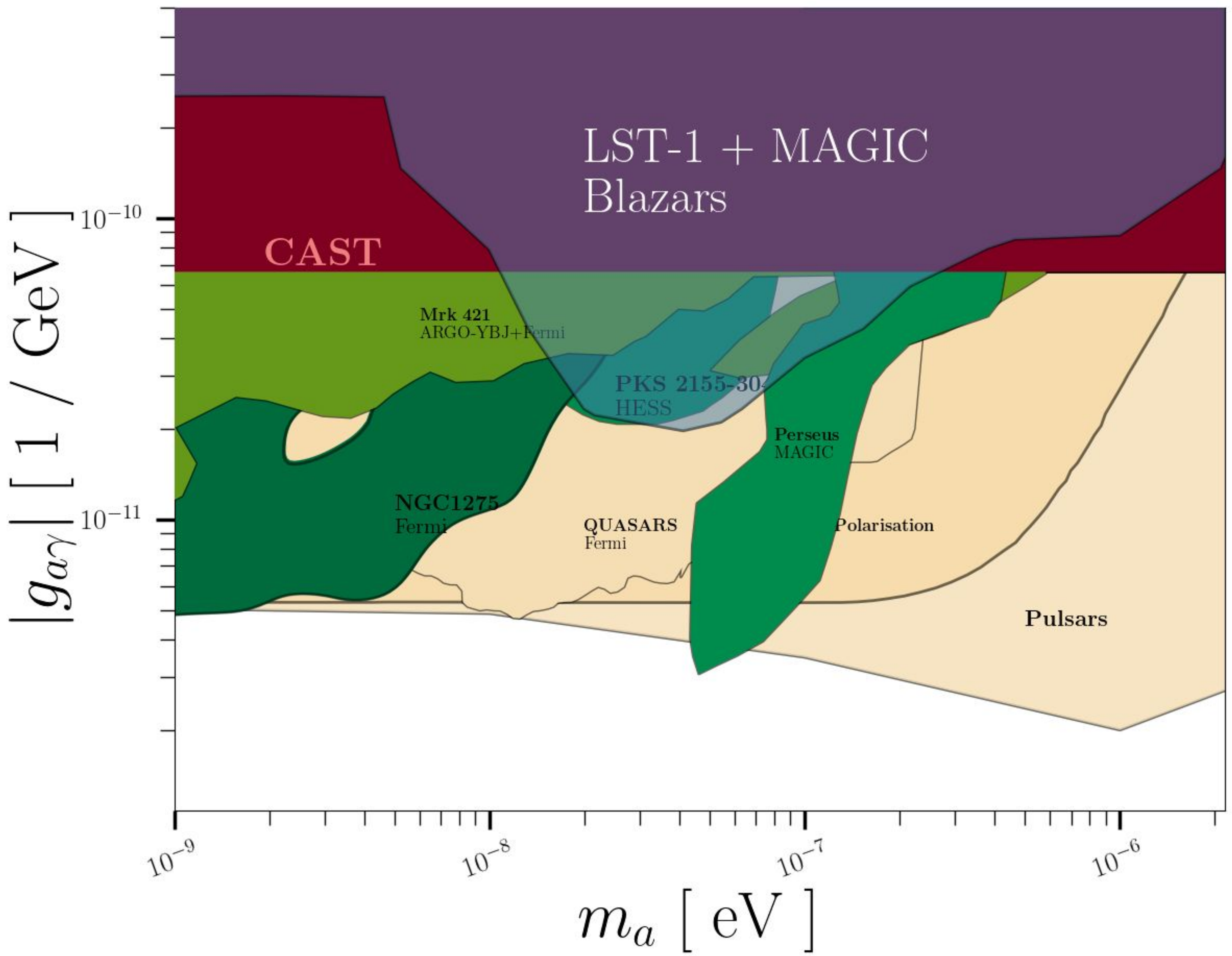
Constraints on axion-like particles from a combined analysis of blazars observed with MAGIC and LST-1

I. Batkovic[a,b], G. D'Amico[c], M. Doro[a,b], F. Schiavone[d]



Structure

Introduction

Axion like particles (ALP) are a common prediction of BSM theories.
VHE photons can oscillate to ALP in magnetic fields (loss of photons), but ALPs can oscillate back to photons (recovery)
ALPs have very low cross-section with matter, so do not interact with EBL during propagation

Oscillation is determined by critical energy dependent on ALP mass, interaction strength

$$E_{\mathrm{crit}} \sim 2.5\ \mathrm{GeV}\ \frac{\left|m_{a,\mathrm{neV}}^2 - \omega_{pl,\mathrm{neV}}^2\right|}{g_{11}\ B_{\mu\mathrm{G}}},$$

Observable:
- around critical energy, interaction is oscillatory and introduce wiggles on the spectrum
- if reconversion, EBL effect may be reduced

Important to study propagation of VHE, ALP through magnetic fields from target to Earth

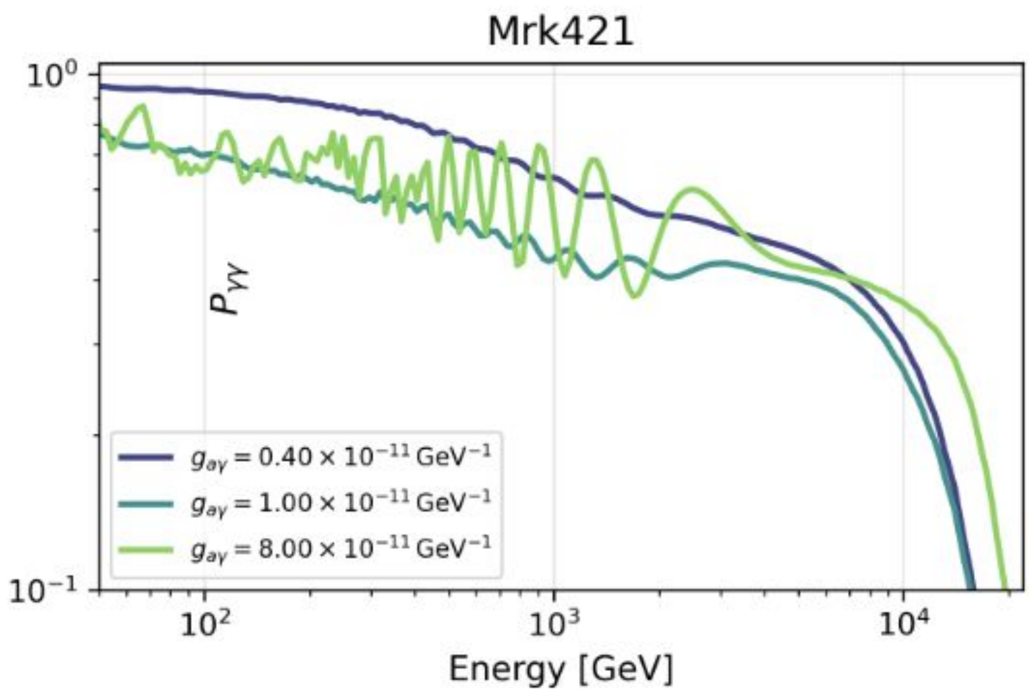


Figure 3: Photon survival probability for tl

Depends on ALP characteristics

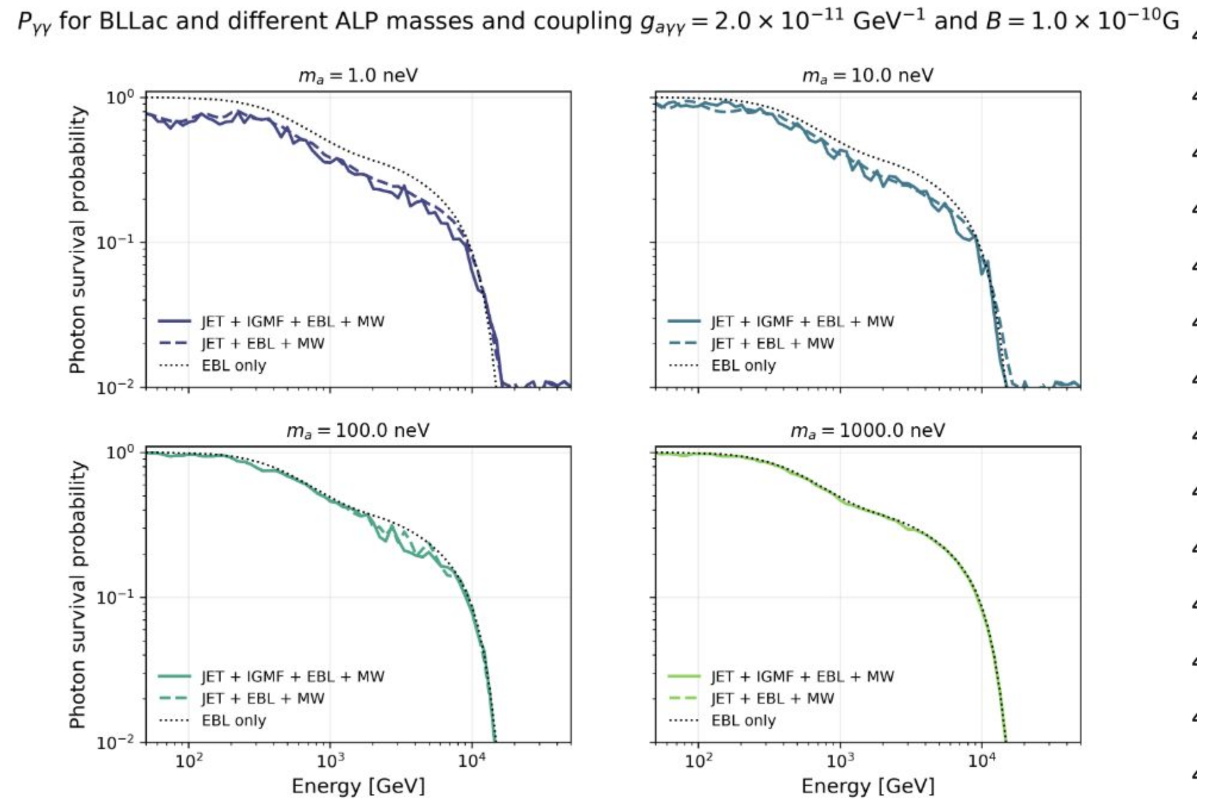


Figure 2: Photon survival probability in the case of BL Lac, assuming $g_{a\gamma\gamma}$ =

Depends on magnetic fields during propagation

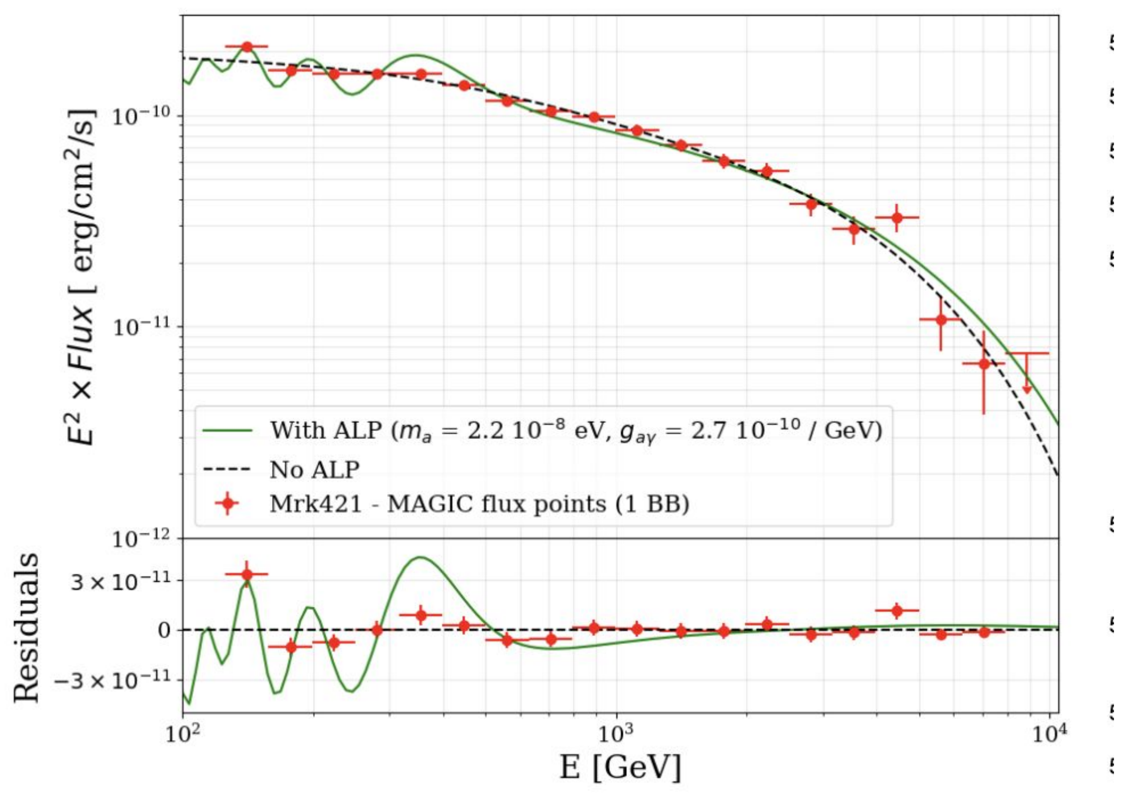


Figure 4: Spectral energy distribution of a representative Bayesian Block

Wiggles on the spectrum: what we study

Magnetic fields along l.o.s.

1. Target blazars: so B-field in the jet
2. Is there a cluster with B-field in the core?
3. Is there IGMF?
4. Is there reconversion in the Galactic B field?

1. ← more relevant
2. ← no, there is not
3. ← the effect is minimal
4. ← well measured, tested different models

Ultrarelativistic jet

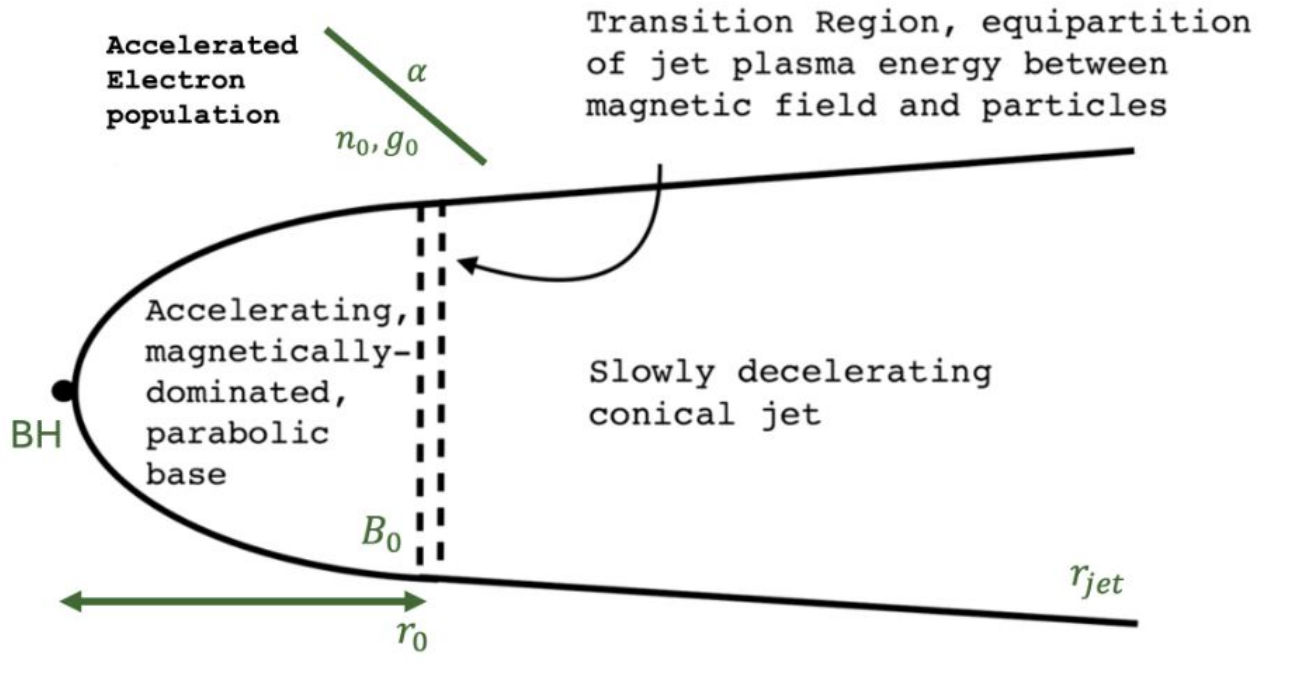


Figure 1: Schematic structure of a relativistic jet (not to scale). The innermost

Model Potter and Cotter, that gives parameters for our targets.

Target	Mrk 421	Mrk 501	BL Lac	1ES 1959+650
r_0 [pc]	6.02	0.3	0.12	0.10
B_0 [G]	0.03	0.8	2.68	1.88
n_0 [cm^{-3}]	$8.5 \cdot 10^3$	$4.5 \cdot 10^4$	$8 \cdot 10^5$	$7 \cdot 10^2$
r_{jet} [10^3 pc]	9.72	3.24	16.20	32.41
g_{min}	9	2	5	4
g_{max}	12	9	8	8
α	1.55	1.68	1.95	1.60

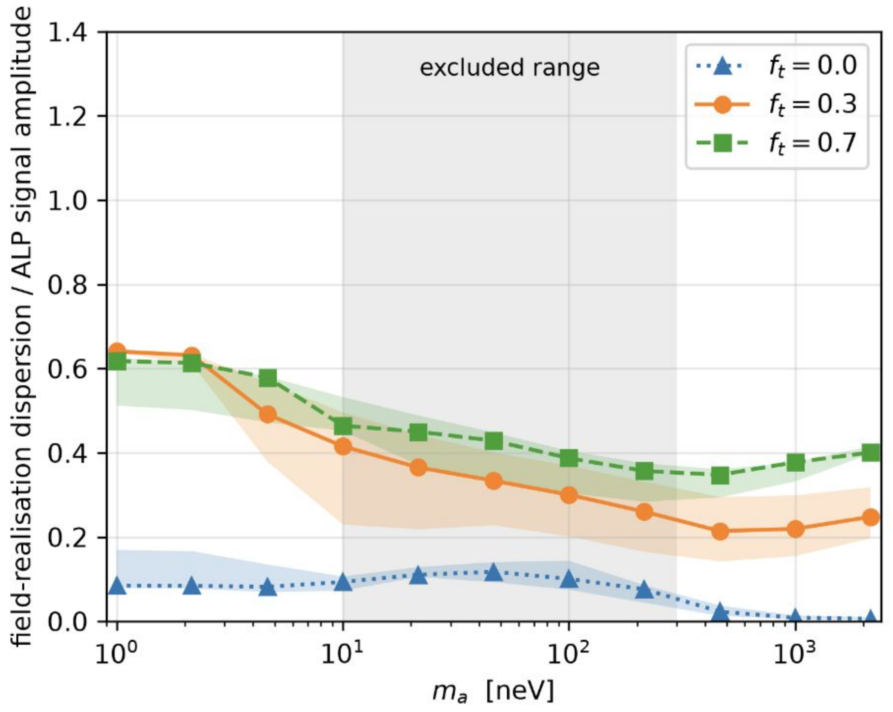
Table 2: Parameters of the jet magnetic field model; radius r_0 where B field is

- PC model is extended and inhomogeneous rather than one-zone: the emission is computed by integrating along the whole jet. r0 is therefore measured rather than assumed
- What governs photon–ALP mixing is the integral of the magnetic field across the propagation length. Thus ALP-photon mixing does not depend on B0 or r0 separately but on their product: for a toroidal field the mixing integral is

$$\int B\,\mathrm{d}l \;=\; B_0\, r_0\, \ln\!\left(\frac{r_{jet}}{r_0}\right)$$

B0 r0 equals 0.18, 0.24, 0.32, 0.19 G.pc for Mrk 421, Mrk 501, BL Lac and 1ES 1959+650 respectively, so that although r0 and B0 vary a lot, across the sample, their product varies by less than a factor of two

- We have also examined how the tangled component affects the ALP signature: a larger tangled fraction does not smear the spectral features
- Also, stochasticity of the tangled component does not significantly impact the results



Other tests in Appendices

Magnetic fields along l.o.s.

Conclusions: all B-fields tested. Results are solid and systematics checked

1. Target blazars: so B-field in the jet
2. Is there a cluster with B-field in the core?
3. Is there IGMF?
4. Is there reconversion in the Galactic B field?

1. ← more relevant
2. ← no, there is not
3. ← the effect is minimal
4. ← well measured, tested different models

Galactic Magnetic Fields

The most recent model is proposed by Unger and Farrar (2023) and consists of eight different realisations. All eight variants are consistent with the present observational data.

The ‘neCL’ realization yields conversion probabilities most similar to those obtained with the (previous standard) Jansson and Farrar (2012) model. Landgraf (2024) notes that the the ‘base’ and ‘expX’ models can be regarded as providing an estimate of the maximal systematic range.

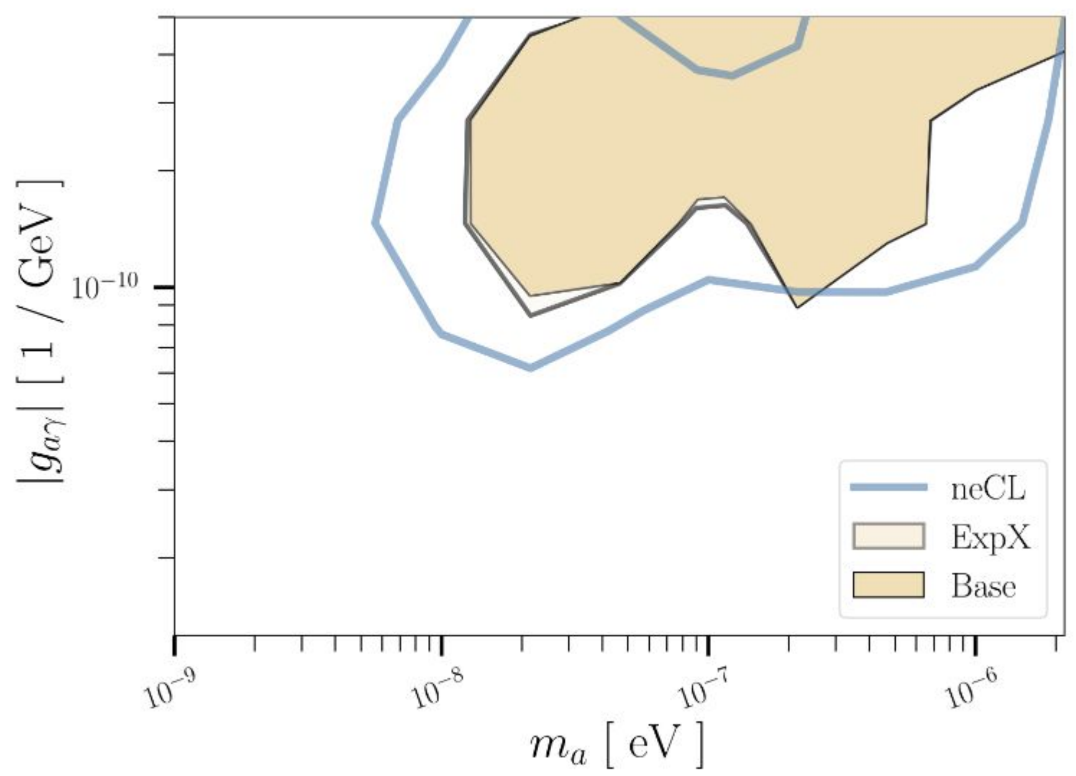


Figure C.15: Impact of the choice of the GMF model realisation on the c

IGMF/EBL

IGMF introduces some additional spectral irregularities in the 1 neV range, while these effects are subleading
EBL, we used Dominguez et al. (2011), however, all sources are close

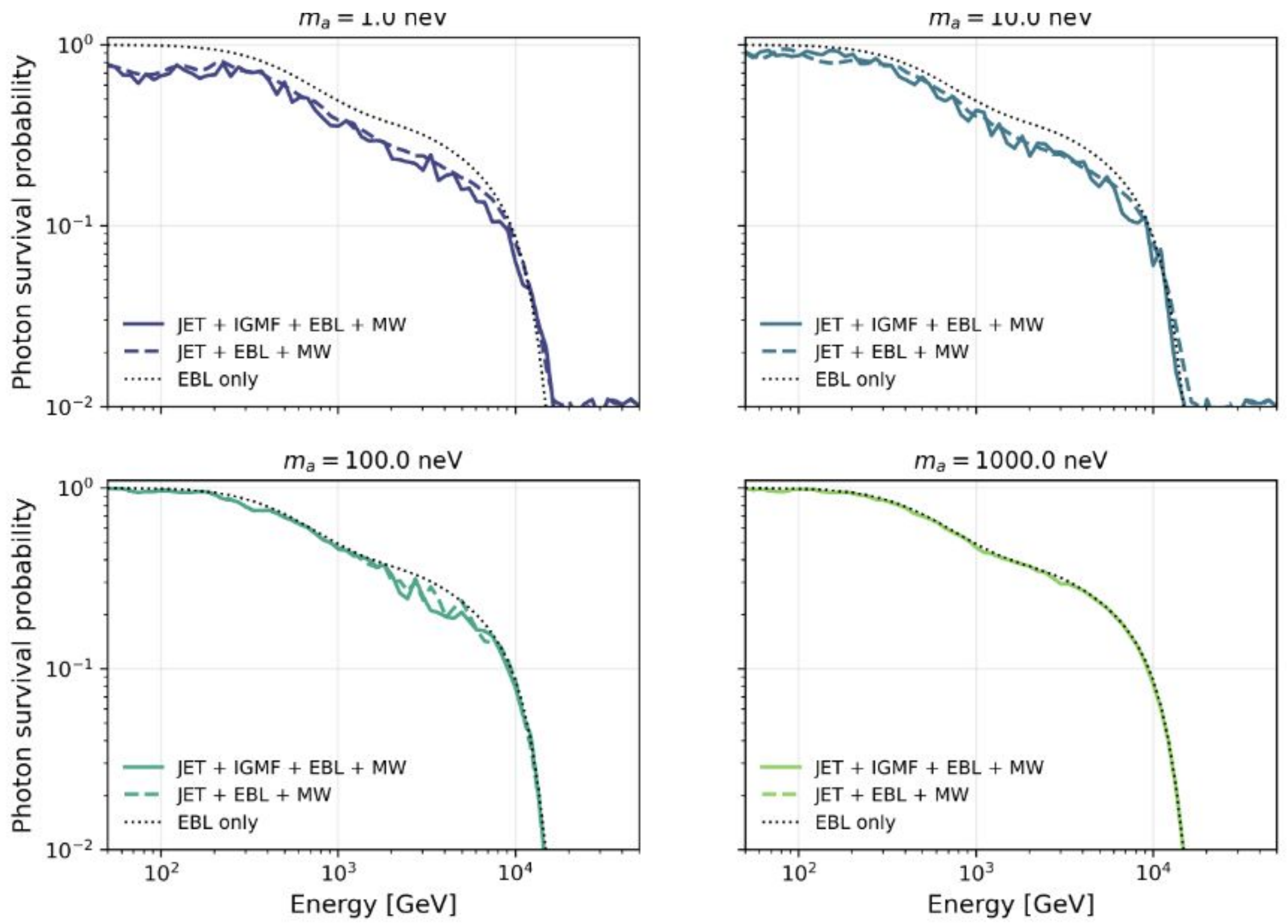


Figure 2: Photon survival probability in the case of BL Lac, assuming $g_{a\gamma\gamma}$ = $2.0 \cdot 10^{-11}\mathrm{GeV}^{-1}$, four different ALP masses and three magnetic-field configurations. The solid lines correspond to photon–ALP mixing in the jet, the IGMF

2.1 Experimental data preparation

Target	Class	z	MAGIC Period	MAGIC Duration	MAGIC σ_{LM}	MAGIC BBs	LST Period	LST Duration	LST σ_{LM}	LST BBs
Mrk 421	HSP	0.031	2014.12–2024.12	263.6 h	438.5	175	2020.12–2024.02	71.7 h	69.5	36
Mrk 501	HSP	0.034	2015.07–2024.09	235.0 h	85.2	55	2020.03–2023.06	55.7 h	17.1	29
BL Lac	LSP	0.069	2015.06–2024.11	90.9 h	101.9	43	2021.08–2022.11	31.8 h	31.6	17
1ES 1959+650	HSP	0.048	2015.03–2024.12	217.1 h	224.2	78	2020.08–2023.09	14.9 h	5.2	9

Table 1: Summary of MAGIC and LST observations. Classes: HSP/LSP denote High/Low Synchrotron Peak blazars. z is the redshift. σ_{LM} is the Test Statistic (global significance in standard deviations relative to the background-only hypothesis), computed following Li and Ma (1983). BBs is the number of Bayesian Blocks (see text).

The observations used in this study combine MAGIC data taken in years 2014-2024 and data from LST-1, overall between 2020–2024. We restrict our analysis to optimal quality data, that is data obtained under dark-sky conditions and with atmospheric transmission above 0.8. Tab. 1 summarizes the complete dataset along with some statistics. This comprises most of the data from the selected targets observed by each experiment, constituting the largest dataset analyzed to date, both for each target and for ground-based gamma-ray ALP searches (Batković et al., 2021). Data sets partially overlapping with those considered here have been used in previous astrophysical studies by the MAGIC and LST-1 collaborations, including, for MAGIC, (e.g. MAGIC, 2015a, 2018, 2020b, 2019, 2020a), and, for LST-1, (LST, 2025).

MAGIC data quality requirements

- max DC 2200 nA
- 5° < ZD < 62°
- TRANS9 ≥ 0.8
- CLOUD < 45
- run duration > 300 s

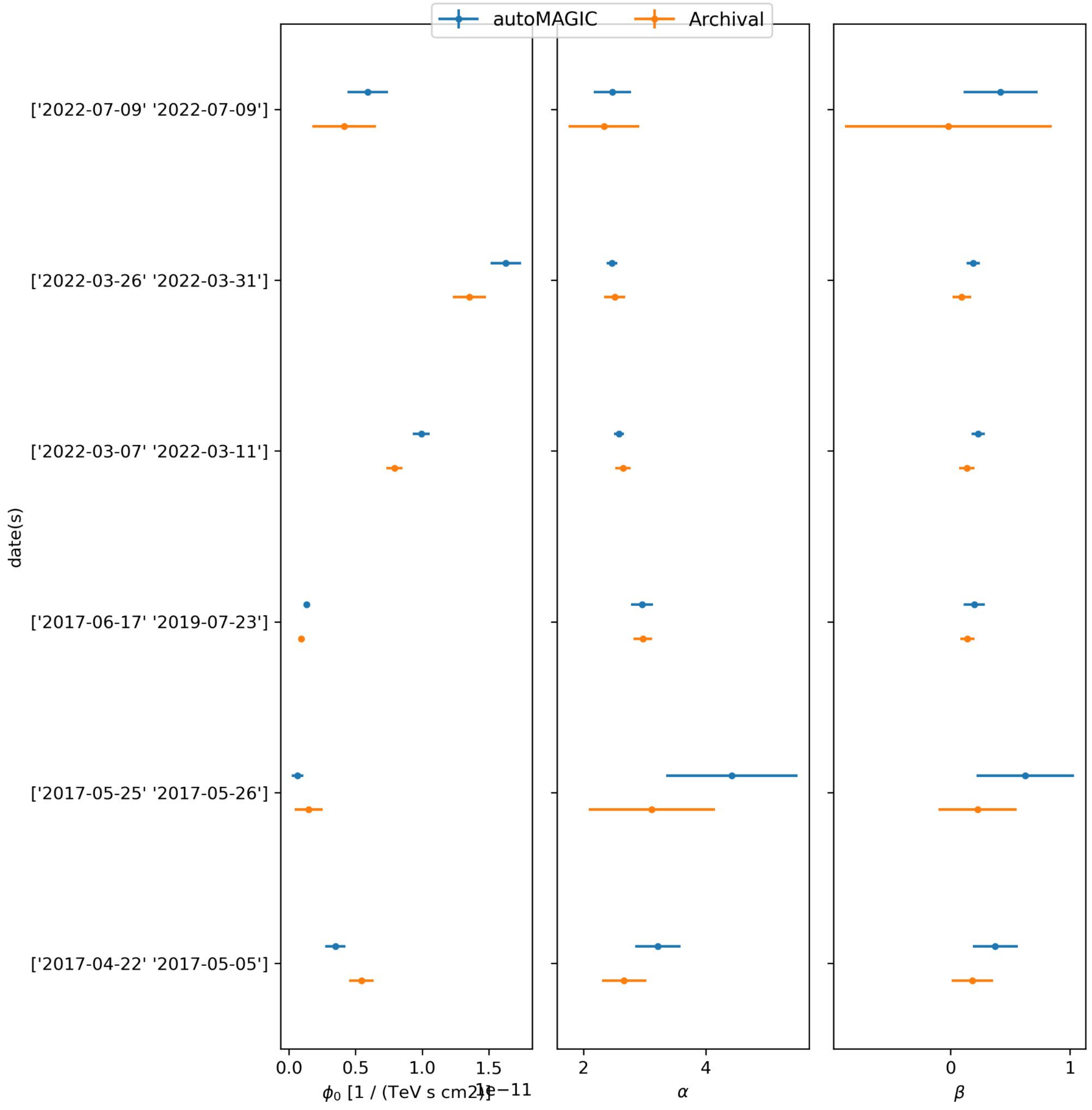
A. The `autoMAGIC` framework

https://gitlab.pic.es/magic/automatic_analysis

The MAGIC data analyses presented in this work employ `autoMAGIC`, a database–driven pipeline for automated MAGIC data reduction, reported in Schubert (2023); Walther et al. (2025). Its design philosophy is to minimize human intervention while ensuring full reproducibility of results. At the core of the system lies a relational SQL database that manages all analysis information through approximately eighty interlinked tables. These include job tables for each MARS executable, data tables for observation runs and Monte Carlo simulations, and logic tables to enforce dependencies and consistency across the workflow.

We cross-checked a subset of `autoMAGIC` datasets against published archival data analyzed with the regular MARS pipeline. Results are not in the paper but are available on the Wiki. (e.g. Mrk 501 on the right)

A few bugs fixed: MR#149, MR#153



2.1 Experimental data preparation

Target	Class	z	MAGIC Period	MAGIC Duration	MAGIC σ_{LM}	MAGIC BBs	LST Period	LST Duration	LST σ_{LM}	LST BBs
Mrk 421	HSP	0.031	2014.12–2024.12	263.6 h	438.5	175	2020.12–2024.02	71.7 h	69.5	36
Mrk 501	HSP	0.034	2015.07–2024.09	235.0 h	85.2	55	2020.03–2023.06	55.7 h	17.1	29
BL Lac	LSP	0.069	2015.06–2024.11	90.9 h	101.9	43	2021.08–2022.11	31.8 h	31.6	17
1ES 1959+650	HSP	0.048	2015.03–2024.12	217.1 h	224.2	78	2020.08–2023.09	14.9 h	5.2	9

Table 1: Summary of MAGIC and LST observations. Classes: HSP/LSP denote High/Low Synchrotron Peak blazars. z is the redshift. σ_{LM} is the Test Statistic (global significance in standard deviations relative to the background-only hypothesis), computed following Li and Ma (1983). BBs is the number of Bayesian Blocks (see text).

The observations used in this study combine MAGIC data taken in years 2014-2024 and data from LST-1, overall between 2020–2024. We restrict our analysis to optimal quality data, that is data obtained under dark-sky conditions and with atmospheric transmission above 0.8. Tab. 1 summarizes the complete dataset along with some statistics. This comprises most of the data from the selected targets observed by each experiment, constituting the largest dataset analyzed to date, both for each target and for ground-based gamma-ray ALP searches (Batković et al., 2021). Data sets partially overlapping with those considered here have been used in previous astrophysical studies by the MAGIC and LST-1 collaborations, including, for MAGIC, (e.g. MAGIC, 2015a, 2018, 2020b, 2019, 2020a), and, for LST-1, (LST, 2025).

LST data **(partially overlaps with the AGN ZOO paper, but larger):**

- quality requirements (standard)

- 0° < ZD < 60°
- max_ped_std=2.3 (dark)
- run duration > 120 s check

2.1 Experimental data preparation

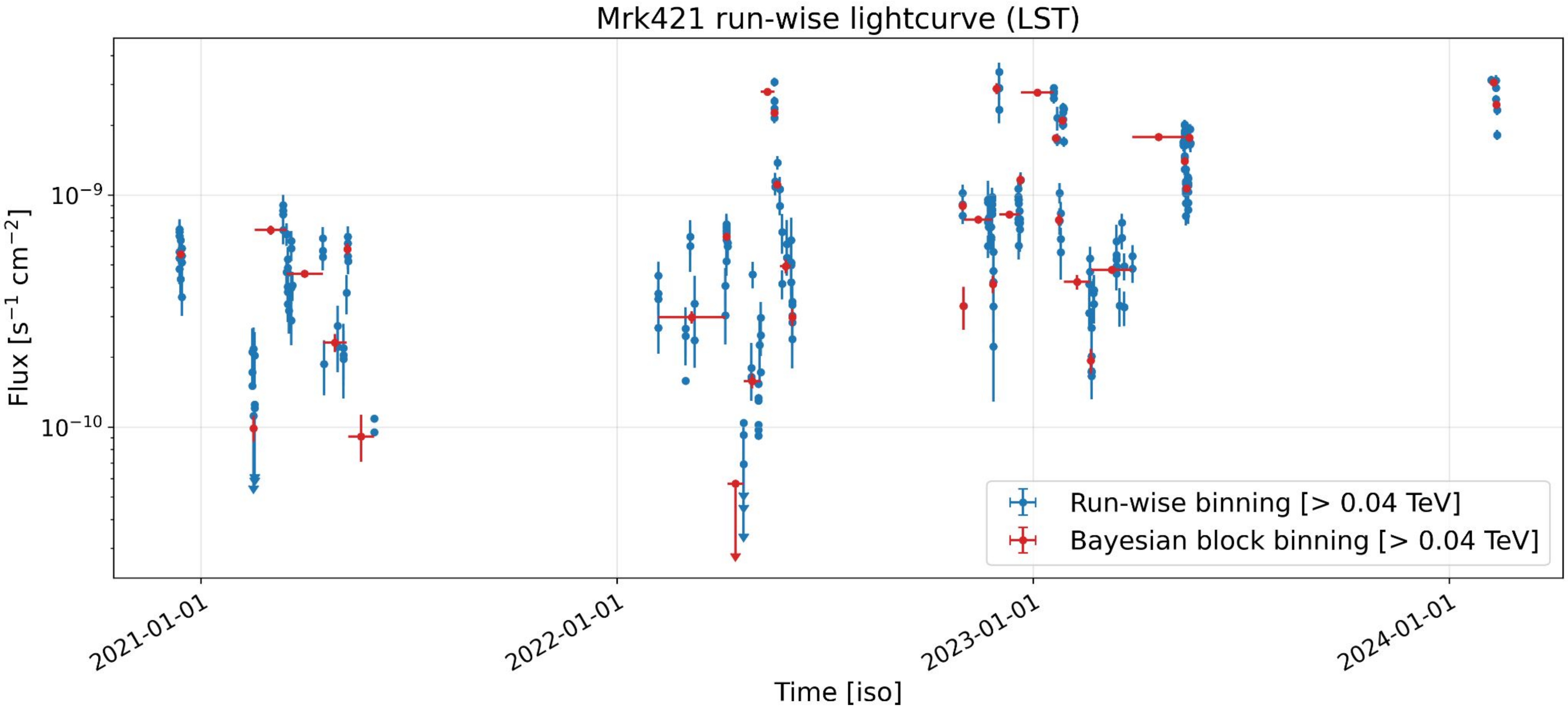
- In total, we are analysing:
  - Mrk 421: 277 runs
  - Mrk 501: 210 runs
  - BL Lac: 121 runs
  - 1ES1959+650: 59 runs
- Initially processed using lstchain v0.10.7 starting from official DL2 and official IRF files available on IT-cluster → after the presentation at the lst-reco meeting in January, due to slight miss-match in the BL Lac data, **we decided to use officially released LST-1 DL3 files**

Spectral analysis

- Energy range [10 GeV, 100 TeV], 10 bins per decade
- Used **`ReflectedRegionsBackgroundMaker`** method with $\alpha = 1$
- Run-wise and night-wise LCs are computed, and used as input for Bayesian Blocks tool in Astropy:
  - fitness: **`measures`**,
  - false_prob: **0.0027**, 3 standard deviations in a normal distribution
  - Groups of data are defined as separated by a minimum amount of 2 months - to avoid long BBs with large runs-less gaps

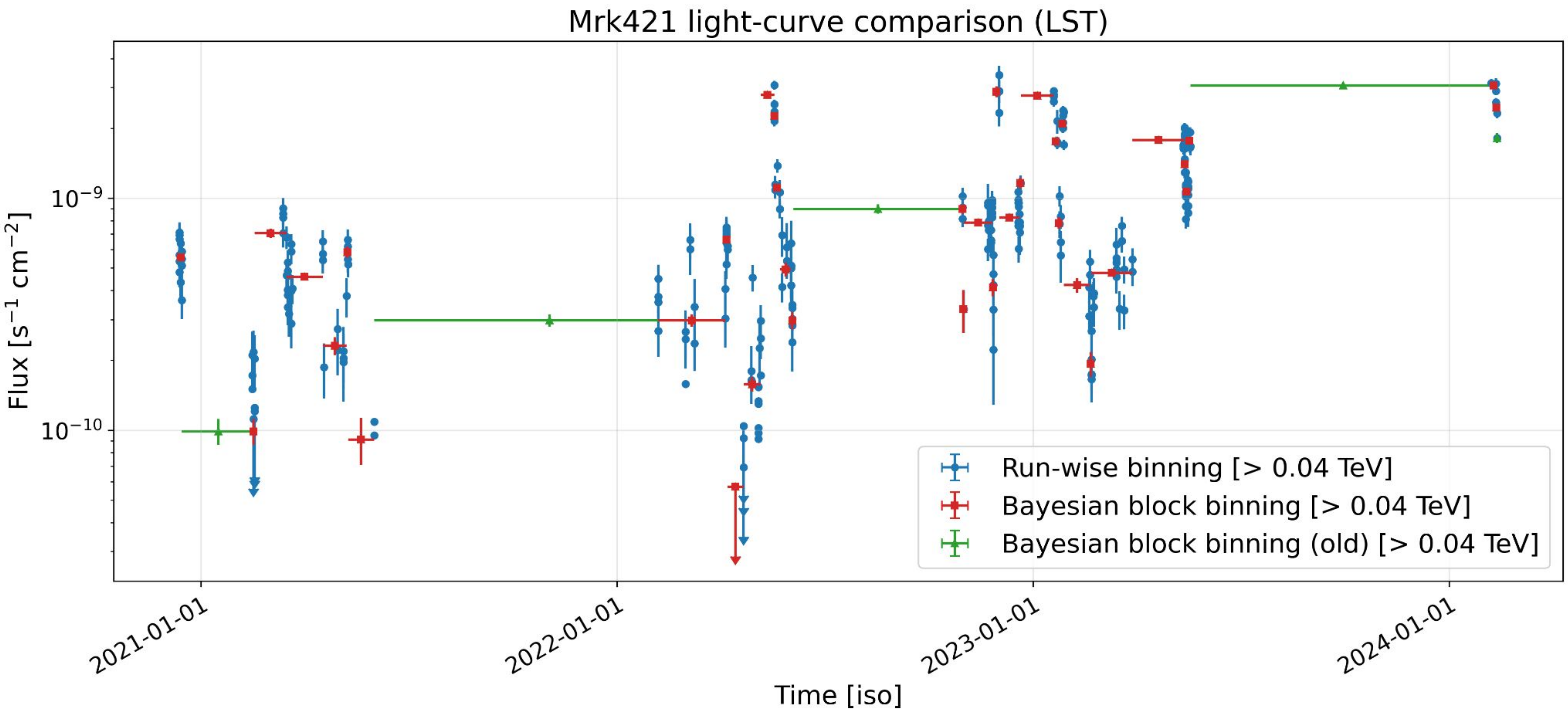
2.1 Experimental data preparation

- The data in each BB were fit to a PWL, ECPL or LP model, accounting for EBL
- Results are found stable under different choices of energy or temporal binning
- Resulting BB are used as the input for the ALP analysis



2.1 Experimental data preparation

- The data in each BB were fit to a PWL, ECPL or LP model, accounting for EBL
- Results are found stable under different choices of energy or temporal binning
- Resulting BB are used as the input for the ALP analysis



2.4 Signal modeling

$$\frac{d\Phi_{\rm est}}{dE} = \frac{d\Phi_{\rm int}}{dE} \cdot P_{\gamma\gamma}\left(E_\gamma; m_a, g_{a\gamma}, B, z\right)$$

$$\boxed{\mathrm{EPWL:}\quad \phi(E) = \phi_0 \left(\frac{E}{E_0}\right)^{-\Gamma} \exp\left(-\frac{E}{E_c}\right)}$$

$$\boxed{\mathrm{LP:}\quad \phi(E) = \phi_0 \left(\frac{E}{E_0}\right)^{-\alpha-\beta\log(E/E_0)}}$$

For each BB, the EPWL or LP is selected as the spectral fit with the best □2 after including EBL absorption and ALP mixing.

Energy range zenith-dependent: the same used for the LST-1 AGN-zoo paper

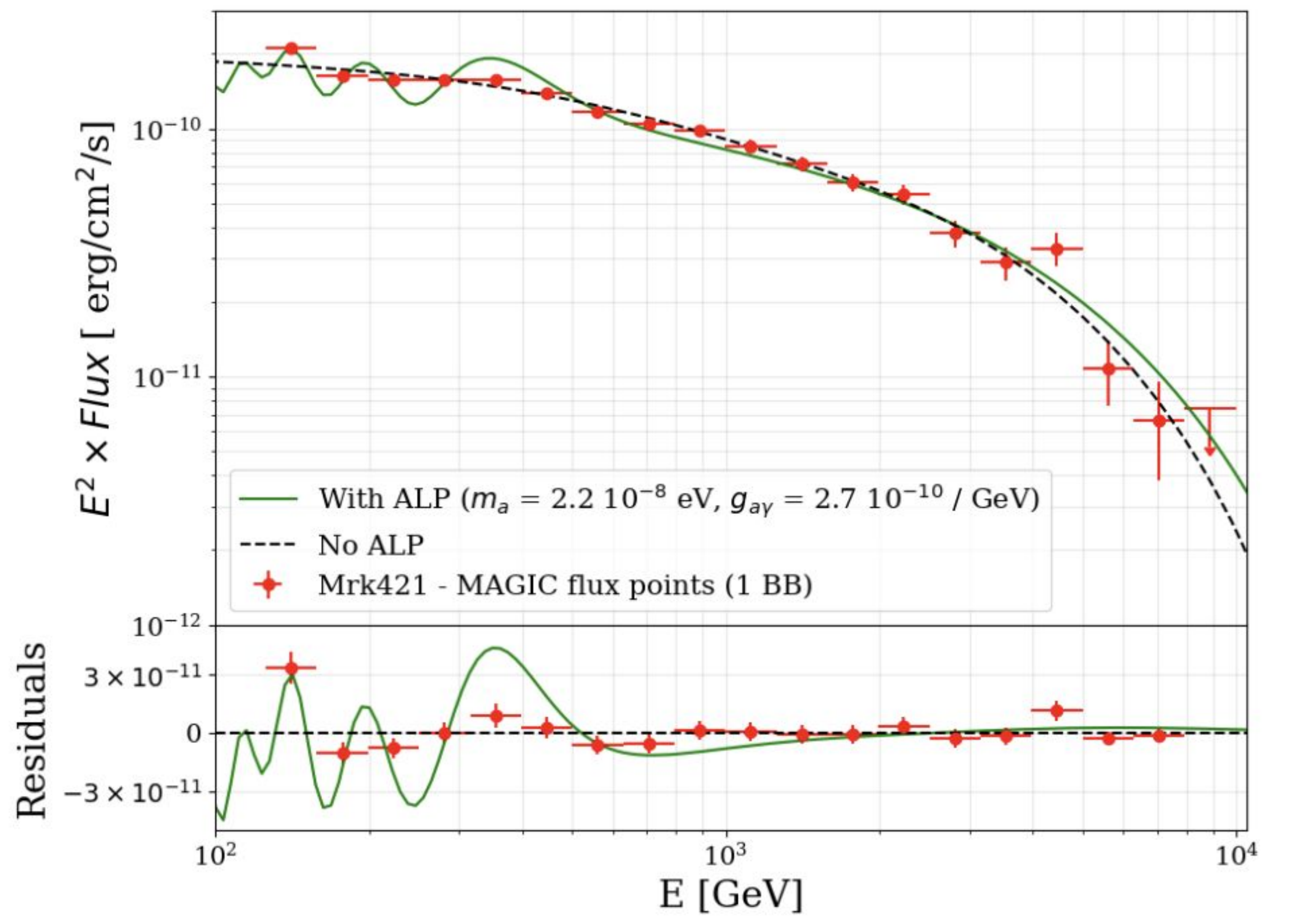


Figure 4: Spectral energy distribution of a representative Bayesian Block from the MAGIC observations of Mrk 421. The red points show the reconstructed flux measurements with their statistical uncertainties. The black dashed line represents the best-fit intrinsic spectral model of Eq. (3) after EBL attenuation, assuming the EBL model of Dominguez et al. (2011) and no ALPs. The green curve shows the corresponding observed spectrum of Eq. (4) including photon–ALP mixing for the benchmark point $(m_a, g_{a\gamma})$ = (2.2 10^{-8} eV, 2.7 10^{-10} /GeV), corresponding to the ALP hypothesis yielding the largest discrepancy with respect to the null hypothesis for this Bayesian Block. The lower panel displays the residuals with respect to the null-hypothesis model.

2.5 Statistical analysis

On/Off Poisson Likelihood

$$\mathscr{L}_i(g_{a\gamma}, m_a; \mu_i) = \prod_{j=1}^{N_i} \left[\frac{\left(s_{ij} + \alpha_i b_{ij}\right)^{n_{ij}}}{n_{ij}!} e^{-(s_{ij}+\alpha_i b_{ij})} \times \frac{b_{ij}^{m_{ij}}}{m_{ij}!} e^{-b_{ij}} \right]$$

i: BB index

j: energy bin index

Expected signal counts

$$s_{ij} = \int_{\Delta E_j} dE \int dE' \, \Phi^i_{\mathrm{int}}(E'; \mu_i) \, P_{\gamma\gamma}(E'; g_{a\gamma}, m_a) \, \mathrm{IRF}^i(E \mid E')$$

2.5 Statistical analysis

On/Off Poisson Likelihood

$$\mathscr{L}_i(g_{a\gamma}, m_a; \mu_i) = \prod_{j=1}^{N_i} \left[\frac{(s_{ij} + \alpha_i b_{ij})^{n_{ij}}}{n_{ij}!} e^{-(s_{ij}+\alpha_i b_{ij})} \times \frac{b_{ij}^{\,m_{ij}}}{m_{ij}!} e^{-b_{ij}} \right]$$

i: BB index

j: energy bin index

Expected signal counts

$$s_{ij} = \int_{\Delta E_j} dE \int dE' \, \Phi^i_{\rm int}(E'; \mu_i) \, P_{\gamma\gamma}(E'; g_{a\gamma}, m_a) \, {\rm IRF}^i(E \mid E')$$

W-stat statistics or “goodness of fit”

$$W_i(g_{a\gamma}, m_a) \equiv -2 \log \mathscr{L}_i(g_{a\gamma}, m_a; \hat{\mu}_i) + 2 \log \hat{\mathscr{L}}_i =$$

$$= 2 \sum_j^{N_i} \left(s_{ij} + (1+\alpha_i)\hat{b}_{ij} - n_{ij} \log(s_{ij} + \alpha_i \hat{b}_{ij}) - m_{ij} \log(\hat{b}_{ij}) - C \right)$$

The statistic $W_i(g_{a\gamma}, m_a)$ is χ^2-distributed with $d_i(g_{a\gamma}, m_a)$ degrees of freedom

$$\boxed{\hat{b}(s) = \frac{n + m - (1 + 1/\alpha)s + \sqrt{(n + m - (1 + 1/\alpha)s)^2 + 4(1 + 1/\alpha)s\, m}}{2(1+\alpha)}}$$

$$\boxed{C = n_{ij} + m_{ij} - n_{ij} \log(n_{ij}) - m_{ij} \log(m_{ij})}$$

Global statistics

$$W(g_{a\gamma}, m_a) = \sum_i W_i(g_{a\gamma}, m_a)$$

with $d(g_{a\gamma}, m_a) = \sum_i d_i(g_{a\gamma}, m_a)$ degrees of freedom

2.5 Statistical analysis

On/Off Poisson Likelihood

$$\mathscr{L}_i(g_{a\gamma}, m_a; \mu_i) = \prod_{j=1}^{N_i} \left[\frac{(s_{ij} + \alpha_i b_{ij})^{n_{ij}}}{n_{ij}!} e^{-(s_{ij}+\alpha_i b_{ij})} \times \frac{b_{ij}^{m_{ij}}}{m_{ij}!} e^{-b_{ij}} \right]$$

i: BB index

j: energy bin index

Expected signal counts

$$s_{ij} = \int_{\Delta E_j} dE \int dE' \, \Phi^i_{\rm int}(E'; \mu_i)\, P_{\gamma\gamma}(E'; g_{a\gamma}, m_a)\, \mathrm{IRF}^i(E \mid E')$$

W-stat statistics or “goodness of fit”

$$W_i(g_{a\gamma}, m_a) \equiv -2 \log \mathscr{L}_i(g_{a\gamma}, m_a; \hat{\mu}_i) + 2 \log \hat{\mathscr{L}}_i =$$

$$= 2 \sum_j^{N_i} \left(s_{ij} + (1+\alpha_i)\hat{b}_{ij} - n_{ij} \log(s_{ij} + \alpha_i \hat{b}_{ij}) - m_{ij} \log(\hat{b}_{ij}) - C \right)$$

The statistic $W_i(g_{a\gamma}, m_a)$ is χ^2-distributed with $d_i(g_{a\gamma}, m_a)$ degrees of freedom

$$\boxed{\hat{b}(s) = \frac{n + m - (1+1/\alpha)s + \sqrt{(n+m-(1+1/\alpha)s)^2 + 4(1+1/\alpha)s\, m}}{2(1+\alpha)}}$$

$$\boxed{C = n_{ij} + m_{ij} - n_{ij}\log(n_{ij}) - m_{ij}\log(m_{ij})}$$

Global statistics

$$W(g_{a\gamma}, m_a) = \sum_i W_i(g_{a\gamma}, m_a)$$ with $d(g_{a\gamma}, m_a) = \sum_i d_i(g_{a\gamma}, m_a)$ degrees of freedom

A given point in the ALP parameter space $(g_{a\gamma}, m_a)$ can be excluded with a statistical significance quantified by the p-value:

$$p(g_{a\gamma}, m_a) = Q\left(\frac{d(g_{a\gamma}, m_a)}{2}, \frac{W(g_{a\gamma}, m_a)}{2} \right) \rightarrow \sigma(g_{a\gamma}, m_a) = \sqrt{2}\, \mathrm{erf}^{-1}\left(1 - \mathrm{p}(\mathrm{g}_{a\gamma}, \mathrm{m_a})\right)$$

$$\boxed{Q(s,x) = \frac{1}{\Gamma(s)} \int_x^\infty t^{s-1} e^{-t} dt}$$ regularized upper incomplete gamma function

Appendix B.2 How do we get the degrees of freedoms ?

MC analysis done for each BB and ALP point

1. **Fit to real data.** We perform the fit described in App. 2.5, for the chosen $(g_{a\gamma}, m_a)$, obtaining the expected signal counts s_{ij} and background estimates $\hat{b}_{ij}(s_{ij})$. These represent the model prediction under the hypothesis that the ALP parameters $(g_{a\gamma}, m_a)$ are the true ones.
2. **Generate MC realizations.** We generate synthetic ON/OFF observations by drawing

$$n_{ij}^{\mathrm{MC}} \sim \mathrm{Poisson}(s_{ij} + \alpha_i \hat{b}_{ij}), \qquad m_{ij}^{\mathrm{MC}} \sim \mathrm{Poisson}(\hat{b}_{ij}),$$

   i.e. we simulate what would be observed if the ALP hypothesis were true.
3. **Refit each MC realization.** For each simulated dataset, we refit the SED parameters $\boldsymbol{\mu}_i$ and compute the corresponding statistic $W_i^{\mathrm{MC}}(g_{a\gamma}, m_a)$.
4. **Extract the effective d.o.f.** The resulting distribution of W_i^{MC} is well described by a χ^2 distribution, as illustrated in Fig. B.10, where both the null and an ALP hypothesis are shown for comparison. Its effective degrees of freedom $d_i(g_{a\gamma}, m_a)$ are given by the expected value of the distribution,

$$d_i(g_{a\gamma}, m_a) \equiv \mathrm{E}\left[W_i^{\mathrm{MC}}(g_{a\gamma}, m_a)\right],$$

   as predicted for a χ^2 random variable where $\mathrm{E}[\chi^2] = d$ is the number of degrees of freedom.

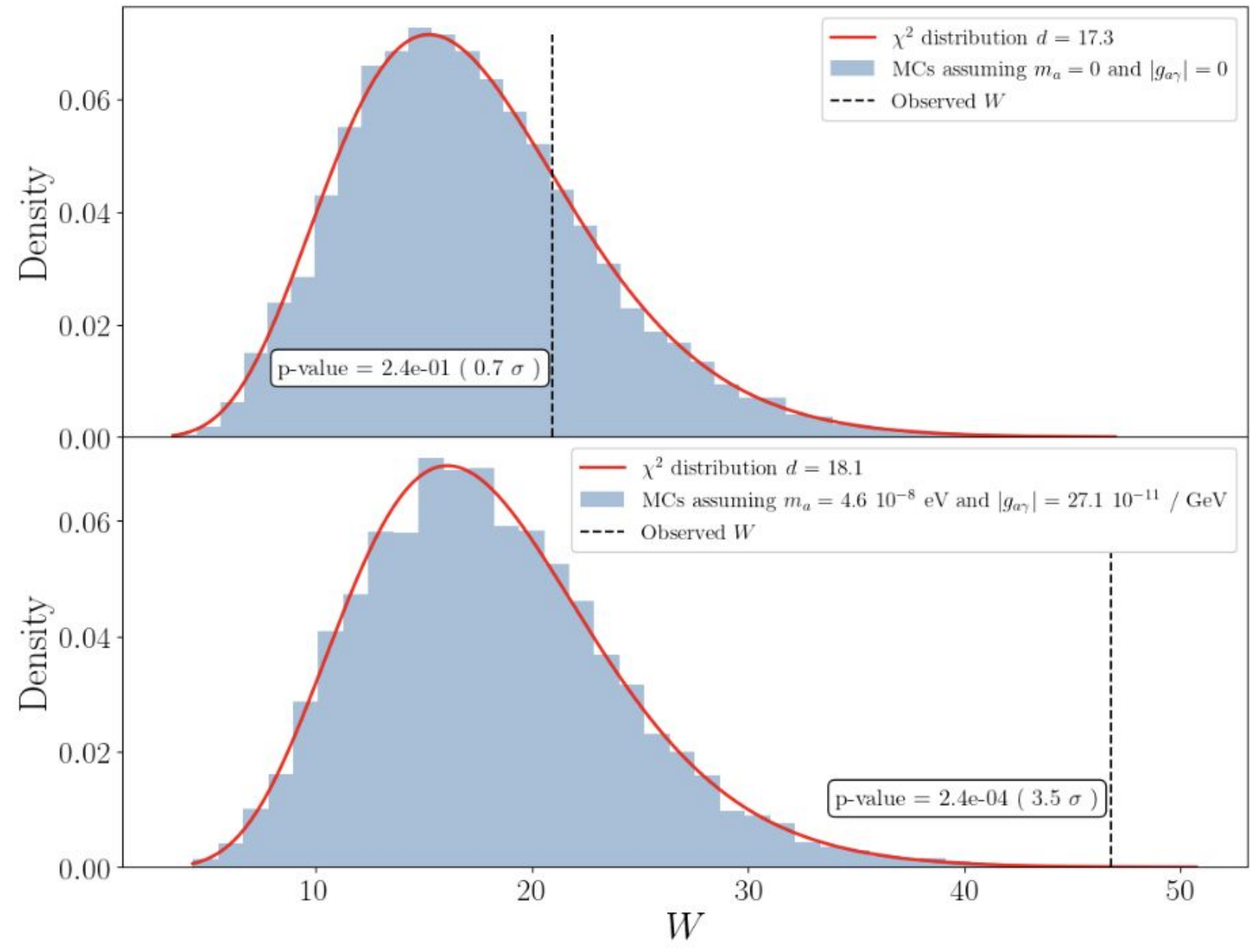


Figure B.10: Distribution of the MC-simulated test statistic W_i^{MC} for a given BB under two different hypotheses. **Top:** Null hypothesis with $m_a = 0$ and $g_{a\gamma} = 0$. **Bottom:** ALP hypothesis with $m_a = 4.6 \cdot 10^{-8}\,\mathrm{eV}$ and $|g_{a\gamma}| = 27.1 \cdot 10^{-11}\,\mathrm{GeV}^{-1}$. The blue histograms show the W_i^{MC} values obtained from pseudo-experiments, while the red curves represent the best-fit χ^2 distributions. The vertical dashed lines indicate the observed W_i value from real data. The reported p-values and corresponding Gaussian significances quantify the level of disagreement between the model and data.

Appendix B.2 Treatment of blocks with poor intrinsic SED fits

Among the 442 BBs we have (as expected) encountered some BBs whose intrinsic SED fit gives a “bad” chi-square value

Possible solution: discard those “bad” BBs whose null-hypothesis fit showed a discrepancy larger than 2σ with respect to the χ^2 expectation

Criticism: discarding BBs requires defining what constitutes a “bad” block (e.g. adopting a 2σ threshold), which is an intrinsically arbitrary choice

The **shift method** avoids this ambiguity and incorporates the observed discrepancy in a well-defined and continuous manner
(the resulting exclusion limits were nevertheless consistent with both approaches)

Appendix B.2 Treatment of blocks with poor intrinsic SED fits

Among the 442 BBs we have (as expected) encountered some BBs whose intrinsic SED fit gives a “bad” chi-square value

Possible solution: discard those “bad” BBs whose null-hypothesis fit showed a discrepancy larger than 2σ with respect to the □2 expectation

Criticism: discarding BBs requires defining what constitutes a “bad” block (e.g. adopting a 2σ threshold), which is an intrinsically arbitrary choice

The **shift method** avoids this ambiguity and incorporates the observed discrepancy in a well-defined and continuous manner
(the resulting exclusion limits were nevertheless consistent with both approaches)

First, we verify that ALPs do not improve the fit, in other words, **the mismatch is due to the spectral model, not to missing ALP effects**

Then, we add this **mismatch** that we observed in the data to the simulation as well

$$W^{\mathrm{MC}}(g_{a\gamma}, m_a) \rightarrow W^{\mathrm{MC}}(g_{a\gamma}, m_a) + \Delta$$

$$\Delta = W^{\mathrm{data}}(0{,}0) - \bar{W}^{\mathrm{MC}}(0{,}0)$$

Global statistics under the null hypothesis from the data with the “bad” BBs

Mean global statistics under the null hypothesis from the MC simulation, which by default cannot have bad BB

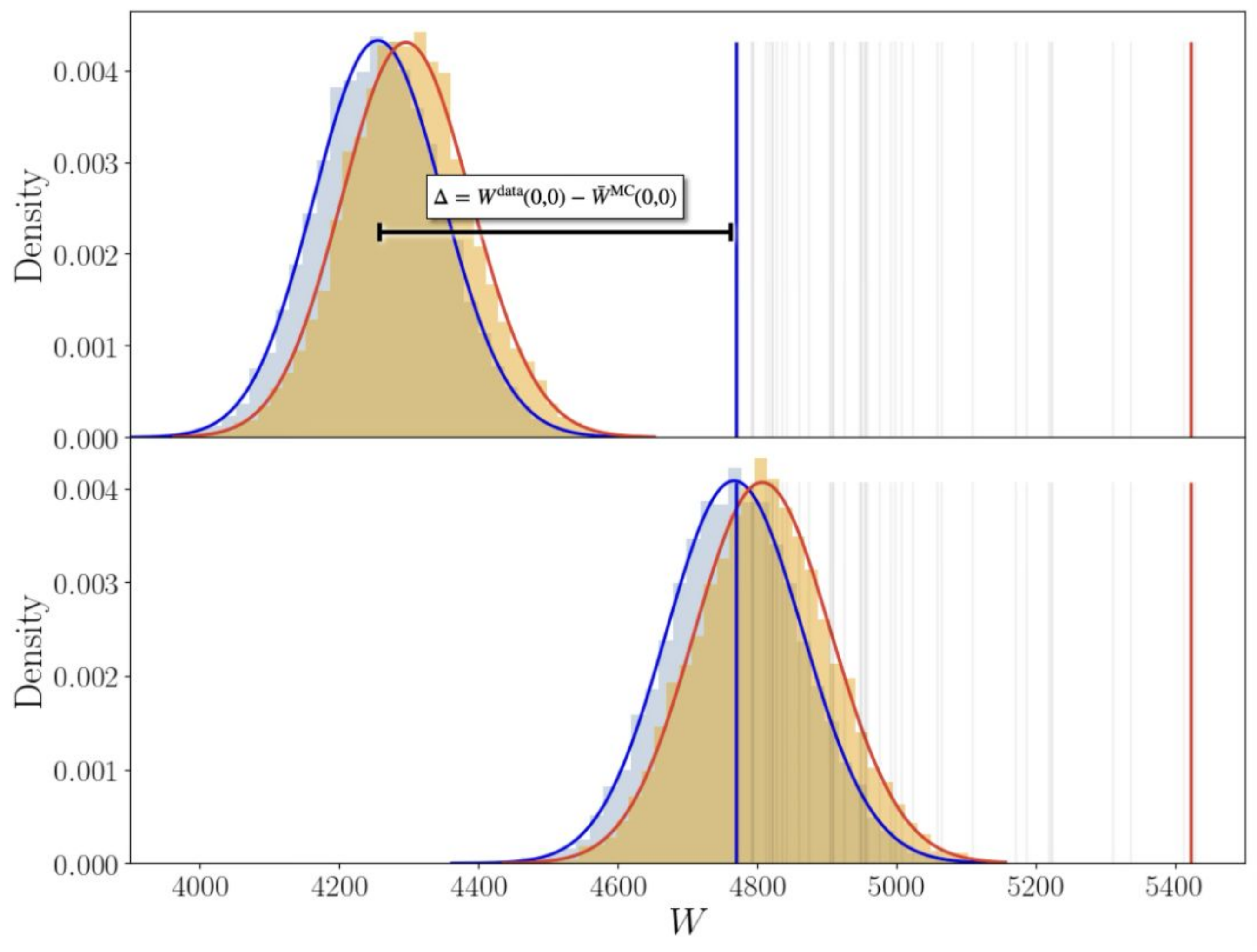


Figure B.11: Illustration of the correction applied to MC-simulated distr butions of the test statistic W. The histograms show the distribution of W^{M} values obtained under two hypotheses: the null hypothesis (blue) with m_a = and $g_{a\gamma} = 0$, and an ALP benchmark hypothesis (red) with $m_a = 2.2 \cdot 10^{-8}$ e and $|g_{a\gamma}| = 0.4 \cdot 10^{-11}\ \mathrm{GeV}^{-1}$. Solid curves represent the best-fit χ^2 distr butions, while vertical lines indicate the corresponding W^{data} values for eac hypothesis (blue for null, red for benchmark ALP). Light grey vertical line show observed W values for other ALP hypotheses. **Top:** Original (unco rected) MC distributions. **Bottom:** Same distributions after applying a shi $\Delta = W^{\mathrm{data}}(0,0) - \bar{W}^{\mathrm{MC}}(0,0)$ that aligns the simulated null distribution with th observed value.

Validation of the shift prescription

We first generated ON/OFF counts under the null hypothesis of no ALP signal and performed the complete analysis on this fake dataset, obtaining the corresponding exclusion

We then repeated the procedure after artificially increasing or decreasing the signal counts in selected energy bins, thereby introducing a spectral mismatch analogous to that of a BB with a poor intrinsic SED fit. From this, we obtain exclusions using the “shift” procedure.

The two exclusion regions were found to be **consistent**

Appendix C Systematic uncertainty in the effective collection area

In this appendix, we tried to quantify how uncertainties in the collection area may affect the ALP limits
As for other systematic studies we tested it on the 10 “best” BBs

Additional nuisance parameter $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma_\varepsilon)$ $\quad s_{ij} \longrightarrow s_{ij}(1 + \varepsilon_{ij})$

i: BB index

j: energy bin index

New likelihood:

$$\mathcal{L}_i(g_{a\gamma}, m_a; \mu_i) = \prod_{j=1}^{N_i} \left[\frac{\left(s_{ij}(1+\varepsilon_{ij}) + \alpha_i b_{ij}\right)^{n_{ij}}}{n_{ij}!} e^{-s_{ij}(1+\varepsilon_{ij}) - \alpha_i b_{ij}} \times \frac{b_{ij}^{m_{ij}}}{m_{ij}!} e^{-b_{ij}} \right] \exp\left(-\frac{\varepsilon_{ij}^2}{2\sigma_\varepsilon^2}\right)$$

New statistics:

$$W_\epsilon(g_{a\gamma}, m_a) = -2 \ln \mathcal{L}\left(g_{a\gamma}, m_a; \hat{\mu}, \hat{\varepsilon}\right) + 2 \ln \hat{\hat{\mathcal{L}}}$$

$$\boxed{\alpha\, \hat{\varepsilon}^3 + A\, \hat{\varepsilon}^2 + B\, \hat{\varepsilon} + C = 0}$$

Repeating the analysis with the new statistic assuming an uncertainty of 10% per energy bin in the effective area (i.e. σ=0.1)

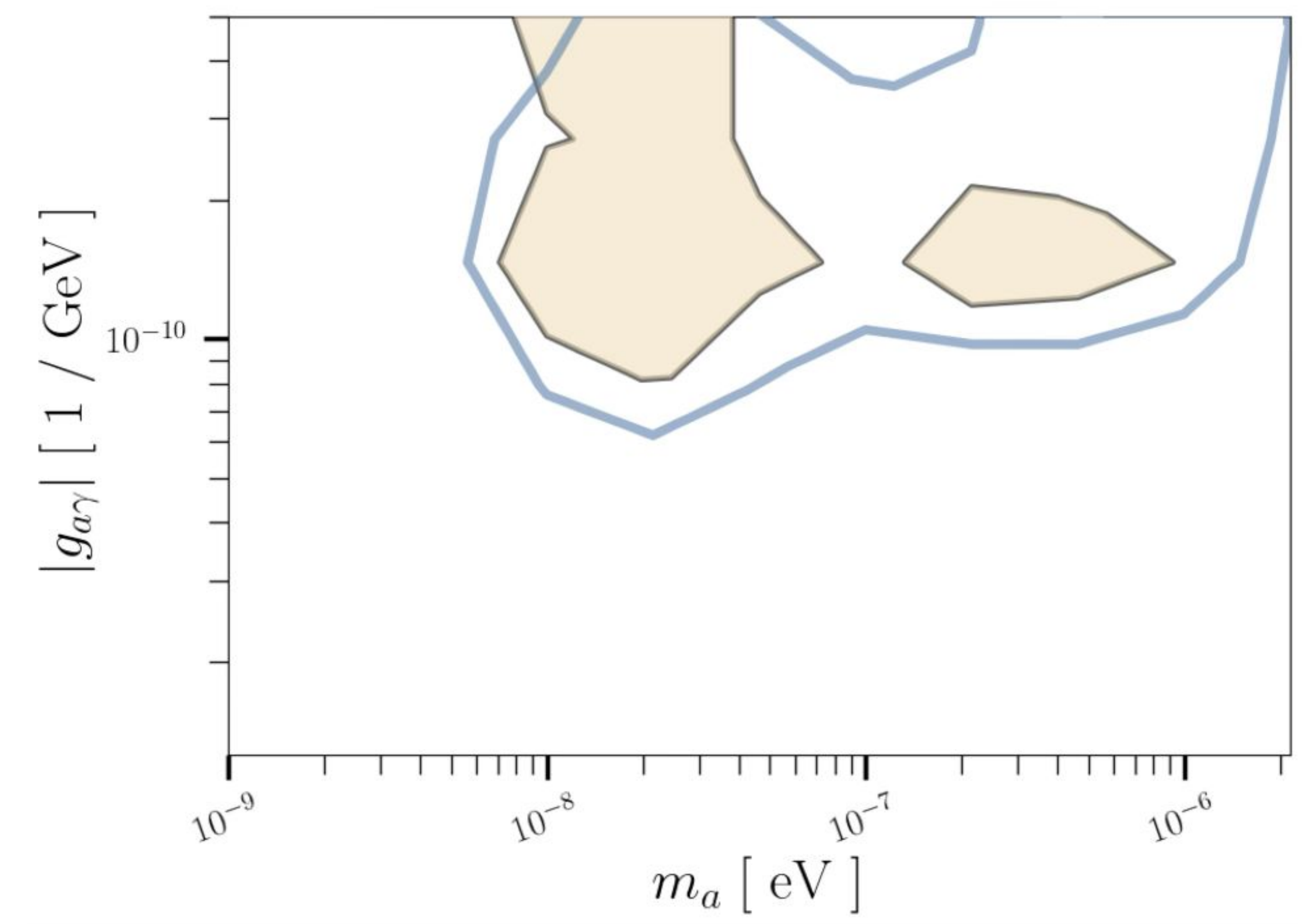


Figure D.16: The blue contour shows the nominal exclusion obtained with the standard likelihood, while the orange contour is obtained using the modified likelihood of Eq. (D.3), assuming an independent 10% uncertainty ($\sigma = 0.1$) on the expected signal yield in each energy bin.

3. Results and Discussion

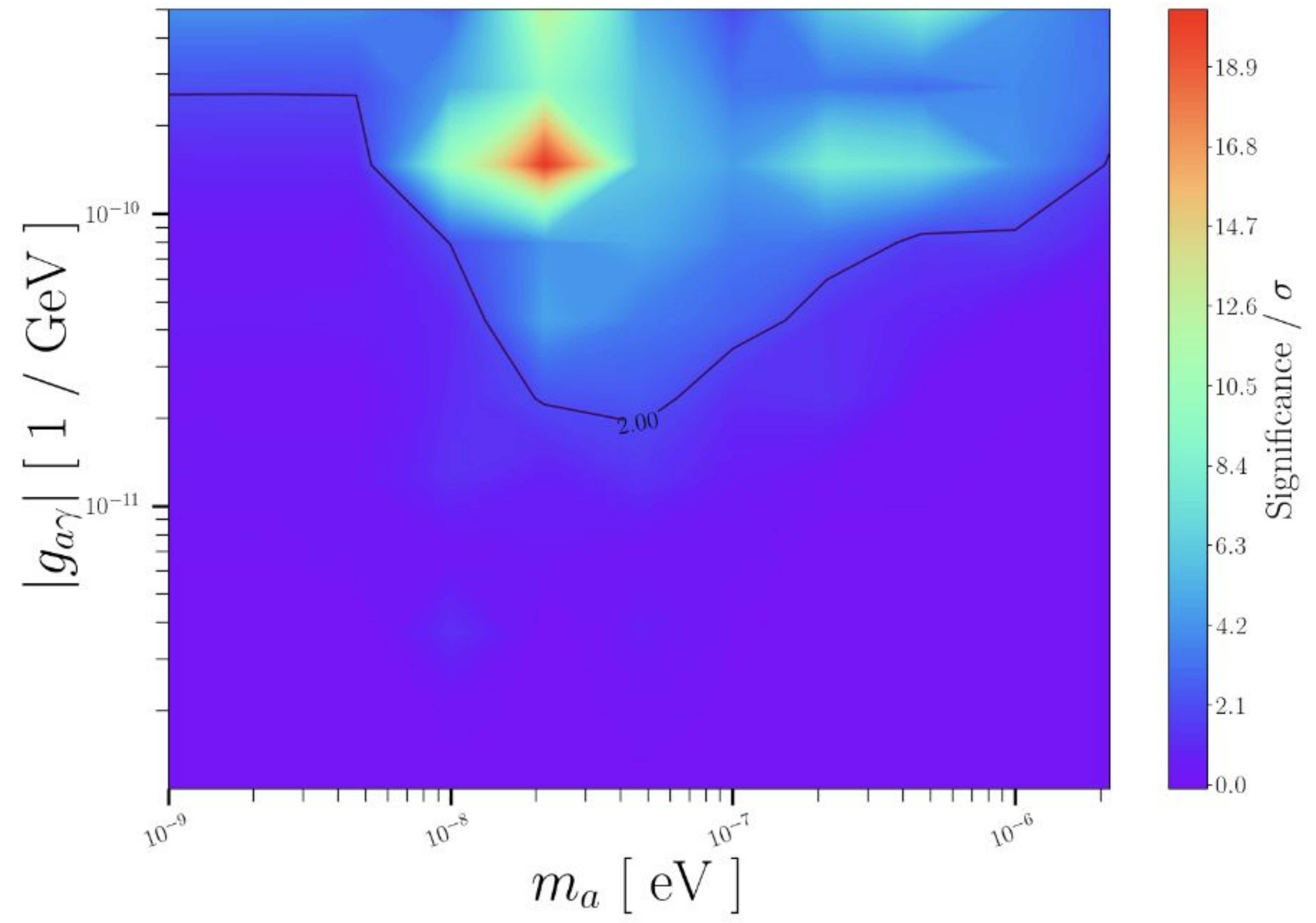
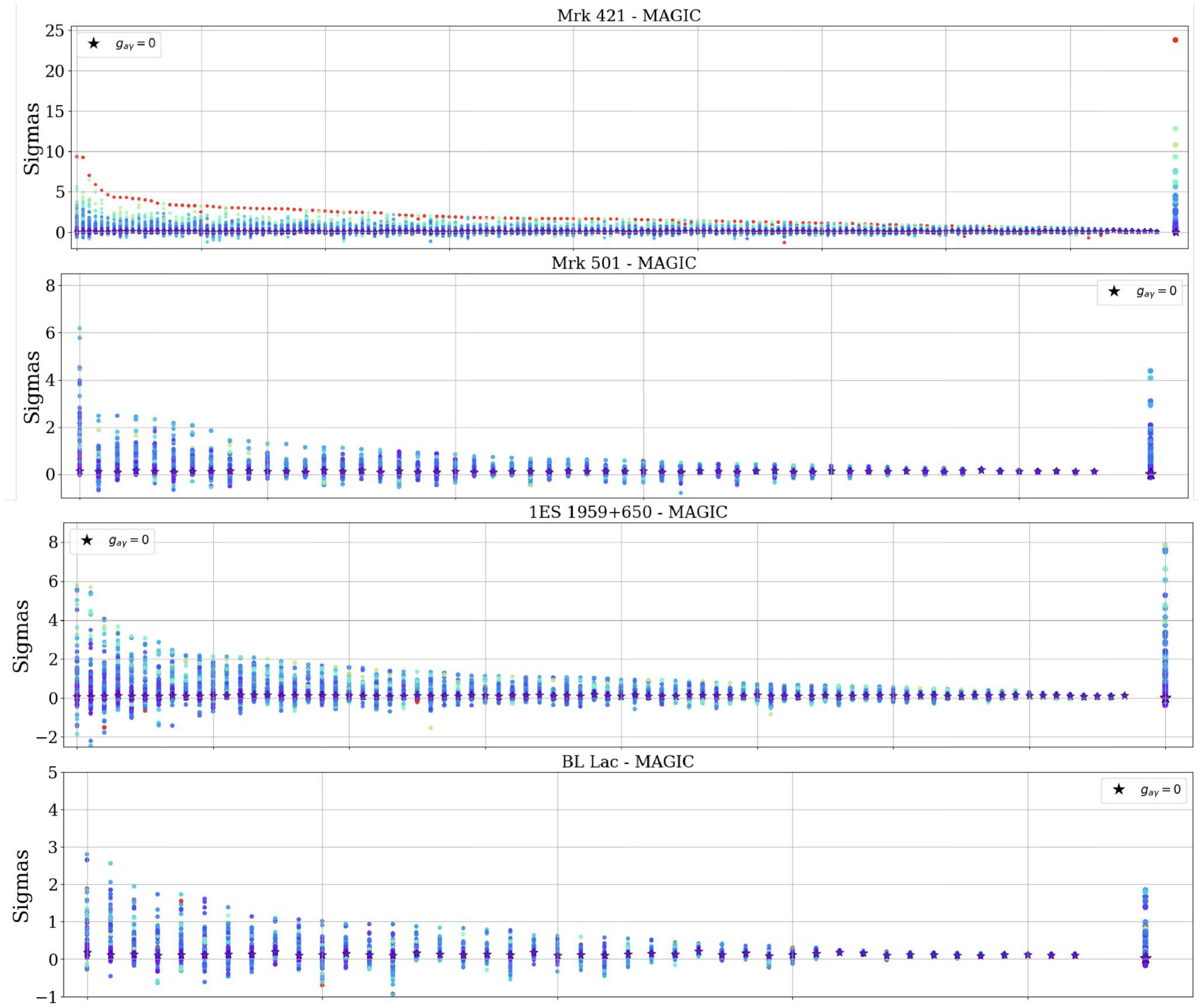


Figure 5: Exclusion significance, expressed in Gaussian standard deviations, over the ALP parameter space obtained by combining all MAGIC and LST-1 observations listed in Tab. 1. The solid black line marks the 95% confidence-level exclusion contour. The significance is evaluated on a grid of 121 ALP parameter points, while the color map is obtained with `Matplotlib contourf` function by interpolation between the sampled points.

121 $(g_{a\gamma}, m_a)$ points

Figure 6: contribution of individual BBs (for MAGIC)



3. Results and Discussion

Which BBs most contribute to the final exclusion?

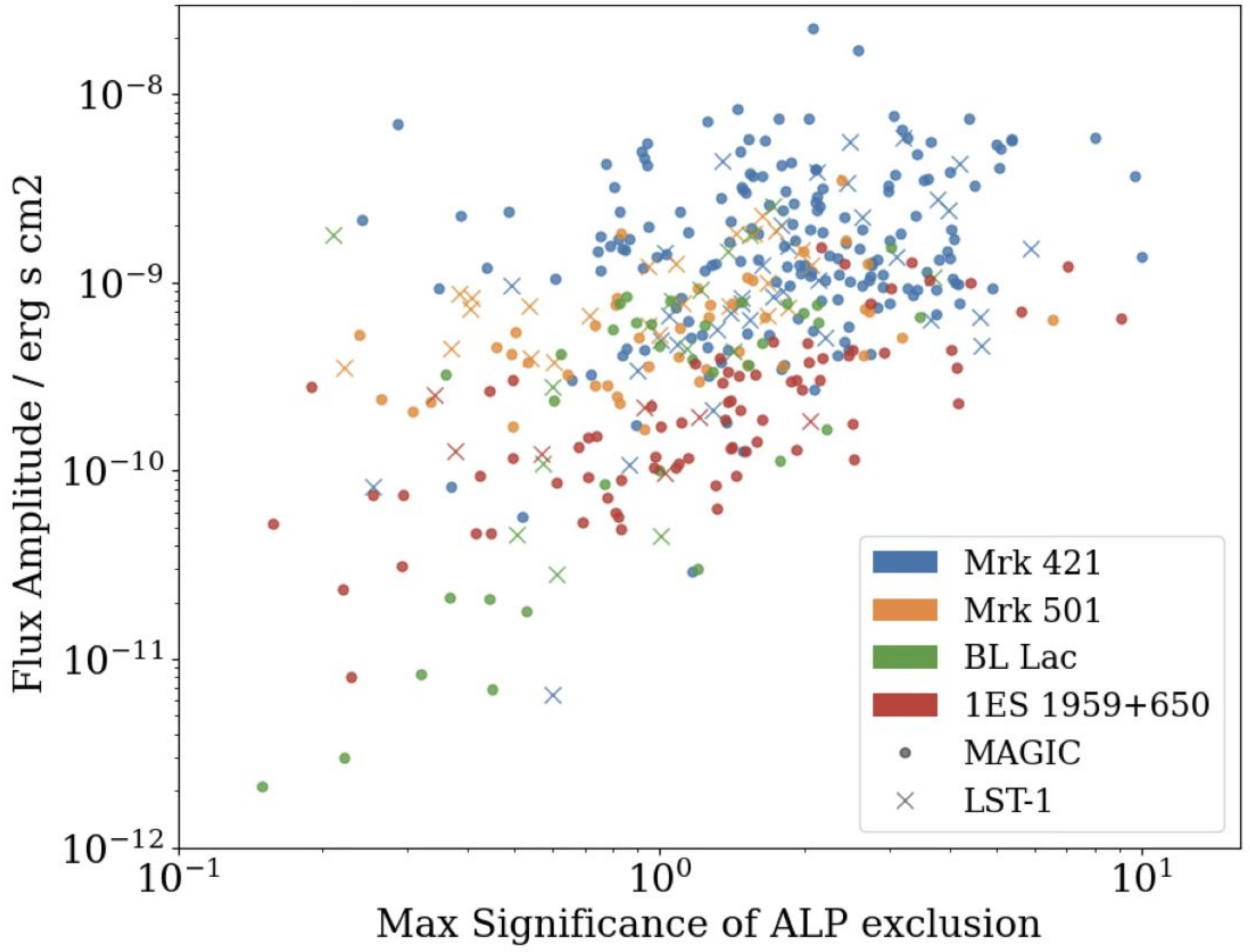


Figure 7: Distribution of the Bayesian Blocks (BBs) used in this analysis as a function of their intrinsic flux normalization and their constraining power on the ALP parameter space. Each point corresponds to one BB. The vertical axis shows the best-fit intrinsic flux normalization ϕ_0 (at the reference energy $E_{\mathrm{ref}} = 0.5$ TeV) obtained by fitting the data with the intrinsic spectral model including only EBL attenuation. The horizontal axis reports the maximum exclusion significance, expressed in Gaussian standard deviations, reached by that BB over the explored ALP parameter space. Marker colors identify the source, while circles and crosses denote MAGIC and LST-1 observations, respectively.

How do the limits change by removing sources and BBs?

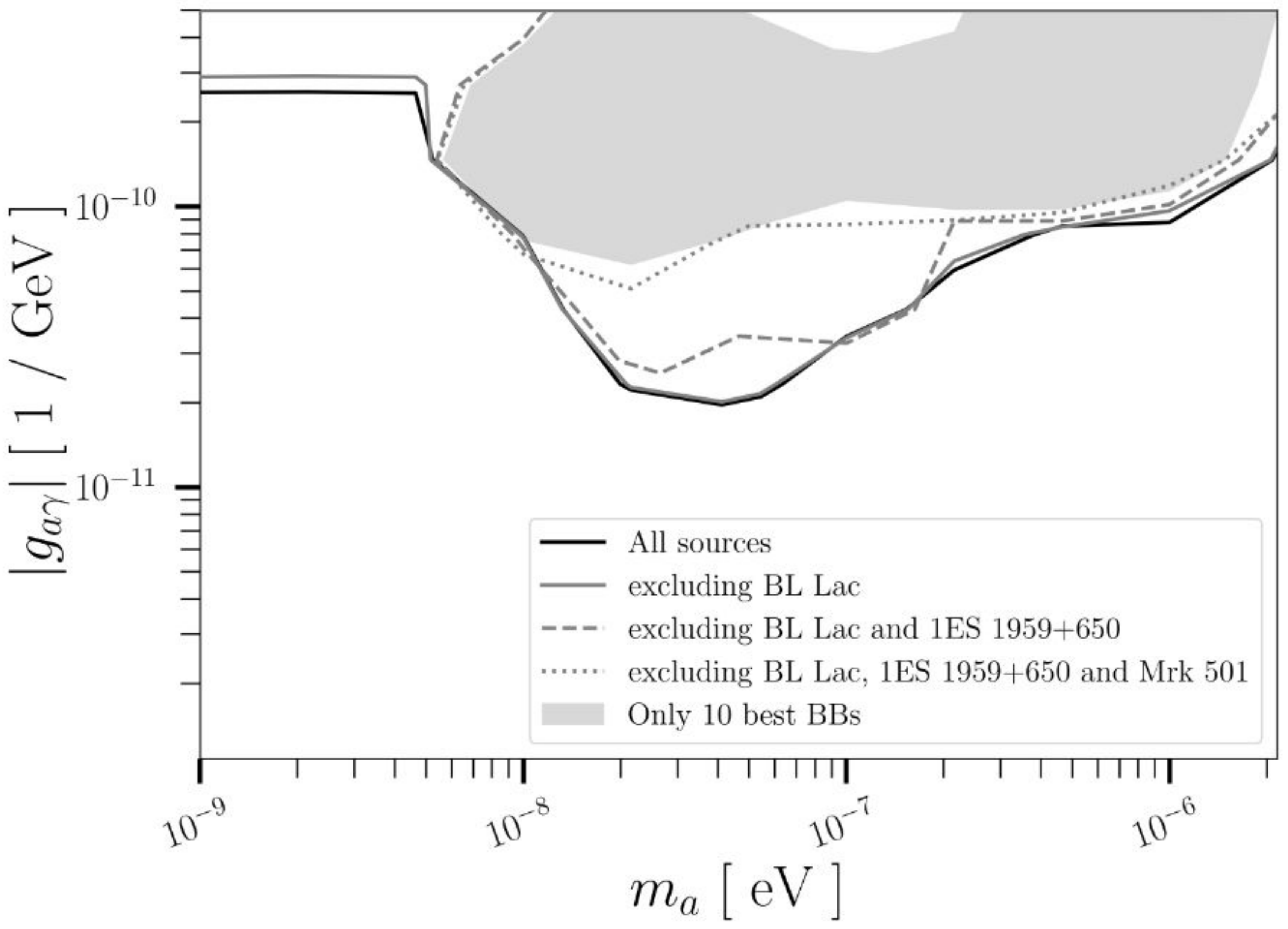


Figure 8: Comparison of the 95% CL exclusion obtained with different subsets of the dataset. The solid black curve shows the exclusion derived using all MAGIC and LST-1 observations. The gray curves show the result after progressively removing the least constraining sources, ultimately leaving only Mrk 421. The shaded gray region corresponds to the exclusion obtained using only the ten most constraining Bayesian Blocks of Mrk 421. These ten BBs provide a good approximation of the full Mrk 421 exclusion and are therefore used in the systematic studies presented in the appendix, where repeating the full analysis on all 442 BBs would be computationally prohibitive.

Conclusion:

The exclusion is driven primarily by a relatively **small number** of **bright** flaring states

Rare flaring episodes matter more than long integrations in the baseline state

4. Conclusion

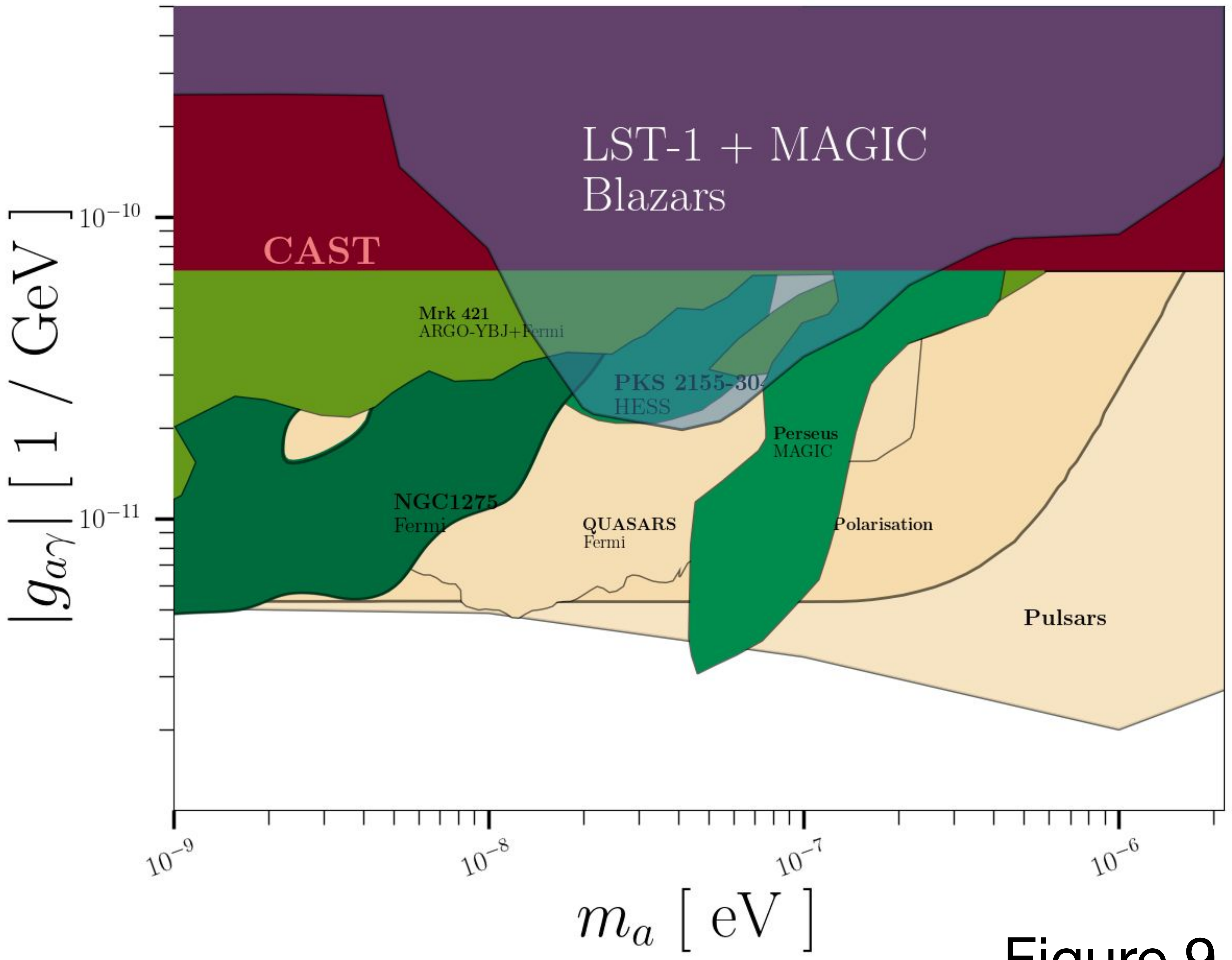


Figure 9

Limits consistent with previous Blazars-based analysis for instance:

- Flaring states of PKS 2155–304 measured by H.E.S.S (13 hours)
- Roughly ~110 hours of Mrk 421 and PG 1553+113 measured by MAGIC and Fermi-LAT (*Hai-Jun Li et al 2022 Chinese Phys. C 46 085105)*

Previous multi-epoch analyses of blazars have obtained similar limits and likewise shown that **only a subset** of activity periods provides substantial constraining power.

Other studies (lik Davies et al. 2023 and Li, Chao & Zhou 2024) that obtained much stronger limits assume:

- a larger B_0
- a larger emission distance r_{VHE}
- a coherent toroidal field

4. Conclusion

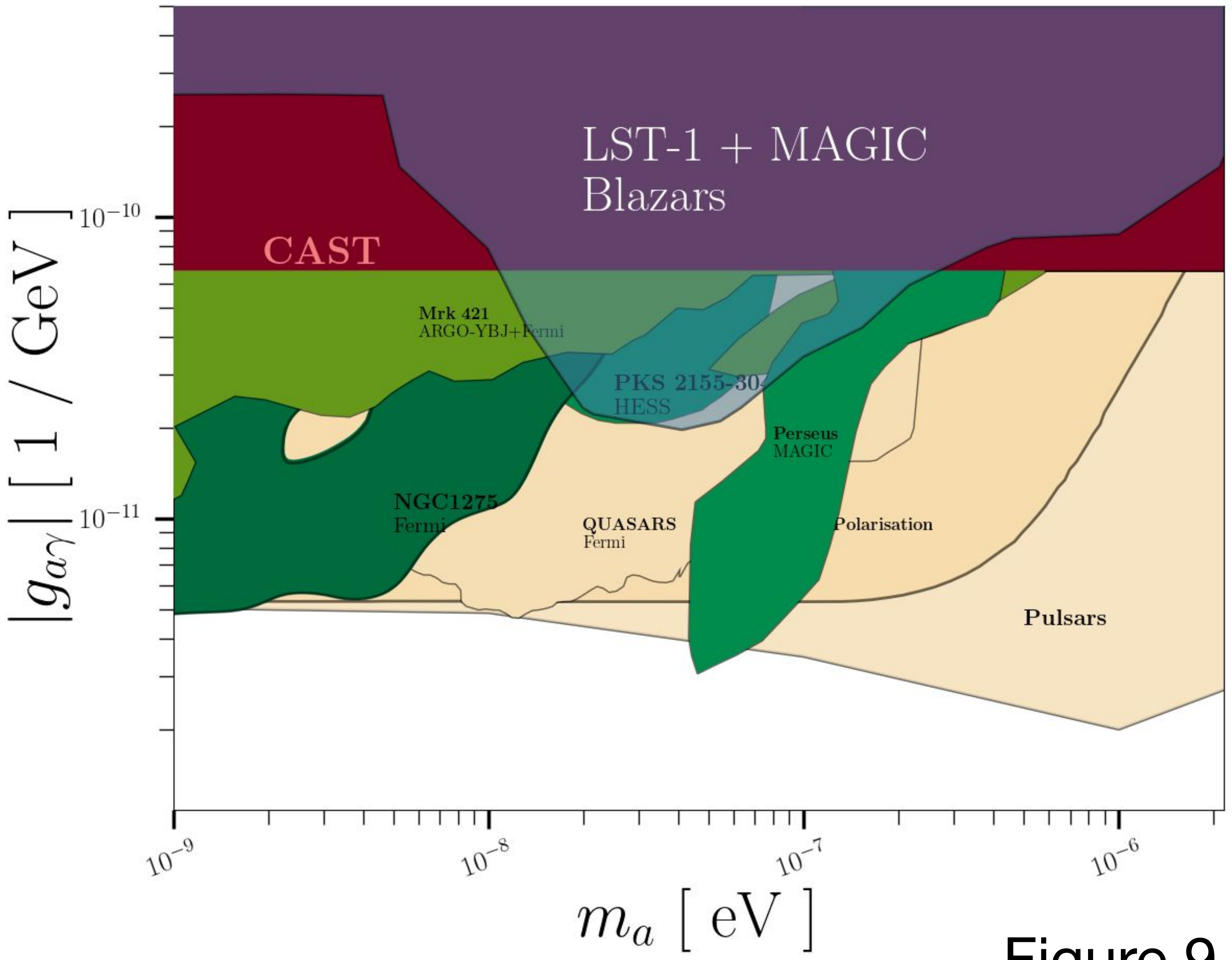


Figure 9

Selling points we explicitly wrote in the conclusion:

- The exclusion derived in this work is limited by the current **understanding of AGN jet magnetic fields**
- The limits derived, as in previous blazar-based ALP searches, should be read as **conditional on the adopted jet-field model**
- We identify the statistical framework, rather than a new ALP bound, as the main methodological contribution: **likelihood-level combination of Bayesian Blocks** provides a scalable approach for incorporating new observations without reanalyzing existing data
- Given that a handful of flaring blocks carries most of the sensitivity, the most effective gain will come not from longer exposures but from **catching more of the rare, bright states**