional polarization averaging of the seed synchrotron flux and computing the final Compton polarization.

Our jet is conical with opening angle θ_{op} pointing θ_{obs} from our line of sight. The cross sectional radius at launch R_0 is determined by the jet power W_j , initial magnetic field strength B_0 and bulk Lorentz factor Γ (all values set in the jet frame) assuming an equipartition fraction = 1. The jet is segmented into slices (sections) along the jet, each made up of multiple zones i , which share the same bulk Γ but have different θ_{obs_i} . The B-field orientation varies, typically randomly, between zones. Alternatively a subset (assumed to be the jet core) has a coherent helical B-field during

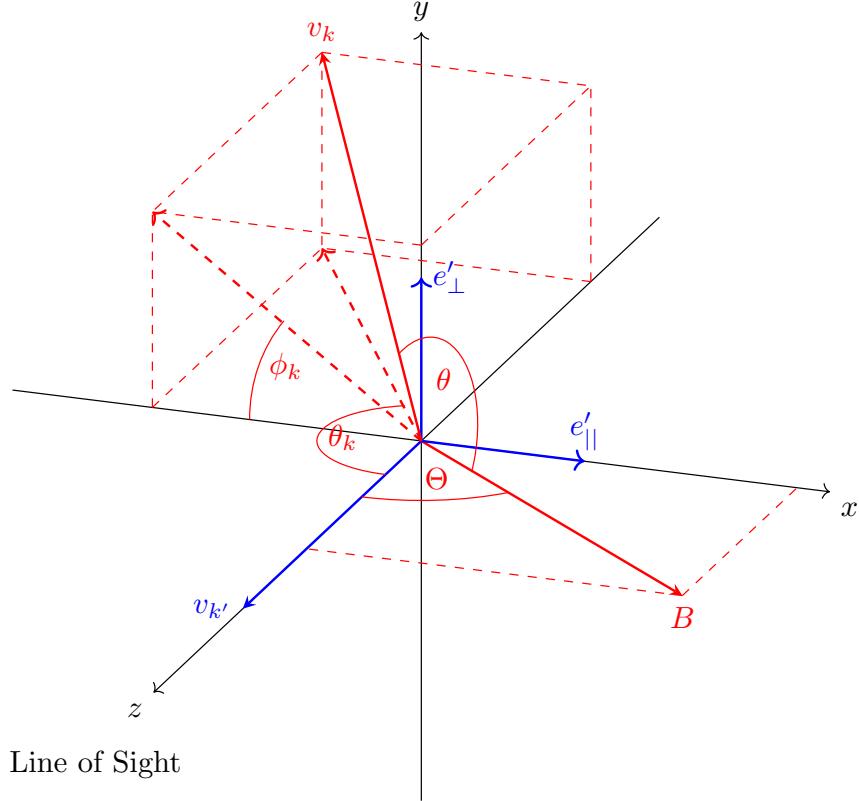


Figure 5.1: Sketch of the SSC scattering geometry to illustrate the definition of angles. The coordinate system is chosen such that the magnetic field B lies in the $(x-z)$ plane and the scattered photon direction $\mathbf{v}_{k'}$ is along the z axis.

polarization rotation epochs. Each zone has an initial electron population set by power law index α and exponential cutoff set by γ_{\max} . As a given slice moves down the jet, at each step dx the polarized synchrotron emission from each zone is calculated using expressions from Rybicki and Lightman [93]. B , R and the electron populations are evolved at each step. Applying relativistic PA rotation (RPAR) to the emission of each zone and summing the Stokes parameters gives the final SED and polarization.

To focus on overall geometrical trends our base calculations assume that all zones are identical except for field orientation. This allows us to ignore electron migration between zones in a given slice; we also ignore small losses associated with complete escape from the jet. Variation in zone efficiency should introduce additional variability, diluting but preserving the geometric trends described here. We do discuss (§3.5) cases when a subset of zones dominate the synchrotron emission, since the ICS emission can be sensitive to their disposition across the jet.

As we are most interested in X-ray and lower energies, we treat SSC in the Thomson limit, which Zhang and Böttcher [105] show is valid up to at least 500MeV for a relativistic jet with Doppler factor $D \gtrsim 10$. The incoming and outgoing photons have momentum unit vectors \mathbf{v}_k , $\mathbf{v}_{k'}$ and frequencies ϵ , ϵ' respectively. BCS show that photons with original polarization direction \mathbf{e} (perpendicular to the magnetic field for synchrotron seed photons) scatter to energy ϵ' with powers $P_{||}^{SSC}$ and P_{\perp}^{SSC} for polarization along or perpendicular to the projection of the B-field onto the plane orthogonal to $\mathbf{v}_{k'}$:

$$P_{||}^{SSC}(\epsilon') = C\epsilon' \int \frac{d\epsilon}{\epsilon} d\Omega_k E_{\min} n(\epsilon) q(\theta) \cdot (Z_{||}(\Sigma_1 + \Sigma_2) + \Sigma_2) \quad (5.1)$$

$$P_{\perp}^{SSC}(\epsilon') = C\epsilon' \int \frac{d\epsilon}{\epsilon} d\Omega_k E_{\min} n(\epsilon) q(\theta) \cdot (Z_{\perp}(\Sigma_1 + \Sigma_2) + \Sigma_2). \quad (5.2)$$

Here $C = \pi(\frac{e^2}{4\pi})^2 \frac{c}{m_e c^2}$ in c.g.s units,

$$E_{\min} = \sqrt{\frac{\epsilon'}{2\epsilon(1 - \cos\theta_k)}} \quad (5.3)$$

is the minimum electron energy required for scattering of a photon from ϵ to ϵ' ,

$$Z_{\mathbf{e}'} = \left(\mathbf{e} \cdot \mathbf{e}' + \frac{(\mathbf{v}_k \cdot \mathbf{e}')(\mathbf{v}_{k'} \cdot \mathbf{e})}{1 - \cos\theta_k} \right)^2 \quad (5.4)$$

where $Z_{||}$ and Z_{\perp} are found by selecting $\mathbf{e}'_{||}$ or \mathbf{e}'_{\perp} respectively (fig. 5.1), and the solid angle of the photon direction before scattering is

$$d\Omega_k = d\cos\theta_k d\phi_k. \quad (5.5)$$

Σ_1 and Σ_2 are integrals over the electron population doing the scattering, with maximum electron energy E_2 and minimum E_1 :

$$\Sigma_1 = \int_{\beta_2}^{\beta_1} dE \frac{N_e(E)}{E^4} \left(\frac{E_{\min}^2}{E^2} - \frac{E^2}{E_{\min}^2} + 2 \right) \quad (5.6)$$

$$\Sigma_2 = \int_{\beta_2}^{\beta_1} dE \frac{N_e(E)}{E_{\min} E^6} (E^2 - E_{\min}^2)^2 \quad (5.7)$$

where E is the electron energy and

$$\beta_1 = \begin{cases} E_{\min} & E_{\min} > E_2 \\ E_2 & E_{\min} < E_2 \end{cases} \quad (5.8)$$

$$\beta_2 = \begin{cases} E_{\min} & E_{\min} > E_1 \\ E_1 & E_{\min} < E_1 \end{cases} \quad (5.9)$$

$n(\epsilon)$ and $q(\theta)$ denote the synchrotron seed photon spectrum, split into an energy and angle dependent part, where the angle θ is given by

$$\cos\theta = \cos\Theta\cos\theta_k + \sin\Theta\sin\theta_k\cos\phi_k \quad (5.10)$$

from fig. 5.1. We take $q(\theta) \propto \sin^{\frac{p+1}{2}}\theta$, given an isotropic distribution of electron pitch angles. $n(\epsilon)$ is calculated self-consistently from the multizone model at each step of the jet.

The framework described above assumes a 100% polarized seed photon population. BCS treated only power law electron populations with synchrotron polarization independent of ϵ (i.e. $\Pi(\epsilon) = \text{const.}$). For partly polarized seed photons Bonometto, Cazzola, and Saggion [100] ignored energy dependence, simply re-scaling the final SSC polarization fraction $(P_{\perp}^{\text{SSC}} - P_{\parallel}^{\text{SSC}})/(P_{\perp}^{\text{SSC}} + P_{\parallel}^{\text{SSC}})$. In our case the electron population cools, so that E_c and the photon spectrum evolve, meaning that we cannot assume constant $\Pi(\epsilon)$. Thus we split $n(\epsilon)$ into $n_{\perp}(\epsilon)$ and $n_{\parallel}(\epsilon)$ (synchrotron photon populations with polarization parallel and perpendicular to the projection of the B-field in the plane orthogonal to \mathbf{v}_k) where $\Pi(\epsilon) = (n_{\perp}(\epsilon) - n_{\parallel}(\epsilon))/(n_{\perp}(\epsilon) + n_{\parallel}(\epsilon))$. Evaluating these separately using (1) and (2), we sum their Stokes' parameters to get the SSC polarization for arbitrary $\Pi(\epsilon)$.

With jets having significant bulk Γ , we expect blazar emission to be affected by RPAR ([1, 86]). This relativistic aberration strongly changes our effective line of sight, thus rotating the PA we observe as a function of Γ for fixed θ_{obs} . We have shown RPAR to be relevant in both stochastic and rotation phases in blazar synchrotron polarization [1]; we expect it to be even more important here since SSC polarization is strongly dependent on the component of the B-field to our line of sight [101]; note the Θ dependence in Eqn. 10 and the \mathbf{e} vector. We include the effects of RPAR in our model by rotating the jet frame \mathbf{B}' to the effective magnetic field observed in the lab frame for each zone when calculating the SSC emission and final Stokes' parameters.

$$\mathbf{B}_{\text{eff}} = \mathcal{R}(\Theta_{\text{rot}}) \cdot \mathbf{B}' \quad (5.11)$$

and

$$\Theta_{\text{rot}} = \arccos \left(\frac{\cos\theta_{vl} - \beta}{1 - \beta\cos\theta_{vl}} \right) - \theta_{vl}, \quad (5.12)$$

where θ_{vl} is the angle between the zone's velocity vector $\beta c\hat{\mathbf{v}}$ and our line of sight in the lab frame. The rotation takes place along the plane containing $\beta c\hat{\mathbf{v}}$ and the line of sight in the lab frame. This is a simpler more intuitive form of the RPAR equations given in [1, 86].

This prescription gives us the observed synchrotron and SSC polarized emission from a single B-field zone, assuming that all Compton upscatter is only from local synchrotron seed emission ('on the spot' approximation). This is, of course, the approximation used in single zone models. Instead we

expect that the true seed photon field will be strongly dependent on the inhomogeneous surrounding zones §3. Further, since the jet electron populations evolve, the seed photons seen at a given zone are also dependent on light travel effects, which we discuss in §4. In particular, since the low energy seeds dominating the upscattering tend to be dominated by the cooled population, this can be especially important for low energy (e.g. X-ray) SSC emission. Nevertheless for some initial insight, we start by computing emission from a multi-zone jet, with SSC independently computed for each zone, as above.

We proceed by computing the SSC Stokes components for each B-field zone individually, evolving the electron population and jet parameters by calculating the total electron energy losses at each dx step, then summing the Stokes' flux across the full length of the jet. In this evolving, but isolated, zone example the energy density depends only on R and the instantaneous emitted synchrotron power, and so is the same at all points in a given jet cross-section. Figure 5.2 shows a simulation slice with isolated zones and typical blazar parameters for synchrotron + SSC. Note the sharp rise in Π_{Sync} and Π_{SSC} (and EVPA shift) at the upper end of each component. This is more fully explored in §3.

Before we extend to the interacting zone model, we mention some general results already visible in these sums. First, as noted by Bonometto, Cazzola, and Saggion [100], unlike Thomson scattering, Compton scattering does not create polarization. Thus inevitably $\Pi_{SSC} < \Pi_{Sync}$ (although for multi-zone seed photon mixing, this is not always true, §5). Next, a $\Pi_{Sync} = 1$ beam scattering off an e^- powerlaw of index α will produce $\Pi_{SSC} = \Sigma_1 + \Sigma_2 / (\Sigma_1 + 3\Sigma_2)$, with the EVPA reflected in the \mathbf{v}_k and $\mathbf{v}_{k'}$ plane. For typical values of $\alpha \sim 1 - 3$, $\Pi_{SSC} \sim 0.5 - 0.75$. The modest $\Pi_{Sync} \approx 0.03 - 0.1$ of real jets indicates many N_{eff} emission zones with uncorrelated B field orientations. While Π_{SSC} will depend on the particular B orientations of a given realization, as N_{eff} increases, the result tends to an isotropic average. For a single isolated zone averaged over many isotropic B-field realizations we find $\frac{\Pi_{SSC}}{\Pi_{Sync}} \approx 0.35$, (in good agreement with the Bonometto and Saggion [101] result for $\alpha \sim 2$). Note that Π_{Sync} is the polarization of the typical seed photons (e.g. 0.01 – 0.1eV for for X-ray SSC), discussed in §5.

5.3 Multi-zone Effects

In a conical θ_{op} , multizone, fixed- Γ jet model, each B-field zone i is observed at a different θ_{obs_i} and thus has Doppler factor $D_i = \frac{1}{\Gamma(1-\beta\cos\theta_{obs_i})}$. The observed (lab) power of each zone is then $F_{lab}(\nu) = D_i^3 \cdot F_{jet}(\nu)$ where two powers of D_i come from relativistic aberration and the other from time dilation. Additionally, the frequencies in the lab frame are blue-shifted, $\nu_{lab} = D_i \cdot \nu_{jet}$. For a power law photon spectrum $F(\nu) \propto \nu^{-\beta}$, this provides an additional factor of $D_i^{-\beta}$ to the energy spectrum. For the Doppler boosting computations, we assume that the jet is structured as a set of discrete blobs, as observed at VLBI scale. Thus non-cylindrical geometry alone, with Doppler

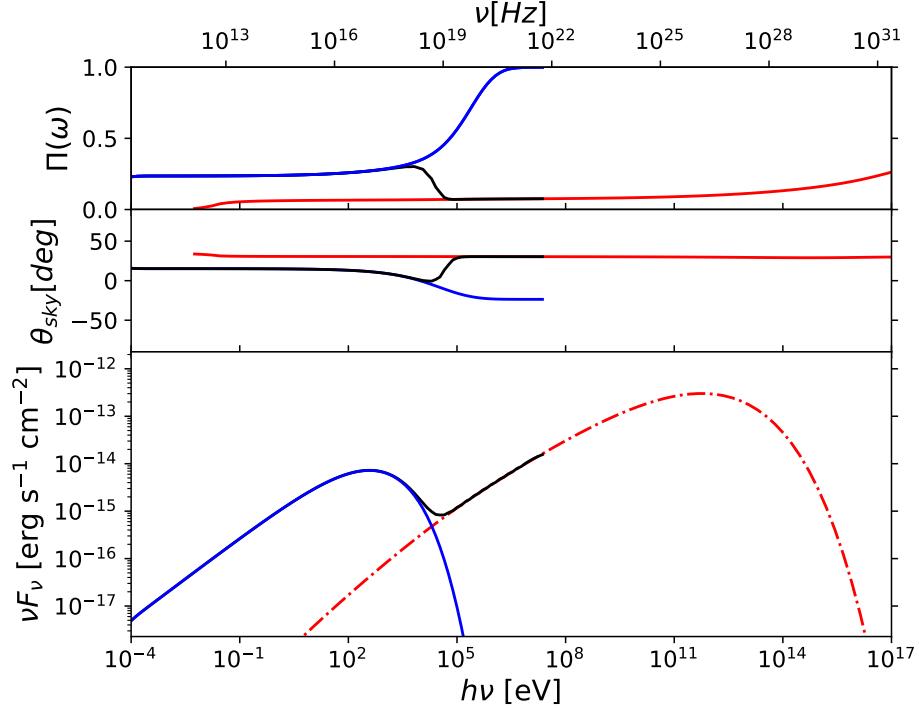


Figure 5.2: The polarization fraction, observed PA angle and SED for a jet with 19 independent B-field zones jet for a single dx step (i.e. no cooling). The SSC emission is calculated using only local synchrotron seed photons (i.e. evaluating Eq. (17) instantaneously). Published models typically use such a single zone. Red denotes SSC, blue synchrotron and black their combination. The jet parameters are tabulated in the Appendix.

boosting, guarantees that identical jet zones contribute differently to the the observed synchrotron peak (fig. 5.3) and to the seed photon population seen by other zones.

Note that with increasing observed photon energy one samples further into the exponential tail of the individual zone spectra. Thus a decreasing number of zones contribute until a single zone dominates. The result is an increasing Π_{Sync} and a gradual EVPA evolution, converging on that of the most boosted zone (in the few-zone regime this behavior may not be monotonic). The top panel of fig. 5.2 showcases this effect, while fig. 5.3 shows an example of individual zone spectra. A closely related effect occurs when the zones themselves have different γ_{max} . For example Marscher [17] assumes in a jet-shock model that γ_{max} depends on B field-shock inclination angle, for an injected electron spectrum of index α , giving

$$N(E_e) \propto E_e^{-\alpha} e^{\frac{-E_e}{\gamma_{\text{max}} m_e c^2}} \quad (5.13)$$

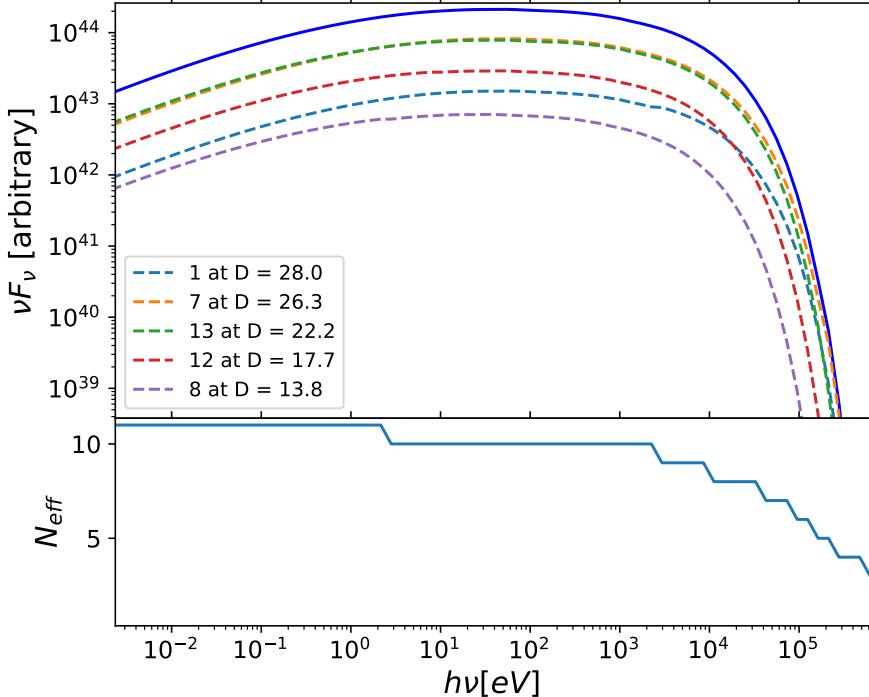


Figure 5.3: Plot of an example full jet synchrotron SED (blue solid) made up of emission from 19 B-field zones. The zones are grouped into 5 different Doppler factor sets. Dashed lines show the SEDs from zone sets with differing θ_{obs} . We can see that the total emission (blue solid) has contributions from fewer total zones at high energy. Once the SED is exponentially dominated, the zone with the highest γ_{max} dominates the observed flux.

with

$$\gamma_{\text{max}} \propto \left(\frac{B_{||}}{B} \right)^2. \quad (5.14)$$

and $B_{||}$ the zone's B-field component parallel to the shock normal. Thus this version is sensitive to the shock geometry. Alternatively we might imagine that shock turbulence gives rise to the same γ_{max} distribution as (14) but with orientation independent of the shock geometry, as for our fully random B distribution. In any case, the geometric Doppler effect combines with the intrinsic γ_{max} effect to disproportionately weight a subset of the zones. Figure 5.4 shows the zone spectra when both effects are present. Figure 5.5 gives the effect on net polarization, and its dependence on jet parameters. Here we define N_{eff} as the number of zones contributing half of the integrated flux. Clearly when γ_{max} scales with the shock $B_{||}$ one finds the highest polarization fraction, since the dominating zones have B-fields nearly aligned (although we note that when the shocks are transverse, this large Π is strongly dependent on RPAR effects). Interestingly for large θ_{op} , Doppler boosting alone can produce close to the same Π rise as a γ_{max} spread.

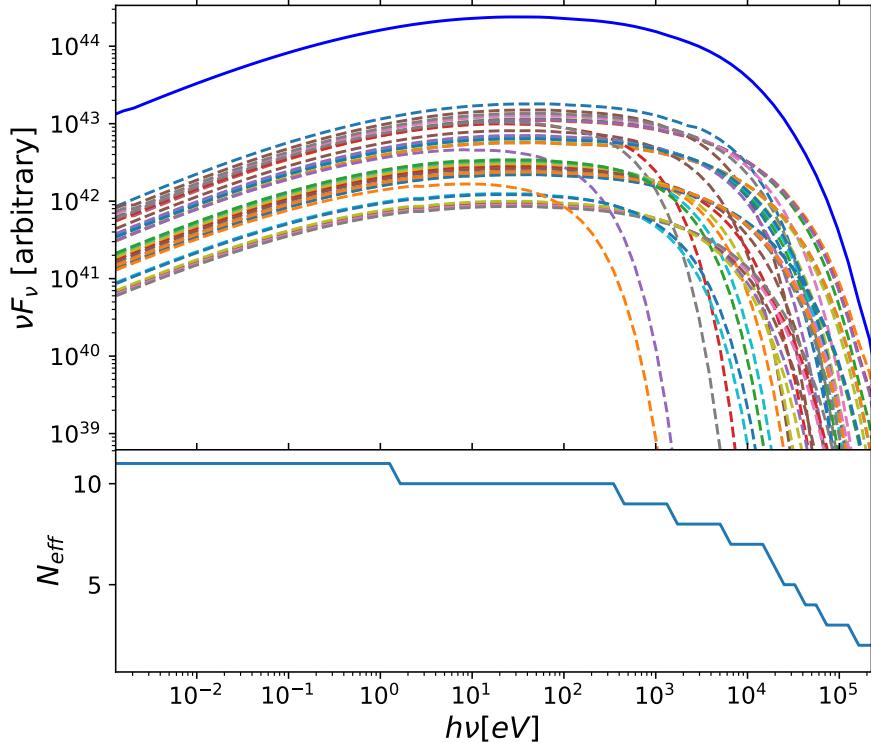


Figure 5.4: Effect of combined Doppler boosting and γ_{\max} spread at the synchrotron SED. Each dashed line now represents an individual zone. Again, N_{eff} drops at high energy.

Even without any zone differences, a small increase in Π is expected at frequencies emitted by electrons above γ_{\max} due to the deviation from a perfect power law. Thus Π tends to 1, not ~ 0.75 , when a single zone dominates on the exponential tail.

5.4 Light Travel Effects

Since our multizone jets have evolving (cooling) electron populations and since B may also be a function of time or distance along the jet, we need to consider how the finite light travel time (and the slower jet speed) affect the emission observed at any one moment. This is especially important when considering propagation between jet zones in building up the SSC seed photons. First, we should recall that our multizone model has a spatial coherence scale across the jet, the distance over which the magnetic field (and possibly γ_{\max}) decorrelate. With N zones this is $\approx 2R_0/\sqrt{N}$ for the stochastic magnetic field. Assuming isotropic jet turbulence, this should also be dx_{ch} , the decorrelation length *along* the jet in the jet frame – this gives a decorrelation timescale $\sim dx_{\text{ch}}/c$. For

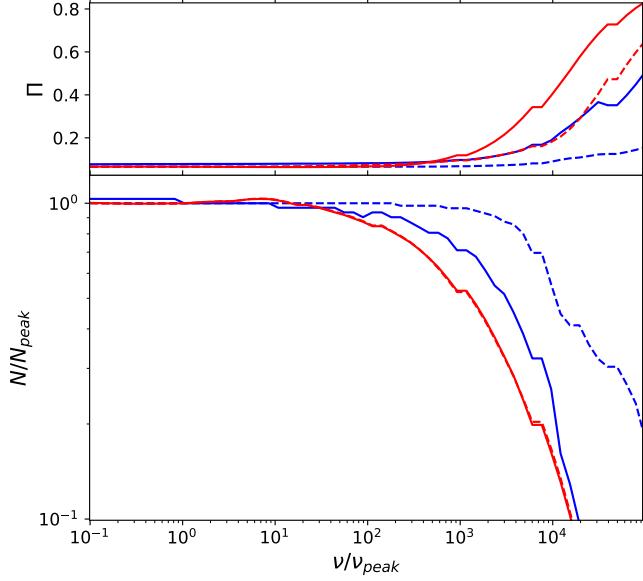


Figure 5.5: *Bottom Panel:* The number of zones contributing half of the observed flux (normalized to the number contributing at the SED peak) for a typical full jet blazar synchrotron SED. The blue lines represent the dependence for the Doppler effect of fig. 5.3. The solid line shows a jet with large θ_{op} viewed near its edge (large Doppler spread) while the dashed line shows a small θ_{op} viewed near the axis (small Doppler spread). Red curves give the case where γ_{max} varies between zones. The solid red curve is the scenario proposed in Marscher [17] where γ_{max} depends on $B_{||}$, the shock-aligned field (here we assume a transverse shock). The dashed red curve represents the same γ_{max} distribution as the solid line case, but with random B orientation. *Top Panel:* The corresponding polarization fractions; colors and lines as below.

a range of blazar parameters we find $dx_{\text{ch}} \sim 10^{15} - 10^{16}$ cm giving polarization variability timescales of $\sim 0.5 - 5$ d, in agreement with the stochastic optical variability measured by RoboPol [5].

For rotation-dominated epochs with the helical B fields we can instead associate dx_{ch} with the timescale of a typical observed optical rotation, roughly 360° per month [82], so we take $dx_{\text{ch}} \approx 1\text{d}c/\beta$. For many typical blazar parameters these stochastic and helical characteristic length scales are of similar size. We consider how the observed spectrum and polarization are sensitive to this coherence scale.

5.4.1 Finite Bulk Lorentz factor

For infinite Γ , the jet particles and their emitted radiation would be co-spatial for their entire radiation history (and we would detect this flux only along the jet axis). However with finite Γ the photons outrun the jet particles. If at some energy the dominant radiation is produced sufficiently far downstream it will lag behind the radiation produced closer to the jet base by that same zone.

Thus at a distance $x_p = \beta dx_{\text{ch}}/(1 - \beta)$ the radiation from our designated zone will not yet have reached the Earth observer; we will instead measure the flux of the *preceding* zone (along the same jet flow line). This preceding zone will in general have different B and particle population properties. Further downstream additional zones can also contribute.

However in practice for $\Gamma > 5$ and $N \leq 150$ we find that the bulk of the observed emission at all frequencies of interest has been radiated before x_p (as can be confirmed with the bottom panel of fig. 5.7). Thus we can infer that the radiation from a single slice is co-eval, except for the most extreme jet parameters.

In this picture the field orientation at the jet base is frozen in and $dx_{\text{ch}} \sim \text{const}$, so that zones expand only transversely. If in contrast the zones stay quasi-spherical (e.g. due to turbulent cascading along the jet), a longer variability timescale and a decorrelation in polarization compared to higher frequencies can result for late jet emission (radio).

5.4.2 Non-zero Viewing Angle of a Conical Jet

For a diverging jet viewed off axis, the increasing width of an observer time slice includes an increasing range of jet distances (i.e. larger range of emission times for the jet particles). This is shown in fig. 5.6. In our zonal picture, this means that once $2R\tan\theta_{\text{obs}}/\sqrt{N} > dx_{\text{ch}}$ zones from more than one slice contribute to the emission. For a given N zones in a jet, expansion stretches the zone horizontally, but not radially. Thus with a tilt, the increased radial range incorporates more B -field zones at a given observer time slice, and the polarization decreases. In practice, this is dependent on the jet opening angle θ_{op} through both the expansion rate of R and the individual zone Doppler factors that control dx_{ch} .

For typical blazar parameters $\theta_{\text{obs}} \sim 0.5^\circ - 7^\circ$, $\theta_{\text{op}} \sim 5^\circ - 45^\circ$, $N \sim 1 - 100$ (in a stochastic phase) and $dx_{\text{ch}} \approx 2R_0/\sqrt{N}$. Then in the worst case, $2R\tan\theta_{\text{obs}}/\sqrt{N} > dx_{\text{ch}}$ when $R \approx 2.3R_0$. Beyond this N_{eff} increases while polarization fraction and variability decrease. These radii contribute most to the radio-microwave range of the synchrotron peak. Through SSC this also de-polarizes the very low end of the Compton component. For rotating B-field zones the effect is similar, but slice mixing not only lowers Π but smooths over rotational phase in the helix structure. However the large viewing angle conditions required for significant slice mixing will make the rotation less prominent, in any case.

5.4.3 Finite Travel Time at Large Jet Radius: Seed Photon Build-up

The two geometrical effects above increase N_{eff} slightly at large jet distance x_p (affecting low energy synchrotron and SSC). But finite light travel time affects the SSC of all zones since the seed photon population in a given zone is made up of contributions from all other zones. The furthest zones are a significant light travel time away and this means that their emission represents a smaller x and earlier time in the slice evolution down the jet.

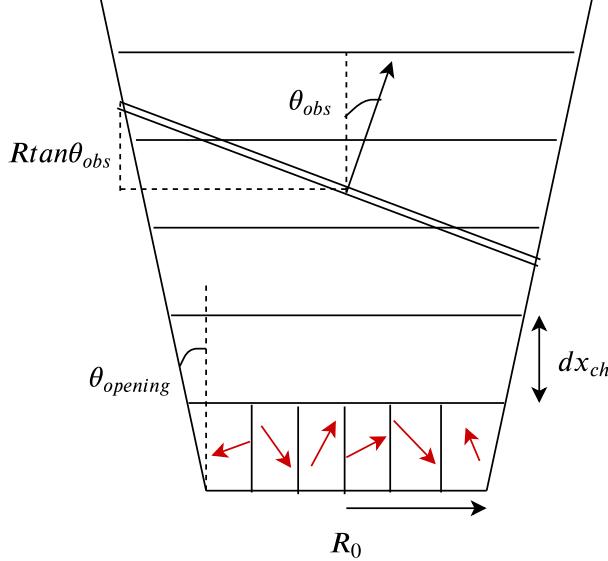


Figure 5.6: Simplified schematic displaying the light travel effect described in §4.2. The jet is split into horizontal slices of width dx_{ch} , the scale in the jet frame over which the magnetic field vectors decorrelate. These B -field vectors are denoted by the red arrows. Note that for jets viewed at large off-axis angles multiple slices are viewed at a given time.

The total synchrotron energy in the co-moving jet frame at any point i on the jet cross section is given by:

$$E^i(t) = \int \frac{P_{Sync}^i(r, \phi, t)}{\pi R(t)^2} r dr d\phi dz dt \quad (5.15)$$

where $P_{Sync}^i(x, \phi)$ is the total synchrotron power per unit length in the jet cross section. The functional form varies depending on the cross section point i . To a good approximation, the emitted synchrotron radiation from the zones is co-moving with its jet slice (§4.1). This reduces the radiative transfer to a 2D sum, allowing us to set $dt = dx/c$ and $dr = dx$ where x is the distance the jet cross section has travelled along the length of the jet. So the energy density is:

$$\rho_E^i(x) = \frac{dE^i(x)}{dV} = \int_0^x \int_0^{2\pi} \frac{P_{Sync}(x, \phi)}{2\pi^2 c R(x)^2} dx d\phi \quad (5.16)$$

For a cylindrical jet with no cooling and $x_{max} \gg R$, choosing i to be the point in the center of the jet (16) reduces to:

$$\rho_E(x) = \frac{P_{Sync}}{\pi c R^2} \int_0^R x dx = \frac{P_{Sync}}{\pi c R}, \quad (5.17)$$

which is a familiar expression for the energy density at the center of a 2D emitting disk.

To treat the polarization, we sum up the energy density coming from all zones in a given slice. This is done by simply evaluating the distance of every $P_{Sync}(x, \phi)dx$ annulus from each of the zones

in the jet cross section. Using the mutual displacement vectors between zones and their individual $\mathbf{B}_i(x)$ we can construct the total seed photon polarization and energy density at every point in the jet. For this one must compute the correct solid angle subtended by the scattering zone and the effects of RPAR rotation on the polarization vectors. Given the finite numerical nature of our simulation, we expect it to be asymptotically more accurate for a higher number of zones.

The top panel of fig. 5.7 show this resulting seed photon population, computed using Equation (16), showing the energy density as a function of distance x along the jet. The line types show the difference between edge and central zones for a conical jet. We also show how beyond a critical distance other zones in the slice dominate over self-emission in the seed photon density; this occurs later at the jet edge. Its effect can be seen on the SSC EVPA (see §5).

SSC photons of a given energy are, of course produced by a range of seed photons, so care must be taken in comparing the observed polarizations. Figure 5.8, shows the effective seed photon SED for X-ray (keV) and soft γ -ray (MeV) Compton emission. X-ray polarization measurements by upcoming missions are thus best compared with synchrotron observations in the mm-optical band.

5.5 SSC with All Effects

As a concrete example, we compute with B_i randomly drawn for all zones at the jet base. Orientations are frozen thereafter, e.g. fixed as the zones propagate down the jet, but magnitude can vary. As in Potter and Cotter [106] for a ballistic conical jet we assume that the only energy loss mechanism is radiation, so magnetic energy U_B is conserved. Thus magnetic flux density decreases as the jet expands. The magnetic/particle energy ratio increases slightly along the jet but remains ~ 1 . We assume here a single fixed γ_{\max} in all zones. Thus geometrical (Doppler boosting) effects dominate the prominence of individual zones. Indeed, geometric parameters (Γ , θ_{open} , θ_{obs}) have the largest effect on polarization. Other parameters (W_j , α , E_{\max} , B_0) primarily affect the shape of the SED. As expected, polarization thus serves as an excellent (and largely independent) probe of jet geometry, although we do note when other (spectral) parameters have a large effect.

In contrast to the treatment of BCS, who assumed a simple power law electron spectrum and uniform (energy independent) synchrotron polarization, we need to consider how all seed photon energies contribute to the observed Compton radiation at a given energy in computing the ratio $\Pi_{\text{SSC}}/\Pi_{\text{Sync}}$. Since our electron population evolves along the jet (§3, §4.2), different seeds dominate at different locations along the jet. Nevertheless, we can give a qualitative picture of the seed spectrum. To connect X-ray SSC with observed synchrotron fluxes, we focus on the synchrotron seeds in the optical and mm range (see fig. 5.8). These are computed in simulations using all effects described above. The simulations employ various zone multiplicities (1, 7, 19, 37) to illustrate the effect the increasing the zone averaging on both the synchrotron and SSC polarization amplitudes. The principal effect is, of course a diminution $\Pi_{\text{Sync}} \propto N_{\text{eff}}^{-1/2}$. We compute 200+ realizations of

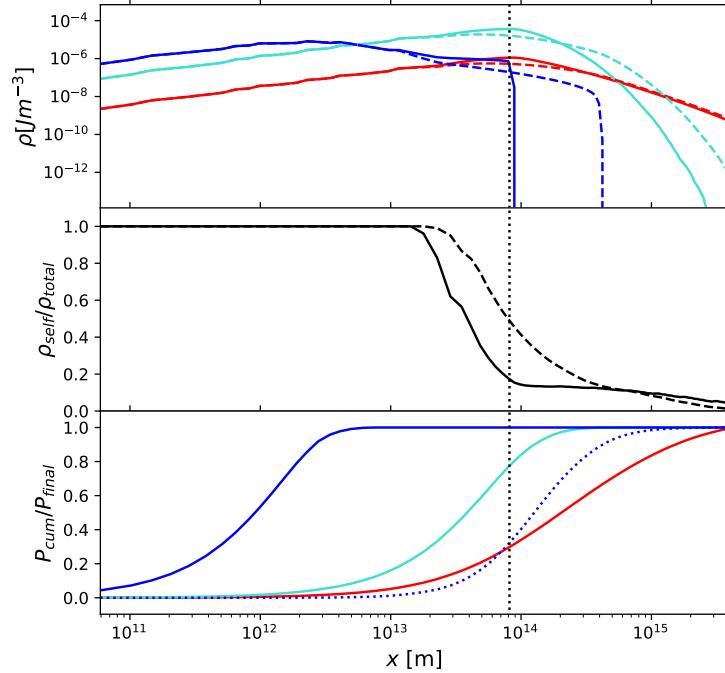


Figure 5.7: *Top panel:* Synchrotron energy density as a function of distance along the jet x for three different energy bands (red: mm wavelength, cyan: optical, blue: 10 keV). Solid lines are for the jet center, while the dashed lines are for the edge. The black vertical dotted line represent $x = R_0$. *Middle panel:* Fraction of the integrated synchrotron density due to the local zone (solid line central zone, dashed line edge zone). Neighboring zones dominate for $x > R_0$. *Bottom panel:* Build up of the on-axis synchrotron seed photon population at three energies. The dotted blue line represents the corresponding 1-10 keV SSC flux. These are fluxes measured in the jet frame; jet parameters are as in §5.

each configuration to average down these fluctuations and display Π_{Sync} , Π_{SSC} trends.

5.5.1 Effective number of emission zones

In §3 we discussed how the number of effective emission zones, N_{eff} , affects the net polarization, with a large increase at synchrotron cutoff energies. N_{eff} effects can help explain both synchrotron and SSC polarization behavior across the whole SED. Synchrotron polarization is controlled solely by N_{eff} , since every zone emits its synchrotron independently. Averaging over a large number of isotropic B-field iterations, we expect $\Pi_{\text{Sync}} \propto N_{\text{eff}}^{-1/2}$. However, the relationship between Π_{SSC} and N_{eff} is not *a priori* obvious since each single zone scatters synchrotron seed photons from all the other zones in the jet, weighted by their proximity and power.

As in §3 we define here N_{eff} as the minimum number of zones that contribute half of the flux. This is an imperfect estimate since N_{eff} depends on the weighted contribution of all zones; Π_{Sync} itself

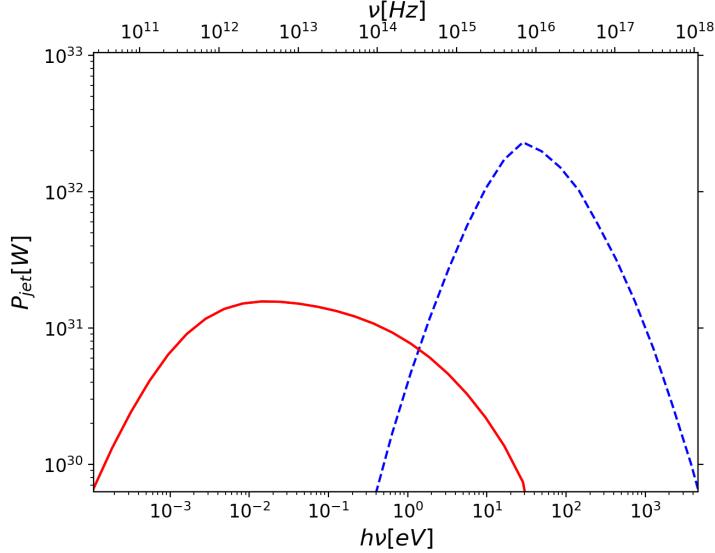


Figure 5.8: SSC power in the jet frame as a function of the (lab frame) seed photon energy – i.e. the seed photon spectrum for a given Compton energy. The solid line is for keV SSC emission while the dashed line shows the seed spectrum producing MeV emission. For the overall SED and assumed jet parameters, see §5.1.

provides the best metric for N_{eff} . In any event, the underlying behavior is adequately approximated with $N_{\text{eff}} \propto \Pi_{\text{Sync}}^{-2}$.

When the observed brightness of N total zones is controlled purely by viewing angle-determined Doppler weighting, jet geometry determines N_{eff}/N . In practice measuring the synchrotron polarization fraction well below the cutoff (and above radio energies affected by light travel time) provides a direct measure of N_{eff} , and thus constrains jet geometry. Figure 5.9 shows the relationship between N_{eff}/N and $\theta_{\text{obs}}/\theta_{\text{op}}$ for several jet geometries (blue). For low (radio) energies (which radiate at large x) and large off-axis angles, finite light travel time effects in a widely diverging jet (§4.2, fig. 5.6) increase N_{eff} . Red lines in fig. 5.9 shows this effect (at $\theta_{\text{op}} = 40^\circ$) for long wavelength synchrotron.

While we expect that Doppler weighting of zone flux is always present, the observed flux from individual zones may vary for other reasons. These might include differences in acceleration efficiency between zones [17], with electron density, γ_{max} and Γ variation - or variation in relative sizes of zones. Such variation may be more prominent during flaring events. To illustrate such effects, we plot a simple model where the lab frame flux from individual zones is allowed to vary randomly (as opposed to being controlled purely by jet geometry). The results are compared with a Doppler weighted jet in fig. 5.10 for total zone numbers $N = 19, 37$.

Figure 5.10 displays the expected $\Pi_{\text{Sync}} \propto N_{\text{eff}}^{-1/2}$ trend. The N, N_{eff} behavior of SSC is more complex. $N = 19$ is more polarized than $N = 37$ for the same N_{eff}/N for both synchrotron and SSC. In the synchrotron case this is simply due to the change in N_{eff} . In the SSC case, N_{eff} zones

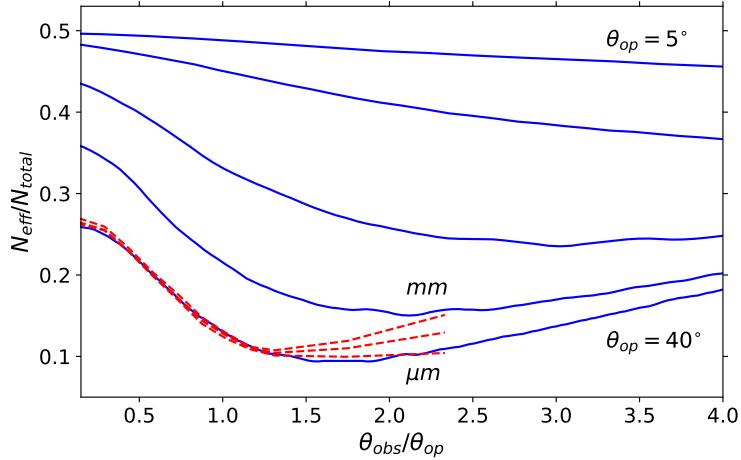


Figure 5.9: The geometrical effects of Doppler boosting on the number of effective emitting zones. Jet frame θ_{op} increases from top to bottom; the fraction of zones within our $1/\Gamma$ effective viewing circle decreases (blue curves). As θ_{obs} increases off-axis, the fraction of the zones in the $1/\Gamma$ viewing cone also decreases, until $\theta_{obs} > 1/\Gamma$ and more (distant) zones are boosted at a similar level, causing N_{eff} to grow again. The exact functional form is controlled by the weighting function, in this case the Doppler factor D^4 , and the jet geometry, in this case a cone. We note that treating N_{eff} as the number of zones that dominates half of the flux is not a perfect proxy, the true N_{eff} is controlled by a weighted distribution. The red lines show the same relationship for a $\theta_{op} = 40^\circ$ jet affected by LT effects (§4.2). At low energies (e.g. mm wavelengths), the emission occurs at sufficiently large radii (see fig. 5.7) that light travel effects substantially increase N_{eff} for jets viewed off-axis.

are boosted and scatter the jet frame synchrotron emission from all N zones. A larger N further averages down the seed photon polarization (for the same N_{eff}), so $\Pi_{\text{SSC}} \propto N^{-1/2}$. This ensures that the dot point SSC curves lie below the triangle points in fig. 5.10. However Π_{SSC} also depends on N_{eff} (apparent in fig. 5.10 from the non-zero slope), since with larger N_{eff} one has more scatterers sampling the angular distribution of the synchrotron radiation of the N total zones. This averaging decreases Π_{SSC} as N_{eff} grows. The scaling depends on how the N_{eff} zones are chosen/weighted. We can characterize the dependence using a simple power law:

$$\Pi_{\text{SSC}} \propto N_{\text{eff}}^{-p} N^{-1/2}. \quad (5.18)$$

The Doppler-boosted cases have similar slopes $p \approx 0.3$ for both $N = 19$ and $N = 37$, with the decreased Π_{Sync} driving down the coefficient of the latter. Randomly selected zones (dotted lines in fig. 5.10) give $p \approx 0.25$. However these are not universal power laws; the averaging over the seed photon's angular polarization distribution depends on the particular weighting scheme.

This can also be seen in the right panel of fig. 5.10. Since $p < 0.5$, we have a slow increase in the SSC/Sync polarization ratio as N_{eff} increases, most obvious for the random zones. For pure Doppler boosting, geometrical effects complicate this trend. As N_{eff} grows (θ_{obs} decreases), the ratio growth

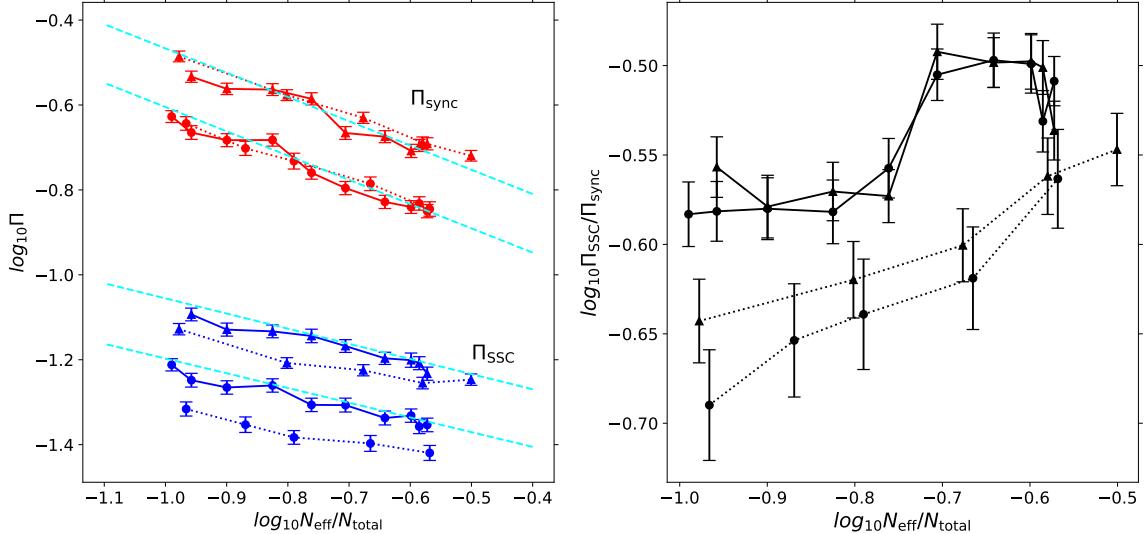


Figure 5.10: *Left Panel:* The optical synchrotron (red) and X-ray SSC (blue) polarization fraction plotted against N_{eff}/N . Triangle points are for an $N = 19$ zone jet, circles for $N = 37$. Solid lines show a Doppler-dominated jet. For the dotted lines zone intensities are randomly distributed. The synchrotron cyan dashed lines show $N_{\text{eff}}^{-1/2}$, while the SSC equivalents show the best fit power laws ($N=37$ gives $p = 0.29$; $N=19$, $p = 0.31$). *Right Panel:* The SSC/synchrotron polarization ratio versus N_{eff}/N (line types as in left panel). Note that the SSC angle-averaging is more effective for a randomly distributed jet, leading to a lower ratio. Values are averaged over many simulations with isotropic B-fields; error bars show the residual errors on the mean. The jet parameters used in the simulations are given in the Appendix Table.

is slow until one starts viewing near the jet edge (or near $1/\Gamma$). Here the increasing N_{eff} zones are together at the near edge of the jet; these all receive synchrotron emission from the N emitting zones at similar angle. This coherence gives rise to poor averaging over the synchrotron beam and, on average, larger SSC polarization, giving an abrupt rise to the ratio. Conversely the largest N_{eff} occur when viewing close to the jet axis. There the most strongly boosted zones dominating the flux are nearly uniformly surrounded by their N synchrotron sources, leading to better angle averaging of the synchrotron field and a drop of SSC polarization and the polarization ratio for jets viewed nearly on axis.

Other complications are also present. For example, mm wavelengths and below are emitted on average downstream from X-ray SSC (fig. 5.7), so $\Pi_{\text{SSC}}/\Pi_{\text{Sync}}$ will be larger (~ 0.4) for such comparison. For some jet parameters (high θ_{obs} , high θ_{op}) energies up to the optical synchrotron and X-ray SSC can also be affected by light-travel time induced N_{eff} increase.

In addition to these statistical trends there is a strong correlation between Π_{SSC} and Π_{Sync} in individual realizations (fig. 5.11). The magnitude of the correlation depends on N_{eff}/N : for high N_{eff}/N (black points) both Π and PA are more highly correlated (small θ_{obs} ; Spearman $r = 0.7$)

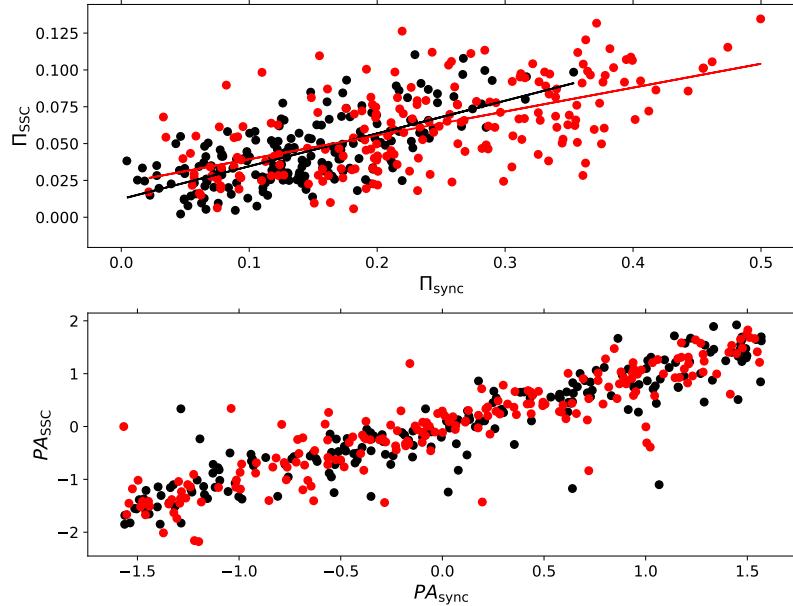


Figure 5.11: Correlation plot of the X-ray SSC polarization fraction (top) and PA (bottom) against the corresponding optical synchrotron for two different viewing angles (N_{eff}) (red: $\theta_{\text{obs}} = 4.5^\circ$, black: $\theta_{\text{obs}} = 0.1^\circ$). Note Π_{SSC} can persist even when chance cancellation makes Π_{Sync} for the Earth line-of-sight; this leads to high $\Pi_{\text{SSC}}/\Pi_{\text{Sync}} \gg 1$ and a non-zero intercept. EVPAs are strongly correlated, but show scatter due to similar differences in the zone-sampling for the Earth line-of-sight.

than the lower N_{eff}/N case (red points, Spearman $r = 0.6$). Of course with the smaller θ_{obs} /larger N_{eff} case (black) we seldom achieve very high Π_{Sync} . But when we do, we have good confidence that the SSC polarization will also be high. Notice that the intercepts are not zero; The different SSC sample can display polarization, even when the observed synchrotron polarization happens to average to near zero. This cautions us to avoid strong conclusions from one large $\Pi_{\text{SSC}}/\Pi_{\text{Sync}}$ measurement in a weakly polarized source, e.g. that hadronic processes are present.

Overall, for synchrotron seed photons emitted early in the jet (optical for the example blazar parameters) we expect to see $\Pi_{\text{SSC}}/\Pi_{\text{Sync}} \sim 0.3$ when $N_{\text{eff}} \sim N/2$. The synchrotron and SSC polarization fractions will be strongly correlated. As N_{eff} decreases relative to N both the ratio and correlation will decrease, controlled in detail by the zone weighting system. For mm (Π -suppressed) seed photons we expect higher ratios $\Pi_{\text{SSC}}/\Pi_{\text{Sync}} \sim 0.4$ but lower correlations. For the typical blazar parameters determined by the inferred opening angles [107, 108], Γ_{bulk} and θ_{obs} are such that $N_{\text{eff}} \sim N/2$, and we expect in most cases $\Pi_{\text{SSC}}/\Pi_{\text{Sync}} \sim 0.3$. However for other viewing geometries the ratio can be lower.

Since boosting is the only feasible way to change N_{eff} without changing N , we expect that non-Doppler weighting effects will preserve $N_{\text{eff}} \sim N/2$ and thus $\Pi_{\text{SSC}}/\Pi_{\text{Sync}} \sim 0.3$. We note that for

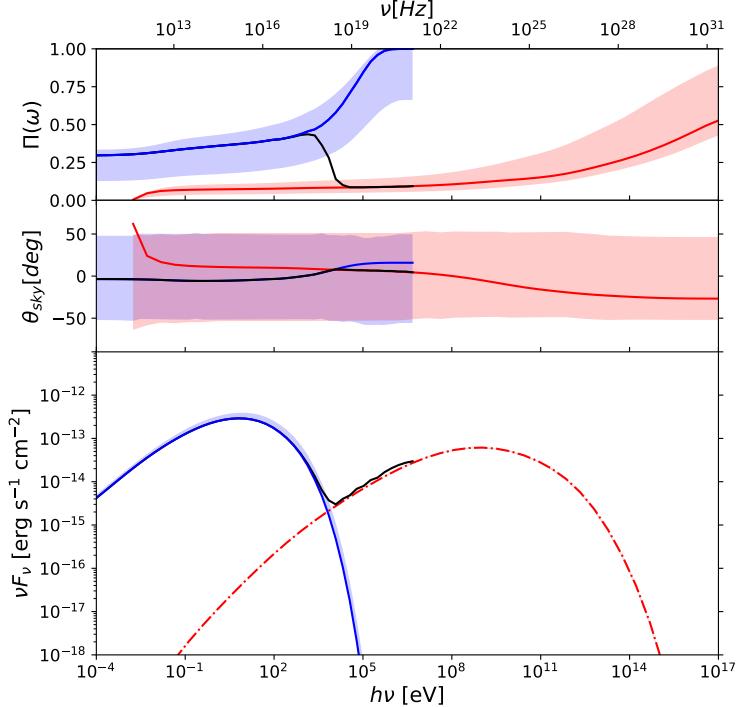


Figure 5.12: Example fully treated jet SED for a single random B-field draw, with $\theta_{\text{obs}} = 6.0^\circ$ and the remaining jet parameters as in fig. 5.10. This simulation uses 19 B-field zones and serves to display some typical polarization behavior. Note the substantial jump in EVPA between the Sync and SSC components. Note also the substantial Π increase above the synchrotron and SSC peaks due to the decreased N_{eff} and the average Π_{Sync} decrease for sub-infrared energies due to §4.2. The shaded bands show the standard deviations in individual realization over many isotropic B-field draw; EVPA average has no preferred direction for either component, with an average value 0.

>MeV energies, SSC is emitted before neighbouring zones dominate the seed photon population (fig. 5.7). In this case the polarization amplitudes will scale as $N_{\text{eff}}^{-1/2}$ and Sync and SSC will be highly correlated, with $\Pi_{\text{SSC}}/\Pi_{\text{Sync}} \sim 0.35$ for the appropriate Π_{Sync} in fig. 5.8. Finally, all results shown here are for electron power laws $\alpha \sim 2$. For similar systems with different power laws we expect our results for $\Pi_{\text{SSC}}/\Pi_{\text{Sync}}$ can be scaled with α as in BCS.

5.6 Conclusion

We have shown that for a multizone relativistic conical jet, the averaging effects that control the final net polarization are sensitive to the jet opening angle and viewing geometry. In general when viewed at larger off axis angle, fewer zones contribute to the observed radiation and the residual polarization is higher. This is countered to some extent for low frequency (e.g. radio) emission, where the electrons cool slowly enough that emission comes from a large range of radii and the finite

Fig.	N	B_0 [G]	L_{jet} [m]	θ_{op} [°]	θ_{obs} [°]	Γ
5.2	19	1.0	1.8×10^{10}	40.0	1.5	14.0
5.10	19, 37	1.0	5.3×10^{15}	40.0	0.1 – 4.5	14.0
5.11	37	1.0	5.3×10^{15}	40.0	0.1, 4.5	14.0
5.12	19	1.0	5.3×10^{15}	40.0	6.0	14.0

Table 5.1: Parameters used for the plotted jet models. All models additionally assume jet power $W_j = 1.3 \times 10^{37} W$, electron spectral index $\alpha = 1.85$, and electron energy range $\gamma_{\min} = 10$ to $\gamma_{\max} = 3.3 \times 10^4$.

light travel time can let zones from different radii across the jet contribute at a given observation epoch – the increase in the number of effective zones decreases Π_{Sync} . Note that these trends are guaranteed by the differential Doppler effect across a conical jet, but will be obscured if electron power, γ_{\max} or Γ fluctuations dominate zone brightness variations.

One particularly interesting effect is the increased dominance of a few zones as one observes at energies well above the synchrotron peak. There the tail of the synchrotron emission is necessarily dominated by a few zones, selected either by Doppler boosting or extreme γ_{\max} , and Π_{Sync} increases. This also means that the EVPA converges to a direction controlled by that dominant zone, which can be quite different to that of the (lower energy) jet average. A similar effect occurs at the upper extreme of the Compton component. Thus we expect a rapid increase in polarization, and a rapid jump in EVPA, until the SSC flux overwhelms the synchrotron component, and one jumps to new SSC values (see fig. 5.12 for an example). This is of particular interest for ‘Intermediate Peak’ blazars (ISP) which can have this synchro-Compton transition in the soft X-ray regime; Liodakis, Peirson, and Romani [5] describe this effect and its importance in selecting targets detectable to *IXPE* and similar X-ray polarization missions. For example, the ISP S50716+714 has an X-ray flux of $10^{-11} - 10^{-10}$ erg/s/cm². Using its measured optical polarization and fig. 5.5 we estimate its X-ray polarization fraction to vary between 12 – 30%. At *IXPE*’s nominal sensitivity of 5.5% MDP₉₉ for 10^{-11} erg/s/cm² in 10 days, we should obtain a 99% significance detection in ~ 100 ks exposure or less. Thus variability should not strongly degrade the single epoch polarization, although longer exposures or multiple visits should see variation in Π and EVPA. However the synchrotron emission is steeply falling in the X-ray band and detailed measurement of the polarization variation across the band may require a higher sensitivity future facility.

For SSC polarization, the seed photons are drawn from a variety of jet zones with different B -field orientations. This decreases the average polarization of the seed population and hence the final Compton polarization. Since in a conical jet different jet sectors have different angles to the Earth line-of-sight and hence different boosting, the averaging is dominated by a sub-set of the jet zones and the final effects are sensitive to RPAR effects. Nevertheless an overall trend of $\Pi_{\text{SSC}} \approx 0.3\Pi_{\text{Sync}}$ (compared to optical photons) can be expected, for both Doppler zone and random zone dominated jets.

Overall, the simulations show the danger of drawing conclusions from any one realization: The scatter in Π is comparable to Π itself, and expected geometrical and spectral trends are only recovered when averaging over many realizations. One should also recall that external seed photons are expected to be largely unpolarized so that any EC flux will dilute the high energy polarization signal. We see that Compton polarization is understandably less powerful as a probe of jet geometry than the synchrotron signal. Nevertheless X-ray SSC polarization can be large enough to be detected in favorable cases, where comparison with the instantaneous Π_{Sync} can give (at least statistically) information on the seed fields and scattering geometry. Finally, large positive correlations between the SSC and seed synchrotron Π make low energy polarization monitoring a useful tool for monitoring fluctuations and aiding in Π_{SSC} detection.

Chapter 6

X-ray Polarization Prospects for Blazars

6.1 Introduction

Polarization can be an important tool for probing the physics of the acceleration zone, especially in characterizing the magnetic field structures that control expected shocks and induce synchrotron radiation. VLBI polarization maps have long been effective at measuring jet fields at pc-scale (e.g., [109]) while more recently optical polarization has provided new information on the field orientation and variability in the unresolved core [82]. X-ray polarization, to be measured by the approved Imaging X-ray Polarimetry Explorer (*IXPE*, [27], launched 2021) and Enhanced X-ray Timing and Polarization mission (*eXTP*, [25], launch \sim 2025), offers new opportunities to probe the jet fields and radiation physics, even closer to the acceleration site. In particular, polarization can help answer whether leptonic or hadronic process dominate in a given band (e.g., [98]).

However, the sensitivities of the near-future missions are modest and long exposures will be required, so in light of the variability and limited low energy polarization information one must choose the expected targets with care. In this chapter, we explore such choice based on the synchro-Compton model developed in chapter 4, 5. We begin with basic predictions for all classes of blazars, section 6.2, followed by in depth modelling for ISP blazars, section 6.3.

6.2 Detecting X-ray polarization in blazar jets

In section 6.2.1 we characterize the X-ray variability of sources observed in optical polarization monitoring programs, in section 6.2.2 we use our model to predict X-ray polarization levels (Π_X),

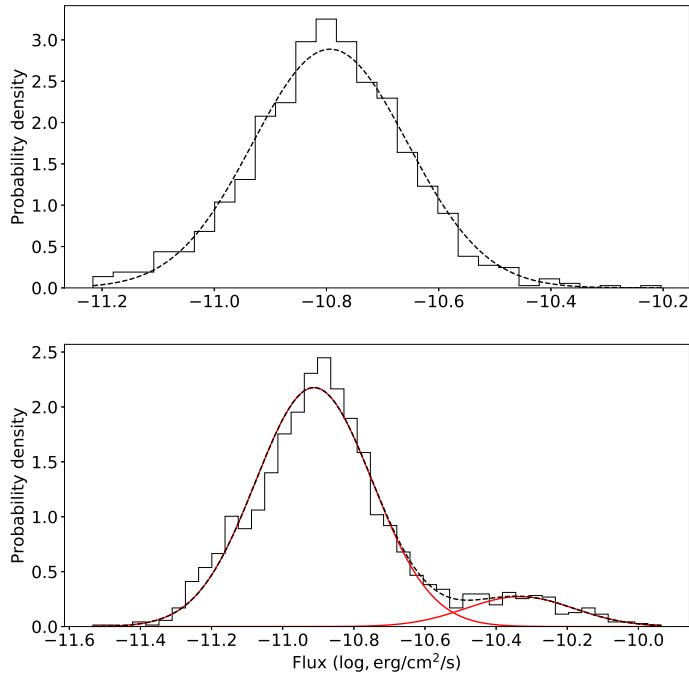


Figure 6.1: Example of a unimodal (top panel) and a bimodal (bottom panel) flux distribution in the 2-10 keV energy range in log-space. The black dashed line shows the best-fit model. In the bimodal case, the red lines show the individual components.

while in section 6.2.3 we combine these factors to quantify the success probability of an *IXPE* measurement for reasonable exposure in an untriggered observation, identifying a list of prime targets, and suggesting other X-ray bright sources that can also be of interest if they exhibit strong optical polarization. We conclude by discussing new measurements that can improve these predictions and monitoring campaigns that could make additional sources, and additional classes of polarization behavior, accessible in the X-ray band.

6.2.1 X-ray variability

Since X-ray polarization measurements are in general sensitivity limited, source flux variability plays a key role in the prospects for a secure, 99% confidence, measurements of a given expected polarization level. We must therefore characterize the variability of likely targets, the great majority of which turn out to be sources detected by the *Fermi* LAT [110]. We use 2-10 keV flux measurements from 2005-2017 measured by the Neil Gehrels *Swift* Observatory's (hereafter *Swift*) LAT source

monitoring program¹ [111] supplemented by 2-10 keV fluxes (from 1995-2012) from the RXTE AGN timing and spectral database² [112]. Strohmayer and Kallman [111] analyze the individual *Swift* observations; we employ their mean spectral parameters tabulated for each source to convert epoch count rates to erg/cm²/s (2-10 keV) using WebPIMMS. 35 sources (19 LSPs, 2 ISPs, 13 HSPs and one unclassified source³) have at least 20 observations so that we can attempt a detailed variability analysis. For the remainder, we characterize their flux variability with a simple mean and standard variation.

Blazar high energy variability has been modeled as a log-normal distribution (e.g., [113, 114]), which may reflect disk-driven fluctuations [115] or variations in the jet particle acceleration [116]. This suffices for some of our sources, but others show wider variability. This may indicate multiple jet states (e.g. quiescent and active flaring episodes), which can be represented by a double normal (Gaussian mixture) model (in log-space) [117]. Of course if we have not sampled the full range of a source's variability the two log-normals might be subsets of a broader single log-normal. Here, using the historical fluxes, we represent our flux distribution functions as either single or double log-normal models without attaching physical significance to the single or double-mode behavior.

The likelihood function for the single Gaussian model is defined as

$$l_{\text{obs}} = \frac{1}{\sqrt{2\pi(\sigma_q^2 + \sigma_{\text{obs}}^2)}} \exp \left[-\frac{(S_q - S_{\text{obs}})^2}{2(\sigma_q^2 + \sigma_{\text{obs}}^2)} \right], \quad (6.1)$$

where S_q and σ_q are mean and standard deviation of the underlying distribution and S_{obs} and σ_{obs} are the observed fluxes and their uncertainties (in log-space). For the Gaussian mixture the likelihood is defined as

$$\begin{aligned} l_{\text{obs}} = & \frac{1-f}{\sqrt{2\pi(\sigma_q^2 + \sigma_{\text{obs}}^2)}} \exp \left[-\frac{(S_q - S_{\text{obs}})^2}{2(\sigma_q^2 + \sigma_{\text{obs}}^2)} \right] \\ & + \frac{f}{\sqrt{2\pi(\sigma_a^2 + \sigma_{\text{obs}}^2)}} \exp \left[-\frac{(S_a - S_{\text{obs}})^2}{2(\sigma_a^2 + \sigma_{\text{obs}}^2)} \right]. \end{aligned} \quad (6.2)$$

where we add mean and standard deviation S_a and σ_a for a brighter ‘active’ state which is realized a fraction f of the observed samples. With such a model, we can draw an arbitrary number of samples from the modeled distribution. To choose between models for a given source we use the Bayesian Information criterion (BIC). Figure 6.1 shows examples of best-fit models for PKS 0558-504 (top panel) and BL Lacertae (bottom panel). There are 15 sources best described as unimodal, 20 sources prefer a bimodal distribution. LSPs show no preference while HSPs slightly more commonly match a bimodal distribution (8 versus 5). The parameters of the best-fit distributions for all the sources

¹<https://www.swift.psu.edu/monitoring/>

²We have included all sources in the RXTE database classified as either BL Lac object (BL Lac) or Flat Spectrum Radio Quasar (FSRQ) with at least 20 observations. <https://cass.ucsd.edu/~rxteagn/>

³SED classification from the 3rd *Fermi* AGN catalog [110], <https://www.ssdc.asi.it/fermi3lac/>

are given in Table 6.1. For sources with < 20 observations, we simply record the mean and variance of the log of the flux, which can be used to form a log-normal distribution.

Measurements are easiest for high polarization Π_X sources in bright large S_{obs} states. Since both quantities are highly variable, we should test if they correlate, as might be expected from e.g. shock-driven flares (e.g., [118]). Of course we lack Π_X , so we use the optical polarization (Π_O) from the Robopol and Steward observatory monitoring programs which have significant temporal overlap with the *Swift* data for 17 sources. We use optical polarization as a tracer of the energetic electrons at the jet base that may also contribute X-ray synchrotron; radio polarization can be dominated by downstream emission. While we cannot make meaningful statements about individual sources, we can check the major source classes by stacking all contemporaneous observations from, say, the HSP. We find that both the LSP and HSP have a mild positive correlation (Spearman's $\rho \sim 0.12$, significance $p - \text{value} < 0.05$). The two ISP showed no correlation. Thus for LSPs and HSPs we draw a flux at a given level in their cumulative distribution function (CDF, e.g. a flux in the top X%) and then draw from that source's polarization CDF at the same top X% level. For the ISP we will assume random uncorrelated draws (see §6.2.3). In practice, we find that this makes a small $< 20\%$ difference to the source detectability, so this assumption is not critical. However it should be tested with future monitoring campaigns.

6.2.2 Expected X-ray polarization

We must use the lower energy (optical) polarization degree Π_O to predict the polarization in the X-ray band. For the HSP and some ISP, the X-rays come from the same (synchrotron) component, while for the LSP and many ISP, they come from the low energy end of the high energy (here assumed to be Compton) peak. Particularly interesting are the ISP for which the synchro-Compton transition occurs within the *IXPE* band. To quantify this connection we adopt a multizone jet picture, chapter 4, where the observed modest Π_O are the result of incoherent averaging of N_{eff} effective emission zones, each of which radiates with the $\Pi_{\text{max}} \approx 70\%$ expected for a uniform field, for a power-law population of electrons with index ≈ 2 , producing the observed synchrotron spectrum. From observed polarization levels we typically infer $N_{\text{eff}} \approx 30-100$ for the emission cone contributing to the Earth line-of-sight. In practice, the zones have different angles to the line-of-sight and different characteristic particle energies γ_{max} so the number of zones, and thus Π , becomes a function of the observation frequency (e.g. [95, 17], and chapter 5).

This multizone picture, with distributions in γ_{max} and B field orientation, generally improves the match to observed blazar SEDs over that of a single-zone model. It also means that the νS_y for the individual zones vary and so the number of zones N_{eff} contributing half of the integrated flux is a function of frequency. A computation with N_{eff} related to the peak of the integrated spectrum, assuming typical jet beaming parameters, $B = 0.1$ G and a uniform squared distribution of γ_{max} randomly distributed among the zones is shown in the inset of Figure 6.2. The consequence is

$\Pi \approx \Pi_{\max}/2\sqrt{N_{\text{eff}}}$, with a small increase from the incoherently averaged half of the flux from the remaining zones (see chapter 5 for details). Thus $N_{\text{eff}}(\nu/\nu_{\text{peak}})$ lets us relate the polarization at different frequencies across the synchrotron component. Note that the N_{eff} decrease and Π increase can be dramatic for $\nu \sim 10^3\nu_{\text{peak}}$; some ISPs can be in this regime.

The behavior in Figure 6.2, where the γ_{\max} range is more important than the effective Doppler factor variation, is slightly conservative. In some models, such as the shock model of Marscher [17], γ_{\max} may depend on the angle of B to a shock front and hence to the jet axis; this organized variation further decreases N_{eff} when one is well above the synchrotron peak. We find that this effect is only important for $3 < \log(\nu/\nu_{\text{peak}}) < 4$, but there the polarization increase can be as much as an additional $\sim 2\times$; a few ISPs may have synchrotron X-rays from this extreme regime.

For HSP we can directly convert the optical band polarization level to the X-ray band using the square root of the ratios of the $N_{\text{eff}}(\nu)$. We truncate at $N_{\text{eff}} = 1$ since our statistical estimate breaks down anyway. For some ISP, the Π increase can be substantial as long as the Compton component contributes weakly at 1-10 keV. For ISPs, we used the Space Science Data Center (SSDC) tools⁴ to construct the SED of each source and determine whether the X-ray emission is synchrotron dominated. If so, we expect a substantial Π increase compared to the optical.

For LSP (and ISP with hard X-ray spectra) our model assumes that we observe Compton X-ray flux. This will only show polarization if the seed photon population is highly polarized (e.g., synchrotron emission). chapter 5 find that for isotropic, many-zone scattering in typical jet geometries the resulting Compton polarization is $0.2 - 0.36\times$ that of the seed photons. This does depend on the viewing angle, opening angle, and Lorentz factor of the jet (see chapter 5 for details). However for the typical jet parameters assumed in the present work [119] the retained polarization is near maximal, so we will assume $\Pi_{\text{Comp}} = 0.36 \times \Pi_{\text{seed}}$. To get the latter, we scale from Π_O using Fig. 6.2. For X-ray Compton emission typical seed photons are in the mm-band, we will assume here ~ 100 GHz, but the dependence on the weighted effective seed photon frequency is weak. Note that we are assuming that all seed photons are synchrotron. If external photons contribute to the seed photon population Π_{Comp} will be lower. This means that our estimates of the LSP polarization may be optimistic. This is useful since any observed LSP polarization *higher* than our estimate indicates that the emission should be non-Compton in nature (e.g., proton synchrotron).

With these two effects we predict a Π_X for each source (Table 6.2). Figure 6.2 shows the shift from Π_O to Π_X for a few sources in each class, with an inset showing the N_{eff} dependence on frequency.

6.2.3 Blazar Detectability

Our prime target candidates are the sources measured with the RoboPol program⁵ [120, 82] and the Steward observatory⁶ [121]. For some of these we have *Swift* and/or RXTE monitoring, and so can

⁴<https://tools.ssdc.asi.it/SED/>

⁵<http://robopol.org/>

⁶<http://james.as.arizona.edu/~psmith/Fermi/>

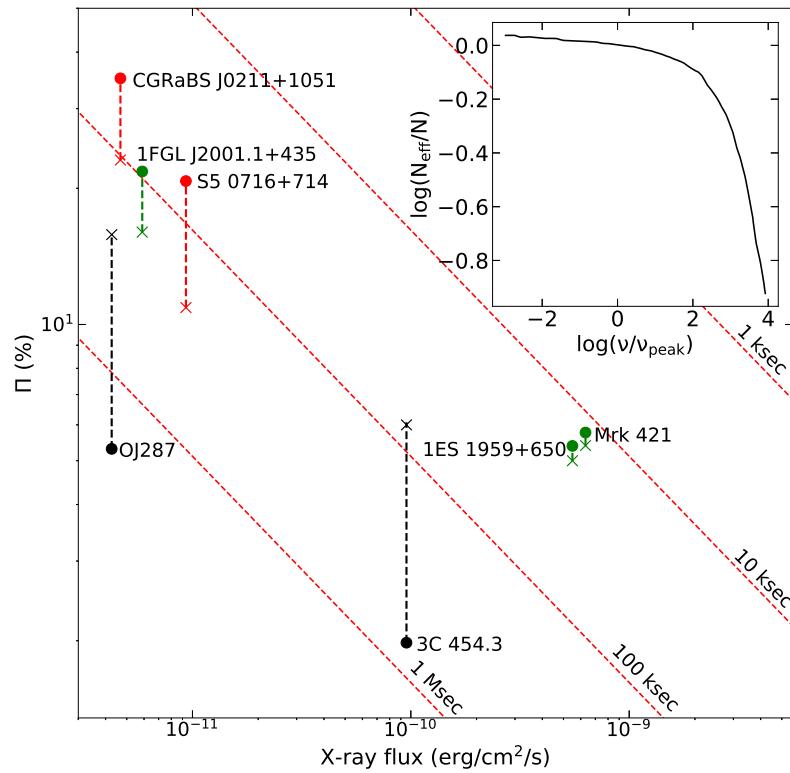


Figure 6.2: The dashed lines show the shift from observed (optical, “x”) to predicted (X-ray, “dots”) Π for a few objects in each source class: LSP (black), ISP (red) and HSP (green). The diagonal lines show the *IXPE* sensitivity for a source with photon index ≈ 2 in a given exposure time. Inset shows the frequency dependence of the effective zone number using the median Lorentz factor and viewing angle from [119].

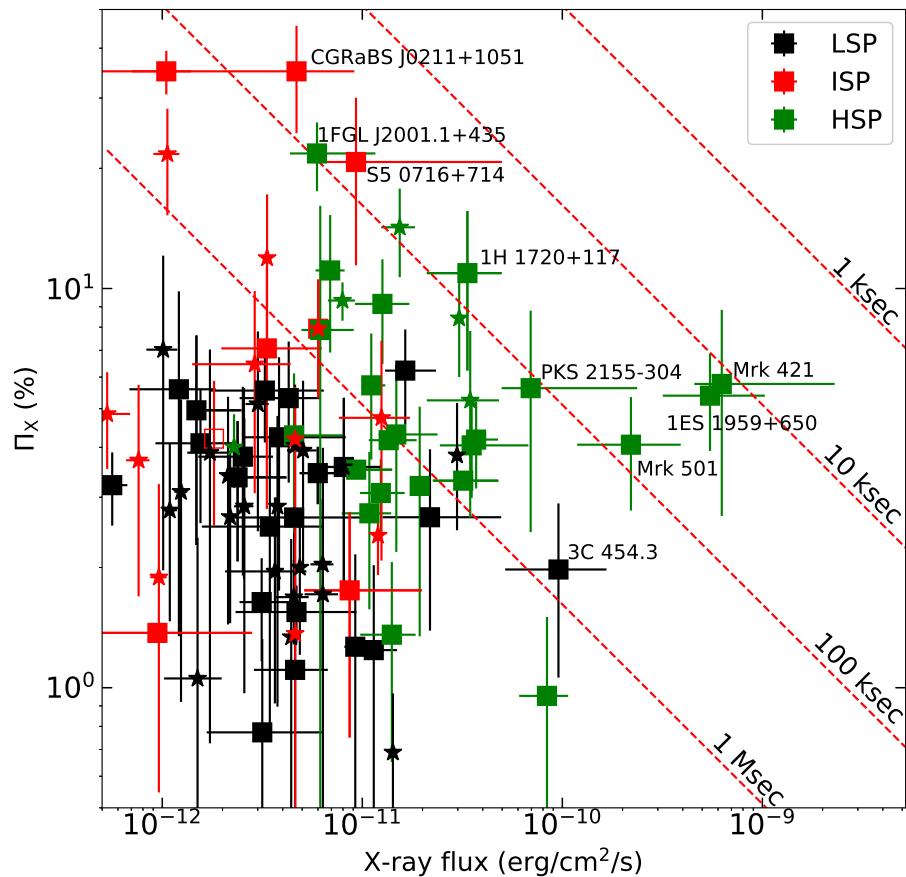


Figure 6.3: Predicted X-ray polarization degree versus the X-ray flux: LSP (black), ISP (red) and HSP (green). Lines show the 1σ variability range in each quantity. Sources lacking at least 3 observations each in S_X and Π_O are shown by open squares. Stars have > 3 measurements in one quantity, solid squares have > 3 in both. Red dashed lines show simple estimates for *IXPE* sensitivity for a source with photon index ≈ 2 in a given exposure time.

construct a detailed variability model (§6.2.1). For the remainder we collect typical fluxes from the HEARSAC⁷ database. There are 103 sources with known SED class and at least one measurement in optical polarization and X-ray flux.

Armed with estimates of the X-ray flux and polarization Π_X and their variability we can make predictions for detectability. We focus on *IXPE* as the most imminent facility, whose sensitivity is estimated using the dedicated online tools⁸ [122, 36]. Typical exposures will be ~ 100 ksec, although the longest may approach a Msec. Figure 6.3 shows the X-ray flux and predicted X-ray polarization degree using the median and 1σ confidence intervals from the PDFs for each source. These PDFs are estimated following §6.2.2. Note that with our assumed correlated fluctuations, HSP and LSP will vary diagonally (UR to LL) within these error bands. As expected, several HSPs are detectable at 100 ksec while a few (e.g., Mrk 421) might give significant measurements in 10 ksec, allowing a detailed variability study. Only few LSP sources are detectable, even with Msec exposures, under typical conditions. A few (e.g., 3C 454.3) are occasionally accessible in shorter time when bright and highly polarized.

Detectability Duty Cycle

We must consider the substantial flux (§6.2.1) and polarization variability (e.g., [96, 123]) when predicting the success of an X-ray polarization search. The uncertainty ranges in Figure 6.3 already give some idea of these effects. But some sources vary well outside these ranges, especially in occasional large S_X flares and less often in polarization increases. Thus we use distribution function models to characterize the full variability range. For the X-ray variability we either use the parameters in Table 6.1 to construct a flux distribution function or use the mean and standard deviation to define a single log-normal model (see §6.2.1). For the optical polarization, we use distribution functions from the maximum likelihood modeling results of [96] for RoboPol sources; for Steward Observatory-monitored sources we use their empirical CDF [121]. As noted above for the ISPs we draw randomly for the CDFs, while for LSPs and HSPs we draw an S_X and then adopt Π_X from the same probability level. We consider only sources with at least three observations in both optical polarization and X-ray flux and estimate the joint detection probability (DP) by computing the MDP_{99} in a given exposure time and comparing with the predicted Π_X . We consider a simulated observation as a detection if $\Pi_X > MDP_{99}$. By repeating this calculation 10^4 times we estimate the fraction of trials a source was detected. Dropping the flux-II correlations results in $< 20\%$ decrease in the LSP, HSP detectability estimates. For the RXTE and *Swift* monitored sources we use the average spectral parameters and WebPIMMS to estimate the MDP_{99} from the drawn flux value. For the remaining sources we use a photon index of 1.5 for inverse Compton and 2.5 for synchrotron emitting sources. In any case, assuming different spectral parameters results in only $\sim 5\%$ change

⁷<https://heasarc.gsfc.nasa.gov/>

⁸<https://ixpe.msfc.nasa.gov/cgi-aft/w3pimms/w3pimms.pl>

in DP. Table 6.3 gives these detection probability values for an assumed 100 ksec *IXPE* exposure. They can be interpreted as the chance of success for a random observation at this exposure, or as the duty cycle for a triggered (by e.g., flux and/or Π_O monitoring) campaign. Of course, if one wants to measure a particular source, one can obtain more acceptable detection odds by increasing the exposure duration. While several HSPs have reasonable detection probabilities, only one ISP (CGRaBs J0211+1051) and one LSP (3C 454.3) are detected at $>10\%$ duty cycle. Thus long monitoring campaigns to allow bright trigger thresholds and/or longer *IXPE* exposures will be needed to reliably detect these source classes. It should be noted that source Π_O can vary by $2\times$ over a few days so longer exposures are not strictly ‘snapshots’ as computed here. Intraday variability is seen, but is uncommon enough to leave our ~ 1 day detectability estimates unaffected.

Sources without measured optical polarization

While many of the best and brightest candidates have been observed in existing optical polarization campaigns, there are other blazars that might be of interest. For example we find 208 blazars from the BZ catalog [124] present in the *Swift* master catalog, 97 of which have $S_X > 5 \times 10^{-13}$ and a known spectral νS_ν class, so that we can evaluate their observability as a function of the unknown optical polarization level. With the observed X-ray flux we estimate the MDP₉₉ (accounting for the different source spectra as in section 6.2.3) as a function of exposure time. We convert this to expected optical polarization using the relation in Fig. 6.2. Figure 6.4 and Table 6.4 show the best prospects from this exercise. Table 6.4 also lists the minimum optical polarization that we would require for ~ 100 ksec *IXPE* detections. This suggests that several more HSP and a few ISP are accessible in reasonable exposures, although one should obtain reconnaissance Π_O measurements first.

6.2.4 Summary

We have used the archival SEDs of bright blazars along with observed optical polarization levels, to predict the expected 1-10 keV X-ray polarization in a basic synchro-self Compton model. This estimate, together with the historical X-ray flux level lets us evaluate the detectability of X-ray polarization for a given mission sensitivity. Including the flux and polarization variability as estimated by cumulative distribution functions modeled from historical data, lets us assess the probability that an exposure of given duration will achieve success. Equivalently, this gives the duty cycle for observations triggered by a monitoring campaign to be successful at a given exposure level. We compute these values for the characteristic *IXPE* mission sensitivity, giving a list of top candidate sources, useful for planning an observing campaign.

Unsurprisingly, HSP dominate the easily detectable sources, but a few ISPs with X-ray emission well above the synchrotron peak are surprisingly observable. In contrast few LSP can be accessed, and then only with long exposures. Recalling that our LSP estimate assumes correlated S_X/Π_O

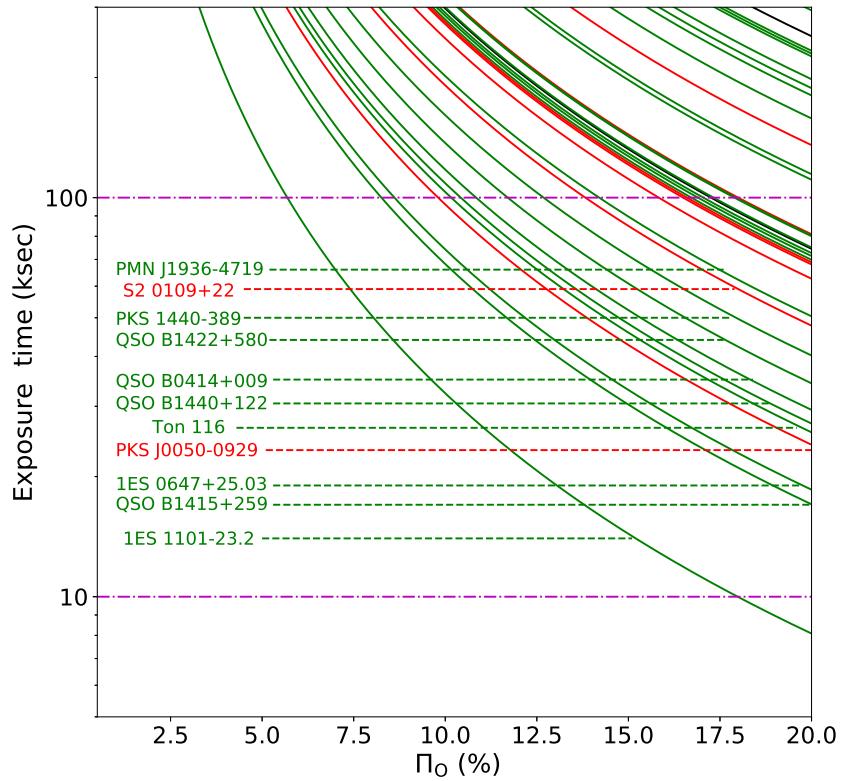


Figure 6.4: Required *IXPE* exposure time as a function of (presently unknown) optical polarization level. This Π_O has been corrected to an expected Π_X , using the sources SEDs. The best prospects are labeled. The sources are colored according to their SEDs: LSPs (black), ISPs (red) and HSPs (green).

Table 6.1: X-ray flux modeling results. The X-ray fluxes are all in erg/cm²/s (log).

Name	Alt. name	SED	f	S _q	σ_q	S _a	σ_a
J0152+0147	1RXS J015240.2+01	HSP	-	-11.19	0.11	-	-
J0210-5101	PKS 0208-512	LSP	0.91	-11.76	0.12	-11.27	0.08
J0222+4302	3C 66A	HSP	-	-11.16	0.06	-	-
J0232+2017	1ES 0229+200	HSP	0.17	-10.89	0.08	-10.64	0.08
J0238+1636	PKS 0235+164	LSP	0.47	-11.84	0.09	-11.46	0.27
J0324+3410	1H 0323+342	HSP	-	-10.85	0.14	-	-
J0530+1331	PKS 0528+134	LSP	-	-11.44	0.24	-	-
J0539-2839	PKS 0537-286	LSP	-	-11.39	0.11	-	-
J0559-5026	PKS 0558-504	-	-	-10.79	0.14	-	-
J0721+7120	S5 0716+714	ISP	0.16	-11.1	0.35	-10.21	0.15
J0831+0429	PKS 0829+046	LSP	-	-11.46	0.08	-	-
J0830+2410	QSO B0827+243	LSP	0.79	-11.70	0.13	-11.08	0.12
J0841+7053	4C 71.07	LSP	-	-10.82	0.09	-	-
J0854+2006	OJ 287	LSP	-	-11.37	0.18	-	-
J1103-2329	1ES 1101-232	HSP	0.2	-10.57	0.06	-10.25	0.02
J1104+3812	Mrk 421	HSP	0.44	-9.52	0.46	-8.94	0.28
J1159+2914	4C 29.45	LSP	0.23	-11.46	0.09	-11.08	0.03
J1221+2813	W Com	ISP	0.73	-12.07	0.06	-11.80	0.22
J1229+0203	3C 273	LSP	0.09	-9.9	0.13	-9.89	0.02
J1256-0547	3C 279	LSP	0.83	-11.24	0.09	-10.94	0.21
J1408-0752	1Jy 1406-076	LSP	-	-12.25	0.09	-	-
J1428+4240	1H 1430+423	HSP	0.52	-10.64	0.2	-10.08	0.12
J1512-0905	PKS 1510-089	LSP	-	-11.07	0.13	-	-
J1555+1111	PG 1553+113	HSP	0.41	-10.89	0.04	-10.65	0.07
J1626-2951	PKS 1622-297	LSP	0.55	-11.52	0.14	-10.96	0.11
J1635+3808	1Jy 1633+38	LSP	-	-11.46	0.24	-	-
J1653+3945	Mrk 501	HSP	-	-9.66	0.24	-	-
J1733-1304	NRAO 530	LSP	-	-11.45	0.11	-	-
J1959+6508	1ES 1959+650	HSP	0.79	-9.93	0.22	-9.27	0.25
J2009-4849	PKS 2005-489	HSP	0.51	-11.01	0.09	-10.03	0.35
J2158-3013	PKS 2155-304	HSP	0.61	-10.62	0.25	-9.94	0.25
J2202+4216	BL Lacertae	LSP	0.1	-10.91	0.16	-10.33	0.15
J2232+1143	CTA 102	LSP	-	-11.02	0.08	-	-
J2253+1608	3C 454.3	LSP	0.93	-10.69	0.1	-10.01	0.22
J2347+5142	1ES 2344+514	HSP	-	-10.47	0.23	-	-

variability, and that no external seed photon flux dominates the up-scatter to the X-ray band, these LSP predictions should be considered optimistic for a Synchro-Compton model. However, other emission scenarios (e.g. proton synchrotron) for the high energy component can produce large Π_X , so a few LSP observations, especially when hadronic emission is indicated, would be desirable.

While our evaluation includes many of the brightest blazars, we have also identified a set which may be interesting targets, if the typical polarization level is sufficiently large. Optical reconnaissance to measure these Π_O and evaluate as possible targets for *IXPE* and/or *eXTP* are strongly encouraged.

6.3 Detecting X-ray polarization in ISP blazars

Both leptonic and hadronic emission processes may contribute to blazar jet emission; which dominates in blazars's high energy emission component remains an open question. Some intermediate synchrotron peaked blazars transition from their low to high energy emission components in the X-ray band making them excellent laboratories to probe both components simultaneously, and good

Table 6.2: X-ray flux and polarization. The X-ray fluxes are all in erg/cm²/s (log). Polarization degree is in %. The table lists sources with > 0.5% X-ray polarization and X-ray flux > 5 × 10⁻¹³erg/cm²/s. The table lists only the first 10 sources. The table is published in its entirety in the machine-readable format. A portion is shown here for guidance regarding its form and content.

Name	Alt. name	Redshift	SED	ν_{peak}	S	σ_S	Π_O	σ_{Π_O}	Π_X
J0017-0512	CGRaBSJ0017-0512	0.227	LSP	13.69	-11.66	0.01	7.99	3.66	2.68
J0035+5950	1ES0033+595	0.086	HSP	17.12	-10.5	0.26	3.1	0.01	3.3
J0045+2127	RXJ00453+2127	—	HSP	16.0	-10.52	0.0	7.4	2.13	8.43
J0102+5824	PLCKERC217G124.4	0.664	LSP	12.94	-11.52	0.06	15.9	8.27	5.14
J0108+0135	PKS0106+01	2.099	LSP	13.18	-11.81	0.25	12.47	4.6	4.09
J0136+4751	S40133+47	0.859	LSP	13.08	-11.59	0.02	11.5	5.76	3.75
J0152+0146	1RXSJ015240.2+01	0.080	HSP	15.46	-11.21	0.29	6.2	6.49	7.87
J0211+1051	CGRaBSJ0211+1051	0.200	ISP	14.12	-11.33	0.41	23.1	6.93	35.0
J0217+0837	PLCKERC217G156.1	0.085	LSP	13.79	-11.44	0.19	5.8	3.09	1.96
J0222+4302	3C 66A	0.340	HSP	15.09	-11.16	0.36	7.8	2.94	11.1

Table 6.3: Detectability duty cycle. The table is sorted according to detection probability and lists only sources with DP> 1%.

Name	Alt. name	SED	ν_{peak}	Det. Prob. (%)
J1959+6508	1ES 1959+650	HSP	16.86	72.9
J1725+1152	1H 1720+117	HSP	16.01	60.6
J2001+4352	1FGL J2001.1+435	HSP	15.21	60.3
J1104+3812	Mrk 421	HSP	17.07	58.5
J0211+1051	CGRaBs J0211+1051	ISP	14.12	49.2
J2158-3013	PKS 2155-304	HSP	15.97	42.1
J1653+3945	Mrk 501	HSP	16.12	30.7
J0222+4302	3C 66A	HSP	15.09	17.0
J1555+1111	PG 1553+113	HSP	15.47	14.4
J2253+1608	3C 454.3	LSP	13.34	10.2
J0721+7120	S5 0716+714	ISP	14.6	5.6
J1838+4802	GB6J1838+4802	HSP	15.8	4.5
J2347+5142	1ES 2344+514	HSP	15.87	3.4
J0958+6533	S4 0954+658	LSP	13.49	3.4
J2202+4216	BL Lac	LSP	13.61	2.5
J1642+3948	3C 345	LSP	13.23	1.8
J1256-0547	3C 279	LSP	13.11	1.8
J0957+5522	4C 55.17	ISP	14.23	1.5

Table 6.4: Sources without optical polarization. The X-ray fluxes are all in erg/cm²/s (log). Column $\Pi_{O,\text{min}}$ lists the minimum optical polarization degree (%) required for an IXPE detection at 100ksec.

Name	Alt. name	Redshift	SED	ν_{peak}	S	σ_S	$\Pi_{O,\text{min}}$
J0050-0929	PKS J0050-0929	0.635	ISP	14.61	-11.09	0.01	9.76
J0112+2244	S2 0109+22	0.265	ISP	14.32	-11.62	0.02	13.76
J0416+0105	QSO B0414+009	0.287	HSP	16.64	-10.7	0.02	10.79
J0650+2502	1ES 0647+250	0.203	HSP	16.42	-10.51	0.01	8.57
J1103-2329	1ES 1101-23.2	0.186	HSP	17.19	-10.07	0.01	5.66
J1243+3627	Ton 116	1.066	HSP	16.15	-10.67	0.02	10.12
J1417+2543	QSO B1415+259	0.236	HSP	15.45	-10.6	0.01	8.22
J1422+5801	QSO B1422+580	0.635	HSP	17.72	-10.73	0.03	11.65
J1442+1200	QSO B1440+122	0.163	HSP	16.35	-10.68	0.02	10.38
J1443-3908	PKS 1440-389	0.065	HSP	15.68	-10.93	0.01	12.63
J1936-4719	PMN J1936-4719	0.265	HSP	16.52	-10.94	0.03	14.14

targets for the newly launched Imaging X-ray Polarimetry Explorer. We characterize the spectral energy distributions for three such blazars: CGRaBS J0211+1051, TXS 0506+056, and S5 0716+714, predicting their X-ray polarization behavior by fitting a multizone polarized leptonic jet model. We find that a significant detection of electron synchrotron dominated polarization is possible with a 300 ks observation for S5 0716+714 and CGRaBS J0211+1051 in their flaring states, while even 500 ks observations are unlikely to measure synchrotron self-Compton polarization. Importantly, non-leptonic emission processes like proton synchrotron are marginally detectable for our brightest ISP, S5 0716+714, during a flaring state. Improved *IXPE* data reduction methods or next generation telescopes like *eXTP* are needed to confidently measure SSC polarization.

Blazars are often classified by the peak frequency (ν_{Sy}) of the low-energy component [13]. Here we focus on the “Intermediate Synchrotron Peak” (ISP) blazars, whose synchrotron emission peaks in optical/UV and have their spectral energy distributions (SEDs) dropping toward eventual high-energy component dominance in the hard X-ray/ γ -ray band. In particular, we focus on a subclass of ISPs whose X-ray emission lies in the valley formed by the combination of the two spectral components.

While the origin of the high-energy component is still unknown, the recent launch of the Imaging X-ray Polarimetry Explorer (*IXPE*, [125]) offers a new diagnostic tool to probe the jet physics, composition, and acceleration of particles. In Liodakis, Peirson, and Romani [5] we used a multizone jet model [17, 1, 2] and optical polarization results from the RoboPol survey [19] to make predictions for the X-ray polarization degree of blazars. In a synchrotron self-Compton (SSC) scenario, we expect substantial polarization from the electron synchrotron and much lower polarization levels from the Compton component. On the other hand, the polarization degree of proton-synchrotron and synchrotron from hadron initiated pair cascades is expected to be much higher than that of IC emission and comparable to that of the primary electron synchrotron component. Interestingly, as one observes further out on the electron synchrotron cut-off tail, fewer jet emission zones have sufficient particle energies and Doppler factors to produce the detected radiation – this means that there is decreased polarization angle (PA) averaging between emission zones and thus larger net synchrotron polarization degree and higher variability [1, 2]. Polarization measurements in the transition region between low- and high-energy components can thus be a powerful tool to probe not only for the high-energy emission processes, but also the jet and magnetic field structure. Coincidentally, recent hybrid (also known as lepto-hadronic) blazar models for the high-energy neutrino emission suggest the existence of subdominant synchrotron components from proton initiated pair cascades that might only be detectable in the transition valley where any primary lepton emission is minimized [e.g., 15]. All of the above suggest that ISPs, whose transition regions lie in the 1-10 keV band are particularly attractive targets for current and future X-ray polarization missions.

We have identified three such sources, namely CGRaBS J0211+1051, TXS 0506+056, and S5 0716+714. CGRaBS J0211+1051 and S5 0716+714 are first year *IXPE* targets, while CGRaBS

J0211+1051 and TXS 0506+056 are potential neutrino emitters [126, 127]. Our goal is twofold: (1) understand the polarization behaviour of the jet across the transition region; (2) make predictions for *IXPE* and future missions to understand the high-energy emission signatures from blazars. In §6.3.1 we describe the data and jet models, in §6.3.2 we make predictions for *IXPE*, and in §6.3.3 we discuss our findings.

6.3.1 Multiwavelength observations and modeling

The non-X-ray multiwavelength data for all sources are taken from the Space Science data center archive⁹. The data are not contemporaneous and include both flaring and quiescent periods of each source. Since the latter may allow improved polarization measurements, we analyze quiescent and flaring data separately. We bin the observations in frequency bins of 0.1 dex and treat the resulting SED as an “average” SED of the source in the given state. Examples are shown in Fig. 6.5.

For the X-ray observations we used publicly available Neil Gehrels Swift Observatory (*Swift*), *NuSTAR* and XMM-Newton data from the High Energy Astrophysics Science Archive Research Center (HEASARC) browse interface. The X-ray observations for the source in normal and flaring (f) states are drawn from the following date ranges: CGRaBS J0211+1051(f) (MJD 55260–55886), CGRaBS J0211+1051 (MJD 59250), S5 0716+714(f) (MJD 57046), S5 0716+714 (MJD 54864–55920), TXS 0506+056(f) (MJD 58025). For CGRaBS J0211+1051 in quiescence we augment with a new ~ 65 ks XMM-Newton observation (AO-19, ID number: 0861840101, MJD 59250). All *Swift* data used were contemporaneous (within 1-2 days) of exposures from one of the other facilities. The data were processed using the standard HEASARC tools and recommended analyses.

Previous fits to these data have typically used absorbed broken powerlaw models for TXS 0506+056 [126] and S5 0716+714 [128]. Instead, we fit the extracted spectra, using *XSPEC*, with a more physically motivated model: the sum of two powerlaw components, subject to absorption by a Galactic neutral hydrogen column N_H . For CGRaBS J0211+1051, the publicly available *Swift* snapshots did not provide sufficient signal-to-noise to unambiguously determine the shape of the 1-10 keV spectrum. We thus tried three models, a single power-law, a power-law with an exponential cut-off, and a sum of two power-laws. We then used the Akaike information criterion to select the model that best describes the data. Again the sum of two power-laws is preferred. The best-fit model parameters are given in Table 6.5.

SED modeling

Our joint SED and polarization modeling uses the polarized leptonic jet emission model developed in Peirson and Romani [1, 2], inspired partly by Potter and Cotter [106] and Marscher [17]. It assumes an initial power-law electron population propagating along a relativistic conical jet. The jet cross

⁹<https://tools.ssdc.asi.it/SED/>

Table 6.5: X-ray spectral parameters. The N_H values are given in units of $\times 10^{22} \text{ cm}^{-2}$.

Name	N_H	Γ_2	Γ_1
CGRaBS J0211+1051	0.16	1.18 ± 0.19	2.68 ± 0.4
TXS 0506+056	0.25	1.83 ± 0.2	3.88 ± 1.0
S5 0716+714	0.031	1.56 ± 0.33	2.36 ± 0.14

section is divided into multiple magnetic field zones, with isotropically distributed field orientations. These magnetic fields are comoving with the jet material.

Polarized synchrotron emission is self-consistently calculated as the electron population propagates and cools. SSC emission is computed, including the propagation of synchrotron photons from downstream emission into each Comptonizing zone. For quasi-spherical magnetic field zones, as often assumed in turbulent scenarios, the model resolves a decorrelation timescale (which depends on the initial jet parameters) of 0.5–5 days. This is the timescale over which steady state jet emission is expected to fluctuate.

An important feature of this jet model is variable Doppler boosting of the zones, since those directed closest to the line of sight increasingly dominant in the observed flux as the SED steepens [2]. This guarantees an increasing expected polarization degree and larger polarization variability above the synchrotron peak. The X-ray SSC polarization degree is typically $\sim 0.2 - 0.35 \times$ that of the synchrotron peak polarization; the components’ PAs are highly correlated. In this model the observed synchrotron and SSC polarization behavior depends significantly on the geometric jet parameters, such as jet opening angle, observation angle, and Lorentz factor.

Our leptonic jet emission model is essentially independent of particle acceleration method since it follows zones downstream of any acceleration region. The assumed ‘chaotic’ disordered magnetic fields and relativistic boosting effects should be present in many magnetic reconnection and shock acceleration scenarios. We also assume steady-state emission, probing polarization variability by re-seeding the magnetic field zones. We briefly discuss how these simplifying assumptions may be violated in other models found in the literature in §6.3.3.

In order to constrain our model’s jet parameters, we fit the multiwavelength SED observations of each blazar state. Due to the chaotic magnetic field zones, our model is stochastic: the same jet parameters can result in different observed SEDs. A stochastic optimization method is necessary to fit such a model to fixed observations. We use a simple variant of the cross-entropy method [129, 130]. At each step, this samples n sets of jet parameters from a multivariate Gaussian and re-fits a new Gaussian using k samples with the lowest χ^2 . Steps are repeated until convergence, when the mean and covariance matrix of the Gaussian no longer change significantly between steps. We use $n = 80, k = 20$. Since the multiwavelength SEDs for each blazar are not simultaneous and the true SEDs can be much more variable than the observational errors imply, we make the simplifying assumption that every observation has the same error. Our jet model has 8 free parameters. We

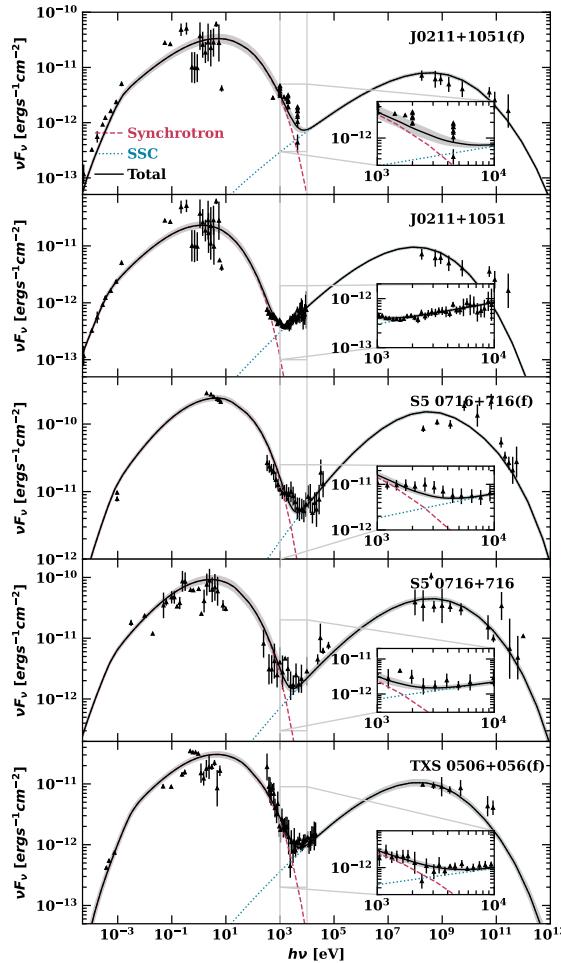


Figure 6.5: Polarized leptonic jet model fits to all blazars and states. ‘(f)’ denotes a flaring state. Black traces show the expected total SED for the best fit jet parameters. Grey shaded regions around the black trace show 1σ model deviations due to different random magnetic field zones. Vertical grey lines denote IXPE’s sensitive band, 1 – 10 keV, and insets show close-ups of this region.

Table 6.6: Polarized jet model best fit parameters. Jet power W_j , electron high energy cutoff before exponential decay E_{\max} , electron population power law index α , full conical jet opening angle in lab frame θ_{open} , bulk Lorentz factor Γ_{bulk} , initial magnetic field strength B_0 , jet observation angle in lab frame θ_{obs} , and initial equipartition fraction A_{eq} . (f) denotes a flaring state. The number of magnetic field zones N_{zones} is selected from [1, 7, 19, 37, 64, 128]. All blazar models shown here use 37 magnetic field zones except CGRaBS J0211+1051, which uses 19.

Name	$W_j [10^{37} W]$	$E_{\max} [10^9 \text{ eV}]$	α	$\theta_{\text{open}} [\text{°}]$	Γ_{bulk}	$B_0 [10^{-5} T]$	$\theta_{\text{obs}} [\text{°}]$	A_{eq}
J(f)	4.94 ± 0.1	13.4 ± 1.0	2.05 ± 0.01	7.16 ± 0.19	14.8 ± 0.64	5.04 ± 0.2	1.95 ± 0.18	0.81 ± 0.01
J	6.48 ± 0.2	9.59 ± 0.1	1.85 ± 0.01	14.1 ± 0.35	7.23 ± 0.15	2.38 ± 0.2	2.31 ± 0.09	0.83 ± 0.02
TXS	5.26 ± 0.5	8.03 ± 0.3	1.89 ± 0.01	4.15 ± 0.26	17.3 ± 0.42	9.63 ± 0.3	1.64 ± 0.11	0.98 ± 0.05
S5(f)	42.4 ± 5.0	13.4 ± 3.0	1.66 ± 0.02	7.35 ± 0.62	13.2 ± 0.51	2.84 ± 0.6	2.51 ± 0.06	1.05 ± 0.02
S5	47.3 ± 11.0	9.49 ± 1.5	1.75 ± 0.09	5.93 ± 0.84	17.0 ± 0.64	3.23 ± 1.5	4.48 ± 0.16	0.84 ± 0.13

open source the code to run our model and reproduce the results¹⁰ [131].

Model fit results are shown in Fig. 6.5. Best fit jet parameters and their respective errors are displayed in Table 6.6. In Fig. 6.6 we show the predicted polarization behavior resulting from the jet model fits displayed in Fig. 6.5. The number of magnetic field zones in the jet model is selected so that the predicted optical polarization fraction matches the median of the observations [19] as closely as possible. Note that individual realizations of the polarization fraction can vary significantly.

It is useful to compare our polarization predictions to previous studies. Zhang et al. [132] model TXS 0506+056 using a single zone leptonic emission model with a uniform magnetic field, matching the observed optical polarization degree with a constant polarization dilution factor. They predict an SSC polarization degree of approximately 5% in the X-ray band, rising to 8% at MeV energies. This represents a slightly higher X-ray polarization fraction, increasing more strongly to high energy. The differences can be attributed to our multizone setup, where multiple magnetic field orientations relative to the line of sight affect the net synchrotron to SSC polarization ratio and its energy dependence [101, 2]; multi-zone models generally predict lower SSC polarization degree. We note that our model also propagates synchrotron seed photons between magnetic field zones, further diluting the SSC polarization and increasing sensitivity to the jet geometry.

6.3.2 IXPE Measurement Simulations

A principal goal of *IXPE*-ISP source measurements is to detect two different X-ray polarizations – a lower energy, electron synchrotron dominated component and a higher energy component. Assuming an SSC spectrum, we explore whether such a measurement is possible for each of our ISPs with typical *IXPE* exposures, using *IXPE*'s standard analysis pipeline processing over a 2-8 keV band.

Using *ixpeobssim*, *IXPE*'s observation simulation software [133], we simulate multiple 300 ks and 500 ks observations for each blazar state assuming polarization and flux are fixed to their expected (average) values (i.e. the black traces in Figs. 6.5, 6.6). We split the simulated 2-8 keV data into

¹⁰<https://github.com/alpv95/SSCPol>

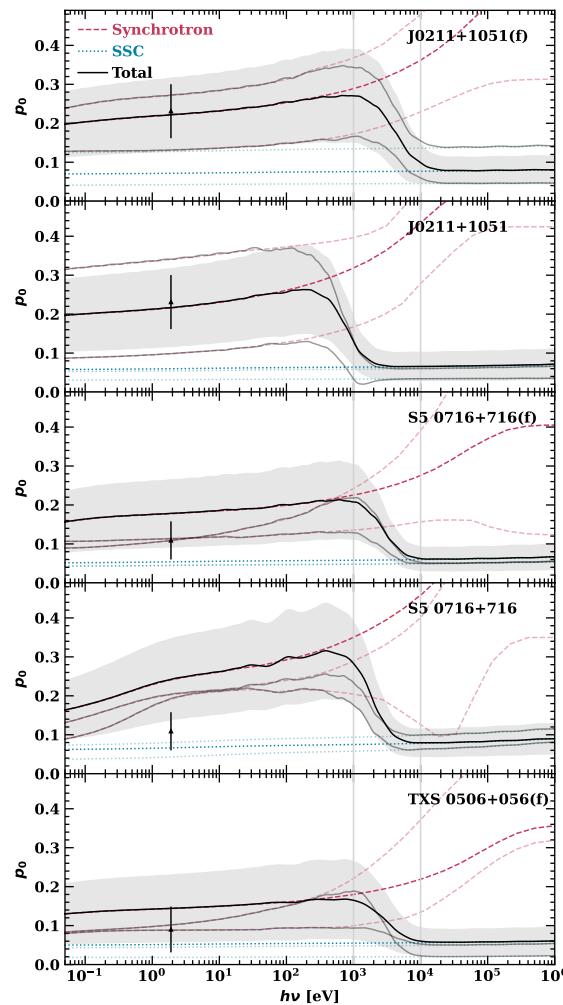


Figure 6.6: Leptonic (SSC) jet model polarization fraction predictions. The jet models used in each panel are the same as those in Fig. 6.5. Black observations denote the average measured optical polarization over multiple epochs (MJD 56432 – 57893 for Robopol measured S5 0716+714 and J0211+1051 [19] and MJD 58019 – 58267 for TXS 0506+056). Lines and shaded regions mean the same as Fig. 6.5 with the addition of two transparent models, which represent randomly selected model realizations.

two energy bins: 2-4 keV and 4-8 keV, extracting the polarization fractions by estimating the Stokes' parameters as in Kislat et al. [38]. Figure 6.7 summarizes the results.

In Fig. 6.7, energy bins where the true polarization fraction distribution (blue, right-hand-side) is fully below the minimum detectable polarization (MDP_{99}) threshold (dotted lines) cannot produce significant ($\gtrsim 3\sigma$) detection of non-zero polarization in the given exposure time. MDP_{99} is the 99th percentile upper confidence bound on polarization fraction for an unpolarized source. Energy bins with some or all of the true polarization distribution above MDP_{99} can have significant detections, if their actual polarization is in the upper portion of the predicted range – the measurement errors would be approximately given by the measured polarization fraction distributions for the most probable p_0 (red, left-hand-side). Planned observation times for first-year *IXPE* ISP targets, including CGRaBS J0211+1051 and S5 0716+714, are expected to range from 200 ks – 400 ks.

For each blazar and state the two energy bins, 2-4 keV and 4-8 keV, contain different relative synchrotron and SSC contributions. Insets in Fig. 6.5 give the relative contributions. In non-flaring states, both energy bins are almost entirely dominated by SSC emission so measurement of the synchrotron cutoff component will not be possible.

Low significance polarization fraction measurements, below MDP_{99} , are strongly biased away from $p_0 = 0$. Strict non-negativity of p_0 forces measurement posteriors (red, Fig. 6.7) to be asymmetric and for $E(\hat{p}_0) > p_0$ (see, esp. CGRaBS J0211+1051 quiescent panel). This highlights the danger of making polarization inferences using low significance point estimates. The measurement bias can be corrected using appropriate p_0 estimators [134].

6.3.3 Discussion

Under a purely leptonic (SSC) jet model for ISPs, we find that simultaneously detecting significant X-ray polarization from both emission components with a ≤ 500 ks *IXPE* observation is impossible, even considering high p_0 fluctuations (see Fig. 6.7). For the assumptions used here, a 2.5 Ms exposure would be required to measure the median predicted SSC polarization in our brightest source, S5 0716+714, during its high state. Unfortunately, blazar polarization variability may preclude such long observation times. Optical polarization measurements [19] suggest that blazar polarization fraction and PA can vary significantly over time periods < 500 ks. This would result in an incoherent averaging of polarization vectors leading to depolarization. Many blazar models [17] including our own (Fig. 6.6) predict polarized X-ray electron synchrotron emission to be more variable than the optical [135].

If external Compton (EC) contributes significantly to a blazar's high energy emission component, the case for measuring its X-ray polarization becomes even more dim. EC emission is usually assumed to be unpolarized [105] since the external photon field being scattered is assumed incoherent, originating in the broad line region or accretion disk. Even a small EC contribution can make observations more difficult because $MDP_{99} \propto 1/\sqrt{N_{ph}}$. A 10% fractional EC contribution would

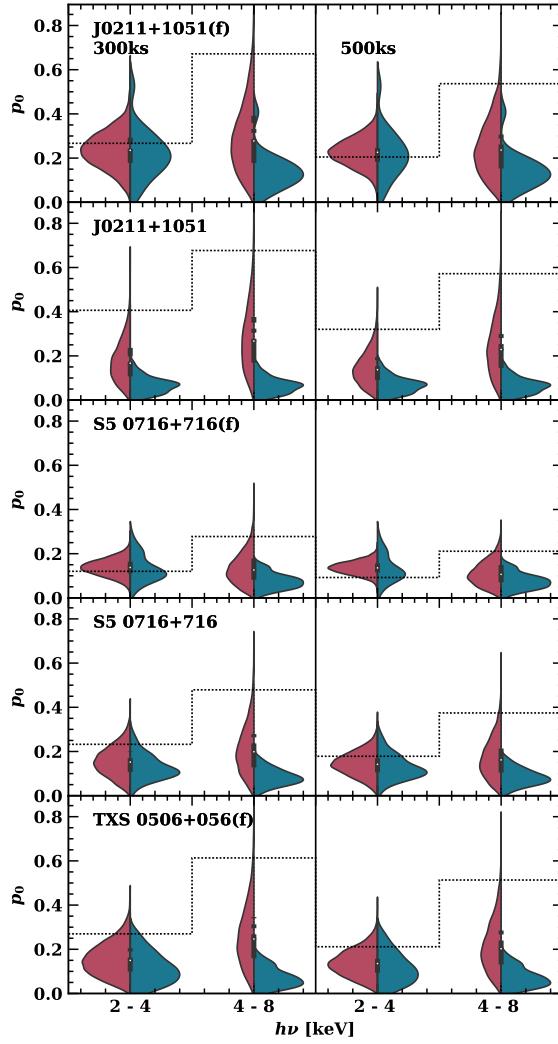


Figure 6.7: Violin plots of the true polarization fraction distribution (blue, left-hand-side) and the measured polarization fraction distribution (red, right-hand-side) for 2 – 4 keV and 4 – 8 keV energy bins, and 300 ks/500 ks exposures. The distributions of true polarization fractions are extracted from our jet model fits (Fig. 6.6) and are the same for both exposure times. Measured polarization fraction distributions assume a single observation with true polarization equal to the expected value. Dashed black traces represent the minimum detectable polarization (MDP₉₉) for each measurement bin.

lower Fig. 6.7 true polarization fractions by 10% and increase required observation times for the same significance by 23%. Luckily, all three sources considered are classified as BL-Lac objects, typically associated with low EC contributions. Padovani et al. [136], however, suggest that TXS 0506+056 is an FSRQ in disguise, in which case there might be significant EC contribution to the high energy component.

For the ISPs in flaring states, a significant polarization measurement of the synchrotron cutoff is feasible although still difficult, requiring the blazar to be in a high polarization state. Along the primary synchrotron cutoff we expect an increased expected polarization fraction and variability compared to the optical SED peak (see Fig. 6.6) as the most Doppler-boosted magnetic field zones increasingly dominate the observed emission [2]. Our model presents the minimal (geometry-induced) increase polarization degree above the primary synchrotron peak; other effects may further increase the dominance of individual zones. For example, in the shock scenario particles are more efficiently accelerated when the magnetic field is aligned along the shock normal [17]. Thus chaotic magnetic field zones will vary in their upper electron energy cutoff and that energy can correlate with the global jet geometry. The highest energy electrons contributing X-ray emission are close to the shock, where cooling is limited and the field orientation (and hence polarization) are more highly correlated. Also, Tavecchio et al. [137] show that immediately downstream of a shock magnetic field compression increases the field perpendicular to the shock normal; this preferential alignment tends to correlate the field orientations and increase net polarization degree, although such correlation decays as turbulence develops downstream. Both these effects may increase polarization at the high energy end of the synchrotron component, improving measurement prospects. In contrast magnetic reconnection scenarios suggest synchrotron cutoff polarization with higher variability but similar net polarization degree to the synchrotron peak emission [138, 137]. Thus comparing X-ray polarization to simultaneous optical polarization degree may be able to distinguish these acceleration scenarios.

Unexpectedly large high energy component polarization arising from non-leptonic jet emission is possible and potentially detectable. Our leptonic (SSC) jet model predicts any high energy component polarization should typically be $0.2 - 0.35 \times$ lower than at the SED optical peak [2] with the decrement sensitive to the jet geometry. Of course, this ratio is variable and can occasionally fluctuate to large values > 0.5 , especially if the peak polarization is low, so only multi-epoch trends or long-term averages have predictive power. In the most optimistic hadronic jet scenario, proton and secondary electron synchrotron dominate the high energy emission component [105, 15]. High energy component X-ray band polarization fractions would be similar to the SED optical peak (Fig. 6.6), corresponding to a X-ray/optical polarization ratio of 1, extremely unlikely in a SSC dominant leptonic emission model. Although this would provide a much needed polarization fraction boost, a two component detection would remain out of reach, even for a ≤ 500 ks *IXPE* observation, for all ISPs except S5 0716+714 in its flaring state. Indeed, S5 0716+714 is the first ISP *IXPE* target, planned for a 300 ks observation on 31st March 2022; a significant polarization detection for both

high and low energy components would be a promising indication of non-leptonic jet emission.

If initial *IXPE* observations do not detect significant polarization from either emission component in any ISPs, it will be difficult to rule out non-leptonic processes. Upper polarization fraction limits based on the MDP_{99s} in Fig. 6.7 will be too high to make any useful inference about the polarization ratio of the two emission components, even with a strong synchrotron detection at 2-4 keV. However, the measurements' sensitivity may be improved. Bayesian neural network analysis of *IXPE* data [3, 4] has been shown to reduce MDP_{99s} by up to 25% compared to the standard *IXPE* analysis pipeline, as well as increasing *IXPE*'s effective energy band to 1-10 keV. We may also tune the energy range of the ‘low’ and ‘high’ energy detection windows for an individual source’s SED, improving our ability to measure or bound the two components’ p_0 . Such improved analysis could, for example, make a flaring 4-8 keV S5 0716+714 SSC polarization detection possible with a 500 ks observation – Fig. 6.7. Although neural network analysis is not yet in production for *IXPE*, a re-analysis of borderline observations could reveal missed discoveries. Looking further ahead, the effective area of future X-ray polarization mission *eXTP* [139] should be four times larger than *IXPE*'s, reducing MDP_{99s} by a factor of 0.5 [140]. Including both improvements, simultaneous measurement of both ISP emission components with a 500 ks observation is well in scope for all the ISPs considered here.

Chapter 7

X-ray Polarization of BL Lacertae in Outburst

7.1 Introduction

We report the first $> 99\%$ confidence detection of X-ray polarization in BL Lacertae. During a recent X-ray/ γ -ray outburst, a 287 ksec observation was taken using the Imaging X-ray Polarimetry Explorer (*IXPE*), together with contemporaneous multiwavelength observations from the Neil Gehrels *Swift* observatory and *XMM-Newton* in soft X-rays (0.3–10 keV), *NuSTAR* in hard X-rays (3–70 keV), and optical polarization from the Calar Alto, and Perkins Telescope observatories. Our contemporaneous X-ray data suggest that the *IXPE* energy band is at the crossover between the low- and high-frequency blazar emission humps. The source displays significant variability during the observation, and we measure polarization in three separate time bins. Contemporaneous X-ray spectra allow us to determine the relative contribution from each emission hump. We find $> 99\%$ confidence X-ray polarization $\Pi_{2-4\text{keV}} = 21.7^{+5.6\%}_{-7.9\%}$ and electric vector polarization angle $\psi_{2-4\text{keV}} = -28.7 \pm 8.7^\circ$ in the time bin with highest estimated synchrotron flux contribution. We discuss possible implications of our observations, including previous *IXPE* BL Lacertae pointings, tentatively concluding that synchrotron self-Compton emission dominates over hadronic emission processes during the observed epochs.

In this chapter we focus on ISPs and in particular BL Lacertae (BL Lac), whose recent γ -ray outburst (see Appendix B.3) briefly boosted its characteristic peak energies, moving it from the LSP into the ISP class. In ISPs the peak of the low-energy component ranges from the near-IR through the UV bands. Thus, the 2–8 keV *IXPE* band may include substantial emission from the falling high-frequency tail of the leptonic synchrotron (Sy) spectrum emitted by the most efficiently accelerated electrons and positrons, by a related higher-frequency peak hadronic component

(synchrotron emission from protons), or by the flatter-spectrum synchrotron self-Compton (SSC) leptonic emission. Indeed, in leptonic models the *IXPE* band may lie in the U-shaped transition region from synchrotron to SSC emission [e.g., 6, fig. 7.1]. Since the polarization is expected to differ between Sy and SSC, *IXPE* ISP observations can probe both radiation processes, and possibly the jet’s composition [6, 105]. The latter is of particular interest, as blazars have been proposed as candidate sources of TeV neutrinos and ultra-high-energy cosmic-rays (UHECR), which would require a significant hadronic component in some blazar jets [15]. The possible 3σ association of ISP blazar TXS 0506+056 with the neutrino IceCube-170922A event motivates this connection [63, 29]. However, the peculiar γ -ray behavior of TXS 0506+056 and the $\sim 40\%$ probability of an atmospheric origin of the neutrino challenges the association. Current blazar models that include neutrino emission assume either lepto-hadronic X-ray emission (e.g., [64]) or subdominant hadronic components where the proton emission only dominates the SED in the transition region where the leptonic component is at a minimum level (e.g., [15]). These factors highlight the importance of distinguishing between leptonic and hadronic emission in an ISP blazar when it comes to X-ray polarization measurements with *IXPE*.

BL Lac is among the few LSP or ISP blazars detected at ≥ 0.1 TeV (very high-energy, VHE) γ -ray energies, and is the 14th brightest AGN at GeV energies in the *Fermi* 4LAC catalog [141]. It is a rapidly variable VHE source on timescales as short as < 1 hour [142, 143]. Moreover, BL Lac has been the focus of a large number of multi-wavelength and polarization studies [e.g., 65, 144, 145, 146]. Because of this blazar’s strong variability, the interpretation of its SED requires simultaneous observations.

In this chapter, we report the first $> 99\%$ confidence detection of X-ray polarization in BL Lac, with redshift $z = 0.0686$ [147] and synchrotron peak frequency $\nu_{\text{syn}} = 1.98 \times 10^{14}$ Hz [148]. Although BL Lac is typically classified as an LSP, where the hard secondary SED component dominates the soft X-ray band, its recent outburst both softened and brightened its X-ray spectrum. The soft X-ray flux density and photon index are highly variable, but typically $< 2 \times 10^{-11}$ erg cm 2 s $^{-1}$ and ≤ 1 , respectively [e.g., 149, 150, 151, 152]. However, during the recent flare the *Swift* XRT finds a flux density $F_{2-8\text{keV}} = (2.77 \pm 0.21) \times 10^{-11}$ erg cm 2 s $^{-1}$, photon index 2.10 ± 0.09 on 2022 November 12, which suggests a significant contribution from the synchrotron component.

In §7.2 we describe the X-ray, optical, and radio polarization observations and data reduction. We discuss our findings in the context of multiwavelength observations in §7.3. Further analysis details can be found in the appendix §B.2, §B.3.

7.2 Data analysis

BL Lac was observed by *IXPE* with an exposure time of 287 ksec on 2022 November 27–30 (MJD 59910.58 – 59913.90). *IXPE*, launched on 2021 December 9, is a joint mission of NASA and the

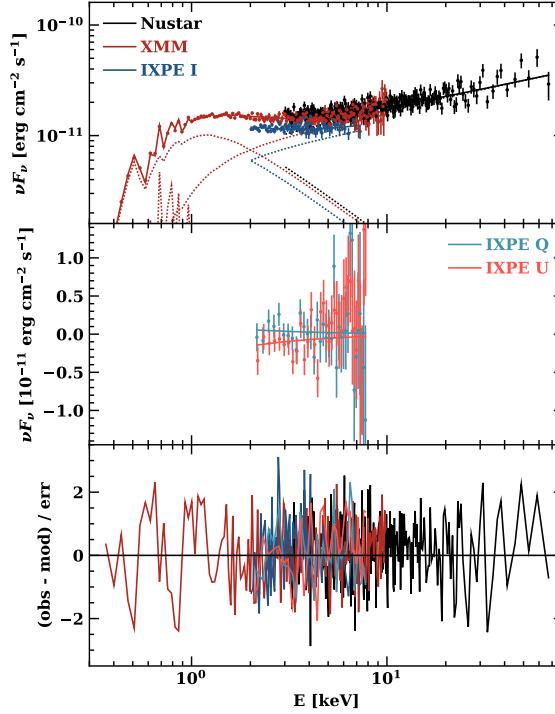


Figure 7.1: Quasi-simultaneous X-ray spectra during the *IXPE* observation. The fitted model is the sum of two absorbed power-laws and an apec component. Dotted lines in the top panel show the individual power-law and apec components. The middle panel shows the polarized spectra with constant polarization for the low-energy component; the high energy component has fixed polarization degree PD=0. The significance of the spectral fit residuals is shown in the bottom panel.

Italian Space Agency (Agenzia Spaziale Italiana, ASI). A description of the instrument is given by Weisskopf et al. [125]. At the $\sim 30''$ angular resolution of *IXPE*, BL Lac is a point-like source. Quasi-simultaneously, BL Lac was observed in the hard and soft X-ray bands with *NuSTAR* (MJD 59911.87 – 59912.31), *XMM-Newton* (MJD 59910.27 – 59910.39) and *Swift* XRT (MJD 59910.00, 59912.78), in linear polarization at optical BVRI bands with Calar Alto, the Nordic Optical Telescope (NOT), the 1.8m Perkins Telescope of Boston University, and the Sierra Nevada Observatory, and at millimeter wavelengths by the Institut de Radioastronomie Millimétrique 30-m Telescope (IRAM-30m) and SubMillimeter Array (SMA). Appendices B.1, B.3 detail the data reduction of these multi-wavelength observations. The *IXPE* data were calibrated and reduced following standard procedures within the *ixpeobssim* pipeline¹ [153].

To characterize the contribution of the two emission components in the *IXPE* 2–8 keV band, we first attempt to fit the broad-band X-ray spectrum averaged over the full observation, including

¹<https://agenda.infn.it/event/15643/contributions/30820/attachments/21780/24810/ixpeobssim.pdf>

Table 7.1: Spectro-Polarimetric XSPEC model fits to joint *NuSTAR*, *XMM-Newton*, and *IXPE* spectra, fig. 7.1. Γ – photon index, Π – polarization fraction, ψ – EVPA. For the sum of two power-laws, the polarization fraction of the second power-law component is fixed to zero. Since the single power-law fit is inadmissible, uncertainties are likely underestimated for this model’s parameters. Upper limits are at 99% confidence.

Model Component	Power-law	+2 Power-laws
χ^2/dof	1898/1087	1102/1085
$N_H [10^{21} \text{ cm}^2]$	1.63 ± 0.23	3.10 ± 0.12
$kT [\text{keV}]$	0.663 ± 0.017	0.345 ± 0.025
Γ_1	2.04 ± 0.009	3.28 ± 0.100
$\Pi_1 [\%]$	< 15.0	27.6 ± 11.6
$\psi_1 [\circ]$	n/a	-34.5 ± 11.6
Γ_2	–	1.62 ± 0.027

XMM-Newton 0.3–10 keV and *NuSTAR* 3–80 keV data. A simple absorbed power-law model is inadequate (table 7.1). We find the sum of two absorbed power-laws best reproduces the observations, as measured by χ^2 per degree of freedom, table 7.1. *XMM-Newton* measurements below 1 keV require an additional soft spectral component, probably from a hot diffuse plasma, as noted for BL Lac by Middei et al. [152]; we thus add an unpolarized apec component in XSPEC. The temperature and normalization of this apec component are free parameters in the fit. For the double power-law case, the best-fit absorption column density, $N_H = 3.1 \pm 0.12 \times 10^{21} \text{ cm}^{-2}$, is consistent with values from previous studies [e.g., $N_H = 2.8 \pm 0.05 \times 10^{21} \text{ cm}^{-2}$; 144]. Figure 7.1 shows the results of the double power-law fit, including the individual power-law and apec components. We note that the soft power-law component comprises 36.7% (2–4 keV) or 26.1% (2–8 keV) of the total flux density.

By restricting the softer power-law to have constant linear polarization, while assuming that the harder power-law is unpolarized, we find $\Pi_X = 27.6\% \pm 11.6\%$ (with $\Pi_X < 57.6\%$ at 99% confidence), $\psi_X = -34.5^\circ \pm 11.6^\circ$, fig. 7.2. We do not consider this a significant detection, since it does not meet our 99% confidence of non-zero polarization requirement. We discuss the rationale for assuming an unpolarized high-energy component in §7.3. Allowing both power-laws to have independent, constant polarization results in high, nearly orthogonal values, which cancel the net polarization over most of the spectrum; these values are poorly constrained owing to the poor statistics.

7.2.1 Time variability

The null result is based on the assumption of constant polarization with time of both power-law spectral components. However, the source is clearly variable during the present *IXPE* epoch, which may affect the derived polarizations. To assess this, we split the observation into three equal time bins based on the the *IXPE* count spectrum and optical observations. *XMM-Newton* and *NuSTAR* observations straddle the first time bin, and three short ~ 1 ksec *Swift* XRT exposures are available,

two within the second time bin and the other in the final time bin. To assess the contribution of the individual power-law components to each time bin, we fit the absorbed sum of a double power-law model (table 7.1) separately to each appropriate set of *IXPE* plus *Swift* XRT time-binned spectra, fixing the photon indices Γ_1 and Γ_2 to their *NuSTAR* plus *XMM-Newton* fitted values. This allows the power-law normalizations to vary in time while their slopes remain fixed. We also analyze the data over two energy bins, measuring separately the low- (2–4 keV) and full-energy (2–8 keV) ranges. Estimated fractions of the lower-energy power-law (synchrotron) contribution for each time and energy bin are displayed in table 7.2. While the similar variation of the fluxes in the two energy bins (fig. 7.3) shows that both power-laws vary together (as expected for synchrotron and Compton components), we find that the soft-component contribution is largest in the first time bin, as might be expected from observations at the high-energy tail of a cooling synchrotron flare.

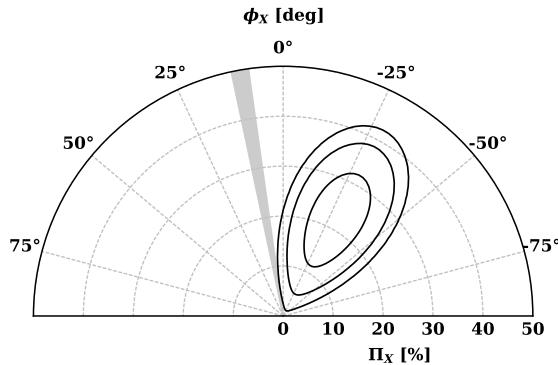


Figure 7.2: Polarization fraction and EVPA confidence levels (68%, 95%, 99%) for the first time bin over the 2–4 keV energy range, fig. 7.3. Gray shaded region represents the VLBI-determined jet axis projection on the plane of the sky [20].

To check whether the polarization degree and electric vector position angles (EVPA) vary with energy, supporting a multi-component interpretation, we measure the polarization signal in each time and energy bin. These X-ray polarization measurements (fig. 7.3, table 7.2) represent the average values for the combined spectral components in the six time/energy bins, rather than a single source component as in table 7.1. Interestingly, in the low energy channel within the first time bin, where our spectral fits indicate that the soft component contributes the maximum flux, we detect linear polarization exceeding minimum detectable polarization at 99% confidence (MDP99), fig. 7.2. When considered as a single measurement, we find 99.3% confidence in non-zero polarization. Considering all six binned polarization measurements jointly we find a 98.8% confidence in non-zero polarization. As noted previously, we do not exceed the 99% threshold over the full time interval.

We have attempted full spectro-polarimetric double power-law fits to each time bin as in fig. 7.1, but none produced a high-significance detection for the soft power-law component. Instead, we use these fits to estimate the lower-energy power-law (synchrotron) contribution fractions in table 7.2.

Table 7.2: Time-binned polarization results presented in fig. 7.3. MDP99 is the minimum detectable polarization, i.e., the 99% confidence level for non-zero polarization fraction. All errors are 68.3% (1σ) confidence intervals. Error on polarization fraction Π_X is non-symmetric because $0 \leq \Pi_X \leq 1$. PL1 fraction is the fraction of flux density in the bin from the lower energy (synchrotron) power-law component. For polarization measurements less than the MDP99 we display upper limits at 99% confidence.

Time [MJD]	Energy [keV]	Π_X [%]	ψ_X [$^\circ$]	MDP99 [%]	PL1 fraction [%]
1 (59910.58 – 59911.69)	2–4	$21.7^{+5.6}_{-7.9}$	-28.7 ± 8.7	20.0	44.1
	2–8	< 27.9	n/a	22.5	32.6
2 (59911.69 – 59912.80)	2–4	< 18.5	n/a	22.8	28.3
	2–8	< 22.4	n/a	23.6	19.5
3 (59912.80 – 59913.90)	2–4	< 23.3	n/a	20.6	22.1
	2–8	< 22.2	n/a	23.2	14.8
Full (59910.58 – 59913.90)	2–4	< 19.8	n/a	12.2	36.7
	2–8	< 14.3	n/a	13.6	26.1

The results for all such fits are detailed in Appendix B.2.

7.3 Discussion & Conclusions

In order to fit our *XMM-Newton* and *NuSTAR* spectra, (Fig. 7.1), we find that a two-component spectrum is required for BL Lac. A power-law fit with low-energy photon index $\Gamma_1 = 3.28$ and high-energy value $\Gamma_2 = 1.62$, with equal flux density at 2 keV, is statistically preferred over a simple power-law. Our polarization results thus probe the cross-over region.

Pure leptonic models predict a significant decrement in polarization of the hard (Compton) component relative to the optical synchrotron value, with $\Pi_{SSC}/\Pi_{Sy} \approx 0.3$ [102, 2]. When the high-energy tail of the synchrotron emission reaches into the *IXPE* band, we might expect the soft-component polarization even to exceed the optical value if the X-ray emission arises from a region of more highly-ordered magnetic field.

One possibility is that the X-ray synchrotron emission occurs mainly in the acceleration and collimation zone of the jet, where the magnetic field is expected to have a well-ordered helical geometry [e.g., 154], as inferred previously in BL Lac [155]. Alternatively, the emission could take place in a region of the jet with a turbulent magnetic field, with acceleration of particles occurring only within a small volume, for example by magnetic reconnection. Since X-ray synchrotron radiation requires extremely high-energy electrons that are subject to strong radiative energy losses, the X-ray emission would be confined to locations close to the site of particle acceleration, while lower-frequency emission from lower-energy electrons, which can travel farther from the acceleration site before their fractional energy loss becomes high, would occur over a larger volume. In a turbulent magnetic field,

the net field is more ordered (but randomly oriented) over smaller volumes, hence the polarization is higher [see, e.g., 17, 1, for discussions and estimates]. The random value of ψ agrees with our finding that the observed EVPA does not appear related to the direction of the jet.

If the turbulent plasma encounters a shock, the magnetic field becomes partially aligned with the shock front [e.g., 156, 17, 137], at which X-ray emitting particles can be accelerated. By the same argument as above, the synchrotron X-ray polarization should then be higher than at longer wavelengths, but instead of a random orientation of the EVPA, ψ should be oriented along the jet direction, contrary to the observations of BL Lac presented here.

During a synchrotron X-ray flare, the more highly polarized synchrotron component contributes more to the X-ray spectrum at lower X-ray energies than does the flatter-spectrum Compton component. As this steep-spectrum component fades across the *IXPE* band, the lower Compton polarization should dominate at higher energies. Although the synchrotron peak and SSC peak EVPAs are expected to be correlated [2], turbulence and energy-dependent emission volumes can cause differences.

To best constrain the soft component polarization we assume that the high energy power-law component is negligibly polarized, as the logical approximation of the expected polarization decrease noted above. The polarization $\Pi_X = 27.6\% \pm 11.6\%$ measured over all *IXPE* times and energies, while a $> 2\sigma$ detection, does not reach the threshold for high significance. However, noting that the soft component varies in strength, we find that a time and energy bin analysis allows a significant $\Pi_X = 21.7^{+5.6\%}_{-7.9\%}$ detection at low energies in the first time bin, when the spectral analysis indicates that the soft component accounts for the largest fraction of the total flux. Note that this polarization is substantially higher than the simultaneous optical polarization $\Pi_O = 13.1\% \pm 2.4\%$ for the same time bin. Both the optical and X-ray EVPA are at $\sim 40^\circ$ to the projection of BL Lac's jet axis on the plane of the sky at 43GHz [20].

Even the first time/energy bin is diluted by the high energy component. Under the assumption of relatively constant photon indices at the *XMM-Newton* and *NuSTAR* values, the high-energy component still comprises 56% of the flux density in this 2–4 keV time bin (table 7.2). If the component is unpolarized, correction for this flux implies $\Pi_{Sy} \approx 49\%$ for the soft component; if the high-energy component corresponds to Compton scattering, polarized at $\sim 0.3 \times$ the synchrotron value with approximately the same EVPA, the polarization of the low energy component is $\Pi_{Sy} \approx 36\%$. These polarization degrees are consistent with the upper limit provided by the double power-law spectral fit in table B.3 of the appendix. These large values dramatically exceed the optical polarization, arguing that we are probing the high-frequency end of the synchrotron component, which fades during our observation window.

Previous *IXPE* observations of BL Lac [157], which occurred during a low state, where the X-ray spectrum was apparently dominated by SSC, produced upper limits on Π_X that were significantly below the simultaneous optical polarization. However, in the data from the third epoch presented

here, even the total (time and energy) averaged *IXPE* polarization signal is comparable to the optical values. After correction for hard-component contamination, the soft component substantially exceeds the optical polarization degree.

Other scenarios have different polarization signatures. In simple single-zone lepto-hadronic models, the polarization should be similar across the transition region (fig. 7.1). Synchrotron emission from primary leptons producing the low-energy hump yields to higher-energy synchrotron radiation from protons and secondary leptons produced in $p\gamma$ cascades [105], all in the same magnetic field environment. As noted above, in more realistic jet models, turbulent magnetic fields, differential cooling times [89] and relativistic boosting [2] cause a significant increase in polarization and changes in EVPA as one approaches the quasi-exponential cutoff of the low-energy peak, where only the highest energy electrons from the most efficient particle acceleration zones contribute synchrotron emission. If a hadronic emission component becomes dominant above this energy, it should be averaged over the lower-energy particle population typical of particle acceleration in the jet, and thus should display polarization similar to that of the primary (optical) synchrotron peak. In models including hadronic synchrotron emission, we expect the polarization fraction at the upper end of the transition region to meet or exceed the optical value.

Our past measurements of $\Pi_X < \Pi_O$ [157] and the persistent low polarization at the upper end of the *IXPE* band therefore argues against a lepto-hadronic emission model. Furthermore, the detection of total polarization at low energies, with $\Pi_{2-4\text{keV}} \gg \Pi_O$, indicates that we are sampling the upper cut-off of the synchrotron spectrum. This is even more striking after correction for dilution by the hard, weakly polarized component. The decreasing relative flux of the soft component and the resulting loss of detection of polarization are consistent with the tail end of a jet flare, as suggested by the γ -ray and X-ray light curves.

While these results provide tentative evidence for a fully leptonic emission model in BL Lac at our *IXPE* observation epoch, variability allows significant hadronic emission at other times. Blazars are particularly variable in polarization [82], and broadband or long-exposure measurements can make polarization results difficult to interpret properly [158, 6]. Thus, repeated polarization measurements of blazars, with attendant contemporaneous multiwavelength intensity and polarization observations, are needed to fully explore blazar emission. Most interesting would be measurements during flares associated with neutrino events. With the plausible neutrino connection, and the ability to probe the critical Sy/SSC transition regions, BL Lac and other ISPs will be prime targets for such studies.

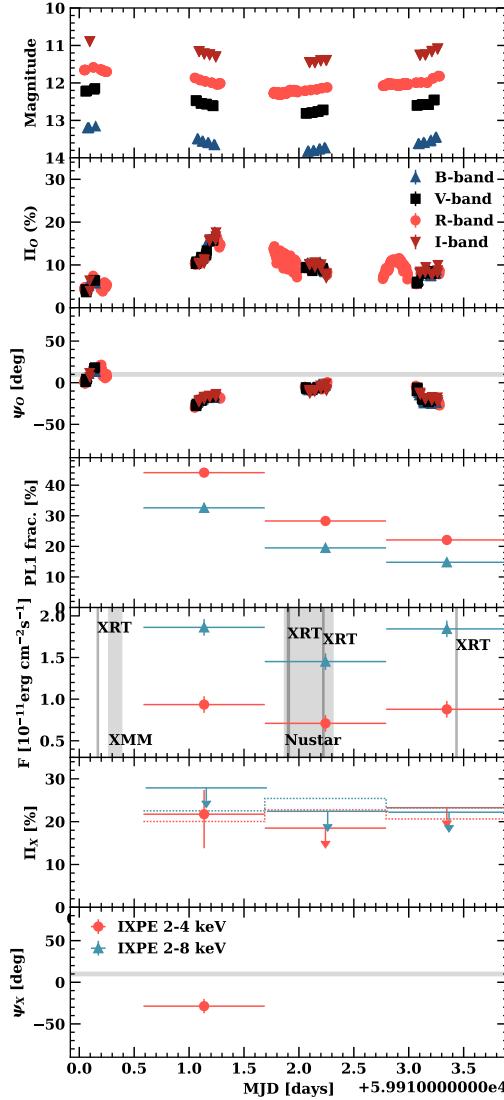


Figure 7.3: Light curves, polarization fractions, and EVPAs in the optical (top three panels) and X-ray (bottom three panels) bands. X-ray soft power flux contribution fraction is shown in the middle panel. In the X-ray panels, the data are split into two energy bins and three equal time bins. Times of observations from satellites other than IXPE are indicated in the flux panel. Errors are 68.3% (1σ) confidence intervals and upper limits are 99% confidence. Shaded horizontal regions in EVPA panels represent the jet axis projection on the plane of the sky [20]. Dotted lines in the X-ray polarization-fraction panel represent the MDP99 level for each bin individually. For X-ray polarization measurements, only the first time bin at 2–4 keV exceeds MDP99.

Chapter 8

Geometric Model for X-ray Polarization Rotation in Mrk 421

8.1 Introduction

Blazars are known to exhibit steady rotations in their electric vector position angle (EVPA), sometimes exceeding many factors of π . While these rotations have been typically observed at optical frequencies and below, the Imaging X-ray Polarimetry Explorer (IXPE) has recently detected the first blazar polarization rotation in the soft X-ray band (1–10 keV), originating in BL Lac blazar Mrk 421. We present a basic geometrical model for the observed rotation. By fitting directly to the data, we can reproduce the observed stepping behavior and recover the jet axis projection on the plane of the sky. Treating the Mrk 421 rotation as two separate events, we find a steady rotation with an underlying constant polarization aligned along the jet axis explains the first rotation. This is consistent with the expected predictions of a spine-sheath jet model. To increase the significance of our result, we re-analyze the IXPE observations using state-of-the-art polarization reconstruction methods.

Blazars have long been known to be substantially polarized. Much recent effort has been dedicated to studying the optical polarization properties of blazars. Polarization fractions (Π) and EVPAs tend to be mostly stochastic, with some preference for EVPAs to align along the jet axis and for the average polarization fraction to increase as a function of frequency [96]. Alongside this typical stochastic behavior, some blazars have been found to display moderately steady EVPA rotations [65], sometimes lasting weeks or months and $> 360^\circ$. Some of these rotations are associated with flares in total intensity, γ -ray flares, and drops in Π [16]. Notably, low synchrotron peaked (LSP) blazar sources tend to dominate observed rotations [159].

While many shorter length rotations (90° – 180°) are consistent with stochastic variability models

[17, 123] or magnetic reconnection [138], longer rotations require a deterministic explanation. Proposed models include a spiraling jet [67, 90], a helical kink propagating along a conical jet [68], an oscillating Doppler factor [145], and a shock moving through a cylindrical jet with a helical magnetic field [160]. Although all of these explanations are plausible, few, if any, have been fit directly to observed rotations.

The connecting theme between deterministic explanations for long blazar rotations is a helical geometry with relativistic boosting. There is significant evidence for helical magnetic field structure in jets at the VLBI scale [161] and strong motivation from postulated jet launching mechanisms [62]. The rotational effect of boosting on observed EVPAs has been studied carefully [86] and its resulting implications for helical geometries, like variations in rotation rate, described by, e.g., Lyutikov, Pariev, and Gabuzda [162] and Peirson and Romani [1].

The advent of X-ray polarimetry with the recent launch of IXPE [125] has allowed us to probe closer to blazar acceleration regions than ever before. Between June 4-6 and 7-9 2022, IXPE observed the high synchrotron-peaked blazar Mrk 421, discovering a relatively steady EVPA rotation of $\sim 360^\circ$ (Di Gesu et al., submitted). owing to the fact that no rotation was observed simultaneously at lower frequencies, the authors attribute the observed X-ray rotation to emission from a fast moving jet spine with a helical magnetic field [163], with the lower frequency emission produced in a slower moving sheath without a strongly imposed magnetic field geometry. However, the authors do not attempt to explain the variations in rotation rate during the two observations.

Here, we distill the essential mechanisms of deterministic rotations into a geometric toy model and fit directly to the observed Mrk 421 rotation. Based on the fit, our model predicts the jet axis projection on the plane of the sky, which can be directly compared with independent VLBI observations. We compare our hypothesised model with a steady rotation and a steady rotation with separate constant polarization component. The novelty in our approach is that, to our knowledge for the first time, we directly fit a very general deterministic rotation model to observations and verify a physical observable: the jet axis projection. To more carefully evaluate the evidence for competing rotation models, we enhance the IXPE analysis with a simulation based-inference approach [3, 7, 4].

In §8.2 we describe the geometric rotation models; in §8.3 we evaluate the competing rotation models on the two IXPE Mrk 421 observations; in §8.4 we discuss the implications of our results and in §8.5 we provide our conclusions.

8.2 Rotation Models

Every proposed deterministic blazar long rotation mechanism appeals to either magnetic field geometry, relativistic boosting, or both. Explanations fall into two categories: a fixed magnetic field with a varying Doppler factor, or a varying magnetic field with a fixed Doppler factor. In fact, in terms of observed EVPA, these two approaches are equivalent. Both rely on a changing effective

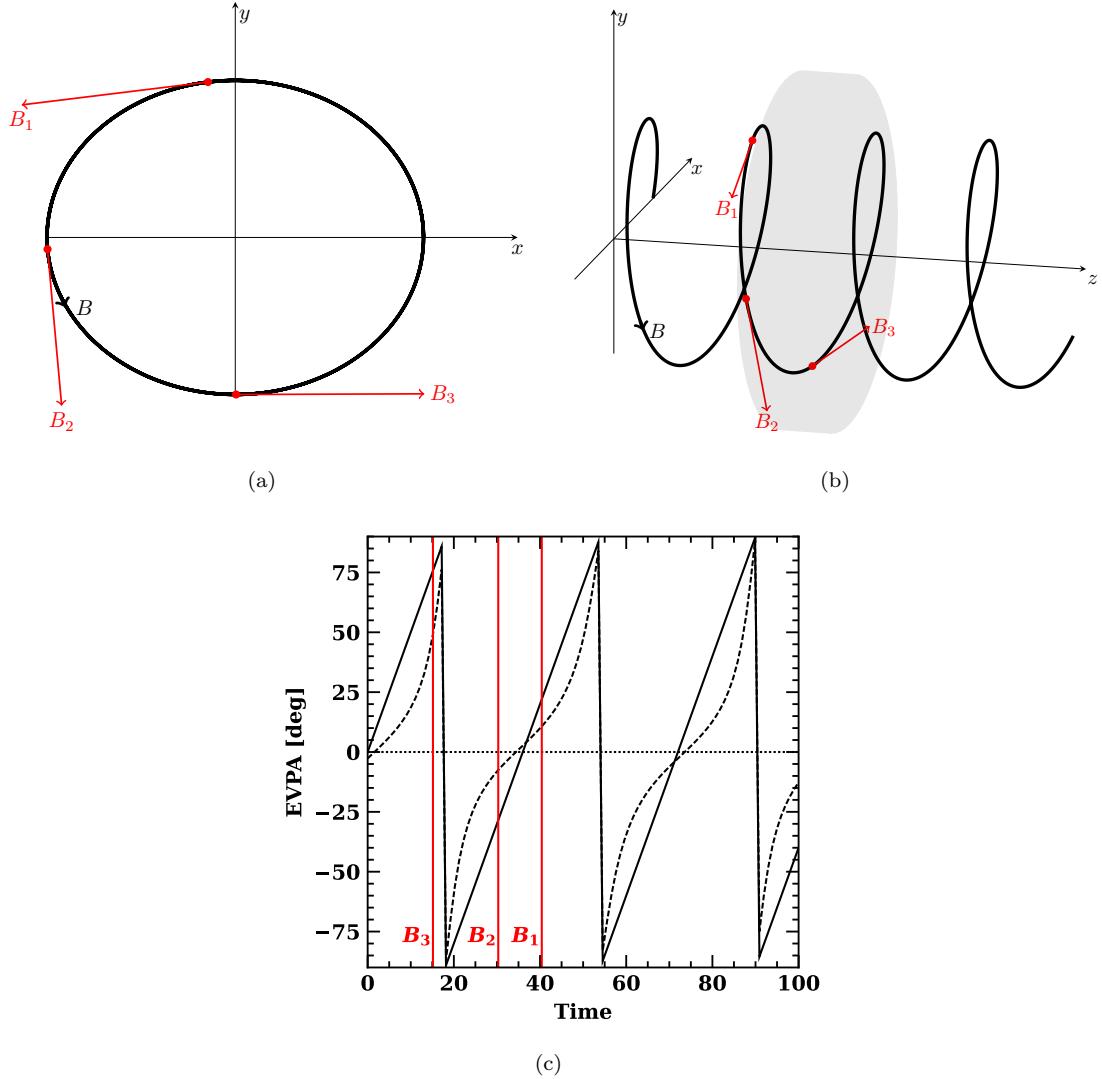


Figure 8.1: (a) A helical magnetic field moving through a stationary acceleration/emission region. The helix advances along $+z$ with B_3 entering the active region first. (b) The same helix as in (a) viewed from 70° . (c) Observed EVPA rotation pattern from synchrotron emission in the stationary acceleration zone as viewed from (a) solid, (b) dashed. Labelled red lines denote specific phase points highlighted in (a) and (b).

viewing direction of the magnetic field vector where the polarized synchrotron emission originates.

In the jet frame, a power law distribution of leptons (power law exponent α) with isotropic distribution of pitch angles in a uniform magnetic field B' produces linearly polarized synchrotron

emission with polarization fraction:

$$\frac{\alpha + 1}{\alpha + 7/3} \quad (8.1)$$

and EVPA perpendicular to the projection of B in the plane perpendicular to the line of sight [94]. For a jet with bulk Lorentz factor Γ and velocity β , when viewed in the laboratory frame at an angle θ from the jet axis, the polarization can be calculated as if emitted from an angle θ' in the jet frame, where

$$\cos \theta' = \frac{\cos \theta - \beta}{1 - \beta \cos \theta}. \quad (8.2)$$

This is equivalent to a rotation of the magnetic field vector in the plane formed by the line of sight and the jet axis [2, 162]. For highly relativistic jets, even small changes in θ can cause large changes in the observed EVPA. Furthermore, a constant EVPA rotation in the jet frame produced by a helical magnetic field viewed on axis in the jet frame can become highly variable in rotation rate in the laboratory frame. See fig. 8.1 for an example. Importantly, a natural axis of symmetry for this effect is the jet axis, hence any rotation fitting with this class of model will constrain it.

8.2.1 Helical field in a relativistic jet

Since all models based on helical fields and boosting are essentially equivalent, we propose the simplest model variant: a helical field moving through a stationary emission region in a relativistic cylindrical jet, fig. 8.1. This model is defined by six parameters:

1. ω [deg/day] (the rotation rate or forward phase velocity of the helix through the transition region. In this work $-100^\circ < \omega < 100^\circ$.)
2. $\Gamma\theta$ (a dimensionless quantity that represents the effective viewing angle. For blazars $0 \leq \Gamma\theta \lesssim 1$.)
3. ϕ_0 [deg] (phase offset, helix phase at $t = 0$.)
4. ψ [deg] (EVPA at jet axis alignment, i.e., jet axis angle on the plane of the sky.)
5. Π [%] (polarization fraction, constant throughout the rotation.)
6. Pitch [deg] (the pitch angle of the helix. In this work we assume the helix is fully toroidal, so pitch is fixed at 90° .)

For a constant polarization fraction and EVPA, inferring the polarization of a source given a set of observations involves 2 degrees of freedom: the polarization fraction Π and EVPA ψ . Assuming a steadily rotating EVPA with constant polarization fraction adds one additional degree of freedom, the rotation rate ω . A constant rotation rate model has no sensitivity to the jet axis and no need for phase offset (the role of phase offset is captured by the EVPA). So the boosted helix with fixed pitch angle adds two additional degrees of freedom compared to a constant rotation, $\Gamma\theta$ and ϕ_0 .

We fix the helix pitch angle since the Mrk 421 data aren't sufficient to constrain it. As mentioned in §8.1, blazars tend to have EVPAs aligned with jet axis suggesting a highly toroidal magnetic field mostly perpendicular to the jet axis [162] – therefore we fix the pitch angle to 90° .

Our simple helix model does not predict variation in the polarization fraction during the rotation; an extension to a conical jet would do so. However, it does predict variation in the observed flux density during the rotation, since for synchrotron emission in a uniform magnetic field the observed flux density depends on the angle χ' between the jet frame line of sight and the jet frame magnetic field as $(\sin \chi')^{(\alpha+1)/2}$.

8.2.2 Rotation with underlying constant polarization

We test a third competing hypothesis for the Mrk 421 rotation: a steady rotation on top of a fixed constant polarization. This toy model is plausible in, for example, a spine-sheath scenario where emission comes from a relatively constant but low polarization sheath and a rapidly rotating spine. More generally, since blazar polarization fractions are rarely close to the synchrotron limit, eq. (8.1), observed emission is likely averaged over mostly disordered magnetic fields. If a rotation occurs over such a background, and the background remains relatively constant during the rotation, then such a model could arise.

This third hypothesis adds three extra degrees of freedom on top of a constant rotation, compared to the two for the boosted helix. These are:

1. Π_0 (the fixed constant polarization fraction.)
2. ψ_0 (the fixed constant EVPA.)
3. Δ (the fraction of constant flux to rotating flux, $0 \leq \Delta \leq 1$.)

Such a model can produce similar apparent rotation rate variations as the boosted helix, fig. 8.1(c), but instead of being tied to the jet axis it can select its own axis of symmetry by varying ψ_0 . Importantly, and unlike the boosted helix, this model also predicts sinusoidal polarization fraction variation. When the rotating component EVPA aligns with ψ_0 , the observed polarization fraction increases.

8.3 Data Analysis

8.3.1 Polarization reconstruction

IXPE observations result in level 1 event files made up of measured photoelectron track images produced by source X-ray photons. The initial photoelectron directions, inferred from the track images, follow a known distribution that depends on the source polarization parameters. By assembling

the appropriate likelihood function, one can infer source polarization model parameters. See Baldini et al. [10] and Baldini et al. [153] for details on IXPE telescope details and official observation reduction; for polarization model inference see [8].

We employ three separate methods to go from level 1 IXPE event files to likelihood functions for model fitting. The first is the standard moment analysis, currently the default method for the IXPE collaboration. This method reconstructs photoelectron track directions using a moment analysis [10] and uses the standard polarization negative log-likelihood function [38]

$$l(q, u) \propto -\sum_{i=1}^N \log \left(1 + \mu \begin{bmatrix} q \\ u \end{bmatrix}^T \begin{bmatrix} \cos 2\theta_i \\ \sin 2\theta_i \end{bmatrix} \right). \quad (8.3)$$

Here, N is the number of events, θ_i are the reconstructed photoelectron track directions, q and u are the normalized stokes parameters to be inferred, and μ is the modulation factor, a spectrum dependent calibration term. Note that

$$\Pi = \sqrt{q^2 + u^2},$$

$$\psi = \frac{1}{2} \arctan \left(\frac{u}{q} \right).$$

The second method, a weighted moment analysis, employs the same track reconstruction as the standard moment analysis, but weights individual events by their normalized track ellipticities [164]. Such weights attempt to reduce the contribution in the likelihood of poorly reconstructed photoelectron tracks. The likelihood function becomes

$$l(q, u) \propto -\sum_{i=1}^N \log \left(1 + \bar{\mu} \begin{bmatrix} q \\ u \end{bmatrix}^T \begin{bmatrix} w_i \cos 2\theta_i \\ w_i \sin 2\theta_i \end{bmatrix} \right). \quad (8.4)$$

The modulation factor must be re-calibrated for the new method, $\bar{\mu}$

The third method is a machine learning based analysis [3, 4, 7] that reconstructs photoelectron track directions and estimates event weights using neural networks (NNs) trained on a simulated IXPE detector. The likelihood function used has the same form as 8.4, but with differing $w_i, \theta_i, \bar{\mu}$.

Table 8.1 gives the minimum detectable polarizations at 99% confidence (MDP99) for each of the three polarization reconstruction methods across different energy bands for the joint Mrk 421 observations. Weighted NNs is the most sensitive and so should be able to best distinguish the competing model hypotheses. We note that Mrk 421's soft spectrum during the two observations, with photon indices $\alpha = 3.17$ and $\alpha = 2.81$ respectively, could make weighted NN improvements smaller than described in Peirson and Romani [4].

Table 8.1: MDP99 values for different polarization reconstruction methods for the joint Mrk 421 observations.

Method			
Energy [keV]	Mom. [%]	Weighted Mom. [%]	Weighted NNs [%]
2–8	2.55	2.24	2.15
1–10	2.88	2.45	2.09
1–4	3.15	2.78	2.35
2–4	2.81	2.49	2.43
4–10	5.57	5.08	4.48

8.3.2 Model fitting

To fit the null hypothesis, a constant rotation model, we adjust the likelihood functions eq. (8.4), eq. (8.3) as follows,

$$l(q, u, \omega) \propto - \sum_{i=1}^N \log \left(1 + \bar{\mu} \begin{bmatrix} q \\ u \end{bmatrix}^T R_i^T \begin{bmatrix} w_i \cos 2\theta_i \\ w_i \sin 2\theta_i \end{bmatrix} \right), \quad (8.5)$$

where R_i is the rotation matrix

$$R_i(\omega, \phi_0) = \begin{bmatrix} \cos[2(\omega t_i + \phi_0)] & -\sin[2(\omega t_i + \phi_0)] \\ \sin[2(\omega t_i + \phi_0)] & \cos[2(\omega t_i + \phi_0)] \end{bmatrix}. \quad (8.6)$$

For a constant rotation model, ϕ_0 can be set arbitrarily and is not a free parameter of the fit. In the case of the standard moment analysis, eq. (8.3), weights are disregarded. To properly model polarized background emission, we augment the likelihood function to include a polarized background component

$$\begin{aligned} l(q, u, \omega, q_{\text{bkg}}, u_{\text{bkg}}) &\propto \\ &- \sum_{i=1}^N \log \left(1 + \bar{\mu} \left(f \begin{bmatrix} q \\ u \end{bmatrix}^T R_i^T + (1-f) \begin{bmatrix} q_{\text{bkg}} \\ u_{\text{bkg}} \end{bmatrix}^T \right) \begin{bmatrix} w_i \cos 2\theta_i \\ w_i \sin 2\theta_i \end{bmatrix} \right) \\ &- \frac{A_{\text{src}}}{A_{\text{bkg}}} \sum_{j=1}^M \log \left(1 + \bar{\mu}_{\text{bkg}} \begin{bmatrix} q_{\text{bkg}} \\ u_{\text{bkg}} \end{bmatrix}^T \begin{bmatrix} w_j \cos 2\theta_j \\ w_j \sin 2\theta_j \end{bmatrix} \right), \end{aligned} \quad (8.7)$$

where there are M background events labelled by index j and N source events labelled by i . The ratio $A_{\text{src}}/A_{\text{bkg}}$ is the area scaling ratio between source and background exposures, and the source

Table 8.2: Natural log marginal likelihood differences from the constant rotation for weighted NNs. RC: constant rotation with constant polarization component, BH: boosted helix. More strongly preferred models have a higher log marginal likelihood difference. Negative log marginal likelihood means the constant rotation is preferred.

Mrk 421 Observation				
E [keV]	Model	Both	First	Second
2–8	RC	-2.1 ± 0.2	8.4 ± 0.3	2.5 ± 0.4
	BH	3.4 ± 0.2	0.8 ± 0.3	3.9 ± 0.4
1–4	RC	-1.7 ± 0.1	9.3 ± 0.1	-0.0 ± 0.4
	BH	3.1 ± 0.3	1.1 ± 0.2	0.9 ± 0.5
4–10	RC	-1.4 ± 0.2	-1.0 ± 0.2	1.8 ± 0.5
	BH	0.1 ± 0.3	-0.5 ± 0.2	1.7 ± 0.4
1–10	RC	-1.7 ± 0.2	10.3 ± 0.2	4.7 ± 0.3
	BH	4.2 ± 0.2	2.1 ± 0.2	5.8 ± 0.1
2–4	RC	-2.0 ± 0.3	7.6 ± 0.3	-0.1 ± 0.1
	BH	2.4 ± 0.3	0.2 ± 0.3	0.8 ± 0.1

fraction f :

$$f = \frac{(N - \frac{A_{\text{src}}}{A_{\text{bkg}}} M)}{N}. \quad (8.8)$$

Extending the constant rotation likelihood functions eq. (8.5), eq. (8.7) to accommodate a boosted helix requires only changing the rotation matrix R_i , adding a dependence on $\Gamma\theta$: $R_i(\omega, \phi_0, \Gamma\theta)$. We calculate this rotation matrix analytically based on the relationships in §8.2. The steady rotation on top of a fixed constant component needs an additional adjustment to eq. (8.5), eq. (8.7)

$$\begin{aligned} l(q, u, \omega, q_{\text{bkg}}, u_{\text{bkg}}, q_0, u_0, \Delta) &\propto \\ &- \sum_{i=1}^N \log \left(1 + \bar{\mu} \left((1 - \Delta)f \begin{bmatrix} q \\ u \end{bmatrix}^T R_i^T + \Delta \begin{bmatrix} q_0 \\ u_0 \end{bmatrix}^T + (1 - f) \begin{bmatrix} q_{\text{bkg}} \\ u_{\text{bkg}} \end{bmatrix}^T \right) \begin{bmatrix} w_i \cos 2\theta_i \\ w_i \sin 2\theta_i \end{bmatrix} \right) \\ &- \frac{A_{\text{src}}}{A_{\text{bkg}}} \sum_{j=1}^M \log \left(1 + \bar{\mu}_{\text{bkg}} \begin{bmatrix} q_{\text{bkg}} \\ u_{\text{bkg}} \end{bmatrix}^T \begin{bmatrix} w_j \cos 2\theta_j \\ w_j \sin 2\theta_j \end{bmatrix} \right). \end{aligned} \quad (8.9)$$

Note that flux density variation during the rotation is not included in the likelihood function for any of the models.

With the likelihood function described above for each model and reconstruction method, we evaluate the suitability of a model using a Bayesian approach: calculating the marginal likelihood (a.k.a. model evidence) of the Mrk 421 observation data. The marginal likelihood measures the

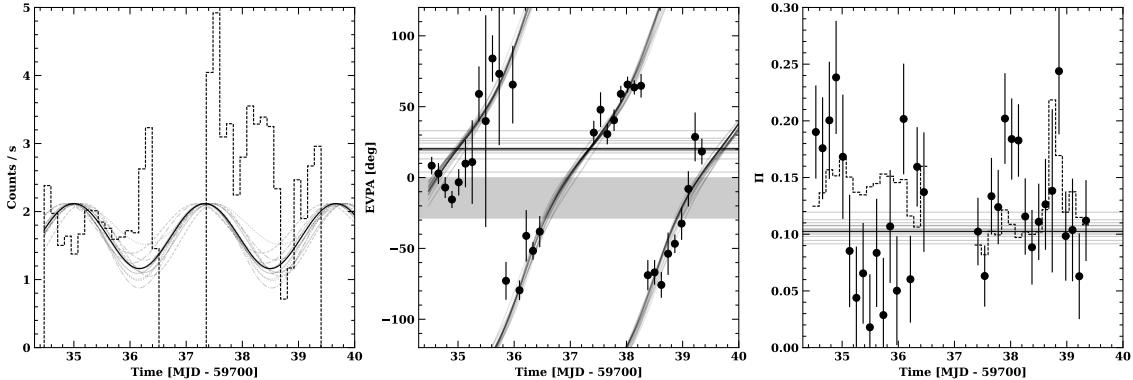


Figure 8.2: Full Mrk 421 rotation fit for the highest marginal likelihood model, the boosted helix. Throughout, model traces are multiple samples from the posterior distribution and points are binned observation derived with weighted NNs. Left panel gives observation binned count spectra with model flux variation overlayed. Middle panel shows binned observed EVPAs with rotation model overlayed. Grey shaded region gives Mrk 421 43GHz jet axis direction [20] and horizontal lines are the fitted model jet axis predictions. Right panel is the binned observed polarization fraction, with a step plot denoting the MDP99s for each bin. Horizontal lines are the model polarization fraction prediction.

probability of the observed data given the model hypothesis, including priors over the model parameters. The ratio of marginal likelihoods for two different models is known as the Bayes factor and signifies how much better one model explains the data compared to the other. We evaluate marginal likelihoods for each model using nested sampling, specifically the *MultiNest* algorithm [165, 166]. In this way, we also get posteriors over the model parameters. Priors are set as the natural, uninformative priors in polarization estimation, meaning uniform priors over polarization fractions and EVPAs $\Pi, \psi, \Pi_0, \psi_0, \Pi_{\text{bkg}}, \psi_{\text{bkg}}$ [38], and uniform priors for the rest covering the standard physical ranges in §8.2.

To ensure consistency of our results, we perform fits over multiple energy ranges and over the two Mrk 421 observations, jointly and individually. By treating the two Mrk 421 observations individually we test whether the preferred rotation model changes with time. Table 8.2 displays the natural log marginal likelihood difference from a constant rotation for each model described in §8.2. The respective tables for the standard moment analysis and weighted moments are attached at the end of the chapter, table 8.4 and table 8.3. The weighted NNs allow better significantly better distinction between models in some cases, especially for the individual rotation fits. In all cases, the 4-10keV energy range has difficulty preferencing any model on account of its relatively low signal-to-noise ratio for these observations, table 8.1 – for this reason we disregard it from the remainder of the discussion. Figure 8.2, fig. 8.3, and fig. 8.4 plot the preferred model for each of the three Mrk 421 datasets respectively.

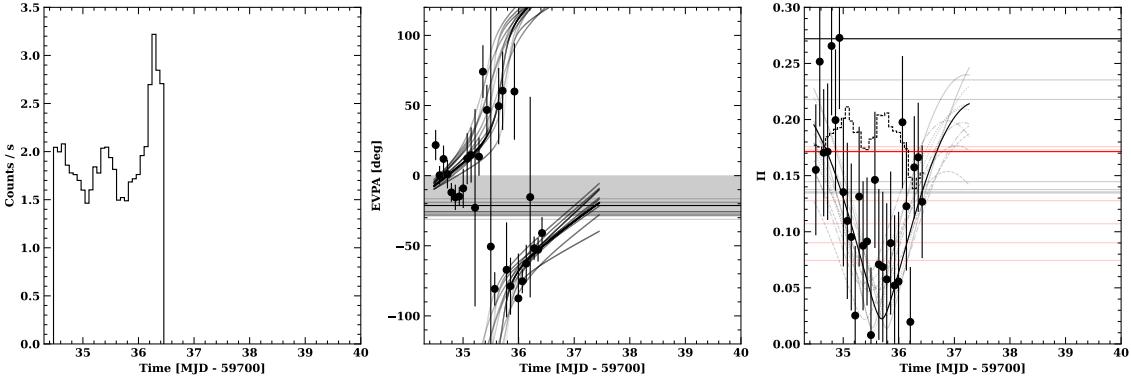


Figure 8.3: Same as fig. 8.2 but for the highest marginal likelihood model on the first Mrk 421 observation only. Here, a rotation with added constant polarization component is strongly preferred. This model predicts a variation in the polarization fraction shown in the right hand panel. Red horizontal lines in the right hand panel are posterior samples of the constant polarization fraction, black lines are the polarization fraction of the rotation component.

8.4 Discussion

When considering the two Mrk 421 observations as a single continuous rotation, a boosted helix model is strongly preferred over a constant rotation and a steady rotation with underlying constant component for all energy ranges, table 8.2. The full posterior distribution for the boosted helix fit is given in fig. 8.5. Bayes factors greater than e^3 are considered moderate to strong evidence. The apparent rotation rate variation predicted by this boosted helix is clear in fig. 8.2. The predicted jet axis projection is $\psi = 19^\circ \pm 8^\circ$, which is relatively close to the 43GHz jet component average over 2008-2018 VLBA-BU-BLAZAR monitoring program: $-14^\circ \pm 14^\circ$ [20]. A boosted helix also predicts sinusoidal flux density variation during the rotation; however, such a simple model does not capture all the exogeneous factors (e.g. acceleration efficiency) that can contribute to flux density so we do not expect it to be predictive.

The polarization fraction during the first observation is quite variable, making the simple assumption of constant polarization in the boosted helix inadequate. While a steady rotation with a constant component (RC) predicts sinusoidal polarization fraction variation, the model is clearly not suitable for the full rotation. On the individual observations, the RC model performs extremely well on the first and strongly on the second, reproducing the sinusoidal polarization fraction variations as well as the apparent EVPA rotation rate variation in both cases. Its poor suitability for the joint rotation is because the inferred constant polarization component EVPAs, ψ_0 , are close to orthogonal between the different observations. For the first observation $\psi_0 = -24^\circ \pm 4^\circ$ and for the second $\psi_0 = 66^\circ \pm 7^\circ$. Notably, for the first observation, ψ_0 is aligned almost exactly with the 43GHz average jet axis. Such a rapid transition from aligned to anti-aligned EVPA with the jet axis is possible with a changing Doppler factor to a pitched toroidal field [162] but is out of scope for our

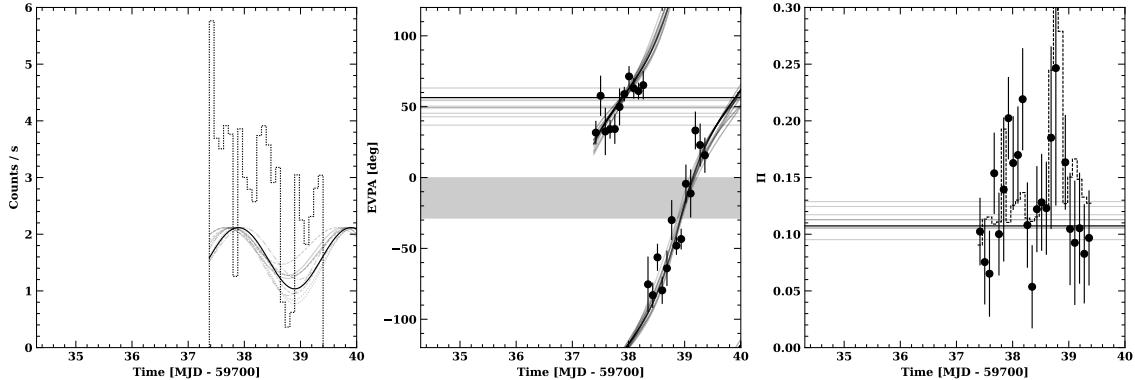


Figure 8.4: Same as fig. 8.2 but for the highest marginal likelihood model on the second Mrk 421 observation only, a boosted helix.

current analysis.

Since short rotations $< 180^\circ$ have numerous deterministic and non-deterministic explanations, we consider only the full rotation. Here a boosted helix is strongly preferred to both alternative models, and the jet axis aligns to within 35° of the 43GHz jet axis. X-rays probe much closer to the core region than radio, thus the core jet axis could differ substantially from the measured 43GHz range. Despite our toy model’s simplicity, it captures rotation rate variations adequately and connects back to a geometric observable: the jet axis.

8.5 Conclusion

Thanks to a weighted NN analysis, we are able to establish a variable rate model based on a helical geometry better fits the full Mrk 421 X-ray polarization rotation, as measured via the Bayesian marginal likelihood. Our toy boosted helix model subsumes typical geometric features in deterministic rotation modelling and allows us to infer the jet axis directly from observation. While our simple model does not explain variations in flux density or polarization fraction, we consider this an initial step in connecting geometric models of blazars to observations. Further observations of rotations with IXPE and the optical rotation data collected by the RoboPol monitoring campaign will allow us to test these geometric model predictions over a large sample of rotating blazars.

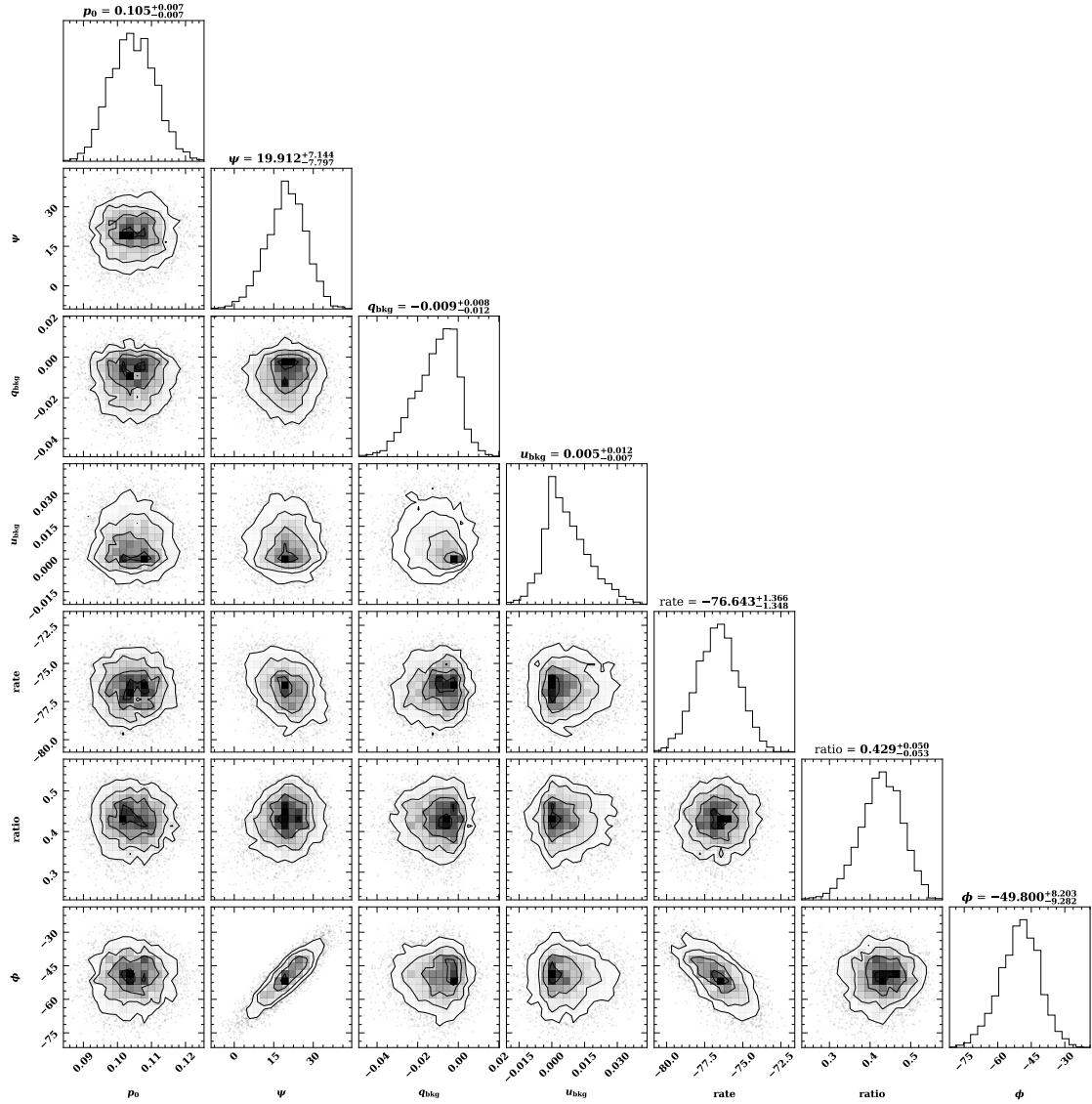


Figure 8.5: Posterior distribution of the preferred boosted helix model for the joint Mrk 421 dataset.

Table 8.3: Log marginal likelihood differences from the constant rotation for weighted moment analysis. RC: constant rotation with constant polarization component, BH: boosted helix. More strongly preferred models have a higher log marginal likelihood difference. Negative log marginal likelihood means the constant rotation is preferred.

Mrk 421 Observation				
E [keV]	Model	Both	First	Second
2–8	RC	-2.3 ± 0.1	6.8 ± 0.4	2.2 ± 0.1
	BH	3.6 ± 0.1	-0.1 ± 0.2	4.2 ± 0.4
1–4	RC	-2.2 ± 0.2	5.7 ± 0.6	-0.5 ± 0.4
	BH	2.3 ± 0.1	2.2 ± 0.3	0.0 ± 0.4
4–10	RC	-1.5 ± 0.2	-0.9 ± 0.5	1.4 ± 0.4
	BH	-0.3 ± 0.2	-0.7 ± 0.3	1.9 ± 0.2
1–10	RC	-2.3 ± 0.1	3.6 ± 0.3	1.7 ± 0.4
	BH	4.1 ± 0.1	1.1 ± 0.3	3.5 ± 0.4
2–4	RC	-2.1 ± 0.1	7.2 ± 0.4	0.4 ± 0.2
	BH	2.5 ± 0.2	0.4 ± 0.1	0.1 ± 0.2

Table 8.4: Log marginal likelihood differences from the constant rotation for the standard moment analysis. RC: constant rotation with constant polarization component, BH: boosted helix. More strongly preferred models have a higher log marginal likelihood difference. Negative log marginal likelihood means the constant rotation is preferred.

Mrk 421 Observation				
E [keV]	Model	Both	First	Second
2–8	RC	-2.1 ± 0.3	3.4 ± 0.3	0.1 ± 0.2
	BH	1.6 ± 0.3	0.0 ± 0.2	2.2 ± 0.3
1–4	RC	-1.3 ± 0.2	3.4 ± 0.7	-1.3 ± 0.3
	BH	2.7 ± 0.1	3.1 ± 0.9	-0.2 ± 0.2
4–10	RC	-1.4 ± 0.3	-1.1 ± 0.2	1.0 ± 0.1
	BH	-0.2 ± 0.2	-0.8 ± 0.2	1.5 ± 0.1
1–10	RC	-2.2 ± 0.5	1.4 ± 0.1	-1.5 ± 0.3
	BH	3.0 ± 0.5	1.5 ± 0.3	0.3 ± 0.3
2–4	RC	-2.5 ± 0.2	3.4 ± 0.6	-1.5 ± 0.2
	BH	1.2 ± 0.3	-0.2 ± 0.1	0.3 ± 0.2

Chapter 9

Conclusions

In this thesis, I have advanced the nascent field of X-ray polarimetry and brought it to bear on longstanding questions about the nature of blazar jets.

In chapters 2 and 3, I develop a simulation-based inference framework for reconstructing polarization from measured photoelectron tracks using simulated IXPE detector data. By using recent ideas for uncertainty quantification and computer vision in machine learning, I improved the reconstruction fidelity of photoelectron tracks and quantified their direction prediction error. Combining these improvements with an appropriate statistical framework, I showed a significant increase in confidence in the estimation of polarization parameters as a function of energy, time, and position on the detector. In the same framework, I also made improvements to IXPE’s energy and spatial resolutions. Future developments in X-ray polarization reconstruction would do well to take advantage of this framework’s resolution gains to reduce GPD position-based systematics like polarization leakage, and energy-based systematics like events converting outside of the detector volume.

In chapters 4 and 5, I construct a conical leptonic jet model to study the geometric effects inherent in blazars and their effects on polarized emission. I establish and emphasize some fundamental predictions, including a rising polarization fraction along the synchrotron cutoff, EVPA rotation with relativistic boosting causing rotation rate variation for helical magnetic fields, and the relationship between synchrotron and synchrotron self-Compton polarization in a jet with disordered magnetic field. My leptonic jet model makes fundamental predictions that, along with accompanying literature, form the guiding principles for recent and upcoming X-ray polarization measurements of blazar jets.

In chapter 6, I use the jet model developed in the previous two chapters to discover the information gain possible with X-ray polarization observations of blazars. First, I predict the expected X-ray polarization and detectability of various blazar source candidates, helping select IXPE’s first year targets. Second, I study the potential for ISP blazars, whose soft X-ray band falls in the synchro-Compton transition, to resolve hadronic vs. leptonic emission processes. I find that future

X-ray polarimetry missions beyond IXPE will likely be required to measure hadronic polarization in a hybrid emission model setting.

In chapter 7, I use IXPE to measure BL Lacertae during a recent short outburst, which fortunately brought its synchro-Compton transition region into IXPE’s sensitive band. Breaking up the observation in time, polarization is detected when the synchrotron component is at peak fractional flux contribution. While I couldn’t make strong conclusions on the nature of the high energy emission hump, the detection result indicates a strongly rising synchrotron polarization fraction along the cutoff, consistent with the predictions in chapters 4 and 5.

In chapter 8, I apply a geometric jet model to IXPE observations of the first X-ray polarization rotation observed in a blazar (Mrk 421). Using the sensitivity improvements from chapters 2 and 3, I find a boosted helix rotation model is strongly preferred over a constant rotation, predicting a jet axis verifiably close to the 43GHz measured value. The chapter adds additional evidence to the literature for helical fields as the deterministic process behind blazar polarization rotations and serves as a rallying cry for direct fitting of these unique events.

Throughout the thesis, I connect recent ideas from machine learning and statistics to standard astrophysical data analysis. While there has been an explosion of interdisciplinary work uniting astrophysics and computer science, rarely has there been such a neat and timely application fit as for X-ray polarimetry. This thesis opens the door to large improvements for all kinds of particle tracking detectors, especially those in astrophysical polarimetry. Thanks to the great success of IXPE, a number of future X-ray and γ -ray polarimetry missions have been proposed [25, 167]; the future looks bright for high-energy polarimetry.

Now that the IXPE collaboration and the perspectives in this thesis have shown what is possible, and tantalizingly what could be possible, with high-energy polarization measurements of blazars, I hope that future investigation will finally unveil the fundamental physics behind blazar jets.

Appendix A

Computational Jet Model

The code used for the work used in chapters 4, 5, and 6 is made publicly available at <https://github.com/alpv95/SSCpol>. Further documentation on how to compile and run can be found there.

The main code consists of a C script, *jet_model.c*, that initializes a single jet slice and follows the evolution of the electron population and the emitted photon spectrum, accumulating the observed Stokes' fluxes for both synchrotron and SSC emission. More detail on the synchrotron emission slice and its application to blazar rotations can be found in Peirson and Romani [1]. The bulk of the CPU time required is, however, spent calculating SSC emission, evaluating the integrals in Eqs. (1) and (2) $N^2 \times$ for each slice step along the jet length. This computation is accelerated using OpenMP. For example, 16 CPU cores runs the N=37 zone model (one random B-field draw, as in fig. 5.12) in ~ 250 minutes.

With the assumptions of §5.5, each slice can be evolved independently (the exception is high observation angle mm); we do not consider synchrotron seed photons from adjacent slices. In this chapter we do not include self-absorption effects which are typically significant in the longer wavelength radio emission; this has essentially no effect on the SSC X-ray fluxes. The jet is assumed to be optically thin at all times.

The algorithm begins by initializing the free jet parameters: total jet power W_j , bulk Lorentz factor Γ_{bulk} , electron exponential energy cutoff γ_{max} , observation angle θ_{obs} , jet opening angle (jet frame) θ_{open} , electron power law index α , initial magnetic flux density B_0 , minimum electron energy γ_{min} , number of B-field zones in jet slice N , the length of the jet L_{jet} , and the number of electron energy and emitted frequency bins desired. From these, the initial jet radius R_0 and electron population $\frac{dN_e}{dE_e}$ discretized in energy bins can be derived following Potter and Cotter [106]. The cross-section is split up into N circular zones, with their position and mutual displacement vectors calculated. Each zone is initialized with a B-field vector direction sampled from an isotropic distribution.

A loop over the jet length $x < L_{jet}$ begins the main calculation. The emitted synchrotron powers per unit length $P_{\perp}^i(\nu)$, $P_{\parallel}^i(\nu)$ for each zone i are calculated assuming an isotropic pitch angle distribution, following Rybicki and Lightman [93]. The synchrotron photon energy density in each zone i contributed by zone j , $\rho_{\perp}^{ij}(\nu)$, $\rho_{\parallel}^{ij}(\nu)$, are calculated using (16). This requires keeping track of emitted synchrotron power for all prior x and accounting for the RPAR between zones in the diverging jet. The integral is treated as a sum over all $x < x_{\text{current}}$. The SSC power per unit length can then be calculated by treating (1) and (2) as discretized sums, resulting in $P_{\perp}^{\text{SSC}}(\nu)$, $P_{\parallel}^{\text{SSC}}(\nu)$ for every zone. The electron energy losses due to emission are found, and the step length dx is set by the cooling time of the highest energy occupied electron bin, with the constraint that $R_{\text{new}} \leq 1.05R$. Then $\frac{dN_e}{dE_e}$, x , R and B are updated and the emitted power for each zone is converted to a Stokes' parameter representation and boosted, using D_i . The loop repeats until $x \geq L_{jet}$. Finally the Stokes' parameters are converted to lab frame quantities $\nu'F_{\nu'}$, $\Pi(\nu')$ and $\theta_{PA}(\nu')$ for SSC and synchrotron separately.

Appendix B

Observations of BL Lacertae

B.1 X-ray observations

Between 2022 November 27–30, BL Lacertae was observed quasi-simultaneously with *IXPE*, *NuSTAR*, and *XMM-Newton*, table B.1. For the *IXPE* data, the cleaned event files and the associated science products were obtained using a dedicated pipeline relying on the Ftools software package and adopting the latest calibration data files from *IXPE* (CALDB 20211118). The source regions for each of the three detector units (DUs) were then selected via an iterative process aimed at maximizing the signal-to-noise ratio (SNR) over the *IXPE* standard energy range of 2–8 keV. In particular, we used circular regions with radius 50 $''$ for all three DUs and annulus regions of size 100–300 $''$ for the background spectra in Stokes parameters I , Q , and U .

To improve the polarimetric sensitivity, we applied a background rejection strategy [168, 164]. The method refines the sensitivity by applying energy-dependent cuts on photo-electron tracks from the level 1 data. This method substantially decreases the background event rate, which are triggered by electrons, positrons, muons, or protons either directly impinging upon the detector or created by high energy interactions in the surrounding satellite structures. The cuts employ (1) the number of pixels: the size of the track region of interest (NUM_PIX); (2) energy fraction: the fraction of the event energy in the track (EVT_FRA); and (3) border pixels: the number of activated pixels along the boundary of the detector (NUM_TRK). Here we eliminated the events that do not satisfy the X-ray photon criteria of Di Marco et al. (2021, submitted). Using the point source as a monitor, we rejected $\sim 30\%$ of the diffuse background, with little impact on the source events. For *IXPE* spectro-polarization analysis, a constant energy binning of 4×0.05 keV PI channels per bin was used for Q and U ; we required 30 counts per bin for the intensity spectra.

The *XMM-Newton* spectra were produced with standard SAS routines and the latest calibration files. The source spectrum was extracted from a circular (radius= 40 $''$) aperture, while the

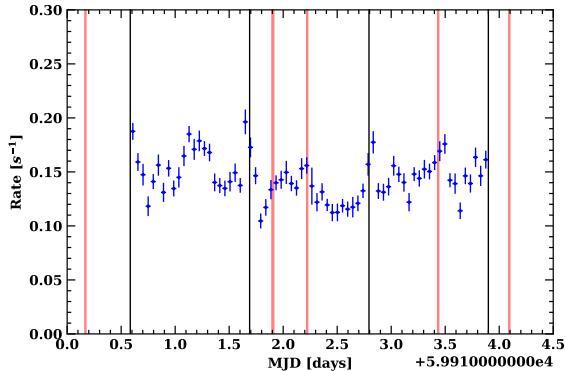


Figure B.1: *IXPE* light curve. Black lines are three equal time bin separators; red shaded regions indicate XRT exposures.

background spectrum was extracted from a blank region on the Epic-pn CCD camera from a circular region of the same size. The resulting spectrum was re-grouped to include at least 30 counts in each bin and to avoid $> 3\times$ oversampling of the spectral resolution. The *NuSTAR* data were calibrated and cleaned with the *NuSTAR* Data Analysis Software (NuSTARDAS7), and employed the nuproducts pipeline using the latest calibration database (v. 20220302). The source spectrum was extracted from a circular radius of $70''$ aperture; a surrounding $270''$ – $370''$ annulus provided the background.

Table B.1: Major quasi-simultaneous X-ray observations related to the 2022 November 27 *IXPE* pointing of BL Lac.

Observatory	Start Time	MJD range	ObsID	Exposure [ksec]
<i>NuSTAR</i>	2022-11-28 20:51:09	59911.87 – 59912.31	90801633002	38.1
<i>XMM-Newton</i>	2022-11-27 06:26:53	59910.27 – 59910.39	0902111801	10.1
<i>IXPE</i>	2022-11-27 14:01:15	59910.58 – 59913.90	02005901	286.4

Swift-XRT exposures were obtained in the context of a monitoring campaign tracking the BL Lac flux level before, during, and after the *IXPE* pointing. Scientific products from the *Swift*-XRT exposures were derived by using the facilities provided by the Space Science Data Center (SSDC8) of the Italian Space Agency (ASI). In particular, the source spectra were extracted from a source-centered $47''$ radius aperture, with a $120''$ – $150''$ concentric annulus providing a background spectrum. The events were grouped to include at least 25 counts in each spectral bin. We modeled each of the four XRT spectra as a simple power-law with Galactic photo-electric absorption. We report the 2–8 keV fluxes and the inferred photon indices in table B.2. The softest spectrum corresponds to the first *IXPE* time bin, fig. 7.3, for which the highest polarization is measured. As discussed in §7.3, softer spectra represent a higher synchrotron fraction.

Table B.2: *Swift*-XRT exposures in the vicinity of the *IXPE* pointing of BL Lac. Figure B.1 displays the exposures superimposed on the *IXPE* light curve. To measure flux density and Γ , single absorbed power-laws were fit to the full XRT 0.3–10 keV range.

Start time	MJD	ObsID	Exposure [s]	Flux density (2–8 keV) [10^{-11} erg cm $^{-2}$ s $^{-1}$]	Γ
2022-11-27T03:56:36	59910.16	00096990016	834	0.827	2.24
2022-11-28T21:30:17	59911.89	00089562001	1474	1.57	1.94
2022-11-29T05:13:35	59912.21	00096990017	826	2.04	2.13
2022-11-30T10:19:01	59913.42	00096990018	895	2.22	1.78
2022-12-01T02:07:36	59914.09	00096990019	849	1.97	1.87

B.2 Time-resolved spectro-polarimetric analysis.

B.2.1 XSPEC analysis

We use XSPEC to perform spectro-polarimetric analyses on the entire *IXPE* observation, including *NuSTAR* and *XMM-Newton* spectra, and each of the three *IXPE* time bins defined in the main text, fig. 7.3. Analysis of the entire observation is discussed in the main text (fig. 7.1, table 7.1), where we find that an absorbed sum of two power-laws with an apec component,

```
constant * tbAbs * (polconst * powerlaw + polconst * (apec + powerlaw))
```

is preferred according to XSPEC model fitting. Here we fix polarization degree of both the high-energy power-law (PL) and apec to zero. Figure B.2 displays a contour plot with the polarization measurement of the low-energy PL for the full *IXPE* observation fit, as displayed in fig. 7.1.

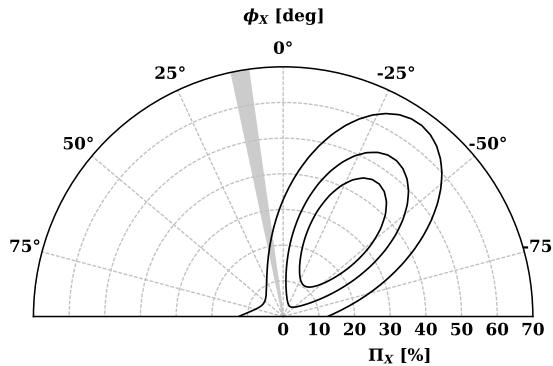


Figure B.2: Polarization fraction and EVPA confidence levels (68%, 95%, 99%) for the low-energy power-law in fig. 7.1. Gray shaded region represents the VLBI-determined jet axis projection on the plane of the sky [20].

In the case of the three *IXPE* time bins, we use the same model as above, but with photon indices and apec temperature (Γ_1, Γ_2, kT) fixed to the values found in the full observation fit, table 7.1.

Hence, only the PL normalization constants and low-energy PL polarization are determined by the fits. We fit this model to the time-binned *IXPE* spectra along with the appropriate XRT observation; see fig. 7.3. The second and third time bins share the same XRT observation, since it straddles the two bins. Table B.3 and fig. B.3 give the results of these fits. We are able to determine the relative flux density contribution of the low-energy PL in table 7.2, which correspond to the PL normalization constants at 1 keV.

Table B.3: Sum of two power-law spectro-polarimetric XSPEC model fits to time binned *IXPE* and *Swift*-XRT spectra, fig. 7.3. Photon indices Γ_1, Γ_2 , apec temperature kT , and N_H are fixed to their full-spectrum fit values; see table 7.1. High-energy PL polarization is fixed to zero. Π – polarization fraction, ψ – EVPA. Relative power-law normalizations in each time bin dictate the low-energy hump fractions in table 7.2.

Model Component	Time bin		
	1	2	3
χ^2/dof	393/386	461/382	406/392
PL1 norm	$(4.33 \pm 0.32) \times 10^{-3}$	$(2.28 \pm 0.23) \times 10^{-3}$	$(2.44 \pm 0.26) \times 10^{-3}$
PL2 norm	$(1.14 \pm 0.062) \times 10^{-3}$	$(1.00 \pm 0.49) \times 10^{-3}$	$(1.31 \pm 0.056) \times 10^{-3}$
$\Pi_1 [\%]$	32.0 ± 16.2	25.9 ± 25.9	53.9 ± 27.9
$\psi_1 [^\circ]$	-28.4 ± 14.3	-21.1 ± 28.8	-38.3 ± 14.4

B.3 Multiwavelength observations

During the *IXPE* observation, BL Lac was contemporaneously observed in polarization by different telescopes at millimeter and optical wavelengths. Millimeter-wave (mm) observations were performed by the Institut de Radioastronomie Millimétrique 30-m Telescope (IRAM-30m) on the 2022 November 28 (MJD 59911.5717) at 1.3 mm (228.93 GHz) and 3.5 mm (86.24 GHz), and by the Submillimeter Array (SMA) at 1.3 mm (225.538 GHz) on the 2022 December 1 (MJD 59914.0). The IRAM-30m observations were obtained and analyzed as part of the POLAMI¹ program [169, 170, 171]. The polarization degree of BL Lac was measured to be $\Pi_R = 8.08 \pm 1.38\%$ along position angle $\psi_R = -6.5^\circ \pm 4.2^\circ$ at 1.3 mm, and $\Pi_R = 7.27 \pm 0.43\%$, $\psi_R = 5.7^\circ \pm 1.5^\circ$ at 3.5 mm. No circular polarization was detected ($< 1.1\%$, 99% C.I.) at 1.3 mm and $< 0.6\%$ (99% C.I.) at 3.5 mm. The SMA [172] observation was taken within the framework of the SMA Monitoring of AGNs with Polarization (SMAPOL) program with the SMA polarimeter [173]. The polarized quantities are derived from the Stokes I , Q , and U visibilities and calibrated with the MIR software package². The measurements yield $\Pi_R = 6.22 \pm 0.88\%$ along $\psi_R = 1.9^\circ \pm 0.3^\circ$, consistent with the contemporaneous POLAMI observation.

¹<http://polami.iaa.es/>

²<https://lweb.cfa.harvard.edu/~cqj/mircook.html>

Table B.4: Multi-wavelength polarization observations of BL Lac during the IXPE pointing. The uncertainties in Π and ψ are either the uncertainty of the measurement or, in the case of multiple measurements, the median uncertainty. σ_{Π} and σ_{ψ} show the standard deviation of the observations.

Telescope	Π (%)	σ_{Π}	ψ (deg.)	σ_{ψ}
POLAMI (3 mm)	7.27 ± 0.43	–	5.7 ± 1.5	–
POLAMI (1.3 mm)	8.08 ± 1.38	–	-6.5 ± 4.2	–
SMA (1.3 mm)	6.22 ± 0.88	–	1.9 ± 0.3	–
Calar Alto & SNO (R-band)	10.9 ± 0.3	1.74	178 ± 0.6	4.41
Perkins (B-band)	10.33 ± 0.22	3.61	172 ± 0.6	11.19
Perkins (V-band)	9.66 ± 0.15	3.21	173 ± 0.4	11.44
Perkins (R-band)	9.28 ± 0.2	4.10	175 ± 0.7	13.51
Perkins (I-band)	10.23 ± 0.2	3.33	168 ± 1	8.08

Optical polarization coverage was provided by the Calar Alto (Calar Alto Faint Object Spectrograph – CAFOS, R-band), Higashi-Hiroshima Observatory (Kanata telescope) with the Hiroshima Optical and Near-Infrared camera (HONIR – R, J-band), the Nordic Optical Telescope (NOT) with the Alhambra Faint Object Spectrograph and Camera (ALFOSC, BVRI), the 1.8m Perkins Telescope (BVRI), and the Sierra Nevada Observatory (SNO, R-band). A detailed description of the observing strategy and data reduction of the aforementioned telescopes can be found in [174, 175, 176, 177, 178, 157], respectively, and references therein. The observations cover the entire duration of the *IXPE* observation, revealing high variability of the polarization degree, from $\sim 3\%$ to $\sim 17\%$, with EVPA fluctuating about the direction of the jet axis, $10^\circ \pm 2^\circ$ [20]. Figure B.4 displays the optical observations, while table B.4 summarizes the radio and optical polarization results for the individual telescopes.

The 3-day binned *Fermi*-LAT γ ray light curve of BL Lac is given in fig. B.5, extracted from the *Fermi*-LAT Light Curve Repository [179], with *IXPE* observation times highlighted. The third *IXPE* observation of BL Lac, reported here, occurred during an outburst.

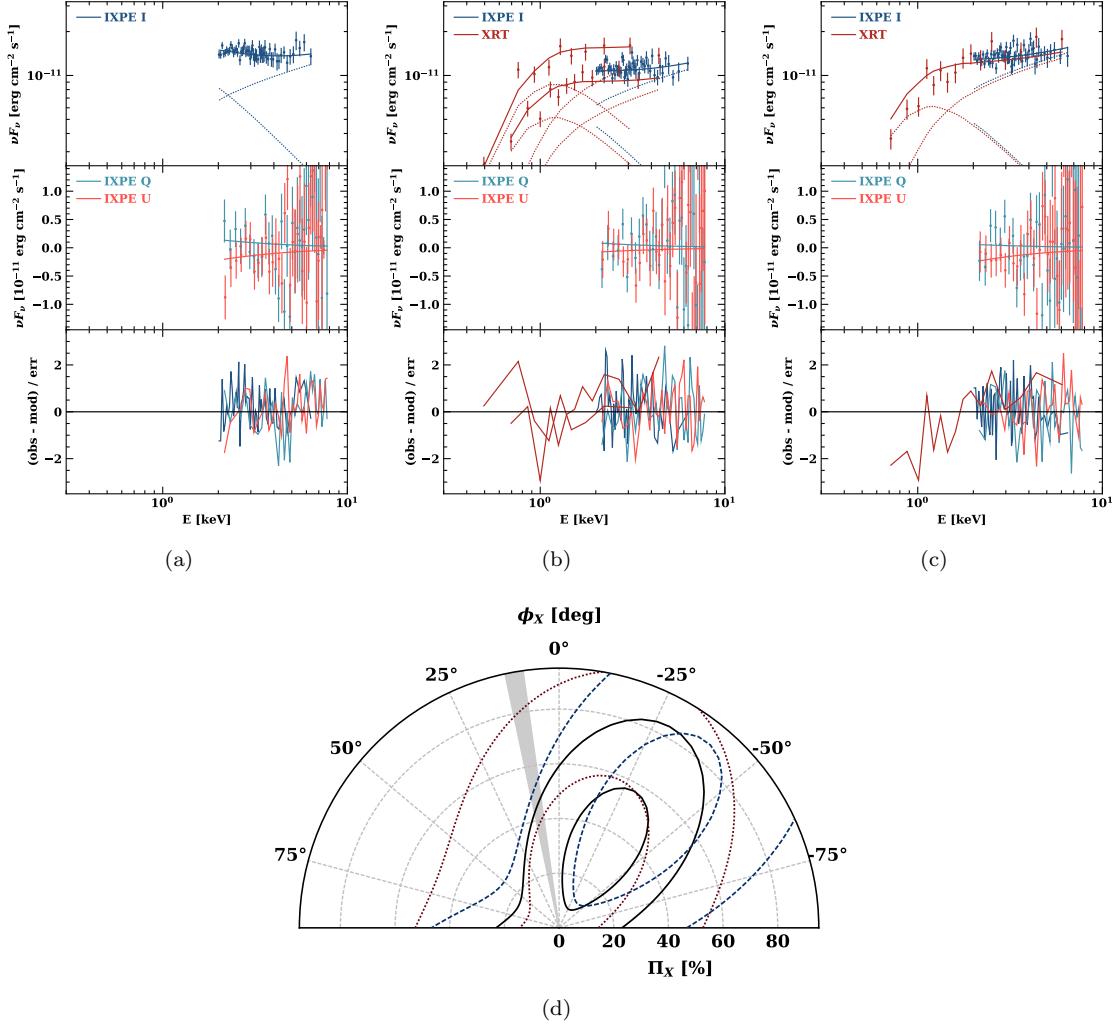


Figure B.3: (a) Absorbed sum of two power-laws with apec fit to first *IXPE* time bin with first *Swift* XRT observation. Photon indices and apec temperature are fixed to values in table 7.1. The high-energy power-law polarization is fixed to zero. (b) Second *IXPE* time bin with second XRT observation. (c) Third *IXPE* time bin with second XRT observation. (d) Polarization fraction and EVPA confidence levels (68%, 99%) of the low-energy PL for each time bin: black – T1, red – T2, blue – T3. Grey shaded region represents the VLBI-determined jet axis projection on the plane of the sky [20].

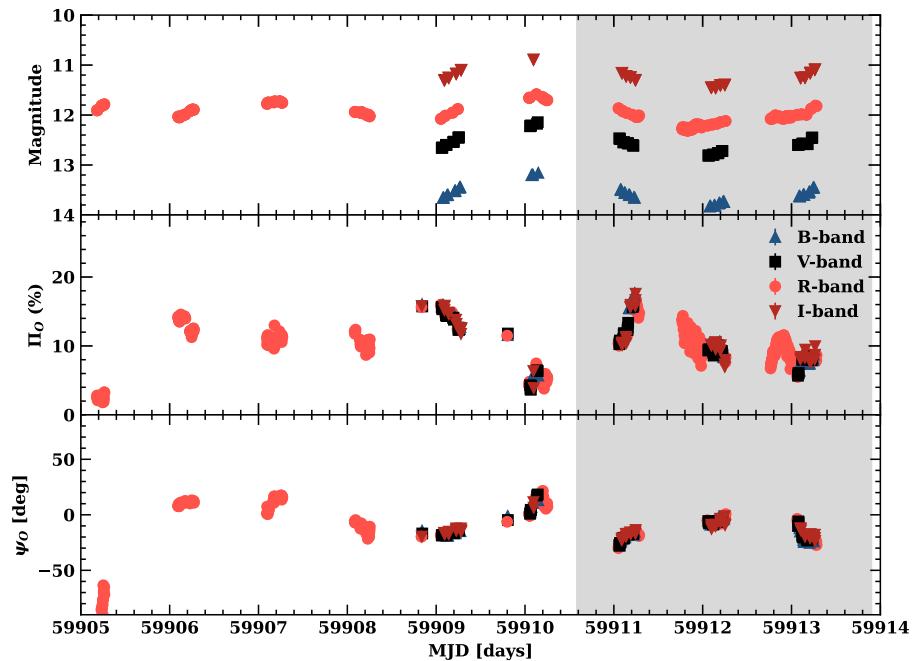


Figure B.4: Optical BVRI observations of BL Lac before and during the 2022 November 27–30 *IXPE* pointing, showing brightness in magnitudes (top panel), polarization degree (middle panel), and polarization position angle (bottom panel). The duration of the *IXPE* observation is marked by the gray shaded area.

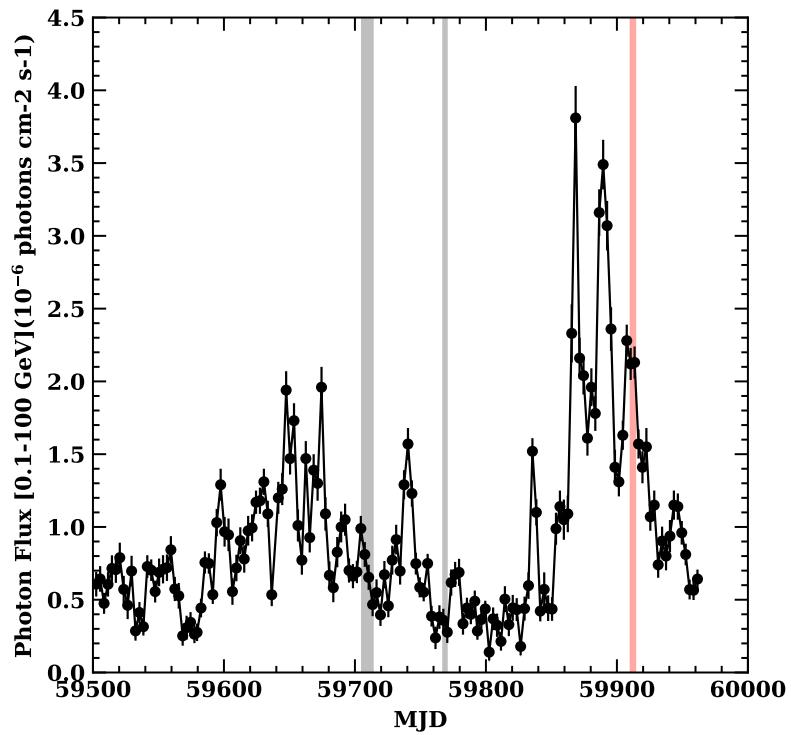


Figure B.5: *Fermi*-LAT 3-day cadence light curve, showing the outburst that triggered the *IXPE* observation. A single power-law model with a free photon index is used to determine the photon fluxes. The two gray-shaded regions mark two previous *IXPE* observations, while red shaded indicates the observation discussed in this paper.

Bibliography

- [1] A. L. Peirson and Roger W. Romani. “The Polarization Behavior of Relativistic Synchrotron Jets.” In: *The Astrophysical Journal* 864.2 (2018), p. 140. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aad69d (cit. on pp. v, 67, 70, 99, 100, 115, 119, 133).
- [2] A. L. Peirson and Roger W. Romani. “The Polarization Behavior of Relativistic Synchrotron Self-Compton Jets.” In: *ApJ* 885.1 (Nov. 2019), p. 76. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab46b1 (cit. on pp. v, 99–101, 103, 107, 114–116, 121).
- [3] A. L. Peirson, R. W. Romani, H. L. Marshall, J. F. Steiner, and L. Baldini. “Deep Ensemble Analysis for Imaging X-ray Polarimetry.” In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 986 (Jan. 2021), p. 164740. ISSN: 0168-9002. DOI: 10.1016/j.nima.2020.164740 (cit. on pp. v, 26–32, 34, 108, 119, 123).
- [4] A. L. Peirson and R. W. Romani. “Towards Optimal Signal Extraction for Imaging X-ray Polarimetry.” In: *arXiv:2107.08289 [astro-ph]* (July 2021). arXiv: 2107.08289 [astro-ph] (cit. on pp. v, 27, 28, 32, 34, 37, 45, 46, 108, 119, 123).
- [5] I. Liodakis, A. L. Peirson, and R. W. Romani. “Prospects for Detecting X-ray Polarization in Blazar Jets.” In: *arXiv:1906.01647 [astro-ph]* (June 2019). arXiv: 1906.01647 [astro-ph] (cit. on pp. v, 75, 85, 99).
- [6] Abel L. Peirson, Ioannis Liodakis, and Roger W. Romani. “Testing High-energy Emission Models for Blazars with X-Ray Polarimetry.” In: *The Astrophysical Journal* 931.1 (May 2022), p. 59. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ac6a54 (cit. on pp. v, 110, 116).
- [7] A. L. Peirson and R. W. Romani. *A Deep Ensemble Approach to X-ray Polarimetry*. Jan. 2022. DOI: 10.48550/arXiv.2111.03047. arXiv: 2111.03047 [astro-ph] (cit. on pp. v, 119, 123).
- [8] Abel L. Peirson. *Neural Network Analysis of X-ray Polarimeter Data*. June 2022. DOI: 10.48550/arXiv.2206.10537. arXiv: 2206.10537 [astro-ph] (cit. on pp. v, 123).

- [9] Brian D. Ramsey et al. “Optics for the Imaging X-Ray Polarimetry Explorer.” In: *Journal of Astronomical Telescopes, Instruments, and Systems* 8.2 (May 2022), p. 024003. ISSN: 2329-4124, 2329-4221. DOI: 10.1117/1.JATIS.8.2.024003 (cit. on pp. xiii, 3).
- [10] L. Baldini et al. “Design, Construction, and Test of the Gas Pixel Detectors for the IXPE Mission.” In: *Astroparticle Physics* 133 (Dec. 2021), p. 102628. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2021.102628 (cit. on pp. xiii, 4, 7, 27, 123).
- [11] Ronaldo Bellazzini et al. “Novel Gaseous X-Ray Polarimeter: Data Analysis and Simulation.” In: *Polarimetry in Astronomy*. Vol. 4843. International Society for Optics and Photonics, Feb. 2003, pp. 383–393. DOI: 10.1117/12.459381 (cit. on pp. xiii, 1, 5, 8, 9).
- [12] Javier Antorán, James Urquhart Allingham, and José Miguel Hernández-Lobato. “Depth Uncertainty in Neural Networks.” In: *arXiv:2006.08437 [cs, stat]* (Dec. 2020). arXiv: 2006.08437 [cs, stat] (cit. on pp. xiv, 21).
- [13] A. A. Abdo et al. “Spectral Properties of Bright Fermi-Detected Blazars in the Gamma-Ray Band.” In: *\apj* 710 (Feb. 2010), pp. 1271–1285. DOI: 10.1088/0004-637X/710/2/1271 (cit. on pp. xiv, 22, 23, 99).
- [14] M. Petropoulou and A. Mastichiadis. “Bethe–Heitler Emission in BL Lacs: Filling the Gap between X-rays and γ -Rays.” In: *Monthly Notices of the Royal Astronomical Society* 447.1 (Feb. 2015), pp. 36–48. ISSN: 0035-8711. DOI: 10.1093/mnras/stu2364 (cit. on pp. xiv, 25).
- [15] Shan Gao, Anatoli Fedynitch, Walter Winter, and Martin Pohl. “Modelling the Coincident Observation of a High-Energy Neutrino and a Bright Blazar Flare.” In: *Nature Astronomy* 3 (Jan. 2019), pp. 88–92. DOI: 10.1038/s41550-018-0610-1 (cit. on pp. xiv, 23, 25, 99, 107, 110).
- [16] D. Blinov et al. “RoboPol: Do Optical Polarization Rotations Occur in All Blazars?” In: *MNRAS* 462 (Oct. 2016), pp. 1775–1785. ISSN: 0035-8711. DOI: 10.1093/mnras/stw1732 (cit. on pp. xvi, 24, 51, 52, 57, 58, 118).
- [17] Alan P. Marscher. “Turbulent, Extreme Multi-zone Model for Simulating Flux and Polarization Variability in Blazars.” In: *ApJ* 780 (Jan. 2014), p. 87. ISSN: 0004-637X. DOI: 10.1088/0004-637X/780/1/87 (cit. on pp. xvii, xviii, 24, 52, 57, 67, 72, 75, 80, 90, 91, 99, 100, 105, 107, 115, 119).
- [18] William J. Potter and Garret Cotter. “Black Hole Jets, Accretion Discs and Dark Energy.” Doctoral dissertation. Oxford University, UK, Jan. 2013 (cit. on pp. xvii, 58, 59).
- [19] D. Blinov et al. “RoboPol: AGN Polarimetric Monitoring Data.” In: *\mnras* 501.3 (Mar. 2021), pp. 3715–3726. DOI: 10.1093/mnras/staa3777 (cit. on pp. xxi, 99, 103–105).

- [20] Zachary R. Weaver et al. “Kinematics of Parsec-scale Jets of Gamma-Ray Blazars at 43 GHz during 10 Yr of the VLBA-BU-BLAZAR Program.” In: *\apjs* 260.1 (May 2022), p. 12. DOI: [10.3847/1538-4365/ac589c](https://doi.org/10.3847/1538-4365/ac589c) (cit. on pp. xxi, xxii, 113, 115, 117, 126, 127, 137, 139, 140).
- [21] R. Novick, M. C. Weisskopf, R. Berthelsdorf, R. Linke, and R. S. Wolff. “Detection of X-Ray Polarization of the Crab Nebula.” In: *The Astrophysical Journal* 174 (May 1972), p. L1. ISSN: 0004-637X. DOI: [10.1086/180938](https://doi.org/10.1086/180938) (cit. on p. 1).
- [22] R. Bellazzini et al. “A Sealed Gas Pixel Detector for X-ray Astronomy.” In: *Nuclear Instruments and Methods in Physics Research A* 579 (Sept. 2007), pp. 853–858. ISSN: 0168-9002. DOI: [10.1016/j.nima.2007.05.304](https://doi.org/10.1016/j.nima.2007.05.304) (cit. on p. 1).
- [23] Hua Feng and Ronaldo Bellazzini. “The X-ray Polarimetry Window Reopens.” In: *Nature Astronomy* 4.5 (May 2020), pp. 547–547. ISSN: 2397-3366. DOI: [10.1038/s41550-020-1103-6](https://doi.org/10.1038/s41550-020-1103-6) (cit. on p. 1).
- [24] C. Sgrò and IXPE Team. “The Imaging X-ray Polarimetry Explorer (IXPE).” In: *Nuclear Instruments and Methods in Physics Research A* 936 (Aug. 2019), pp. 212–215. ISSN: 0168-9002. DOI: [10.1016/j.nima.2018.10.111](https://doi.org/10.1016/j.nima.2018.10.111) (cit. on p. 1).
- [25] S. N. Zhang et al. “eXTP: Enhanced X-ray Timing and Polarization Mission.” In: *Space Telescopes and Instrumentation 2016: Ultraviolet to Gamma Ray*. Vol. 9905. International Society for Optics and Photonics, Jan. 2017, 99051Q. DOI: [10.1117/12.2232034](https://doi.org/10.1117/12.2232034) (cit. on pp. 1, 87, 132).
- [26] Henric Krawczynski et al. “Using X-Ray Polarimetry to Probe the Physics of Black Holes and Neutron Stars.” In: 51 (May 2019), p. 150 (cit. on p. 1).
- [27] Martin Weisskopf. “An Overview of X-Ray Polarimetry of Astronomical Sources.” In: *Galaxies* 6 (Mar. 2018), p. 33. DOI: [10.3390/galaxies6010033](https://doi.org/10.3390/galaxies6010033) (cit. on pp. 1, 87).
- [28] Roger Blandford, David Meier, and Anthony Readhead. “Relativistic Jets from Active Galactic Nuclei.” In: *\araa* 57 (Aug. 2019), pp. 467–509. DOI: [10.1146/annurev-astro-081817-051948](https://doi.org/10.1146/annurev-astro-081817-051948) (cit. on p. 2).
- [29] ICECUBE COLLABORATION et al. “Neutrino Emission from the Direction of the Blazar TXS 0506+056 Prior to the IceCube-170922A Alert.” In: *Science* 361.6398 (July 2018), pp. 147–151. DOI: [10.1126/science.aat2890](https://doi.org/10.1126/science.aat2890) (cit. on pp. 2, 110).
- [30] Alex Graves, Santiago Fernández, Faustino Gomez, and Jürgen Schmidhuber. “Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks.” In: *Proceedings of the 23rd International Conference on Machine Learning*. ICML ’06. Pittsburgh, Pennsylvania, USA: Association for Computing Machinery, June 2006, pp. 369–376. ISBN: 978-1-59593-383-6. DOI: [10.1145/1143844.1143891](https://doi.org/10.1145/1143844.1143891) (cit. on p. 2).

- [31] Tom Young, Devamanyu Hazarika, Soujanya Poria, and Erik Cambria. “Recent Trends in Deep Learning Based Natural Language Processing.” In: *arXiv:1708.02709 [cs]* (Nov. 2018). arXiv: 1708.02709 [cs] (cit. on p. 2).
- [32] Binhu Tang, Zixiang Pan, Kang Yin, and Asif Khateeb. “Recent Advances of Deep Learning in Bioinformatics and Computational Biology.” In: *Frontiers in Genetics* 10 (2019). ISSN: 1664-8021. doi: 10.3389/fgene.2019.00214 (cit. on p. 2).
- [33] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. “ImageNet Classification with Deep Convolutional Neural Networks.” In: *Advances in Neural Information Processing Systems 25*. Ed. by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger. Curran Associates, Inc., 2012, pp. 1097–1105 (cit. on pp. 2, 19).
- [34] Aryeh Brill et al. “Investigating a Deep Learning Method to Analyze Images from Multiple Gamma-ray Telescopes.” In: *2019 New York Scientific Data Summit (NYSDS)* (June 2019), pp. 1–4. doi: 10.1109/NYSDS.2019.8909697. arXiv: 2001.03602 (cit. on p. 2).
- [35] Nicholas Choma et al. “Graph Neural Networks for IceCube Signal Classification.” In: *arXiv:1809.06166 [astro-ph, stat]* (Sept. 2018). arXiv: 1809.06166 [astro-ph, stat] (cit. on p. 2).
- [36] Stephen L. O’Dell et al. “The Imaging X-ray Polarimetry Explorer (IXPE): Technical Overview.” In: 0699 (Aug. 2018), p. 106991X. doi: 10.1117/12.2314146 (cit. on pp. 3, 94).
- [37] Enrico Costa et al. “An Efficient Photoelectric X-ray Polarimeter for the Study of Black Holes and Neutron Stars.” In: *Nature* 411 (June 2001), pp. 662–665. ISSN: 0028-0836 (cit. on pp. 4, 7).
- [38] F. Kislat, B. Clark, M. Beilicke, and H. Krawczynski. “Analyzing the Data from X-ray Polarimeters with Stokes Parameters.” In: *Astroparticle Physics* 68 (Aug. 2015), pp. 45–51. ISSN: 0927-6505. doi: 10.1016/j.astropartphys.2015.02.007 (cit. on pp. 14, 45, 46, 105, 123, 126).
- [39] Martin C. Weisskopf, Ronald F. Elsner, and Stephen L. O’Dell. “On Understanding the Figures of Merit for Detection and Measurement of X-Ray Polarization.” In: *arXiv:1006.3711 [astro-ph]* (July 2010), 77320E. doi: 10.1117/12.857357. arXiv: 1006.3711 [astro-ph] (cit. on p. 15).
- [40] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016 (cit. on pp. 15, 17, 18).
- [41] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer, 2009 (cit. on p. 15).
- [42] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. “Multilayer Feedforward Networks Are Universal Approximators.” In: *Neural Networks* 2.5 (Jan. 1989), pp. 359–366. ISSN: 0893-6080. doi: 10.1016/0893-6080(89)90020-8 (cit. on p. 17).

- [43] John Bridle. “Training Stochastic Model Recognition Algorithms as Networks Can Lead to Maximum Mutual Information Estimation of Parameters.” In: *Advances in Neural Information Processing Systems*. Ed. by D. Touretzky. Vol. 2. Morgan-Kaufmann, 1990 (cit. on p. 17).
- [44] Tom Brown et al. “Language Models Are Few-Shot Learners.” In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin. Vol. 33. Curran Associates, Inc., 2020, pp. 1877–1901 (cit. on p. 18).
- [45] Yann LeCun and Yoshua Bengio. “Convolutional Networks for Images, Speech, and Time Series.” In: *The Handbook of Brain Theory and Neural Networks*. Cambridge, MA, USA: MIT Press, Oct. 1998, pp. 255–258. ISBN: 978-0-262-51102-5 (cit. on p. 19).
- [46] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. “Deep Residual Learning for Image Recognition.” In: *arXiv:1512.03385 [cs]* (Dec. 2015). arXiv: 1512.03385 [cs] (cit. on pp. 19, 30).
- [47] Yu Zhang and Qiang Yang. “A Survey on Multi-Task Learning.” In: *arXiv:1707.08114 [cs]* (Mar. 2021). arXiv: 1707.08114 [cs] (cit. on p. 19).
- [48] Dario Amodei et al. “Concrete Problems in AI Safety.” In: *arXiv:1606.06565 [cs]* (July 2016). arXiv: 1606.06565 [cs] (cit. on p. 20).
- [49] Alex Kendall and Yarin Gal. “What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?” In: *Advances in Neural Information Processing Systems 30*. Ed. by I. Guyon et al. Curran Associates, Inc., 2017, pp. 5574–5584 (cit. on pp. 20, 22).
- [50] Sungjoon Choi, Kyungjae Lee, Sungbin Lim, and Songhwai Oh. “Uncertainty-Aware Learning from Demonstration Using Mixture Density Networks with Sampling-Free Variance Modeling.” In: *arXiv:1709.02249 [cs]* (Sept. 2017). arXiv: 1709.02249 [cs] (cit. on pp. 20, 21).
- [51] Kalvik Jakkala. “Deep Gaussian Processes: A Survey.” In: *arXiv:2106.12135 [cs, stat]* (June 2021). arXiv: 2106.12135 [cs, stat] (cit. on p. 21).
- [52] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. “Simple and Scalable Predictive Uncertainty Estimation Using Deep Ensembles.” In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. NIPS’17. Long Beach, California, USA: Curran Associates Inc., Dec. 2017, pp. 6405–6416. ISBN: 978-1-5108-6096-4 (cit. on pp. 21, 22).
- [53] Fredrik K. Gustafsson, Martin Danelljan, and Thomas B. Schon. “Evaluating Scalable Bayesian Deep Learning Methods for Robust Computer Vision.” In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 318–319 (cit. on p. 21).

- [54] Stanislav Fort, Huiyi Hu, and Balaji Lakshminarayanan. “Deep Ensembles: A Loss Landscape Perspective.” In: *arXiv:1912.02757 [cs, stat]* (Dec. 2019). arXiv: 1912 . 02757 [cs, stat] (cit. on p. 21).
- [55] Yaniv Ovadia et al. “Can You Trust Your Model’s Uncertainty? Evaluating Predictive Uncertainty Under Dataset Shift.” In: *arXiv:1906.02530 [cs, stat]* (Dec. 2019). arXiv: 1906 . 02530 [cs, stat] (cit. on p. 21).
- [56] Alex Kendall, Yarin Gal, and Roberto Cipolla. “Multi-Task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics.” In: *arXiv:1705.07115 [cs]* (Apr. 2018). arXiv: 1705 . 07115 [cs] (cit. on pp. 22, 32).
- [57] C. Megan Urry and Paolo Padovani. “Unified Schemes for Radio-Loud Active Galactic Nuclei.” In: *Publ Astron Soc Pac* 107 (Sept. 1995), p. 803. ISSN: 0004-6280, 1538-3873. DOI: 10 . 1086/133630. arXiv: astro-ph/9506063 (cit. on p. 22).
- [58] L. Maraschi, G. Ghisellini, and A. Celotti. “A Jet Model for the Gamma-Ray Emitting Blazar 3C 279.” In: *ApJ* 397 (Sept. 1992), pp. L5–L9. ISSN: 0004-637X. DOI: 10 . 1086/186531 (cit. on p. 22).
- [59] M. Boettcher. “Modeling the Spectral Energy Distributions and Variability of Blazars.” In: *ArXiv e-prints* (May 2012) (cit. on p. 22).
- [60] M. Ackermann et al. “Minute-Timescale >100 MeV γ -Ray Variability during the Giant Outburst of Quasar 3C 279 Observed by Fermi-LAT in 2015 June.” In: *\apjl* 824 (June 2016), p. L20. DOI: 10 . 3847/2041-8205/824/2/L20 (cit. on p. 22).
- [61] R. Blandford, D. Meier, and A. Readhead. “Relativistic Jets in Active Galactic Nuclei.” In: *arXiv e-prints* (Dec. 2018) (cit. on p. 22).
- [62] R. D. Blandford and R. L. Znajek. “Electromagnetic Extraction of Energy from Kerr Black Holes.” In: *MNRAS* 179 (May 1977), pp. 433–456. ISSN: 0035-8711. DOI: 10 . 1093/mnras/179 . 3 . 433 (cit. on pp. 23, 53, 119).
- [63] T. Kintscher et al. “IceCube Search for Neutrinos from 1ES 1959+650: Completing the Picture.” In: *35th International Cosmic Ray Conference (ICRC2017)*. Vol. 301. International Cosmic Ray Conference. Jan. 2017, p. 969 (cit. on pp. 23, 110).
- [64] M. Cerruti et al. “Leptohadronic Single-Zone Models for the Electromagnetic and Neutrino Emission of TXS 0506+056.” In: *\mnras* 483.1 (Feb. 2019), pp. L12–L16. DOI: 10 . 1093/mnrasl/sly210 (cit. on pp. 23, 110).
- [65] D. Blinov et al. “RoboPol: First Season Rotations of Optical Polarization Plane in Blazars.” In: *MNRAS* 453 (Oct. 2015), pp. 1669–1683. ISSN: 0035-8711. DOI: 10 . 1093/mnras/stv1723 (cit. on pp. 24, 51, 110, 118).

- [66] Ryan S. Lynch et al. “The Green Bank North Celestial Cap Pulsar Survey III: 45 New Pulsar Timing Solutions.” In: *arXiv:1805.04951 [astro-ph]* (May 2018). arXiv: 1805.04951 [astro-ph] (cit. on p. 24).
- [67] Maxim Lyutikov and Evgeniya Kravchenko. “Polarization Swings in Blazars.” In: *MNRAS* 467.4 (June 2017), pp. 3876–3886. ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/stx359. arXiv: 1702.02354 (cit. on pp. 24, 53, 119).
- [68] Krzysztof Nalewajko. “A Model of Polarisation Rotations in Blazars from Kink Instabilities in Relativistic Jets.” In: *Galaxies* 5.4 (Oct. 2017), p. 64. ISSN: 2075-4434. DOI: 10.3390/galaxies5040064. arXiv: 1711.00899 (cit. on pp. 24, 53, 119).
- [69] Martin C. Weisskopf et al. “The Imaging X-ray Polarimetry Explorer (IXPE).” In: *Results Phys* 6 (Jan. 2016), pp. 1179–1180. ISSN: 2211-3797. DOI: 10.1016/j.rinp.2016.10.021 (cit. on pp. 24, 63).
- [70] Niall O’ Mahony et al. “Deep Learning vs. Traditional Computer Vision.” In: *arXiv:1910.13796 [cs]* 943 (2020). DOI: 10.1007/978-3-030-17795-9. arXiv: 1910.13796 [cs] (cit. on p. 26).
- [71] Moloud Abdar et al. “A Review of Uncertainty Quantification in Deep Learning: Techniques, Applications and Challenges.” In: *Information Fusion* 76 (Dec. 2021), pp. 243–297. ISSN: 15662535. DOI: 10.1016/j.inffus.2021.05.008. arXiv: 2011.06225 (cit. on p. 26).
- [72] Takao Kitaguchi et al. “A Convolutional Neural Network Approach for Reconstructing Polarization Information of Photoelectric X-ray Polarimeters.” In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 942 (Oct. 2019), p. 162389. ISSN: 01689002. DOI: 10.1016/j.nima.2019.162389. arXiv: 1907.06442 (cit. on pp. 26, 27).
- [73] S. Agostinelli et al. “Geant4—a Simulation Toolkit.” In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (July 2003), pp. 250–303. ISSN: 0168-9002. DOI: 10.1016/S0168-9002(03)01368-8 (cit. on p. 27).
- [74] Constantin Steppa and Tim Lukas Holch. “HexagDLy - Processing Hexagonally Sampled Data with CNNs in PyTorch.” In: *SoftwareX* 9 (Jan. 2019), pp. 193–198. ISSN: 23527110. DOI: 10.1016/j.softx.2019.02.010. arXiv: 1903.01814 (cit. on p. 29).
- [75] Raveen Doon, Tarun Kumar Rawat, and Shweta Gautam. “Cifar-10 Classification Using Deep Convolutional Neural Network.” In: *2018 IEEE Punecon*. Nov. 2018, pp. 1–5. DOI: 10.1109/PUNECON.2018.8745428 (cit. on p. 30).
- [76] Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger. “Densely Connected Convolutional Networks.” In: *arXiv:1608.06993 [cs]* (Jan. 2018). arXiv: 1608.06993 [cs] (cit. on p. 30).

- [77] Ashish Vaswani et al. *Attention Is All You Need*. Dec. 2017. doi: [10.48550/arXiv.1706.03762](https://doi.org/10.48550/arXiv.1706.03762). arXiv: 1706.03762 [cs] (cit. on p. 30).
- [78] Ilya Sutskever, James Martens, George Dahl, and Geoffrey Hinton. “On the Importance of Initialization and Momentum in Deep Learning.” In: *Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28*. ICML’13. Atlanta, GA, USA: JMLR.org, June 2013, pp. III-1139–III-1147 (cit. on pp. 34, 35).
- [79] Kyle Cranmer, Johann Brehmer, and Gilles Louppe. “The Frontier of Simulation-Based Inference.” In: *Proceedings of the National Academy of Sciences* 117.48 (Dec. 2020), pp. 30055–30062. ISSN: 0027-8424, 1091-6490. doi: [10.1073/pnas.1912789117](https://doi.org/10.1073/pnas.1912789117) (cit. on p. 47).
- [80] S. Jorstad et al. “Multifrequency Polarization Properties of Blazars.” In: *ChJP* 6.S1 (2006), p. 247 (cit. on pp. 52, 56).
- [81] Francesca D. D’arcangelo et al. “Synchronous Optical and Radio Polarization Variability in the Blazar OJ287.” In: *ApJ* 697.2 (2009), p. 985. ISSN: 0004-637X. doi: [10.1088/0004-637X/697/2/985](https://doi.org/10.1088/0004-637X/697/2/985) (cit. on p. 52).
- [82] D. Blinov et al. “RoboPol: Connection between Optical Polarization Plane Rotations and Gamma-Ray Flares in Blazars.” In: *MNRAS* 474.1 (Feb. 2018), pp. 1296–1306. ISSN: 0035-8711. doi: [10.1093/mnras/stx2786](https://doi.org/10.1093/mnras/stx2786) (cit. on pp. 52, 75, 87, 91, 116).
- [83] P. A. Hughes, H. D. Aller, and M. F. Aller. “Synchrotron Emission from Shocked Relativistic Jets. I - The Theory of Radio-Wavelength Variability and Its Relation to Superluminal Motion.” In: *ApJ* 341 (June 1989), pp. 54–79. ISSN: 0004-637X. doi: [10.1086/167471](https://doi.org/10.1086/167471) (cit. on p. 52).
- [84] Haocheng Zhang, Xuhui Chen, Markus Böttcher, Fan Guo, and Hui Li. “Polarization Swings Reveal Magnetic Energy Dissipation in Blazars.” In: *ApJ* 804 (May 2015), p. 58. ISSN: 0004-637X. doi: [10.1088/0004-637X/804/1/58](https://doi.org/10.1088/0004-637X/804/1/58) (cit. on p. 52).
- [85] Denise C. Gabuzda, Éamonn Murray, and Patrick Cronin. “Helical Magnetic Fields Associated with the Relativistic Jets of Four BL Lac Objects.” In: *MNRAS* 351.4 (July 2004), pp. L89–L93. ISSN: 0035-8711. doi: [10.1111/j.1365-2966.2004.08037.x](https://doi.org/10.1111/j.1365-2966.2004.08037.x) (cit. on p. 52).
- [86] M. Lyutikov, V. I. Pariev, and R. D. Blandford. “Polarization of Prompt Gamma-Ray Burst Emission: Evidence for Electromagnetically Dominated Outflow.” In: *The Astrophysical Journal* 597.2 (Nov. 2003), p. 998. ISSN: 0004-637X. doi: [10.1086/378497](https://doi.org/10.1086/378497) (cit. on pp. 54, 70, 119).
- [87] R. D. Blandford and A. Koenigl. “Relativistic Jets as Compact Radio Sources.” In: *ApJ* 232 (Aug. 1979), pp. 34–48. ISSN: 0004-637X. doi: [10.1086/157262](https://doi.org/10.1086/157262) (cit. on p. 54).

- [88] Shan-Jie Qian and Xi-Zhen Zhang. “An Intrinsic Model for the Polarization Position Angle Swing Observed in QSO 1150+812.” In: *Chin. J. Astron. Astrophys.* 4.1 (2004), p. 37. ISSN: 1009-9271. DOI: 10.1088/1009-9271/4/1/37 (cit. on p. 55).
- [89] Alan P. Marscher et al. “Probing the Inner Jet of the Quasar PKS 1510-089 with Multi-waveband Monitoring during Strong Gamma-ray Activity.” In: *ApJ* 710 (Jan. 2010), pp. L126–L131. DOI: 10.1088/2041-8205/710/2/L126 (cit. on pp. 58, 116).
- [90] V. M. Larionov et al. “The Outburst of the Blazar S5 0716+71 in 2011 October: Shock in a Helical Jet.” In: *ApJ* 768 (May 2013), p. 40. ISSN: 0004-637X. DOI: 10.1088/0004-637X/768/1/40 (cit. on pp. 58, 119).
- [91] M. L. Lister et al. “MOJAVE. X. Parsec-scale Jet Orientation Variations and Superluminal Motion in Active Galactic Nuclei.” In: *AJ* 146.5 (2013), p. 120. ISSN: 1538-3881. DOI: 10.1088/0004-6256/146/5/120 (cit. on p. 59).
- [92] J. N. Bregman. “Diffusive Shock Acceleration and Quasar Photospheres.” In: *ApJ* 288 (Jan. 1985), pp. 32–42. ISSN: 0004-637X. DOI: 10.1086/162760 (cit. on p. 60).
- [93] George B. Rybicki and Alan P. Lightman. *Radiative Processes in Astrophysics*. John Wiley & Sons, 1979. ISBN: 978-0-471-82759-7 (cit. on pp. 61, 68, 134).
- [94] Malcolm S. Longair. *High Energy Astrophysics*. Cambridge University Press, Feb. 2011. ISBN: 978-1-139-49454-0 (cit. on pp. 61, 121).
- [95] Alan P. Marscher and Svetlana G. Jorstad. “Rapid Variability of Gamma-ray Emission from Sites near the 43 GHz Cores of Blazar Jets.” In: *arXiv:1005.5551 [astro-ph]* (May 2010). arXiv: 1005.5551 [astro-ph] (cit. on pp. 62, 90).
- [96] E. Angelakis et al. “RoboPol: The Optical Polarization of Gamma-Ray-Loud and Gamma-Ray-Quiet Blazars.” In: *MNRAS* 463.3 (Dec. 2016), pp. 3365–3380. ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/stw2217. arXiv: 1609.00640 (cit. on pp. 62, 94, 118).
- [97] Ryosuke Itoh et al. “Systematic Study of Gamma-ray Bright Blazars with Optical Polarization and Gamma-ray Variability.” In: *ApJ* 833.1 (Dec. 2016), p. 77. ISSN: 1538-4357. DOI: 10.3847/1538-4357/833/1/77. arXiv: 1610.04313 (cit. on p. 62).
- [98] Haocheng Zhang, Wei Deng, Hui Li, and Markus Böttcher. “Polarization Signatures of Relativistic Magnetohydrodynamic Shocks in the Blazar Emission Region. I. Force-free Helical Magnetic Fields.” In: *ApJ* 817 (Jan. 2016), p. 63. ISSN: 0004-637X. DOI: 10.3847/0004-637X/817/1/63 (cit. on pp. 62, 87).
- [99] Fabian Kislat et al. “Optimization of the Design of X-Calibur for a Long-Duration Balloon Flight and Results from a One-Day Test Flight.” In: *JATIS* 4.1 (Mar. 2018), p. 011004. ISSN: 2329-4124, 2329-4221. DOI: 10.1117/1.JATIS.4.1.011004 (cit. on p. 63).

- [100] S. Bonometto, P. Cazzola, and A. Saggion. “Polarization in Inverse Compton Effect.” In: *Astronomy and Astrophysics* 7 (Aug. 1970), p. 292. ISSN: 0004-6361 (cit. on pp. 67, 70, 71).
- [101] S. Bonometto and A. Saggion. “Polarization in Inverse Compton Scattering of Synchrotron Radiation.” In: *Astronomy and Astrophysics* 23 (Feb. 1973), p. 9. ISSN: 0004-6361 (cit. on pp. 67, 70, 71, 103).
- [102] H. Krawczynski. “The Polarization Properties of Inverse Compton Emission and Implications for Blazar Observations with the GEMS X-Ray Polarimeter.” In: *\apj* 744.1 (Jan. 2012), p. 30. DOI: 10.1088/0004-637X/744/1/30 (cit. on pp. 67, 114).
- [103] Juri Poutanen. “Relativistic Jets in Blazars: Polarization of Radiation.” In: *The Astrophysical Journal Supplement Series* 92 (June 1994), pp. 607–609. ISSN: 0067-0049. DOI: 10.1086/192024 (cit. on p. 67).
- [104] Aimee L. McNamara, Zdenka Kuncic, and Kinwah Wu. “X-Ray Polarization in Relativistic Jets.” In: *arXiv:0902.1562 [astro-ph]* (Feb. 2009). DOI: 10.1111/j.1365-2966.2009.14608.x. arXiv: 0902.1562 [astro-ph] (cit. on p. 67).
- [105] H. Zhang and M. Böttcher. “X-Ray and Gamma-Ray Polarization in Leptonic and Hadronic Jet Models of Blazars.” In: *\apj* 774.1 (Sept. 2013), p. 18. DOI: 10.1088/0004-637X/774/1/18 (cit. on pp. 67, 69, 105, 107, 110, 116).
- [106] William J. Potter and Garret Cotter. “Synchrotron and Inverse-Compton Emission from Blazar Jets I: A Uniform Conical Jet Model.” In: *Monthly Notices of the Royal Astronomical Society* 423.1 (June 2012), pp. 756–765. ISSN: 00358711. DOI: 10.1111/j.1365-2966.2012.20918.x. arXiv: 1203.3881 (cit. on pp. 78, 100, 133).
- [107] E. Clausen-Brown, T. Savolainen, A. B. Pushkarev, Y. Y. Kovalev, and J. A. Zensus. “Causal Connection in Parsec-Scale Relativistic Jets: Results from the MOJAVE VLBI Survey.” In: *Astronomy & Astrophysics* 558 (Oct. 2013), A144. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201322203 (cit. on p. 83).
- [108] Svetlana G. Jorstad et al. “Kinematics of Parsec-scale Jets of Gamma-Ray Blazars at 43 GHz within the VLBA-BU-BLAZAR Program.” In: *The Astrophysical Journal* 846.2 (Sept. 2017), p. 98. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aa8407 (cit. on p. 83).
- [109] T. Hovatta et al. “MOJAVE: Monitoring of Jets in Active Galactic Nuclei with VLBA Experiments. VIII. Faraday Rotation in Parsec-scale AGN Jets.” In: *\aj* 144 (Oct. 2012), p. 105. DOI: 10.1088/0004-6256/144/4/105 (cit. on p. 87).
- [110] F. Acero et al. “Fermi Large Area Telescope Third Source Catalog.” In: *\apjs* 218 (June 2015), p. 23. DOI: 10.1088/0067-0049/218/2/23 (cit. on pp. 88, 89).

- [111] T. E. Strohmayer and T. R. Kallman. “ON THE STATISTICAL ANALYSIS OF X-RAY POLARIZATION MEASUREMENTS.” In: *The Astrophysical Journal* 773.2 (July 2013), p. 103. ISSN: 0004-637X. DOI: 10.1088/0004-637X/773/2/103 (cit. on p. 89).
- [112] E. Rivers, A. Markowitz, and R. Rothschild. “Full Spectral Survey of Active Galactic Nuclei in the Rossi X-ray Timing Explorer Archive.” In: *\apj* 772 (Aug. 2013), p. 114. DOI: 10.1088/0004-637X/772/2/114 (cit. on p. 89).
- [113] C. Romoli, N. Chakraborty, D. Dorner, A. Taylor, and M. Blank. “Flux Distribution of Gamma-Ray Emission in Blazars: The Example of Mrk 501.” In: *Galaxies* 6 (Dec. 2018), p. 135. DOI: 10.3390/galaxies6040135 (cit. on p. 89).
- [114] Z. Shah et al. “Log-Normal Flux Distribution of Bright Fermi Blazars.” In: *Research in Astronomy and Astrophysics* 18 (Oct. 2018), p. 141. DOI: 10.1088/1674-4527/18/11/141 (cit. on p. 89).
- [115] Y. E. Lyubarskii. “Flicker Noise in Accretion Discs.” In: *\mnras* 292 (Dec. 1997), p. 679. DOI: 10.1093/mnras/292.3.679 (cit. on p. 89).
- [116] A. Sinha et al. “The Flux Distribution of Individual Blazars as a Key to Understand the Dynamics of Particle Acceleration.” In: *\mnras* 480 (Oct. 2018), pp. L116–L120. DOI: 10.1093/mnrasl/sly136 (cit. on p. 89).
- [117] I. Liodakis et al. “Bimodal Radio Variability in OVRO-40 m-Monitored Blazars.” In: *\mnras* 467 (June 2017), pp. 4565–4576. DOI: 10.1093/mnras/stx432 (cit. on p. 89).
- [118] A. P. Marscher. “The Core of a Blazar Jet.” In: *Extragalactic Jets: Theory and Observation from Radio to Gamma Ray*. Ed. by T. A. Rector and D. S. De Young. Vol. 386. Astronomical Society of the Pacific Conference Series. June 2008, p. 437 (cit. on p. 90).
- [119] I. Liodakis, V. Pavlidou, and E. Angelakis. “Detecting the Elusive Blazar Counter-Jets.” In: *\mnras* 465 (Feb. 2017), pp. 180–191. DOI: 10.1093/mnras/stw2723 (cit. on pp. 91, 92).
- [120] V. Pavlidou et al. “The RoboPol Optical Polarization Survey of Gamma-Ray-Loud Blazars.” In: *\mnras* 442 (Aug. 2014), pp. 1693–1705. DOI: 10.1093/mnras/stu904 (cit. on p. 91).
- [121] P. S. Smith et al. “Coordinated Fermi/Optical Monitoring of Blazars and the Great 2009 September Gamma-ray Flare of 3C 454.3.” In: *ArXiv e-prints* (Dec. 2009) (cit. on pp. 91, 94).
- [122] P. Soffitta. “IXPE the Imaging X-ray Polarimetry Explorer.” In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 10397. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Aug. 2017, p. 103970I. DOI: 10.1117/12.2275485 (cit. on p. 94).

- [123] S. Kiehlmann, D. Blinov, T. J. Pearson, and I. Liodakis. “Optical EVPA Rotations in Blazars: Testing a Stochastic Variability Model with RoboPol Data.” In: *\mnras* 472 (Dec. 2017), pp. 3589–3604. DOI: [10.1093/mnras/stx2167](https://doi.org/10.1093/mnras/stx2167) (cit. on pp. 94, 119).
- [124] E. Massaro et al. “The 5th Edition of the Roma-BZCAT. A Short Presentation.” In: *\apss* 357 (May 2015), p. 75. DOI: [10.1007/s10509-015-2254-2](https://doi.org/10.1007/s10509-015-2254-2) (cit. on p. 95).
- [125] Martin C. Weisskopf et al. “The Imaging X-Ray Polarimetry Explorer (IXPE): Pre-Launch.” In: *Journal of Astronomical Telescopes, Instruments, and Systems* 8.2 (Apr. 2022), p. 026002. DOI: [10.1117/1.JATIS.8.2.026002](https://doi.org/10.1117/1.JATIS.8.2.026002) (cit. on pp. 99, 111, 119).
- [126] The IceCube et al. “Multi-Messenger Observations of a Flaring Blazar Coincident with High-Energy Neutrino IceCube-170922A.” In: *Science* 361.6398 (July 2018), eaat1378. ISSN: 0036-8075, 1095-9203. DOI: [10.1126/science.aat1378](https://doi.org/10.1126/science.aat1378). arXiv: [1807.08816 \[astro-ph, physics:hep-ex\]](https://arxiv.org/abs/1807.08816) (cit. on p. 100).
- [127] T. Hovatta et al. “Association of IceCube Neutrinos with Radio Sources Observed at Owens Valley and Metsähovi Radio Observatories.” In: *åp* 650 (June 2021), A83. DOI: [10.1051/0004-6361/202039481](https://doi.org/10.1051/0004-6361/202039481) (cit. on p. 100).
- [128] A. Wierzcholska and H. Siejkowski. “First Hard X-ray Observations of the Blazar S5 0716+714 with NuSTAR during a Multiwavelength Campaign.” In: *Monthly Notices of the Royal Astronomical Society* 458.3 (May 2016), pp. 2350–2359. ISSN: 0035-8711. DOI: [10.1093/mnras/stw425](https://doi.org/10.1093/mnras/stw425) (cit. on p. 100).
- [129] Reuven Y. Rubinstein and Dirk P. Kroese. *The Cross Entropy Method: A Unified Approach To Combinatorial Optimization, Monte-carlo Simulation (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2004. ISBN: 978-0-387-21240-1 (cit. on p. 101).
- [130] Mykel J. Kochenderfer and Tim A. Wheeler. *Algorithms for Optimization*. Cambridge, MA, USA: MIT Press, Mar. 2019. ISBN: 978-0-262-03942-0 (cit. on p. 101).
- [131] Abel Peirson. *Alpv95/SSCPol: V1.3.1*. Zenodo. Apr. 2022. DOI: [10.5281/zenodo.6450411](https://doi.org/10.5281/zenodo.6450411) (cit. on p. 103).
- [132] Haocheng Zhang et al. “Probing the Emission Mechanism and Magnetic Field of Neutrino Blazars with Multiwavelength Polarization Signatures.” In: *The Astrophysical Journal* 876.2 (May 2019), p. 109. ISSN: 0004-637X. DOI: [10.3847/1538-4357/ab158d](https://doi.org/10.3847/1538-4357/ab158d) (cit. on p. 103).
- [133] Melissa Pesce-Rollins, Niccolò Di Lalla, Nicola Omodei, and Luca Baldini. “An Observation-Simulation and Analysis Framework for the Imaging X-ray Polarimetry Explorer (IXPE).” In: *Nuclear Instruments and Methods in Physics Research A* 936 (Aug. 2019), pp. 224–226. ISSN: 0168-9002. DOI: [10.1016/j.nima.2018.10.041](https://doi.org/10.1016/j.nima.2018.10.041) (cit. on p. 103).

- [134] J. F. L. Simmons and B. G. Stewart. “Point and Interval Estimation of the True Unbiased Degree of Linear Polarization in the Presence of Low Signal-to-Noise Ratios.” In: *Astronomy and Astrophysics* 142.1 (Jan. 1985), pp. 100–106. ISSN: 0004-6361 (cit. on p. 105).
- [135] L. Di Gesu et al. *Testing Particle Acceleration Models for BL LAC Jets with the Imaging X-ray Polarimetry Explorer*. Jan. 2022 (cit. on p. 105).
- [136] P. Padovani, F. Oikonomou, M. Petropoulou, P. Giommi, and E. Resconi. “TXS 0506+056, the First Cosmic Neutrino Source, Is Not a BL Lac.” In: *Monthly Notices of the Royal Astronomical Society: Letters* 484.1 (Mar. 2019), pp. L104–L108. ISSN: 1745-3925, 1745-3933. DOI: 10.1093/mnrasl/slz011. arXiv: 1901.06998 (cit. on p. 107).
- [137] F. Tavecchio, M. Landoni, L. Sironi, and P. Coppi. “Probing Dissipation Mechanisms in BL Lac Jets through X-ray Polarimetry.” In: *\mnras* 480.3 (Nov. 2018), pp. 2872–2880. DOI: 10.1093/mnras/sty1491 (cit. on pp. 107, 115).
- [138] Haocheng Zhang et al. “Radiation and Polarization Signatures from Magnetic Reconnection in Relativistic Jets—I. A Systematic Study.” In: *The Astrophysical Journal* 901.2 (Oct. 2020), p. 149. ISSN: 0004-637X, 1538-4357. DOI: 10.3847/1538-4357/abb1b0. arXiv: 2008.09444 (cit. on pp. 107, 119).
- [139] ShuangNan Zhang et al. *The Enhanced X-ray Timing and Polarimetry Mission - eXTP*. Vol. 62. Dec. 2018. DOI: 10.1007/s11433-018-9309-2 (cit. on p. 108).
- [140] L. Di Gesu et al. “Prospects for IXPE and eXTP Polarimetric Archaeology of the Reflection Nebulae in the Galactic Center.” In: *Astronomy and Astrophysics* 643 (Nov. 2020), A52. ISSN: 0004-6361. DOI: 10.1051/0004-6361/202037886 (cit. on p. 108).
- [141] M. Ajello et al. “The Fourth Catalog of Active Galactic Nuclei Detected by the Fermi Large Area Telescope.” In: *The Astrophysical Journal* 892.2 (Apr. 2020), p. 105. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab791e (cit. on p. 110).
- [142] J. Albert et al. “Discovery of Very High Energy γ -Ray Emission from the Low-Frequency-peaked BL Lacertae Object BL Lacertae.” In: *The Astrophysical Journal* 666.1 (Aug. 2007), p. L17. ISSN: 0004-637X. DOI: 10.1086/521550 (cit. on p. 110).
- [143] T. Arlen et al. “RAPID TeV GAMMA-RAY FLARING OF BL LACERTAE.” In: *The Astrophysical Journal* 762.2 (Jan. 2013), p. 92. ISSN: 0004-637X. DOI: 10.1088/0004-637X/762/2/92 (cit. on p. 110).
- [144] Z. R. Weaver et al. “Multiwavelength Variability of BL Lacertae Measured with High Time Resolution.” In: *\apj* 900.2 (Sept. 2020), p. 137. DOI: 10.3847/1538-4357/aba693 (cit. on pp. 110, 112).

- [145] C. M. Raiteri et al. “The Awakening of BL Lacertae: Observations by Fermi, Swift and the GASP-WEBT.” In: *Monthly Notices of the Royal Astronomical Society* 436.2 (Dec. 2013), pp. 1530–1545. ISSN: 0035-8711. DOI: 10.1093/mnras/stt1672 (cit. on pp. 110, 119).
- [146] C. Casadio et al. “The Jet Collimation Profile at High Resolution in BL Lacertae.” In: *Astronomy & Astrophysics* 649 (May 2021), A153. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202039616 (cit. on p. 110).
- [147] R. C. Vermeulen et al. “When Is BL Lac Not a BL Lac?” In: *The Astrophysical Journal* 452 (Oct. 1995), p. L5. ISSN: 0004-637X. DOI: 10.1086/309716 (cit. on p. 110).
- [148] Liang Chen. “Curvature of the Spectral Energy Distributions of Blazars.” In: *The Astrophysical Journal* 788 (June 2014), p. 179. ISSN: 0004-637X. DOI: 10.1088/0004-637X/788/2/179 (cit. on p. 110).
- [149] Ann E. Wehrle et al. “ERRATIC FLARING OF BL LAC IN 2012–2013: MULTIWAVELENGTH OBSERVATIONS.” In: *The Astrophysical Journal* 816.2 (Jan. 2016), p. 53. ISSN: 0004-637X. DOI: 10.3847/0004-637X/816/2/53 (cit. on p. 110).
- [150] Paolo Giommi et al. “X-Ray Spectra, Light Curves and SEDs of Blazars Frequently Observed by Swift.” In: *Monthly Notices of the Royal Astronomical Society* 507.4 (Nov. 2021), pp. 5690–5702. ISSN: 0035-8711. DOI: 10.1093/mnras/stab2425 (cit. on p. 110).
- [151] N Sahakyan and P Giommi. “A 13-Yr-Long Broad-Band View of BL Lac.” In: *Monthly Notices of the Royal Astronomical Society* 513.3 (July 2022), pp. 4645–4656. ISSN: 0035-8711. DOI: 10.1093/mnras/stac1011 (cit. on p. 110).
- [152] R Middei et al. “The First Hard X-ray Spectral Catalogue of Blazars Observed by NuSTAR.” In: *Monthly Notices of the Royal Astronomical Society* 514.3 (Aug. 2022), pp. 3179–3190. ISSN: 0035-8711. DOI: 10.1093/mnras/stac1185 (cit. on pp. 110, 112).
- [153] Luca Baldini et al. “Ixpeobssim: A Simulation and Analysis Framework for the Imaging X-ray Polarimetry Explorer.” In: *SoftwareX* 19 (July 2022), p. 101194. DOI: 10.1016/j.softx.2022.101194 (cit. on pp. 111, 123).
- [154] N. Vlahakis and A. Königl. “Magnetic Driving of Relativistic Outflows in Active Galactic Nuclei. I. Interpretation of Parsec-Scale Accelerations.” In: *\apj* 605 (Apr. 2004), pp. 656–661. DOI: 10.1086/382670 (cit. on p. 114).
- [155] A. P. Marscher et al. “The Inner Jet of an Active Galactic Nucleus as Revealed by a Radio-to- γ -Ray Outburst.” In: *\nat* 452 (Apr. 2008), pp. 966–969. DOI: 10.1038/nature06895 (cit. on p. 114).
- [156] P. A. Hughes, H. D. Aller, and M. F. Aller. “Polarized Radio Outbursts in BL Lacertae. II. The Flux and Polarization of a Piston-Driven Shock.” In: *\apj* 298 (Nov. 1985), pp. 301–315. DOI: 10.1086/163611 (cit. on p. 115).

- [157] Riccardo Middei et al. “X-Ray Polarization Observations of BL Lacertae.” In: *\apjl* 942.1 (Jan. 2023), p. L10. doi: 10.3847/2041-8213/aca281 (cit. on pp. 115, 116, 139).
- [158] S. Kiehlmann et al. “The Time-Dependent Distribution of Optical Polarization Angle Changes in Blazars.” In: *\mnras* 507.1 (Oct. 2021), pp. 225–243. doi: 10.1093/mnras/stab2055 (cit. on p. 116).
- [159] D. Blinov et al. “RoboPol: Optical Polarization-Plane Rotations and Flaring Activity in Blazars.” In: *MNRAS* 457 (Apr. 2016), pp. 2252–2262. issn: 0035-8711. doi: 10.1093/mnras/stw158 (cit. on p. 118).
- [160] Haocheng Zhang, Xuhui Chen, and Markus Böttcher. “Synchrotron Polarization in Blazars.” In: *ApJ* 789 (July 2014), p. 66. issn: 0004-637X. doi: 10.1088/0004-637X/789/1/66 (cit. on p. 119).
- [161] Carolina Casadio et al. “The Magnetic Field Structure in CTA 102 from High-Resolution Mm-VLBI Observations during the Flaring State in 2016–2017.” In: *Astronomy & Astrophysics* 622 (Feb. 2019), A158. issn: 0004-6361, 1432-0746. doi: 10.1051/0004-6361/201834519 (cit. on p. 119).
- [162] M. Lyutikov, V. I. Pariev, and D. C. Gabuzda. “Polarization and Structure of Relativistic Parsec-Scale AGN Jets.” In: *Monthly Notices of the Royal Astronomical Society* 360.3 (July 2005), pp. 869–891. issn: 0035-8711. doi: 10.1111/j.1365-2966.2005.08954.x (cit. on pp. 119, 121, 122, 127).
- [163] B. Glenn Piner, Niraj Pant, and Philip G. Edwards. “The Jets of TeV Blazars at Higher Resolution: 43 GHz and Polarimetric VLBA Observations from 2005–2009.” In: (Sept. 2010). doi: 10.1088/0004-637X/723/2/1150 (cit. on p. 119).
- [164] Alessandro Di Marco et al. *Handling Background in IXPE Polarimetric Data*. Feb. 2023. doi: 10.48550/arXiv.2302.02927. arXiv: 2302.02927 [astro-ph] (cit. on pp. 123, 135).
- [165] F. Feroz, M. P. Hobson, and M. Bridges. “MultiNest: An Efficient and Robust Bayesian Inference Tool for Cosmology and Particle Physics.” In: *Monthly Notices of the Royal Astronomical Society* 398.4 (Oct. 2009), pp. 1601–1614. issn: 00358711, 13652966. doi: 10.1111/j.1365-2966.2009.14548.x. arXiv: 0809.3437 (cit. on p. 126).
- [166] F. Feroz, M. P. Hobson, E. Cameron, and A. N. Pettitt. “Importance Nested Sampling and the MultiNest Algorithm.” In: *The Open Journal of Astrophysics* 2.1 (Nov. 2019), 10.21105/astro.1306.2144. issn: 2565-6120. doi: 10.21105/astro.1306.2144. arXiv: 1306.2144 (cit. on p. 126).
- [167] Fabian Kislat. “Gamma-Ray Polarimetry with the All-sky Medium Energy Gamma-ray Observatory (AMEGO).” In: *AAS/High Energy Astrophysics Division #16*. Vol. 16. Aug. 2017, p. 103.18 (cit. on p. 132).

- [168] F. Xie et al. “A Study of Background for IXPE.” In: *Astroparticle Physics* 128 (Mar. 2021), p. 102566. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2021.102566 (cit. on p. 135).
- [169] Iván Agudo et al. “POLAMI: Polarimetric Monitoring of Active Galactic Nuclei at Millimetre Wavelengths - III. Characterization of Total Flux Density and Polarization Variability of Relativistic Jets.” In: *\mnras* 473.2 (Jan. 2018), pp. 1850–1867. DOI: 10.1093/mnras/stx2437 (cit. on p. 138).
- [170] Iván Agudo et al. “POLAMI: Polarimetric Monitoring of AGN at Millimetre Wavelengths - I. The Programme, Calibration and Calibrator Data Products.” In: *\mnras* 474.2 (Feb. 2018), pp. 1427–1435. DOI: 10.1093/mnras/stx2435 (cit. on p. 138).
- [171] Clemens Thum et al. “POLAMI: Polarimetric Monitoring of Active Galactic Nuclei at Millimetre Wavelengths - II. Widespread Circular Polarization.” In: *\mnras* 473.2 (Jan. 2018), pp. 2506–2520. DOI: 10.1093/mnras/stx2436 (cit. on p. 138).
- [172] Paul T. P. Ho, James M. Moran, and Kwok Yung Lo. “The Submillimeter Array.” In: *The Astrophysical Journal* 616 (Nov. 2004), pp. L1–L6. ISSN: 0004-637X. DOI: 10.1086/423245 (cit. on p. 138).
- [173] Daniel P. Marrone and Ramprasad Rao. “The Submillimeter Array Polarimeter.” In: 7020 (July 2008), 70202B. DOI: 10.1117/12.788677 (cit. on p. 138).
- [174] Koji S. Kawabata et al. “A New Spectropolarimeter at the Dodaira Observatory.” In: *Publications of the Astronomical Society of the Pacific* 111.761 (July 1999), p. 898. ISSN: 1538-3873. DOI: 10.1086/316387 (cit. on p. 139).
- [175] Hiroshi Akitaya et al. “HONIR: An Optical and near-Infrared Simultaneous Imager, Spectrograph, and Polarimeter for the 1.5-m Kanata Telescope.” In: 9147 (Aug. 2014), 91474O. DOI: 10.1117/12.2054577 (cit. on p. 139).
- [176] T. Hovatta et al. “Optical Polarization of High-Energy BL Lacertae Objects.” In: *ap* 596 (Dec. 2016), A78. DOI: 10.1051/0004-6361/201628974 (cit. on p. 139).
- [177] Alan P. Marscher and Svetlana G. Jorstad. “Frequency and Time Dependence of Linear Polarization in Turbulent Jets of Blazars.” In: *Galaxies* 9.2 (Apr. 2021), p. 27. DOI: 10.3390/galaxies9020027 (cit. on p. 139).
- [178] Ioannis Liodakis et al. “Polarized Blazar X-rays Imply Particle Acceleration in Shocks.” In: *\nat* 611.7937 (Nov. 2022), pp. 677–681. DOI: 10.1038/s41586-022-05338-0 (cit. on p. 139).
- [179] Abdollahi Soheila et al. *The Fermi-LAT Light Curve Repository*. Jan. 2023. DOI: 10.48550/arXiv.2301.01607. arXiv: 2301.01607 [astro-ph] (cit. on p. 139).