Correlation Between Hard X-Ray and Cosmic Neutrino Sources: From Obscured AGN to Blazars

Downloaded from: https://research.chalmers.se, 2026-11-10 05:48 UTC

Citation for the original published paper (version of record):

Kun, E., Bartos, I., Ricci, C. et al (2026). Correlation Between Hard X-Ray and Cosmic Neutrino Sources: From Obscured AGN to Blazars. Publications of the Astronomical Society of the Pacific, 138(7). http://dx.doi.org/10.1088/1538-3873/ae829b

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

Publications of the Astronomical Society of the Pacific, 138:074102 (18pp), 2026 July https://doi.org/10.1088/1538-3873/ae829b
© 2026. The Author(s). Published by IOP Publishing Ltd on behalf of the Astronomical Society of the Pacific (ASP).

OPEN ACCESS

Correlation Between Hard X-Ray and Cosmic Neutrino Sources: From Obscured AGN to Blazars

Emma Kun[1,2,3,4], Imre Bartos[5], Claudio Ricci[6,7], Santiago del Palacio[8], Francis Halzen[9], Julia Becker Tjus[4,8,10], Peter L. Biermann[11,12], and Anna Franckowiak[13]

[1] Department of Astronomy, Institute of Physics and Astronomy, ELTE Eötvös Loránd University, Pázmány Péter sétány 1a, Budapest, Hungary; kun.emma@csfk.org
[2] Konkoly Observatory, HUN-REN Research Centre for Astronomy and Earth Sciences, Konkoly Thege Miklós út 15–17, H-1121 Budapest, Hungary
[3] CSFK, MTA Centre of Excellence, Konkoly Thege Miklós út 15–17, H-1121 Budapest, Hungary
[4] Theoretical Physics IV, Faculty for Physics & Astronomy, Ruhr University Bochum, 44780 Bochum, Germany
[5] Department of Physics, University of Florida, PO Box 118440, Gainesville, FL 32611-8440, USA
[6] Department of Astronomy, University of Geneva, 1205 Geneva, Switzerland
[7] Instituto de Estudios Astrofísicos, Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Av. Ejército Libertador 441, Santiago, Chile
[8] Department of Physics and Astronomy Chalmers University of Technology SE-412 96 Gothenburg, Sweden
[9] Department of Physics, University of Wisconsin, Madison, WI 53706, USA
[10] Ruhr Astroparticle And Plasma Physics Center (RAPP Center), Ruhr University Bochum, 44780 Bochum, Germany
[11] Max Planck Institute for Radio Astronomy, 53121 Bonn, Germany
[12] Department of Physics & Astronomy, University of Alabama, Tuscaloosa, AL 35487, USA
[13] Astronomical Institute, Faculty for Physics & Astronomy, Ruhr University Bochum, 44780 Bochum, Germany
Received 2026 May 14; revised 2026 June 22; accepted 2026 June 25; published 2026 July 17

Abstract

The origin of high-energy astrophysical neutrinos remains a key open question in multimessenger astrophysics. A correlation between unabsorbed hard X-ray and high-energy neutrino luminosity has been reported in six active galactic nuclei with the highest individual IceCube significances, linking neutrino production to compact, photon-rich environments near supermassive black holes. We study whether the threshold-near IceCube excesses associated with seven Nuclear Spectroscopic Telescope Array-observed blazars are statistically consistent with that relation. Calibrating the $L_{\rm hX}$–L_ν relation on the six published sources via a Bayesian regression with errors on both axes, the slope is consistent with $\beta = 1$ and the intrinsic scatter is $\sim$0.6 dex. All seven new blazars are posterior-predictively consistent with this calibration ($\chi^2_7 = 1.58$, $p = 0.98$) under the working hypothesis that the published IceCube $\hat{n}_s$ values reflect the signal. A null-injection test confirms that, at the present calibration sample size, the consistency test does not by itself adjudicate between signal and selected-background origins. A distance-free $L_{\rm hX}/L_\nu$ ratio diagnostic places both populations within the photohadronic prediction band, statistically indistinguishable. Two diagnostics that control the common d_L^2 distance bias, a redshift-partial rank correlation ($\tau|z = 0.69$, $\sim 2.7\sigma$) and a flux-space permutation test on the 13-source joint sample ($p = 6.3 \times 10^{-4}$, 3.23σ), indicate a residual $L_{\rm hX}$–L_ν association beyond the distance-induced trend. We interpret these results as a conditional consistency check; a detection-level statement requires either an enlarged calibration set or an X-ray-weighted IceCube stacking likelihood with internal data.

Unified Astronomy Thesaurus concepts: Active galactic nuclei (16); Blazars (164); Neutrino astronomy (1100); X-ray active galactic nuclei (2035); Seyfert galaxies (1447)

1. Introduction

Active galactic nuclei (AGN) are among the most powerful persistent sources in the Universe (e.g. Begelman et al. 1984; Padovani et al. 2017), powered by accretion onto supermassive black holes (SMBHs). Despite their observational diversity, AGN share a common central engine composed of an SMBH, an accretion disk, and a hot X-ray-emitting corona. In radio-loud systems such as blazars, this structure is accompanied by relativistic jets aligned close to the line of sight with superluminal speeds (e.g. Kellermann et al. 2007), while in radio-quiet Seyfert galaxies, the emission is dominated by the accretion flow and corona. This structural similarity suggests that the fundamental physical processes governing energy dissipation and particle acceleration may be common across AGN classes, even if their observational manifestations differ.

The origin of high-energy neutrinos detected by the IceCube observatory remains one of the central open questions in

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

multimessenger astrophysics (Aartsen et al. 2013, 2020). Blazars have long been considered prime candidates due to their relativistic jets and efficient particle acceleration (e.g., Halzen 2022). However, the lack of a clear one-to-one association between neutrino events and γ-ray flares (e.g. Kun et al. 2023), together with the growing evidence for neutrino emission from non-blazar AGN such as the Seyfert galaxies NGC 1068 and NGC 4151 (Abbasi et al. 2022, 2025), suggests that the production sites of neutrinos may not be limited to large-scale jet regions.

Recent observational and theoretical developments point toward compact regions near the SMBH as promising sites of neutrino production. The hot corona and inner accretion flow provide dense photon fields favorable for photohadronic ($p\gamma$) interactions (Murase et al. 2020); the same photon field that seeds these interactions also renders the source opaque to the accompanying pionic γ rays, which are reprocessed via electromagnetic cascades into hard X-ray emission. Generally, the X-ray emission is dominated by Comptonization in non-beamed AGN (e.g. Haardt & Maraschi 1991), and such cascaded emission might be on top of such baseline in their case. Neronov et al. (2024) found a possible correlation between the hard X-ray and neutrino emission for three Seyfert galaxies. Kun et al. (2024, hereafter K2024) reported a tentative correlation supporting the hard X-ray–neutrino correlation scenario, based on six AGN with strong individual IceCube excesses: the Seyfert galaxies NGC 1068, NGC 4151, NGC 3079, and CGCG 420-015, together with the blazars TXS 0506+056 and GB6 J1542+6129. Since hard X-rays trace both coronal inverse Compton emission (e.g. Haardt & Maraschi 1991) and the cascade contribution from pionic γ rays in γ-opaque environments, this correlation is consistent with neutrino production occurring in the corona or its immediate environment. Khatee Zathul et al. (2025) argued that the IceCube neutrinos from TXS 0506+056 may originate not in the blazar jet but in a gamma-ray-obscured core within $\sim$10–100 Schwarzschild radii of the central black hole, the same mechanism inferred for NGC 1068 and NGC 4151.

A correlation between millimeter and hard X-ray emission has been observed in both beamed and non-beamed AGN (Kawamuro et al. 2022; Ricci et al. 2023), indicating that the millimeter emission probes compact regions associated with the central engine: directly the corona in radio-quiet AGN, and the corona-jet interface in radio-loud ones. The coronal origin of the compact millimeter emission in radio-quiet Seyferts is further supported by its spectral shape, which is consistent with synchrotron emission from the corona (Inoue & Doi 2018). This diagnostic is cleanest in non-beamed or misaligned systems, where the 15–55 keV emission can be cleanly associated with the corona. Our test sample, by contrast, consists of aligned blazars, whose hard X-rays can include a substantial jet inverse-Compton contribution; the calibration relation is thus established on a Seyfert-dominated, non-beamed basis and applied to an aligned-blazar test sample. We address its consequences for the emission-site interpretation and the κ diagnostic in Sections 3.4 and 5.3. Independent evidence linking hard X-ray emission to neutrino production in blazars comes from a statistical association between neutrino-detected sources and elevated hard X-ray flux measured by SRG/ART-XC, Swift/BAT, and INTEGRAL/IBIS, suggesting that the target photons for photohadronic interactions originate in compact, relativistically beamed jet regions (Plavin et al. 2024). The role of relativistic beaming in setting the multi-messenger visibility of such sources is dramatically illustrated by recent VLBA polarimetry of the BL Lac PKS 1424+240, one of the highest peaks in the IceCube 9 yr neutrino sky, which reveals a jet viewing angle $0.^{\circ}6$, Doppler factor $\delta \sim 30$, and a coherent toroidal magnetic-field component (Kovalev et al. 2025).

The IceCube ten-year point-source analysis (Abbasi et al. 2022, 2025) contains many candidate sources with threshold-near excesses (local pretrial significances of order 1σ–2σ) whose individual reality cannot be established from the IceCube likelihood alone. We ask whether such excesses, when identified in blazars and combined with measured hard X-ray luminosities, fall where the established AGN $L_{\rm hX}$–L_ν relation predicts.

We organize the analysis around two distinct questions. *Question (1): Is the underlying* $\mathrm{L_{hX}}$–L_ν *correlation real?* This was addressed by K2024 on the six AGN with the highest IceCube significances noted above ($R = 0.97$); we take the result as given and use the same six sources as our calibration sample throughout this work. *Question (2): Are the threshold-near IceCube excesses on candidate blazars consistent with the established relation?* We test whether seven *NuSTAR*-observed IceCube-candidate blazars are posterior-predictively consistent with the relation calibrated on the six K2024 sources. We are explicit that this is a conditional consistency test, not a statement on the reality of individual IceCube excesses against an independent background null: the seven test sources and the six calibration sources are both pre-selected on having $\hat{n}_s > 0$ from a common IceCube point-source analysis, and a detection-level claim against background fluctuations would require an X-ray-weighted IceCube stacking likelihood with public event-level data, which is not available for the analyzes underlying the published $\hat{n}_s$ values. We characterize the discriminating power of the test as a function of $\hat{n}_s$, and we calibrate it against an explicit background-only null in Section 3.3.

Section 2 describes the NuSTAR spectral analysis, the neutrino flux and luminosity calculation and the combined $L_{\rm hX}$–L_ν sample. Section 3 contains the posterior-predictive test, the null calibration, and the distance-free ratio diagnostic, which together form the central statistical analysis. Section 4 reports the flux-space permutation test as a finite-sample

diagnostic. Section 5 discusses implications and caveats; Section 6 concludes.

2. Analysis and Results

2.1. NuSTAR Spectral Analysis

We analyzed archival observations of the Nuclear Spectroscopic Telescope Array (NuSTAR) for a sample of seven blazars spatially associated with IceCube neutrino source candidates having best-fit neutrino number $\hat{n}_s > 0$ in the published ten-year point-source analyzes (Abbasi et al. 2022, 2025). Data were reduced using standard `NuSTARDAS` procedures, version v2.1.2 and CALDB 20230718, and spectra from FPMA and FPMB were fitted jointly in `XSPEC` (Arnaud 1996). All sources are point-like in NuSTAR. Spectra were extracted within a circular aperture of 60″ and background was subtracted from a source-free region.

For each observation, we adopted an absorbed power-law model of the form

$$\texttt{const} \times \texttt{tbabs} \times \texttt{po}, \tag{1}$$

using the abundance table of Wilms et al. (2000). The Galactic hydrogen column density $N_{\rm H}$ was fixed to values from the HI4PI survey (HI4PI Collaboration et al. 2016). A cross-normalization constant between FPMA and FPMB was included to account for instrumental systematics, while all other parameters were tied between the two detectors.

Unabsorbed fluxes in the 15–55 keV band were derived using the `cflux` convolution model, and luminosities were computed with `clumin`. All uncertainties are quoted at the 1σ level. We note that the Galactic column density is typically $N_{\rm H} \lesssim 10^{21}$ cm^{-2}, so the absorption correction is small and the distinction between observed and unabsorbed flux in the 15–55 keV band is minor. This is in contrast to the heavily obscured Seyfert galaxies in K2024, for which intrinsic column densities can reach $N_{\rm H} \sim 10^{24}$ cm^{-2} and the unabsorbed flux correction is both large and model-dependent. The blazar sample therefore provides hard X-ray luminosities that are relatively free from the dominant absorption-correction systematic present in the calibration sample.

All spectra are well described by a single power law, with reduced χ^2 values close to unity (Table 1). The photon indices span $\Gamma \sim 1.4$–2.4, with flat-spectrum radio quasars (FSRQs) generally exhibiting harder spectra than the BL Lac OJ 287, although the FSRQ range itself spans $\Gamma \sim 1.4$–2.1 across the sample and the IBL S2 0109+22 has the hardest indices ($\Gamma \sim 1.5$–1.6). This spectral diversity within the blazar class reflects the different positions of the SED peaks relative to the 15–55 keV band: the FSRQs and the low-frequency-peaked BL Lac OJ 287 are inverse-Compton–dominated in this band, while the intermediate-peaked S2 0109+22 yields the hardest indices consistent with its IBL classification.

To convert the observed energy flux to isotropic-equivalent luminosity, we applied a K-correction appropriate for a power-law spectrum,

$$L_{\rm hX} = 4\pi d_L^2 \, F_{\rm X} \, (1+z)^{\Gamma_{\rm X}-2}, \tag{2}$$

where $F_{\rm X}$ is the observed 15–55 keV energy flux and $\Gamma_{\rm X}$ is the X-ray photon index from Table 1. The X-ray log-space uncertainty $\sigma_{{\rm X},i}$ combines the measured NuSTAR flux error with the K-correction systematic $\sigma_{K,i} = \Delta\Gamma \cdot \log_{10}(1+z_i)$ ($\Delta\Gamma = 0.10$) added in quadrature (see Barger et al. 2001, for context).

The derived unabsorbed fluxes and luminosities lie in the range

$$F_{15-55\,\rm keV} \sim (0.6 - 27) \times 10^{-12}\,{\rm erg\,s^{-1}\,cm^{-2}}, \tag{3}$$

and

$$L_{15-55\,\rm keV} \sim 7.1 \times 10^{44} - 1.2 \times 10^{47}\,{\rm erg\,s^{-1}}, \tag{4}$$

respectively. Sources with multiple observations (e.g., 3C 454.3, OJ 287, MITG J201534+3710) show variability typically within a factor of ∼2–3. The photon indices remain consistent within uncertainties across epochs, indicating that variability is dominated by flux normalization rather than spectral changes. This epoch-to-epoch spectral stability supports the use of individual observations as independent flux measurements rather than requiring time-averaged spectra.

2.2. Neutrino Fluxes and Luminosities

To calculate the neutrino luminosities, we follow K2024. We assumed a power-law muon-neutrino plus antineutrino spectrum,

$$\phi_{\nu_\mu+\bar{\nu}_\mu}(E_\nu) = \phi_0 \left(\frac{E_\nu}{E_0}\right)^{-\gamma_\nu}, \tag{5}$$

with $E_0 = 1$ TeV. For sources with published best-fit signal-event number $\hat{n}_s$ and spectral index $\hat{\gamma}_\nu$ from the ten-year IceCube analysis (Abbasi et al. 2022), we derived the normalization ϕ_0 from the IceCube effective area and detector live-time by inverting the expression

$$n_s = \tau \sum_i A_{{\rm eff},i} \int_{E_{l,i}}^{E_{u,i}} \phi_{\nu_\mu+\bar{\nu}_\mu}(E_\nu)\, dE_\nu, \tag{6}$$

where the sums run over the tabulated energy bins. The published ten-year point-source analysis covers full-detector IC 86 data from 2011 May 13 to 2020 May 29 with declination-dependent effective areas; following K2024, we adopt the IC 86 reference effective area tables together with the cumulative ten-year exposure τ, which provides a time-averaged $A_{\rm eff} \times \tau$ that approximates the season-summed exposure for the mid-declination sources in our sample to within ∼10%–20%. The 0.3–100 TeV integrated neutrino

Table 1
Source Properties

Source	Type	z	$\hat{n}_s$	$-\log_{10} p_{\rm loc}$	ObsID	Date	Exposure (ks)	Γ	χ^2_ν	$F_{15-55\ \rm keV}$ (erg s^{-1} cm^{-2})
PKS 1441+25 (1)	FSRQ	0.940	3	0.7 (0.9σ)	90101004002	2015-04-25	38.2	2.10 ± 0.15	1.13	$(6.31 \pm 1.41) \times 10^{-13}$
3C 454.3 (2)	FSRQ	0.859	1	1.2 (1.6σ)	60701034002	2022-06-13	20.8	1.46 ± 0.04	0.92	$(1.49 \pm 0.07) \times 10^{-11}$
…	…	…	…	…	60802038002	2023-05-01	40.2	1.48 ± 0.02	0.86	$(2.67 \pm 0.07) \times 10^{-11}$
…	…	…	…	…	60802038004	2023-05-04	22.4	1.45 ± 0.03	0.87	$(2.62 \pm 0.09) \times 10^{-11}$
…	…	…	…	…	60802038006	2023-05-07	20.7	1.44 ± 0.03	0.83	$(2.58 \pm 0.10) \times 10^{-11}$
…	…	…	…	…	60802038008	2023-05-12	23.0	1.47 ± 0.02	0.95	$(2.66 \pm 0.09) \times 10^{-11}$
MITG J201534+3710 (3)	FSRQ	0.858	19	0.7 (0.9σ)	60160732002	2016-05-29	12.7	1.54 ± 0.05	0.82	$(1.17 \pm 0.08) \times 10^{-11}$
…	…	…	…	…	30460021002	2018-11-01	59.6	1.65 ± 0.04	0.83	$(5.02 \pm 0.26) \times 10^{-12}$
S3 0458−02 (4)	FSRQ	2.286	9	0.5 (0.4σ)	60367003001	2018-04-26	20.7	1.73 ± 0.07	0.88	$(3.91 \pm 0.41) \times 10^{-12}$
OJ 287 (5)	BLL	0.306	16	0.5 (0.4σ)	90201054002	2017-04-09	53.0	2.08 ± 0.04	0.95	$(3.60 \pm 0.19) \times 10^{-12}$
…	…	…	…	…	90601616002	2020-05-04	29.6	2.38 ± 0.05	0.91	$(2.04 \pm 0.17) \times 10^{-12}$
…	…	…	…	…	60701023002	2022-03-22	12.6	1.89 ± 0.11	0.75	$(3.06 \pm 0.49) \times 10^{-12}$
…	…	…	…	…	90801634002	2022-12-01	31.9	1.86 ± 0.04	0.96	$(6.72 \pm 0.36) \times 10^{-12}$
…	…	…	…	…	60901019002	2024-04-03	20.5	1.99 ± 0.08	1.01	$(2.59 \pm 0.31) \times 10^{-12}$
S2 0109+22 (6)	IBL	0.265	10	0.7 (0.8σ)	60801029002	2022-11-22	57.3	1.54 ± 0.05	0.74	$(3.73 \pm 0.28) \times 10^{-12}$
…	…	…	…	…	60801029004	2022-11-24	100.6	1.63 ± 0.04	0.94	$(3.74 \pm 0.19) \times 10^{-12}$
Ton 599 (7)	FSRQ	0.725	2	0.2 (0.0σ)	60463037002	2019-05-23	17.0	1.97 ± 0.40	1.32	$(6.11 \pm 4.18) \times 10^{-13}$
…	…	…	…	…	60463037004	2021-06-25	17.6	1.59 ± 0.06	0.95	$(8.21 \pm 0.74) \times 10^{-12}$

Notes. Source name, type, redshift (z), IceCube best-fit neutrino number ($\hat{n}_s$), IceCube local pretrial $-\log_{10} p_{\rm loc}$ (and corresponding local significance), and NuSTAR spectral fit results and derived unabsorbed 15–55 keV fluxes for the seven new blazars analyzed in this work.
Spectra were fitted in `XSPEC` with the model `const*tbabs*po`, adopting the abundance table of Wilms et al. (2000). The Galactic hydrogen column density $N_{\rm H}$ was fixed to the HI4PI survey value for each source (HI4PI Collaboration et al. 2016). The multiplicative constant accounts for cross-normalization between FPMA and FPMB. Unabsorbed 15–55 keV fluxes were derived using `cflux`; luminosities were computed with `clumin` assuming the cosmology described in Section 2.2. Errors are quoted at the 1σ level. Redshifts are from (1) Ahnen et al. (2015), (2) Koss et al. (2022), (3) Ackermann et al. (2015), (4) Sahakyan et al. (2024), (5) Miller et al. (1978), (6) Paiano et al. (2016), (7) Hewett & Wild (2010).

energy flux was then computed as

$$F_{\nu_\mu+\bar{\nu}_\mu} = \int_{0.3\ \rm TeV}^{100\ \rm TeV} E_\nu\ \phi_{\nu_\mu+\bar{\nu}_\mu}(E_\nu)\, dE_\nu. \tag{7}$$

To convert the observed energy flux to isotropic-equivalent luminosity, we applied a K-correction similarly to what has been applied on the hard X-ray fluxes,

$$L_{\nu_\mu+\bar{\nu}_\mu} = 4\pi d_L^2\, F_{\nu_\mu+\bar{\nu}_\mu}\, (1+z)^{\gamma_\nu-2}, \tag{8}$$

where d_L is the luminosity distance and z is the source redshift. The exponent $(\gamma_\nu - 2)$ accounts for the energy dependence of the K-correction; for the typical spectral indices in our sample (best-fit astrophysical spectral indices $\gamma_\nu \sim$ 2–3, Abbasi et al. 2022; IceCube Collaboration 2022), this correction is modest but non-negligible at the highest redshifts.

We emphasize that L_ν in this work is a band-limited luminosity, integrated over the fixed 0.3–100 TeV energy range from the IceCube power-law reconstruction and defined identically for the calibration and test samples. The adopted band is set primarily by consistency with the literature on which the calibration rests: the calibration-Seyfert neutrino fluxes are reported natively in 0.3–100 TeV, Neronov et al. (2024) quote 0.3–100 TeV fluxes for NGC 3079 and NGC 4151 and adopt the NGC 1068 TeV-band flux as their template, so re-deriving them over a different band would introduce an inconsistency with the published values. For the soft best-fit spectra of our sample ($\gamma_\nu \approx$ 2–3; Neronov et al. (2024) find the three Seyfert spectra to be softer than E^{-2}, with most of the neutrino power emitted near the IceCube threshold of several hundred GeV), the integrated energy flux is concentrated well below 100 TeV, so L_ν is insensitive to the upper bound. A hard, $\gtrsim$PeV-peaking spectrum, as expected for pure broad-line-region (BLR) target $p\gamma$ in FSRQs (Dermer et al. 2014), is not what IceCube reconstructs for these sources. We quantify the effect of extending the energy range from 0.3 TeV up to 10 PeV in Appendix B.

The neutrino log-space uncertainty $\sigma_{\nu,i}$ requires more care, and the asymmetric treatment between calibration and test samples is the dominant systematic in this analysis. The published asymmetric uncertainties from K2024 are converted to a symmetric log-space $\sigma_{\log L_\nu}$ via the geometric mean of the upper and lower log-deviations, as required by the Gaussian-error assumption of the linmix regression. For the seven new blazars, the corresponding likelihood profile is not publicly available; we therefore start from the asymptotic Poisson form $\sigma_{\nu,i}^{\rm Poisson} = 1/(\sqrt{\hat{n}_{s,i}}\ \ln 10)$ and apply an empirical scaling factor of 4.47, derived from the mean ratio of spectrum-derived to Poisson uncertainties on the six calibration sources (range 3.7–5.2, standard deviation 0.59). The constancy of this ratio across the calibration sample reflects a $\sigma_{\rm syst} \propto 1/\sqrt{\hat{n}_s}$ scaling,

consistent with the simultaneous determination of $\hat{n}_s$ and $\hat{\gamma}_\nu$ from the same event sample: both the Poisson term and the spectral-index propagation term scale identically with $\sqrt{\hat{n}_s}$, justifying a constant multiplicative form.

We emphasize the regime in which this scaling is calibrated. The six K2024 sources span $\hat{n}_s \in [5, 79]$, with median $\hat{n}_s \approx 25$, where the IceCube profile likelihood is approximately Gaussian. Three of the seven test sources lie outside this regime: 3C 454.3 ($\hat{n}_s = 1$), Ton 599 ($\hat{n}_s = 2$), and PKS 1441 +25 ($\hat{n}_s = 3$). For these sources, the asymptotic Poisson approximation is itself less accurate, and the asymmetry of the underlying profile likelihood is more pronounced. We treat the test-sample $\sigma_{\nu,i}$ values obtained from the 4.47 × scaling as Gaussian throughout the analysis; the impact of this choice on the consistency test is quantified in Section 3.2 through the high-$\hat{n}_s$ versus low-$\hat{n}_s$ split, and the symmetry assumption is shown not to drive the qualitative outcome.

Throughout this work, luminosity distances were calculated assuming a flat ΛCDM cosmology with $H_0 = 69.6$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.286$, $\Omega_\Lambda = 0.714$, and $T_{\rm CMB,0} = 2.72548$ K, consistent with K2024.

2.3. The Joint L_{hX}–L_ν Sample

Figure 1 shows the unabsorbed hard X-ray luminosity versus the neutrino luminosity for the seven new blazars analyzed here, together with the six sources from K2024. The seven new blazars are visually consistent with the trend established by the calibration sample and extend it toward higher luminosities. We emphasize that this figure is descriptive: it displays the joint sample but does not re-derive the correlation. The K2024 relation, calibrated on the six AGN with the highest IceCube significances, is taken as established for the rest of this work. For descriptive purposes, the inverse-variance-weighted rms scatter of the joint 13-source sample around the one-to-one line $L_{\rm hX} = L_\nu$ is $\sigma_{\rm scat}^{\rm w,obs} = 0.321$ dex.

3. Posterior-predictive Consistency Test

The central question of this paper is whether the seven new candidate-blazar IceCube excesses, all near-threshold significances, are consistent with being drawn from the established AGN $L_{\rm hX}$–L_ν relation. We address this in four steps:

(i) We recalibrate the relation using only the six K2024 sources. This calibration uses the same data as K2024 but serves a different purpose: rather than testing whether the correlation exists, we use it to construct a posterior predictive distribution for $\log L_\nu$ at any given $\log L_{\rm hX}$, which the new blazars can then be tested against.

(ii) For each new blazar, we compute the predictive residual $\Delta_i = \log L_\nu^{\rm obs}{}_i - \log L_\nu^{\rm pred}{}_i$, where $\log L_\nu^{\rm pred}{}_i$ is drawn from the posterior predictive distribution at the source's measured $\log L_{\rm hX}$. We then ask whether Δ_i is small compared to the predictive uncertainty $\sigma_{{\rm pred},i}$, which combines intrinsic scatter and measurement errors; the relation-parameter covariance is not propagated in the plug-in form used here and is assessed separately in Appendix B. We also characterize where the test retains discriminating power: in the high-$\hat{n}_s$ sources, where the predictive band is set by $\sigma_{\rm int}$, the test can meaningfully fail; in the low-$\hat{n}_s$ sources, the measurement error $\sigma_{\nu,i}$ dominates and the test is uninformative by construction.

(iii) We calibrate the test against an explicit background-only null by injecting $\hat{n}_s$ values drawn from the IceCube background-only distribution at the seven blazar positions, propagating them through the same posterior-predictive pipeline, and asking whether the resulting $\chi^2_{\rm null}$ distribution is statistically distinguishable from the observed χ^2_7. This addresses the concern that a small calibration sample ($N_{\rm cal} = 6$) may render the posterior-predictive band wide enough to accommodate signal and selected-background hypotheses alike.

(iv) We check the distance-free $\log(L_{\rm hX}/L_\nu)$ ratio diagnostic. The luminosity-luminosity plane shares a common d_L^2 factor on both axes, which can artificially induce a correlation among sources spanning a wide redshift range. Working in $L_{\rm hX}/L_\nu$ cancels d_L^2 exactly, leaving a source-frame quantity sensitive only to the underlying physics. If both populations cluster around the same κ value within the photohadronic prediction band, the correlation is robust against this geometric artifact.

3.1. Calibration of the Kun+2024 Relation

We model the hard X-ray–neutrino relation as

$$\log L_\nu = \alpha + \beta \log L_{\rm hX} + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma_{\rm int}^2), \tag{9}$$

where α and β are the intercept and slope of the linear relation, and ϵ is a Gaussian intrinsic scatter term with variance $\sigma_{\rm int}^2$ that accounts for source-to-source physical variation not captured by the deterministic $L_{\rm hX}$–L_ν relation. We treat $\log L_{\rm hX}$ as the predictor variable (i.e., the independent variable in the regression), and $\log L_\nu$ as the response. This direction reflects the physical scenario: in the photohadronic reprocessing framework, the hard X-ray luminosity traces the ambient photon field that drives both the $p\gamma$ interaction rate and the $\gamma\gamma$ cascade fraction, so $L_{\rm hX}$ causally influences L_ν rather than vice versa. The hard X-ray luminosity is also the better-measured of the two quantities for our sample. We calibrate Equation (9) on the six K2024 sources only, using the Bayesian linear regression of Kelly (2007) with measurement errors on both axes, as implemented in the publicly available `linmix` package (J. Meyers, https://github.com/jmeyers314/linmix). We use the spectrum-derived $\sigma_{\nu,i}$ on $\log L_\nu$ for the calibration sample (see Section 2.2), four parallel

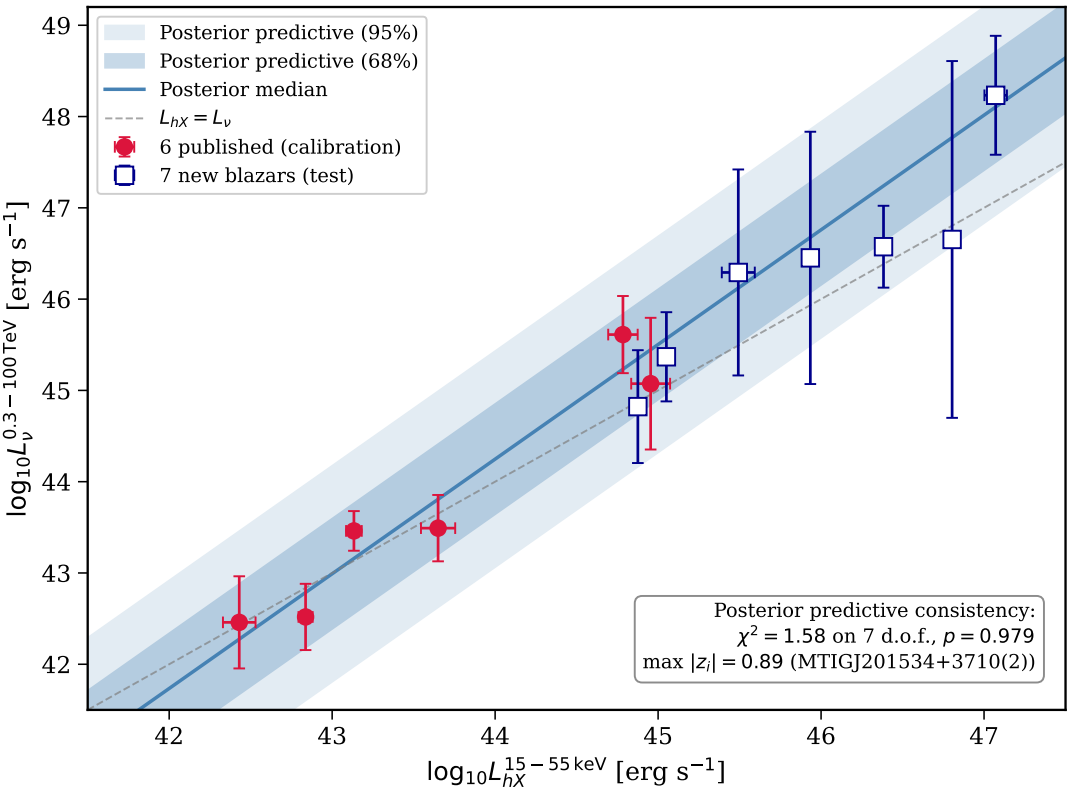

Figure 1. Posterior-predictive consistency test. The blue line shows the plug-in posterior-predictive median of the K2024 relation calibrated on the six published sources only (shown as red filled circles), with shaded 68% (dark) and 95% (light) bands set by $\pm\sigma_{\mathrm{int,med}}$ and $\pm 1.96\,\sigma_{\mathrm{int,med}}$ around the median line. Open blue squares show the seven new blazars analyzed in this work. The dashed line marks $L_{\mathrm{hX}} = L_\nu$. All seven new sources fall within the 95% band; per-source predictive z-scores satisfy $|z_i| \leqslant 0.89$ (largest at MITG J201534+3710), and the global $\chi^2_7 = 1.58$ corresponds to $p = 0.980$. We caution that the apparent consistency must be interpreted in light of the null calibration of Section 3.3: with $\sigma_{\mathrm{int}} = 0.60$ dex calibrated on six sources, the predictive band is wide enough that background-only realizations also yield median $\chi^2_{\mathrm{null}} \sim 2.5$ on this same pipeline.

chains, and a $K = 2$ Gaussian mixture prior on the latent $\log L_{\mathrm{hX}}$ distribution. Convergence is established by the multivariate Gelman–Rubin statistic $\hat{R} < 1.01$ on $(\alpha, \beta, \log \sigma^2_{\mathrm{int}})$. The new blazars are not used at this stage.

The sampler returns posterior chains for $(\alpha, \beta, \sigma_{\mathrm{int}})$, with median values and 68% credible intervals $\beta = 1.26^{+0.40}_{-0.38}$, $\sigma_{\mathrm{int}} = 0.60^{+0.78}_{-0.33}$ dex, $\alpha_{\mathrm{pivot}} = 43.51^{+0.34}_{-0.35}$, where α_{pivot} is the relation value at $\log L_{\mathrm{hX}} = 43.39$ (the median of the calibration sample). The recovered slope is consistent with $\beta = 1$ at $\sim 0.7\,\sigma$, in agreement with the $L_{\mathrm{hX}} \sim L_\nu$ scaling reported by K2024. The intrinsic scatter of 0.60 dex is comparable to the calibration sample's spread along the one-to-one line and is the dominant component of the predictive uncertainty per source for the high-$\hat{n}_s$ test sources. We note that the credible intervals on β and σ_{int} are wide given $N_{\mathrm{cal}} = 6$: β admits values from ~ 0.9 to ~ 1.7 and σ_{int} from ~ 0.3 to ~ 1.4 dex at 1σ. The use of the K2024 relation as "established" for the rest of this work is therefore provisional in the literal statistical sense, and an enlarged calibration sample remains the single most impactful improvement identifiable for follow-up work.

3.2. *Per-source Predictive Residuals and the High-$\hat{n}_s$ versus Low-$\hat{n}_s$ Split*

Having calibrated the relation on the six K2024 sources, we now ask whether the seven new blazars fall where the calibration predicts. The natural quantity to compute is the posterior predictive distribution of $\log L_\nu$ at each new blazar's measured $\log L_{\mathrm{hX}}$: that is, the range of $\log L_\nu$ values the calibrated relation predicts, accounting for the intrinsic scatter σ_{int} and the measurement error on the predictor $\log L_{\mathrm{hX}}$. In the plug-in form adopted below (Equation (12)) the relation parameters $(\alpha, \beta, \sigma_{\mathrm{int}})$ are fixed at their posterior medians and their covariance is not propagated; the impact of including the (α, β) covariance is assessed in Appendix B. For each new blazar i we draw M samples from the plug-in posterior predictive distribution

$$\log L^{\mathrm{pred}}_{\nu,i} \,|\, \log L^{\mathrm{obs}}_{X,i} = \alpha_{\mathrm{med}} + \beta_{\mathrm{med}}\,\xi_i + \mathcal{N}(0, \sigma^2_{\mathrm{int,med}}),$$
$$\xi_i \sim \mathcal{N}(\log L^{\mathrm{obs}}_{X,i}, \sigma^2_{X,i}), \qquad (10)$$

where ξ_i marginalizes over the latent true X-ray luminosity given the measurement error, and $(\alpha_{\mathrm{med}}, \beta_{\mathrm{med}}, \sigma_{\mathrm{int,med}})$ are

Table 2
Posterior-predictive Consistency of the Seven New Blazars with the K2024 Calibration

Source	$\hat{n}_s$	$\log L_{\rm hX}^{\rm obs}$	$\log L_{\nu}^{\rm obs}$	$\log L_{\nu}^{\rm pred}$	z_i
3C 454.3	1	46.80	46.65	47.77 ± 0.60	−0.55
S2 0109+22	10	44.88	44.82	45.34 ± 0.60	−0.60
PKS 1441+25	3	45.49	46.29	46.12 ± 0.61	+0.13
MITG	19	46.38	46.57	47.24 ± 0.60	−0.89
J201534+3710	…	…	…	…	…
S3 0458−02	9	47.07	48.23	48.11 ± 0.60	+0.14
OJ 287	16	45.05	45.37	45.57 ± 0.60	−0.26
Ton 599	2	45.93	46.45	46.68 ± 0.60	−0.15

Full sample: $\chi^2_7 = 1.58$, $p = 0.980$
High-$\hat{n}_s$ ($\hat{n}_s \geqslant 9$, 4 sources): $\chi^2_4 = 1.24$, $p = 0.87$
Low-$\hat{n}_s$ ($\hat{n}_s \leqslant 3$, 3 sources): $\chi^2_3 = 0.34$, $p = 0.95$

Note. Predicted $\log L_\nu$ values are posterior-predictive medians with 68% intervals; $z_i = \Delta_i/\sigma_{{\rm pred},i}$ with $\Delta_i = \log L_{\nu,i}^{\rm obs} - \log L_{\nu,i}^{\rm pred}$, and $\sigma^2_{{\rm pred},i} = \sigma^2_{\nu,i} + \sigma^2_{\rm int,med} + \beta^2_{\rm med}\sigma^2_{X,i}$ in the plug-in posterior-predictive formulation (Section 3.2). The 68% interval on $\log L_\nu^{\rm pred}$ is dominated by $\sigma_{\rm int,med} = 0.60$ dex, which is independent of source for the plug-in choice. The high-$\hat{n}_s$ subsample is where the test retains discriminating power; on the low-$\hat{n}_s$ subsample $\sigma_{\nu,i}$ dominates and the test cannot meaningfully fail (see Section 3.2).

fixed at their posterior medians from the calibration chain (the plug-in choice; see below for justification). We adopt this plug-in form because the posterior of $\sigma_{\rm int}$ at $N_{\rm cal} = 6$ has a heavy upper tail under the Jeffreys-style prior on $\sigma^2_{\rm int}$; full marginalization over the calibration chain would let that tail dominate $\sigma_{{\rm pred},i}$ and render the test uninformative, whereas the plug-in choice retains the representative scatter recovered by the regression (see Hogg et al. 2010). The predictive residual is

$$\Delta_i = \log L_{\nu,i}^{\rm obs} - \log L_{\nu,i}^{\rm pred}, \tag{11}$$

and the full predictive variance per source combines the observed neutrino measurement error and the spread of the predictive distribution,

$$\sigma^2_{{\rm pred},i} = \sigma^2_{\nu,i} + \sigma^2_{\rm int,med} + \beta^2_{\rm med}\sigma^2_{X,i}. \tag{12}$$

We summarize consistency through the predictive z-score $z_i = \Delta_i/\sigma_{{\rm pred},i}$ and the two-sided p-value from the standard normal. A global consistency test is given by the joint statistic

$$\chi^2 = \sum_{i=1}^{N_{\rm new}} z_i^2, \tag{13}$$

which, under the null hypothesis that the new blazars are drawn from the calibrated relation, follows $\chi^2_{N_{\rm new}}$.

The per-source results are given in Table 2 and visualized in Figure 1. All seven new blazars satisfy $|z_i| \leqslant 0.89$, the global $\chi^2_7 = 1.58$ corresponds to $p = 0.980$, and every source falls within the 95% posterior-predictive band.

The discriminating power of the test, however, is unevenly distributed across the sample, and inspection of the dominant term in Equation (12) makes this explicit. Equation (12) fixes (α, β) at their medians and therefore omits the relation-parameter covariance, which at the high-luminosity end is significant given the ∼1–3 dex extrapolation from the calibration pivot. We show in Appendix B that a robust treatment of the (α, β) covariance broadens $\sigma_{{\rm pred},i}$ at the high-luminosity sources (from ∼0.7 to 1.4–2.4 dex) without altering the consistency conclusion, while full marginalization over the heavy-tailed posterior is uninformative. With $\sigma_{\rm int,med} = 0.60$ dex and the 4.47 × scaled Poisson form of $\sigma_{\nu,i}$, the four sources with $\hat{n}_s \geqslant 9$ (MITG J201534+3710 with $\hat{n}_s = 19$, OJ 287 with $\hat{n}_s = 16$, S2 0109+22 with $\hat{n}_s = 10$, S3 0458–02 with $\hat{n}_s = 9$) have $\sigma_{\nu,i} \in [0.45,\ 0.65]$ dex, comparable to or smaller than $\sigma_{\rm int}$, so the predictive variance is $\sigma_{\rm int}$-dominated and a deviation from the calibrated relation would in principle be visible. The three sources with $\hat{n}_s \leqslant 3$ (3C 454.3, Ton 599, PKS 1441+25) instead have $\sigma_{\nu,i} \in [1.1,\ 2.0]$ dex, two to three times larger than $\sigma_{\rm int}$, so their predictive variance is dominated by $\sigma_{\nu,i}$ and any luminosity assignment compatible with $\hat{n}_s > 0$ trivially passes the consistency test. Restricted to the high-$\hat{n}_s$ subsample, the test gives $\chi^2_4 = 1.24$ ($p = 0.87$, $\max|z_i| = 0.89$); restricted to the low-$\hat{n}_s$ subsample, $\chi^2_3 = 0.34$ ($p = 0.95$). The full $\chi^2_7 = 1.58$ should therefore be read primarily as the χ^2_4 contribution of the four high-$\hat{n}_s$ sources, with the three low-$\hat{n}_s$ sources contributing $\sigma_{\nu,i}$-dominated noise rather than discriminating information.

We emphasize that the relevant quantity for the consistency test is $\hat{n}_s$ (through $\sigma_{\nu,i} \propto 1/\sqrt{\hat{n}_s}$), not the local IceCube significance, and the two are not monotonically related: the local test statistic depends on the fitted spectral index, the declination-dependent atmospheric background, and the energies of the contributing events, so a few hard-spectrum, high-energy events in a low-background regime can yield a higher local significance than many soft-spectrum, threshold-near events. This is why, for example, 3C 454.3 ($\hat{n}_s = 1$) reaches a higher local significance (1.6σ) than MITG J201534+3710 ($\hat{n}_s = 19$, 0.9σ; Table 1). The discriminating power of our test, set by $\sigma_{\nu,i}$, therefore resides in the high-$\hat{n}_s$ sources even though these carry the lowest local significances; there is no contradiction, since the local significance measures detection against background while $\sigma_{\nu,i}$ measures the precision of the inferred neutrino luminosity.

A separate consequence of this structure is that the $\hat{n}_s > 0$ selection cut applied to the test sample acts as a survival filter on precisely the quantity being tested. Sources whose underlying $\hat{n}_s$ fluctuated below zero under noise are excluded by construction; sources at $\hat{n}_s = 1$ or 2 pass the cut with $\sigma_{\nu,i}$ large enough that any L_ν within ∼1–2 dex of the prediction looks acceptable. Both effects push the test toward apparent

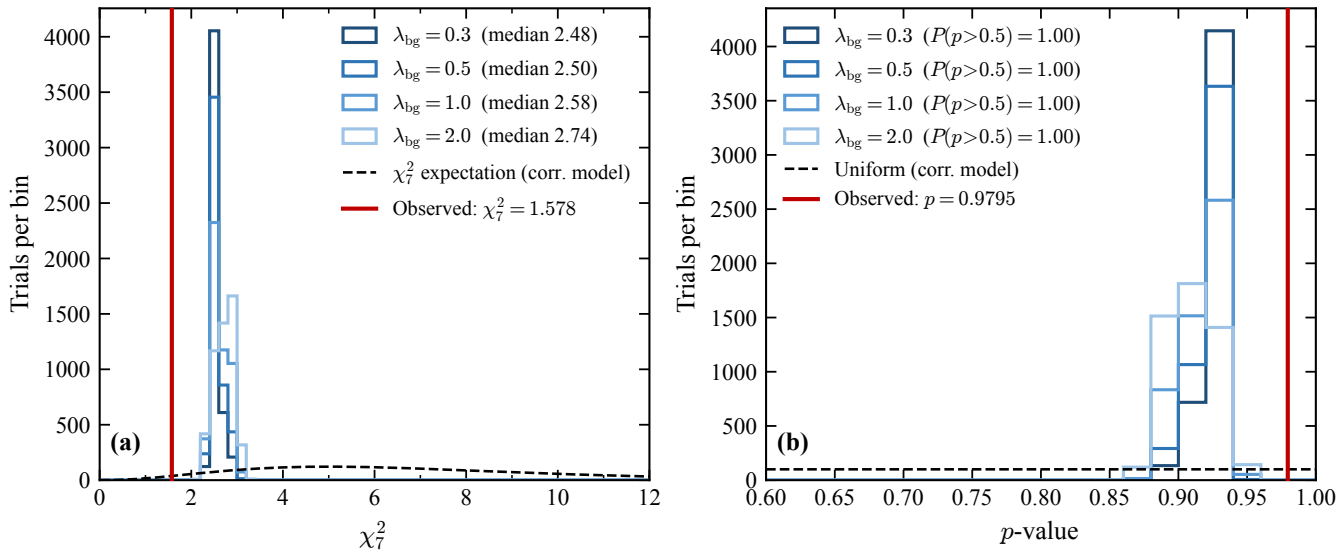

Figure 2. Null calibration of the posterior-predictive consistency test. Left: Distribution of χ^2_7 under a background-only null in which $\hat{n}_s$ values are drawn from Poisson(λ_{bg}) truncated to $\hat{n}_s > 0$, for four background expectations $\lambda_{bg} \in \{0.3, 0.5, 1.0, 2.0\}$. The black dashed curve is the χ^2_7 distribution that would obtain if the test were correctly calibrated against the K2024 relation (median ∼7). The observed value $\chi^2_7 = 1.58$ (red vertical line) sits below the null medians (∼2.5–2.7). Both the null and the observed value lie far below the correctly calibrated expectation, reflecting the wide predictive band that follows from $\sigma_{int} = 0.60$ dex at $N_{cal} = 6$. Right: Distribution of p-values under the same null. Median p_{null} exceeds 0.9 across all four λ_{bg}; $P(p_{null} > 0.5) = 1.00$ in every case. The observed $p = 0.980$ sits in the upper tail of all four null distributions, but the predictive band is too wide for this to constitute formal rejection of the background hypothesis.

consistency independent of the truth of the signal hypothesis. The null calibration of Section 3.3 quantifies this directly.

3.3. Null-calibration of the Consistency Test

To assess whether $\chi^2_7 = 1.58$ ($p = 0.980$) is itself informative about the reality of the seven IceCube excesses, we calibrate the test against an explicit background-only null. The construction is the following. We hold the X-ray luminosities $\{\log L_{X,i}\}$ and redshifts $\{z_i\}$ of the seven test sources fixed at their observed values, preserving the L_{hX} and z distributions exactly. We then generate $\hat{n}_s$ realizations under an IceCube background-only hypothesis, $\hat{n}_{s,i} \sim$ Poisson(λ_{bg}), and apply the $\hat{n}_s > 0$ truncation that mirrors the post-hoc selection cut imposed on the real test sample. Each null realization is propagated through the same flux pipeline (Section 2.2), the same 4.47 × scaled Poisson $\sigma_{\nu,i}$, and the same plug-in posterior-predictive consistency test against the calibrated K2024 relation, yielding a null distribution of χ^2_7 values. We explore $\lambda_{bg} \in \{0.3, 0.5, 1.0, 2.0\}$, spanning the range of empty-sky background expectations for declination-dependent IceCube point-source analyzes; 5000 realizations are drawn at each λ_{bg}. Use of the same $\hat{\gamma}_\nu$ values as the observed sample is conservative for the null, since true background events would have a more steeply falling spectrum.

The results are summarized in Figure 2 and Table 3. Median null χ^2_7 values lie in the range 2.48–2.74 across the four λ_{bg}, corresponding to median $p_{null} \in [0.91, 0.93]$. The probability that a background-only realization produces $p_{null} > 0.5$ is unity in every case, and $P(p_{null} > 0.9)$ ranges from 0.67 to 0.97. Both the null and the observed distributions sit far below the χ^2_7 distribution that would obtain under a correctly calibrated test of the K2024 relation (median $\chi^2_7 \approx 7$, dashed line in Figure 2, left). This reflects the fundamental property of the predictive band: with $\sigma_{int} = 0.60$ dex calibrated on six sources and propagated as a plug-in into $\sigma_{pred,i}$, the band is wide enough that no realistic seven-source sample, signal or background, populates the upper tail of the χ^2_7 expectation. The test is conservative in both directions: it does not falsely reject the calibrated relation when the relation holds, but it also does not reject background.

Table 3
Null Calibration of χ^2_7 Under Background-only $\hat{n}_s$ Realizations

λ_{bg}	Med. χ^2_{null}	Med. p_{null}	$P(p_{null} > 0.5)$	$P(p_{null} > 0.9)$
0.3	2.48	0.93	1.00	0.97
0.5	2.50	0.93	1.00	0.94
1.0	2.58	0.92	1.00	0.83
2.0	2.74	0.91	1.00	0.67
Obs.	1.58	0.98	…	…

Note. 5000 background-only realizations per λ_{bg}. The observed $\chi^2_7 = 1.58$ sits below the null median for all four values, and $P(\chi^2_{null} \leqslant \chi^2_{obs}) < 10^{-3}$ in every case. The same predictive band that prevents the test from rejecting the calibrated relation also prevents it from rejecting background; both lie in a regime where the predictive variance dominates the per-source residual. A meaningful rejection requires an enlarged calibration sample (which would tighten σ_{int}) or an alternative test statistic that is less band-dominated.

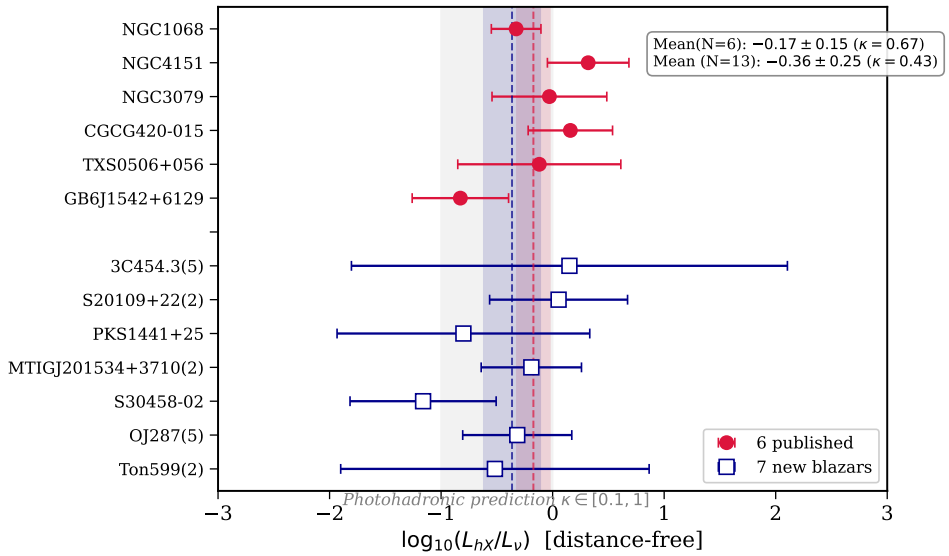

Figure 3. Distance-free $\log(L_{\rm hX}/L_\nu)$ ratio diagnostic. Red filled circles: six K2024 calibration sources; open blue squares: seven new blazars from this work. Vertical dashed lines show inverse-variance-weighted means with shaded 68% bands. The gray band indicates the photohadronic prediction range $\kappa \in [0.1, 1]$. Population-weighted κ values are 0.67 (calibration) and 0.43 (new sample); the two means differ by -0.19 ± 0.29, consistent with zero at $\sim 0.7\,\sigma$. Error bars per source reflect the spectrum-derived $\sigma_{\nu,i}$ for the calibration sample and the empirically scaled Poisson approximation for the test sample (Section 2.2). The agreement is necessary but not sufficient evidence of common physics, since the joint NuSTAR +IceCube selection envelope preferentially populates the same κ window that the photohadronic prediction occupies; see Section 3.4.

We draw two conclusions from this calibration. First, the formal χ^2_7 statistic does not, on its own, distinguish between the hypothesis that the seven IceCube excesses reflect signal and the hypothesis that they are selected background fluctuations: both populate the same low-χ^2, high-p region of the test distribution. Second, the observed value $\chi^2_7 = 1.58$ is nonetheless lower than the null median by approximately one unit, a direction consistent with signal but not statistically distinguishable from background within the precision afforded by the present calibration. The test as constructed is therefore best read as a conditional consistency check: under the assumption that the IceCube excesses reflect signal, the seven blazars are statistically indistinguishable from the K2024 calibration; under the alternative that they are selected background fluctuations, the test does not reject that hypothesis either. Discriminating between these two readings requires either a tighter $\sigma_{\rm int}$ from an enlarged calibration sample, or an independent test of detection-level significance via X-ray-weighted IceCube stacking with public event-level data.

3.4. *Distance-free* L_{hX}/L_ν *Ratio Diagnostic*

A second, independent diagnostic uses the source-frame ratio

$$\log\left(\frac{L_{\rm hX}}{L_\nu}\right) = \log\left(\frac{F_X}{F_\nu}\right) + (\Gamma_X - \gamma_\nu)\log_{10}(1+z), \qquad (14)$$

which cancels the common d_L^2 factor exactly. The photohadronic reprocessing framework predicts $\kappa = L_{\rm hX}/L_\nu \in [0.1, 1]$, set by the cascade development and photon escape fraction (Inoue et al. 2020; Murase et al. 2020; Eichmann et al. 2022); the relevant physics lives entirely in source-frame quantities and is unaffected by sample distance distribution. Figure 3 shows the per-source ratios for the calibration and test samples. Inverse-variance-weighted means are

$$\langle \log(L_{\rm hX}/L_\nu)\rangle_{\rm pub} = -0.17 \pm 0.15 \quad (\kappa \approx 0.67), \qquad (15)$$

$$\langle \log(L_{\rm hX}/L_\nu)\rangle_{\rm new} = -0.36 \pm 0.25 \quad (\kappa \approx 0.43). \qquad (16)$$

Both population-weighted means lie within the photohadronic prediction band $\kappa \in [0.1, 1]$, and the two means differ by -0.19 ± 0.29, consistent with zero at $\sim 0.7\,\sigma$. The seven new blazars are therefore statistically indistinguishable from the published sample in their distance-free X-ray-to-neutrino luminosity ratio.

We caution that, like the χ^2 test, the ratio diagnostic is constructed on samples conditioned on simultaneous NuSTAR detection and IceCube $\hat{n}_s > 0$. This joint sensitivity envelope preferentially populates the very κ range that the photohadronic prediction occupies: a source with κ outside [0.1, 1] at typical distances would fall below one of the two detection thresholds and be absent from any sample so constructed. Population-averaged agreement at $0.7\,\sigma$ within this band is therefore necessary but not sufficient evidence of common physics. The diagnostic places upper limits on outliers, no source in either sample sits more than $\sim$1 dex from the band's edges, but it cannot, by construction, distinguish a coronal/cascade origin from any other emission scenario whose κ predictions overlap the same window. Equivalently, the

comparison of $\langle\kappa\rangle$ between the two samples is physically meaningful only if the hard X-rays are produced by the same mechanism in both (Section 5.3): a jet inverse-Compton contribution to the blazar 15–55 keV band that is unaccompanied by neutrino production would bias $\kappa_{\rm new}$ without reflecting a difference in the underlying photohadronic physics.

4. Flux-space Permutation Test as a Finite-sample Diagnostic

As a finite-sample diagnostic of the combined 13-source structure, we perform a flux-space permutation test in which the neutrino quantities (F_ν, γ_ν, σ_ν) are shuffled jointly across the 13 sources of the joint sample, breaking the observed pairing between each X-ray source and its measured neutrino quantities while preserving the $L_{\rm hX}$–L_ν pair structure under the random reassignment. For each permutation, $\log L_\nu$ is recomputed at the receiving source's redshift,

$$\log L_\nu^{\rm perm}[i] = \log F_\nu^{\rm perm}[i] + \log\left(4\pi d_{L,i}^2\right) + \left(\gamma_\nu^{\rm perm}[i] - 2\right)\log(1+z_i), \quad (17)$$

where $F_\nu^{\rm perm}[i]$ and $\gamma_\nu^{\rm perm}[i]$ are the neutrino flux and spectral index drawn from a random reassignment of the 13 neutrino measurements to the 13 X-ray sources, and $d_{L,i}$, z_i are the luminosity distance and redshift of source i (its original X-ray sky position). This coupled shuffle preserves the d_L^2 factor identically in observation and null, so the test asks whether the observed $L_{\rm hX}$ and L_ν values are paired more tightly than random reassignment within the same IceCube-positive sample. The test statistic is the inverse-variance-weighted rms scatter of $\log L_\nu$ around the weighted forward regression of L_ν on $L_{\rm hX}$, with weights $w_i = 1/(\sigma_{\nu,i}^2 + \sigma_{X,i}^2)$ recomputed per permutation from the shuffled σ_ν. We adopt $L_{\rm hX}$ as the regressor because the X-ray luminosities are substantially better constrained than the Poisson-dominated neutrino luminosities, so this choice minimizes the regression-dilution bias that would otherwise be incurred by using the noisier variable as the predictor. We use $N_{\rm perm} = 10^7$ realizations.

The observed weighted scatter is $\sigma_{\rm scat}^{\rm w,obs} = 0.321$ dex, compared to a null mean $\langle\sigma_{\rm scat}^{\rm w}\rangle_{\rm null} \simeq 1.08$ dex. The p-value is $p = 6.3 \times 10^{-4}$ (3.23σ). For comparison, a naive luminosity-space permutation that scrambles $\log L_\nu$ directly leaves the distance inflation in the observation but breaks it in the null, yielding $\langle\sigma_{\rm scat}^{\rm w}\rangle_{\rm null,lum} \simeq 1.43$ dex, $p = 2.0 \times 10^{-7}$, and an artificially inflated significance of 5.07σ. The gap between the two tests, $\sim 1.8\sigma$, quantifies the d_L^2 contribution that the flux-space construction removes.

The flux-space null distribution is non-Gaussian, with three visible modes that reflect the trimodal $\log d_L^2$ distribution of the joint 13-source sample: four nearby Seyferts at $z \lesssim 0.03$, three mid-redshift blazars at $z \sim 0.3$, and six high-redshift blazars at $z \gtrsim 0.6$. Under a random reassignment, the high-weight neutrino fluxes are shuffled into one of these three distance clusters, producing three regimes of weighted scatter. The p-value is therefore evaluated directly from the empirical CDF of the permutation distribution and does not assume any parametric form for the null.

We treat this permutation result as a secondary, finite-sample diagnostic rather than a detection statistic, in keeping with the framing established in the introduction. The test is conditional on the IceCube-positive sample (it does not establish detection against the blazar population at large), and the flux-space dynamic range available for shuffling is limited by the clustering of the seven new F_ν values near the IceCube threshold. Both effects are addressed in Section 5.4. We also note that this test, unlike the χ^2 test of Section 3, is sensitive to the relative ordering of $L_{\rm hX}$ and L_ν within the joint sample rather than to the absolute predictive band, and the two diagnostics therefore probe partially independent aspects of the structure. Figure 4 shows the observed correlation alongside both null distributions.

As an independent, rank-based check of the distance-controlled structure, we computed the partial rank correlation between $\log L_{\rm hX}$ and $\log L_\nu$ on the joint 13-source sample, controlling for redshift. For rank statistics this is equivalent to controlling for the common $4\pi d_L^2$ factor, since z and d_L are monotonically related. The partial Kendall τ (Padovani 1992; Akritas & Siebert 1996) is $\tau|z = 0.69$ (raw $\tau = 0.90$) and the partial Spearman is $\rho|z = 0.71$ (raw $\rho = 0.96$); a residual (Freedman–Lane) permutation test gives $p = 0.0066$ ($\sim 2.7\,\sigma$) for the partial Spearman, consistent with the 3.23σ of the flux-space test above. The reduction from the raw to the partial coefficients confirms that a substantial part of the luminosity-space correlation is distance-induced, while the surviving partial correlation reflects a real residual association. A Pearson partial controlling for z linearly (0.96) does not remove this confound, $\log(4\pi d_L^2)$ is strongly nonlinear in z across $z = 0.003$–2.29, whereas controlling for $\log d_L^2$ directly gives 0.85; the rank statistics, invariant under the monotonic z–d_L mapping, control the distance factor properly.

5. Discussion

5.1. Physical Interpretation

A simple physical argument motivates the hard X-ray–neutrino correlation within a photohadronic framework. In the Δ-resonance approximation, the threshold condition $E_p' E_\gamma' \simeq 0.3$ GeV^2 combined with the inelasticity $E_\nu' \simeq E_p'/20$ implies $E_\gamma' E_\nu' \simeq 0.015$ GeV^2 in the source frame. Transforming to the observer frame via $E^{\rm obs} = \delta\, E'/(1+z)$ for both photon and neutrino, the target photon energy required for

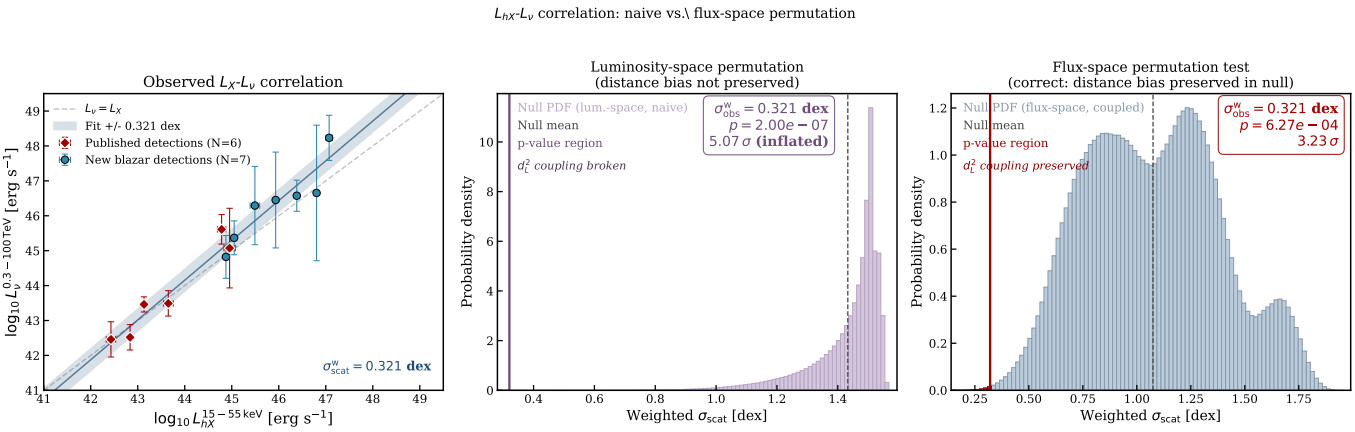

Figure 4. Flux-space permutation test as a finite-sample diagnostic. Left: Observed joint sample with the weighted forward regression and its $\pm\sigma^{w}_{scat}$ band. Middle: Null distribution from a luminosity-space permutation that breaks the d_L^2 coupling in the null but preserves it in the observation, yielding an inflated nominal significance of 5.07σ. Right: Null distribution from the flux-space permutation adopted here, in which (F_ν, γ_ν, σ_ν) are shuffled jointly and L_ν is recomputed at each receiving source's redshift; this preserves the d_L^2 coupling identically in observation and null, and gives the physically meaningful 3.23σ ($p = 6.3 \times 10^{-4}$).

neutrino production at $E_\nu^{\rm obs}$ becomes

$$E_\gamma^{\rm obs} \simeq 1.9\,{\rm keV}\left(\frac{\delta}{10}\right)^2\left(\frac{2}{1+z}\right)^2\left(\frac{E_\nu^{\rm obs}}{200\,{\rm TeV}}\right)^{-1}, \qquad (18)$$

where the $(2/(1+z))^2$ form is chosen so that the prefactor corresponds to the reference values $\delta = 10$, $z = 1$, and $E_\nu^{\rm obs} = 200$ TeV; we caution that the coefficient depends sensitively on the adopted resonance constant and inelasticity factor (values quoted in the literature span ∼1–10 keV at the same reference parameters). Equation (18) demonstrates that for typical neutrino-emitting blazars ($\delta \gtrsim 10$, $z \sim 0.5$–2), the relevant target photons lie in the soft-to-hard X-ray band covered by NuSTAR at the higher-end values of δ, and shift toward the soft X-ray band for moderate Doppler factors. The 15–55 keV NuSTAR band probes the high-δ, high-E_ν corner of the parameter space, with a soft tail that overlaps the Swift/XRT and NuSTAR sensitivity curves.

We stress that Equation (18) illustrates only the X-ray-as-target branch, the kinematics under which soft-to-hard X-ray photons can serve as the $p\gamma$ target for high-Doppler-factor sources, and is complementary to, not the basis of, the $L_{\rm hX}$–L_ν relation. The relation itself rests on the calorimetric reprocessing interpretation (Section 5.3), in which the hard X-rays are the cascade product, $L_{\rm hX} \approx \kappa L_\nu$, independent of the target photon field. The target field is class-dependent: the coronal X-ray field in the Seyferts, and the BLR (UV) field in the FSRQs, for which the neutrino spectrum is expected to peak at $\gtrsim$PeV (Dermer et al. 2014). Equation (18) and Section 5.3 are thus mutually consistent: the hard X-rays trace the cascade reprocessing regardless of which photon field provides the $p\gamma$ target.

For characteristic compact blazar emission regions ($R' \sim 10^{15}$–10^{16} cm, $\delta \sim 10$, $L_{\rm hX} \sim 10^{44}$–10^{46} erg s^{-1}), the photohadronic interaction efficiency is $f_{p\gamma} \sim 10^{-3}$–10^{-1}. The maximum particle energy attainable in such compact regions is set by the magnetic-field energy product Br, which empirically saturates near $\sim 10^{16}$ G cm across nearly ten orders of magnitude in black-hole mass for spin-down-powered systems and may be exceeded in accreting configurations (see also Allen et al. 2024). The accompanying π^0-decay photons experience high optical depths to pair production ($\tau_{\gamma\gamma}/f_{p\gamma} \sim 10^2$–$10^3$), so even moderately efficient sources are opaque to their own pionic γ-ray emission (e.g., Kun et al. 2021; Halzen 2022). The resulting electromagnetic cascades redistribute energy toward lower frequencies, leading to the quasi-calorimetric relation $L_{\rm hX} \sim \kappa L_\nu$, with $\kappa \sim 0.1$–1 set by the cascade development and photon escape fraction (Inoue et al. 2020; Murase et al. 2020; Eichmann et al. 2022). Both our calibration ($\kappa \approx 0.67$) and test ($\kappa \approx 0.43$) populations sit within this predicted band (Section 3.4). We emphasize, however, that the $\kappa \in [0.1, 1]$ window is broad and is not unique to coronal or jet-base photohadronic scenarios: a range of emission models can produce $L_{\rm hX}/L_\nu$ ratios that fall in this window for the parameter combinations relevant to flux-limited samples. The distance-free diagnostic is therefore consistent with, but does not uniquely select, a photohadronic interpretation.

5.2. *Implications for the Reality of Threshold-near IceCube Excesses*

The central observational result of this paper is that the seven IceCube candidate blazars examined here, all with local pretrial significances at the ∼1σ–2σ level, are posterior-predictively consistent with the AGN $L_{\rm hX}$–L_ν relation calibrated independently on the six K2024 sources. This consistency is non-trivial in the descriptive sense: their hard X-ray luminosities span more than two orders of magnitude, and the predicted $\log L_\nu$ values vary correspondingly, yet the

observed $\log L_\nu$ values track these predictions with $|z_i| \leqslant 0.89$ in every case, with the discriminating power concentrated in the four sources where $\sigma_{\rm int}$ rather than $\sigma_{\nu,i}$ dominates the predictive variance.

We emphasize that this result does not, on its own, establish that the threshold-near IceCube excesses are real astrophysical neutrinos rather than selected background fluctuations. The null calibration of Section 3.3 demonstrates that the same χ^2_7 test, when applied to background-only $\hat{n}_s$ realizations propagated through the same flux pipeline and selection cut, yields median $\chi^2_{\rm null}$ in the range 2.5–2.7 and median $p_{\rm null} > 0.9$ across plausible background expectations: the predictive band is wide enough that both signal and selected background populate the same low-χ^2, high-p region of the test distribution. The observed value $\chi^2_7 = 1.58$ sits below the null medians, in the direction expected if the IceCube excesses reflect signal, but the gap is not statistically distinguishable from background within the precision of the present calibration.

What we have therefore established is a conditional consistency between independent multimessenger information, measured NuSTAR luminosities and published IceCube $\hat{n}_s$ values, under the working hypothesis that the AGN hard X-ray–neutrino relation is real and applies to candidate blazars. A detection-level statement against the alternative hypothesis that the threshold-near excesses are background fluctuations requires either (i) an enlarged calibration sample that tightens $\sigma_{\rm int}$ to a value where the test becomes diagnostic of the underlying relation, or (ii) an X-ray-weighted IceCube stacking likelihood with public event-level data, in which $n_s^{\rm exp}$ is defined a priori from the $L_{\rm hX}$–L_ν relation and the test statistic is a likelihood ratio between the signal and background hypotheses rather than a χ^2 against a wide predictive band.

5.3. Implications for the Neutrino Emission Site

Extending the $L_{\rm hX}$–L_ν relation to a pure-blazar test sample is consistent with, though does not uniquely establish, a unifying scenario tied to the central engine of AGN. In both blazars and Seyferts, an SMBH is surrounded by an accretion disk and a hot corona; in this scenario, the corona supplies the dense, photon-rich environment that drives $p\gamma$ interactions, and the resulting electromagnetic cascades from γ-opaque environments redistribute energy into the hard X-ray band (this is the cascade contribution underlying the K2024 relation, not a direct thermal-Comptonization component). The photon field serving as the $p\gamma$ target need not be the same across classes, the coronal X-ray field in the Seyferts, and more plausibly the BLR UV field in the FSRQs, but in either case the hard X-rays entering the relation are the reprocessed cascade product, which is why a single $L_{\rm hX}$–L_ν relation can span both populations. Direct coronal emission via thermal Comptonization (e.g. Oh et al. 2018) contributes in parallel, and is itself coupled to the same coronal properties (electron density, temperature, optical depth) that set the photon-target field for $p\gamma$ interactions, so its luminosity is not statistically independent of L_ν either. The cascade contribution, however, provides the more direct calorimetric link: every $p\gamma$ event reprocesses into hard X-rays via energy conservation in the γ-opaque environment, fixing $L_{\rm hX}^{\rm casc} \approx \kappa\, L_\nu$ with κ set by cascade development rather than by independent coronal microphysics. If the same physical regions drive both, neutrino production becomes a generic outcome of accretion-powered systems, independent of jet presence.

Relativistic beaming in blazars enhances detectability and shifts photon energies into the observed hard X-ray band, but does not by itself break the $L_{\rm hX}$–L_ν correlation: both luminosities are beamed by a common Doppler factor δ^4, so the source-frame ratio $\kappa = L_{\rm hX}/L_\nu$ is invariant under beaming. This δ-invariance, and hence the cross-sample comparison of κ, presumes that the hard X-rays are produced by the same mechanism in both populations, the cascade/coronal channel co-spatial with the neutrino-producing region, so that $L_{\rm hX}$ and L_ν share a common Doppler boost. If part of the blazar 15–55 keV emission arises from jet inverse-Compton scattering that is beamed but not accompanied by neutrino production, κ no longer compares the same physical quantity across the two samples. The fact that both populations cluster within the same κ window in the distance-free diagnostic (Section 3.4) is therefore not trivially expected: if blazar hard X-ray emission were dominated by a beamed jet component that contributes additional X-rays without a matching neutrino output, $\kappa_{\rm blazar}$ would be elevated relative to $\kappa_{\rm Seyfert}$. A proton-loaded jet contributing to both channels would not produce such a shift, so the observed agreement is consistent with the corona supplying the dominant photon target in both populations but does not by itself exclude proton-loaded jet contributions, given also the broad and selection-biased κ window discussed above.

We stress that for FSRQs and BL Lacs, hard X-rays in the 15–55 keV band can be produced by jet inverse-Compton emission, by reflected or coronal emission, or by an electromagnetic cascade contribution following $\gamma\gamma$ absorption of pionic γ rays. Distinguishing between a coronal and a jet-base origin requires observations that can separate the two components, simultaneous broadband SED modeling, Doppler-factor inference, or correlated-variability studies between hard X-ray and γ-ray emission. For TXS 0506+056, the anti-correlation between Fermi-LAT and NuSTAR light curves reported in K2024 provides circumstantial evidence against jet dominance of the 15–55 keV band in that source, but a generalization to the rest of the blazar sample requires source-by-source SED decomposition that is beyond the present scope. The observed consistency therefore constrains the physical conditions in the emission region (photon energy, density, compactness) but does not by itself uniquely determine the emission site.

5.4. Caveats and Limitations

Distance and selection biases in the secondary permutation test. The flux-space permutation test (Section 4) preserves the d_L^2 coupling identically in observation and null, so the distance bias is corrected by construction; the luminosity-space variant (5.07σ) is shown only to quantify this contribution. The IceCube selection bias acts conservatively in the flux-space test: because shuffled neutrino fluxes are drawn from a neutrino-bright pool, the null already exhibits less scatter than a truly random blazar sample would, so the reported 3.23σ is a lower bound on the significance achievable with a blind blazar control. Such a control cannot be constructed from publicly available IceCube data alone; it would require source-specific neutrino flux estimates for an $N \sim 10^2$ blazar sample not preselected as neutrino candidates.

A complementary analysis by Luo et al. (2026) using a 1114-AGN Swift-BAT sample and 10 yr of public IceCube data demonstrates explicitly that the L_X–L_ν correlation can be reproduced from random sky positions through the same d_L^2 coupling that motivates our flux-space permutation test (Section 4). Their Monte Carlo simulations reproduce $r \approx 0.85$–0.93 in luminosity space at TS cuts matching the K2024 sample, but yield $r_F = 0.07$–0.20 in flux space. Our framing as a conditional consistency check, together with the flux-space permutation test and the distance-free κ diagnostic of Section 3.4, is designed to be robust exactly against the selection-effect concern the authors raised.

Neutrino luminosity in the 0.3-100 TeV energy range. L_ν is luminosity integrated over 0.3–100 TeV (Section 2.2); for FSRQs whose physical neutrino SED may peak at $\gtrsim$PeV via BLR target $p\gamma$ (Dermer et al. 2014), it is bound-specific luminosity, a lower bound on the bolometric neutrino luminosity. Because the band is applied identically to the calibration and test samples and the relation is calorimetric (the hard X-rays trace the cascade product; Section 5.3), the band choice does not bias the comparison. Extending the integration to 10 PeV (Appendix B) shifts only the hardest-spectrum (lowest-$\hat{n}_s$) sources and leaves the consistency conclusion unchanged: under the plug-in $\sigma_{\rm pred}$, $\chi^2_7 = 2.46$ ($p = 0.93$); once the relation-parameter covariance is included, the band choice becomes immaterial ($\chi^2_7 = 0.78$ and 0.96 for the 0.3–100 TeV and 0.3–10 PeV bands, respectively).

Non-simultaneity of NuSTAR *and IceCube data.* The IceCube data underlying the published $\hat{n}_s$ and $\hat{\gamma}_\nu$ values (IceCube Collaboration 2022) span 2011 May 13 to 2020 May 29, while several of our NuSTAR epochs (Table 1) post-date this window. For a strict source-frame $L_{\rm hX}$–L_ν correlation in which $L_{\rm hX}$ traces the photon field driving $p\gamma$ neutrino production at the moment of emission, comparing a post-2020 NuSTAR flux to an IceCube excess accumulated over 2011–2020 is physically imperfect. However, the NuSTAR flux variability across multi-epoch sources is typically within a factor of $\sim$2–3 (Section 2; $\sim$0.3–0.5 dex), smaller than the intrinsic scatter $\sigma_{\rm int} = 0.60$ dex that sets the predictive band, so the consistency test is not driven by epoch-to-epoch flux variability. We have verified that excluding the post-2020 NuSTAR epochs and using only contemporaneous observations does not change the qualitative outcome (the four high-$\hat{n}_s$ sources retain $|z| < 1$), but the impact is non-negligible at the source level for the most variable members of the sample, and time-resolved follow-up remains a natural improvement for individual blazar studies.

Asymmetric error treatment between calibration and test samples and regime mismatch of the empirical scaling factor. The calibration uses spectrum-derived $\sigma_{\nu,i}$ from the IceCube profile likelihood (on average $\sim$4.5$\times$ the Poisson form); the seven test sources lack a public profile and use the scaled Poisson form (Section 2.2). The K2024 calibration spans $\hat{n}_s \in [5, 79]$; three test sources have $\hat{n}_s \leqslant 3$ where the Poisson form is less accurate. The high-$\hat{n}_s$ versus low-$\hat{n}_s$ split (Section 3.2) addresses this: the four high-$\hat{n}_s$ sources give $\chi^2_4 = 1.24$; the three low-$\hat{n}_s$ sources are $\sigma_{\nu,i}$-dominated. A sensitivity check reducing the scaling factor to 2.5 (below the calibrated range 3.7–5.2) yields $\chi^2_7 = 2.63$ and $\max|z_i| = 1.02$, all within $|z| < 1.1$, robust to the scaling-factor choice.

Limited statistical power of the consistency test. The intrinsic scatter recovered from the calibration is $\sigma_{\rm int} = 0.60^{+0.78}_{-0.33}$ dex, broad relative to the typical predictive residual. Adopted at its posterior median in the plug-in formulation of Section 3.2, this is the dominant component of $\sigma_{{\rm pred},i}$ for the high-$\hat{n}_s$ test sources and reflects the limited calibration sample size ($N_{\rm cal} = 6$). The $\chi^2_7 = 1.58$ ($p = 0.980$) should therefore be read as "no source is in tension with the predictive band, even where the band is tightest," rather than as a positive measurement of relation tightness beyond the prior. The null calibration of Section 3.3 quantifies this: median $\chi^2_{\rm null}$ under background-only realizations is $\sim$2.5, showing that a low χ^2_7 value is the generic outcome of the present pipeline rather than a signature of signal. Full marginalization over the calibration chain inflates $\sigma_{{\rm pred},i}$ to $\gtrsim$2 dex per source via the heavy upper tail of the $\sigma_{\rm int}$ posterior at $N_{\rm cal} = 6$, which motivates the plug-in choice (Section 3.2).

6. Conclusion

We have addressed the question of whether threshold-near IceCube excesses associated with seven NuSTAR-observed candidate blazars are statistically consistent with being drawn from the hard X-ray–neutrino relation reported by K2024 for the six AGN with the highest IceCube significances, under the working hypothesis that the published $\hat{n}_s$ values reflect signal. We tested the new sample against the K2024 correlation using a posterior-predictive consistency check, characterized the test's discriminating power as a function of $\hat{n}_s$, and calibrated it against an explicit background-only null.

The central findings are:

1. The $L_{\rm hX}$–L_ν relation calibrated on the six K2024 sources only, using a Bayesian regression with errors on both axes (`linmix` implementation of Kelly 2007), gives $\beta = 1.26^{+0.40}_{-0.38}$ (consistent with $\beta = 1$ at $\sim 0.7\sigma$) and $\sigma_{\rm int} = 0.60^{+0.78}_{-0.33}$ dex. The credible intervals are wide given $N_{\rm cal} = 6$.
2. All seven new blazars are posterior-predictively consistent with the calibration: $|z_i| \leqslant 0.89$ for every source, global $\chi^2_7 = 1.58$, $p = 0.980$ (see Figure 1). The discriminating power of the test is concentrated in the four high-$\hat{n}_s$ sources ($\hat{n}_s \geqslant 9$), which give $\chi^2_4 = 1.24$ ($p = 0.87$); the three low-$\hat{n}_s$ sources ($\hat{n}_s \leqslant 3$) give $\chi^2_3 = 0.34$ ($p = 0.95$) and lie in the regime where $\sigma_{\nu,i}$ dominates the predictive variance, rendering the test insensitive there.
3. A null calibration with background-only $\hat{n}_s$ realizations propagated through the same pipeline gives median $\chi^2_{\rm null}$ in the range 2.48–2.74 across plausible background expectations $\lambda_{\rm bg} \in [0.3, 2.0]$, with $P(p_{\rm null} > 0.5) = 1.00$ in every case (see Figure 2). The χ^2_7 test as constructed therefore does not, on its own, distinguish between the signal and selected-background hypotheses; the observed value sits below the null median by ~ 1 unit, in the direction expected for signal but not statistically distinguishable from background at the present calibration precision.
4. The distance-free $\log(L_{\rm hX}/L_\nu)$ diagnostic places both populations within the photohadronic $\kappa \in [0.1, 1]$ band (see Figure 3), with calibration and test means consistent with each other at $\sim 0.7\,\sigma$. We caution that the joint NuSTAR +IceCube selection envelope preferentially populates this same κ window, so the agreement is necessary but not sufficient evidence of common physics.
5. As a finite-sample diagnostic, a weighted flux-space permutation test on the joint 13-source sample gives $\sigma^{\rm w,obs}_{\rm scat} = 0.321$ dex versus a null mean of 1.08 dex; $p = 6.3 \times 10^{-4}$ (3.23σ), d_L^2-bias controlled by construction (see Figure 4). This test probes the relative ordering of $L_{\rm hX}$ and L_ν within the joint sample and is partially independent of the predictive-band-dominated χ^2 test.

We interpret these results as a conditional consistency check: under the working assumption that the IceCube excesses reflect signal, the seven candidate blazars are statistically indistinguishable from the K2024 calibration in posterior-predictive residual, in distance-free luminosity ratio, and in flux-space permutation structure. The analysis is not an independent IceCube detection claim, and the present test does not, on its own, adjudicate between the signal hypothesis and the alternative that the threshold-near $\hat{n}_s$ values are selected background fluctuations: the predictive band is wide enough that both populate the same region of the χ^2 distribution. A detection-level statement for the candidate-blazar population would require either (i) an enlarged calibration sample that tightens $\sigma_{\rm int}$ to a value where the consistency test becomes diagnostic, or (ii) an X-ray-weighted IceCube stacking likelihood with public event-level data, defined a priori from the $L_{\rm hX}$–L_ν relation. We identify the latter as the natural next step.

Future work will benefit from (i) an enlarged calibration sample to sharpen $\sigma_{\rm int}$ and the posterior-predictive band, (ii) targeted NuSTAR observations of additional IceCube candidate blazars currently lacking hard X-ray coverage, ideally contemporaneous with future IceCube exposure windows to address the non-simultaneity caveat, (iii) public release of IceCube point-source likelihood profiles (or a per-source analogue) for the threshold-near sample, which would replace the $\sim 4.5\times$ empirical scaling of the Poisson uncertainty with first-principles $\sigma_{\nu,i}$ values, especially in the low-$\hat{n}_s$ regime where the present scaling is least calibrated, (iv) a blind blazar control sample to quantify selection effects directly, and (v) collaboration with IceCube on a model-weighted stacking analysis with $n_s^{\rm exp}$ defined a priori from the $L_{\rm hX}$–L_ν relation.

Acknowledgments

We thank Ceaser C. Stringfield for carrying out the NuSTAR spectral analyzes of blazars in our sample, and Kaya Mori and Jooyun Woo for their technical and managerial support of that work. E.K. thanks funding from the NKFIH excellence grant TKP2021-NKTA-64. J.B.T., E.K., and A.F. acknowledge support from the German Science Foundation DFG, via the Collaborative Research Center *SFB1491: Cosmic Interacting Matters—from Source to Signal* (grant No. 445052434). Francis Halzen is funded by the U.S. National Science Foundation under grants PHY-2209445 and OPP-2042807. C.R. acknowledges support from SNSF Consolidator grant F01−13252. S.d.P. acknowledges support from ERC Advanced grant 789410. This research has made use of the NuSTAR Data Analysis Software (NuSTARDAS) jointly developed by the ASI Space Science Data Center (SSDC, Italy) and the California Institute of Technology (Caltech, USA). This work made use of data supplied by the UK Swift Science Data Centre at the University of Leicester. This paper makes use of publicly available Fermi-LAT data provided online by the https://fermi.gsfc.nasa.gov/ssc/data/access/ Fermi Science Support Center. This work made use of Astropy[14]: a community-developed core Python package and an ecosystem of tools and resources for astronomy (Astropy Collaboration et al. 2013, 2018, 2022), and the `linmix` package by J. Meyers (Kelly 2007).

Facilities: IceCube, NuSTAR, Fermi (LAT).

[14] http://www.astropy.org

Software: astropy (Astropy Collaboration et al. 2013, 2018, 2022), `linmix` (J. Meyers, port of Kelly 2007), NuSTAR-DAS, XSPEC.

Appendix A
Multiwavelength Light Curves

Figures 5–7 show the monthly Fermi-LAT 0.1–100 GeV photon flux (top panel) and NuSTAR 15–55 keV energy flux (bottom panel) for each source in the blazar sample. Filled circles indicate Fermi-LAT detections with test statistic TS $\geqslant 4$; downward triangles indicate monthly upper limits (TS < 4). The dashed horizontal line marks the quiescent mean flux computed over the full baseline (2008–2025) or over the 2017–2022 period for highly variable sources. Orange shaded bands and dotted vertical lines indicate the epochs of individual NuSTAR observations.

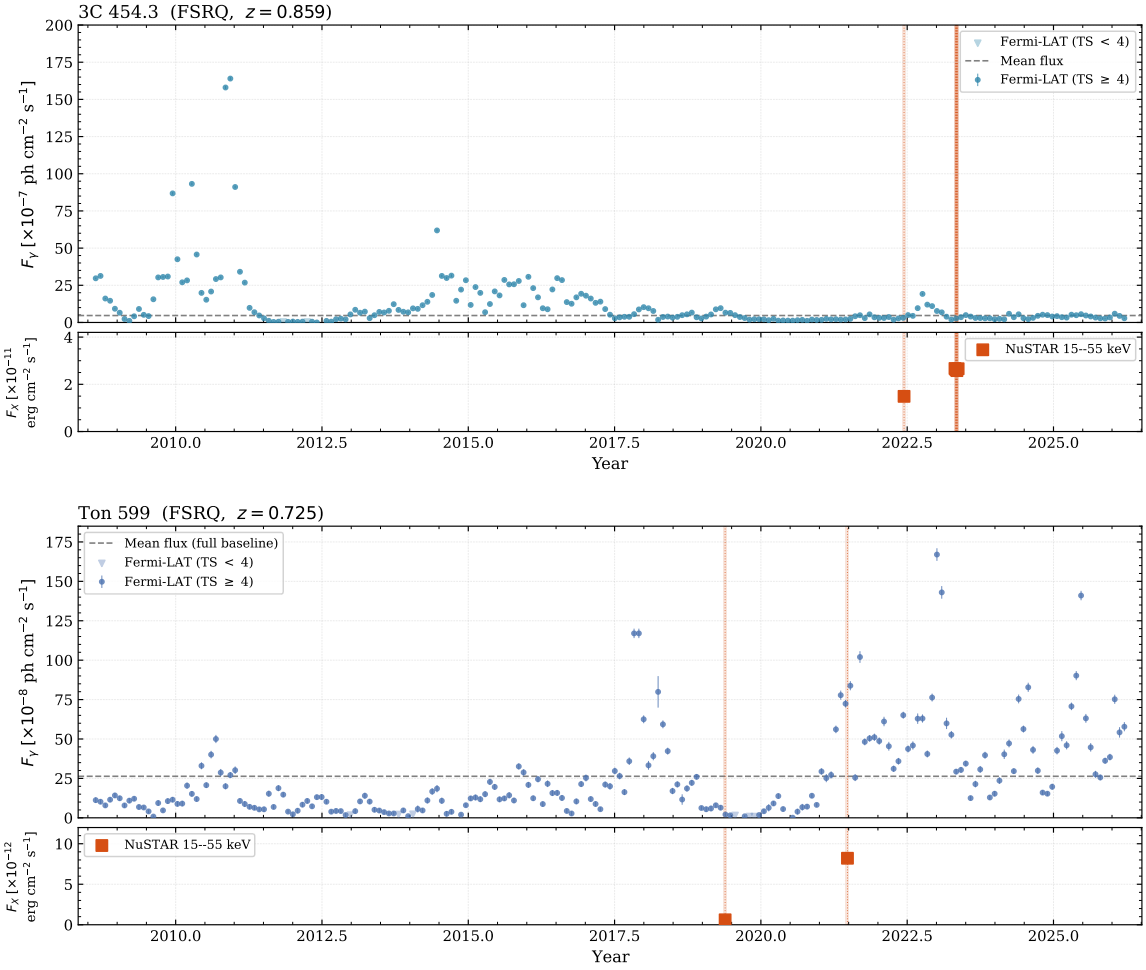

Figure 5. Top: Monthly Fermi-LAT 0.1–100 GeV photon flux (top) and NuSTAR 15–55 keV energy flux (bottom) of 3C 454.3 as a function of time. Filled circles show Fermi-LAT detections (TS $\geqslant 4$); downward triangles indicate upper limits (TS < 4). The dashed line marks the quiescent mean flux over 2017–2022. Orange shaded bands and dotted vertical lines indicate the epochs of NuSTAR observations. Bottom: Monthly Fermi-LAT 0.1–100 GeV photon flux (top) and NuSTAR 15–55 keV energy flux (bottom) Ton 599 (FSRQ, $z = 0.725$). Symbols and lines as in the top panel. Light curves from Abdollahi et al. (2023).

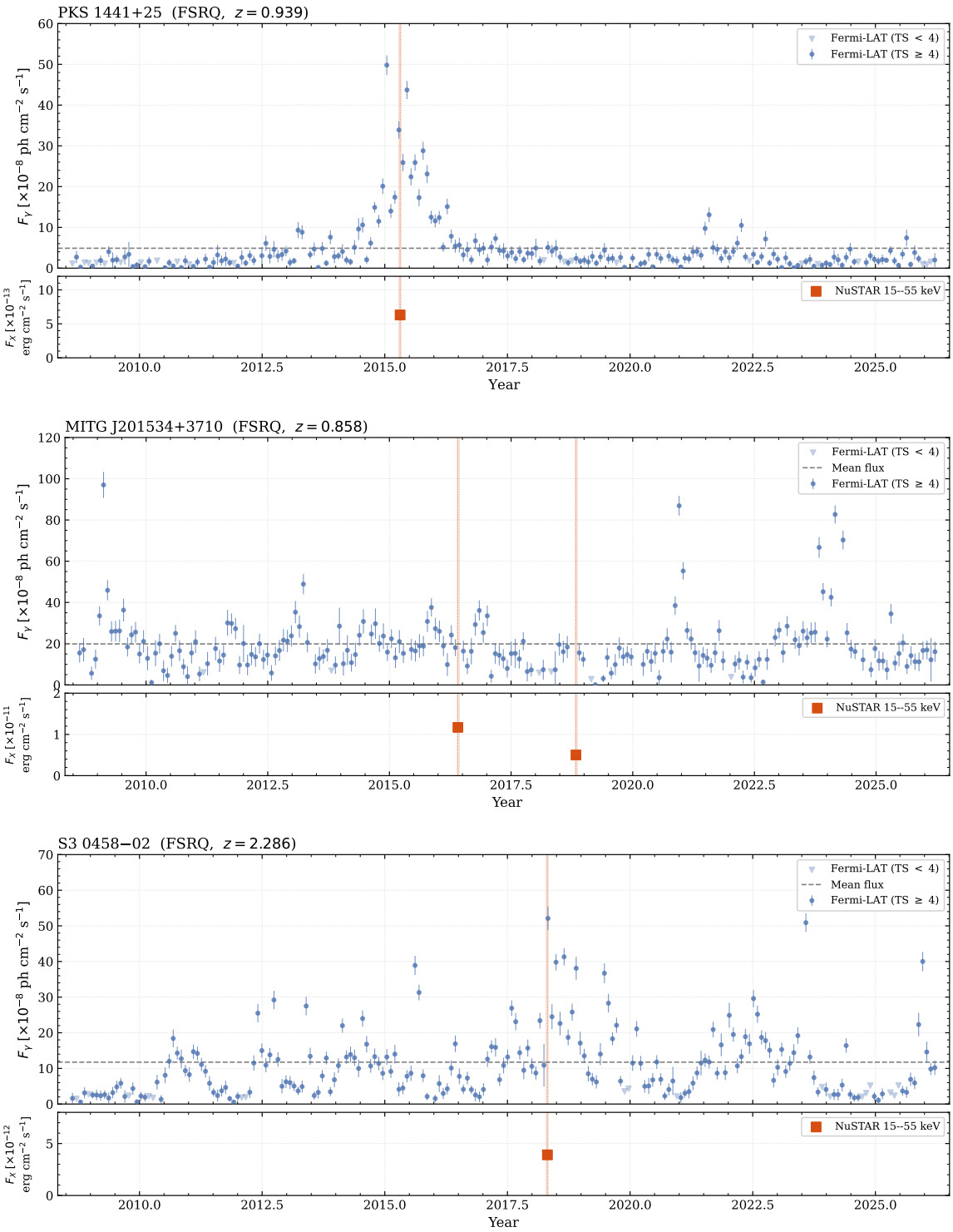

Figure 6. Fermi-LAT and NuSTAR light curves of, from top to bottom: PKS 1441+25 (FSRQ, $z = 0.939$; the neutrino spectral index was fitted freely); MITG J201534+3710 (FSRQ, $z = 0.858$); and S3 0458−02 (FSRQ, $z = 2.286$). Symbols and lines as in Figure 5. Light curves from Abdollahi et al. (2023).

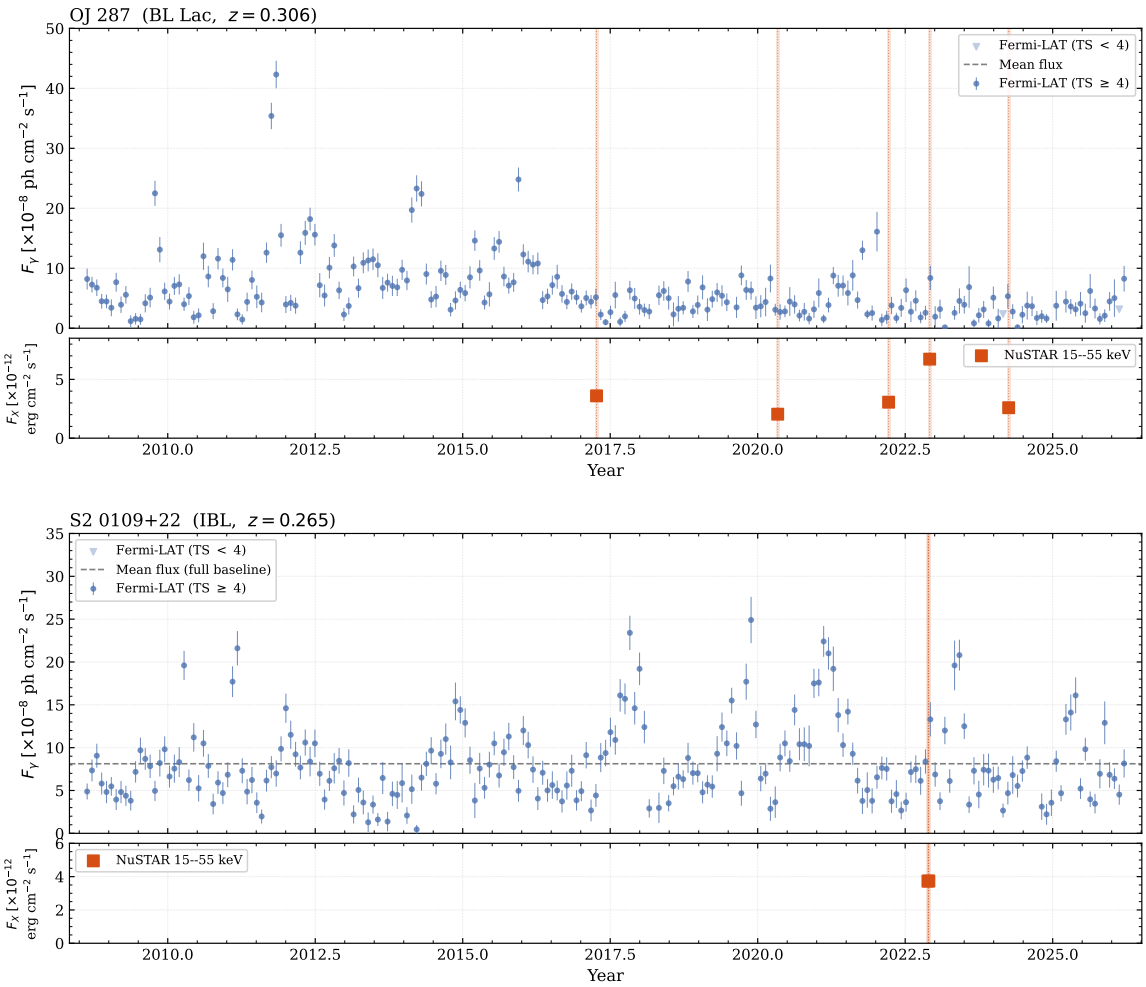

Figure 7. Fermi-LAT and NuSTAR light curves of, from top to bottom: OJ 287 (BL Lac, $z = 0.306$); S2 0109+22 (IBL, $z = 0.265$; the neutrino spectral index was fitted freely). Symbols and lines as in Figure 5. Light curves from Abdollahi et al. (2023).

Appendix B
Robustness: Neutrino Energy Band and Predictive-variance Prescriptions

This appendix quantifies the two robustness checks referenced in the main text: extending the neutrino integration band from 0.3–100 TeV to 0.3–10 PeV, and propagating the relation-parameter (α, β) covariance into the predictive variance.

Energy-band extension. Since the band-integrated energy flux of a power law $E\phi \propto E^{1-\gamma_\nu}$ enters L_ν multiplicatively, raising the upper bound from 100 TeV to 10 PeV rescales L_ν by a factor that depends only on γ_ν (the normalization ϕ_0, the redshift, and the K-correction all cancel). For the soft spectra of our sample the shift is negligible except for the hardest-spectrum sources: in the calibration only TXS 0506+056 ($\gamma_\nu = 2$) shifts appreciably (+0.25 dex), and in the test sample the shifts are concentrated in the two lowest-$\hat{n}_s$ sources (Table 4). Re-running the full calibration and consistency test over the extended band steepens the relation only marginally (β: 1.26 → 1.30, within the ±0.40 credible width) and leaves all sources consistent.

Predictive-variance prescriptions. Naively marginalizing over the full (α, β) posterior is uninformative ($\sigma_{\mathrm{pred},i} = 2$–$5$ dex, $\chi^2_7 = 0.15$): the chain standard deviation of β is 0.65–1.26 (depending on the band), a factor ∼2–3 larger than the 68% credible half-width (≈0.38), so rare steep-slope samples extrapolated 1–3 dex beyond the calibration range dominate σ_{pred}, the same heavy-tail behavior at $N_{\mathrm{cal}} = 6$ that motivates the plug-in treatment of σ_{int} (Section 3.2). We therefore propagate the parameter uncertainty using robust (16–84 percentile) scales in a pivot parametrization at $\log L_{\mathrm{hX}} = 43.39$, where intercept and slope are essentially uncorrelated (corr ≈ −0.04 to −0.08): sd(α_{pivot}) ≈ 0.33,

Table 4
Robustness of the Consistency Test to the Neutrino Energy Band and to the Predictive-variance Prescription, for the Seven Test Sources

Source	$\log L_\nu^{100\,\mathrm{TeV}}$	$\log L_\nu^{10\,\mathrm{PeV}}$	Δ	$\sigma_{\mathrm{pred}}^{\mathrm{plug}}$	z^{plug}	$\sigma_{\mathrm{pred}}^{\mathrm{rob}}$	z^{rob}
3C 454.3	46.65	47.68	+1.02	2.04	−0.55	2.46	−0.46
S2 0109+22	44.82	44.83	+0.01	0.87	−0.60	1.09	−0.50
PKS 1441+25	46.29	46.46	+0.17	1.29	+0.13	1.56	+0.10
MITG J201534+3710	46.57	46.57	0.00	0.75	−0.89	1.42	−0.49
S3 0458−02	48.23	48.23	0.00	0.89	+0.14	1.71	+0.06
OJ 287	45.37	45.37	0.00	0.78	−0.26	1.06	−0.21
Ton 599	46.45	46.45	0.00	1.51	−0.15	1.83	−0.13

Note. $\log L_\nu$ is given for the primary 0.3–100 TeV band and the extended 0.3–10 PeV band (Δ their difference); σ_{pred} and z are for the plug-in (Equation (12)) and robust-covariance (Equation (B1)) prescriptions at the primary band.

Table 5
Global χ^2_7 (p-value) Under Three Predictive-variance Prescriptions and Two Integration Bands

Prescription	0.3–100 TeV	0.3–10 PeV
Plug-in (Equation (12))	1.58 (0.980)	2.46 (0.930)
Robust (α, β) cov. (Equation (B1))	0.78 (0.998)	0.96 (0.995)
Full marg.	uninformative ($\sigma_{\mathrm{pred}} = 2$–$5$ dex; $\chi^2_7 = 0.15$ at 10 PeV)	

$\mathrm{sd}(\beta) \approx 0.38$, with

$$\sigma^2_{\mathrm{pred},i} = \sigma^2_{\nu,i} + \sigma^2_{\mathrm{int,med}} + \beta^2_{\mathrm{med}}\sigma^2_{X,i} + \left[\sigma^2_{\alpha,\mathrm{piv}} + (\log L_{\mathrm{hX},i} - 43.39)^2\ \sigma^2_\beta\right]. \quad \text{(B1)}$$

With this term, $\sigma_{\mathrm{pred},i}$ at the high-luminosity sources grows from ∼0.7 to 1.4–2.4 dex (Table 4), confirming that the plug-in predictions are tight there; the predictive z-scores roughly halve. The sample remains consistent in every prescription and band (Table 5), and the robust-covariance treatment removes essentially all sensitivity of the result to the band choice.

ORCID iDs

Emma Kun https://orcid.org/0000-0003-2769-3591
Imre Bartos https://orcid.org/0000-0001-5607-3637
Claudio Ricci https://orcid.org/0000-0001-5231-2645
Santiago del Palacio https://orcid.org/0000-0002-5761-2417
Francis Halzen https://orcid.org/0000-0001-6224-2417
Julia Becker Tjus https://orcid.org/0000-0002-1748-7367
Peter L. Biermann https://orcid.org/0000-0003-3948-6143
Anna Franckowiak https://orcid.org/0000-0002-5605-2219

References

Aartsen, M. G., Abbasi, R., Abdou, Y., et al. 2013, PhRvL, 111, 021103
Aartsen, M. G., Ackermann, M., Adams, J., et al. 2020, PhRvL, 124, 051103
Abbasi, R., Ackermann, M., Adams, J., et al. 2022, Sci, 378, 538
Abbasi, R., Ackermann, M., Adams, J., et al. 2025, ApJ, 988, 141
Abdollahi, S., Ajello, M., Baldini, L., et al. 2023, ApJS, 265, 31
Ackermann, M., Ajello, M., Atwood, W. B., et al. 2015, ApJ, 810, 14
Ahnen, M. L., Ansoldi, S., Antonelli, L. A., et al. 2015, ApJL, 815, L23
Akritas, M. G., & Siebert, J. 1996, MNRAS, 278, 919
Allen, M. L., Biermann, P. L., Chieffi, A., et al. 2024, APh, 161, 102976
Arnaud, K. A. 1996, ASPC, 101, 17
Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Barger, A. J., Cowie, L. L., Mushotzky, R. F., & Richards, E. A. 2001, AJ, 121, 662
Begelman, M. C., Blandford, R. D., & Rees, M. J. 1984, RvMP, 56, 255
Dermer, C. D., Murase, K., & Inoue, Y. 2014, JHEAp, 3, 29
Eichmann, B., Oikonomou, F., Salvatore, S., Dettmar, R.-J., & Becker Tjus, J. 2022, ApJ, 939, 43
Haardt, F., & Maraschi, L. 1991, ApJL, 380, L51
Halzen, F. 2022, IJMPD, 31, 2230003
Hewett, P. C., & Wild, V. 2010, MNRAS, 405, 2302
HI4PI Collaboration, Ben Bekhti, N., Flöer, L., et al. 2016, A&A, 594, A116
Hogg, D. W., Bovy, J., & Lang, D. 2010, arXiv:1008.4686
IceCube Collaboration 2022, Dataset, doi:10.21234/03fq-rh11
Inoue, Y., & Doi, A. 2018, ApJ, 869, 114
Inoue, Y., Khangulyan, D., & Doi, A. 2020, ApJL, 891, L33
Kawamuro, T., Ricci, C., Imanishi, M., et al. 2022, ApJ, 938, 87
Kellermann, K. I., Kovalev, Y. Y., Lister, M. L., et al. 2007, Ap&SS, 311, 231
Kelly, B. C. 2007, ApJ, 665, 1489
Khatee Zathul, A., Moulai, M., Fang, K., & Halzen, F. 2025, ApJ, 984, 54
Koss, M. J., Ricci, C., Trakhtenbrot, B., et al. 2022, ApJS, 261, 2
Kovalev, Y. Y., Pushkarev, A. B., Gómez, J. L., et al. 2025, A&A, 700, L12
Kun, E., Bartos, I., Becker Tjus, J., et al. 2021, ApJL, 911, L18
Kun, E., Bartos, I., Becker Tjus, J., et al. 2023, A&A, 679, A46
Kun, E., Bartos, I., Tjus, J. B., et al. 2024, PhRvD, 110, 123014
Luo, J.-J., Lu, M.-X., & Liang, Y.-F. 2026, JHEAp, 53, 100639
Miller, J. S., French, H. B., & Hawley, S. A. 1978, in BL Lac Objects, ed. A. M. Wolfe, 176
Murase, K., Kimura, S. S., & Mészáros, P. 2020, PhRvL, 125, 011101
Neronov, A., Savchenko, D., & Semikoz, D. V. 2024, PhRvL, 132, 101002
Oh, K., Koss, M., Markwardt, C. B., et al. 2018, ApJS, 235, 4
Padovani, P. 1992, A&A, 256, 399
Padovani, P., Alexander, D. M., Assef, R. J., et al. 2017, A&ARv, 25, 2
Paiano, S., Landoni, M., Falomo, R., Scarpa, R., & Treves, A. 2016, MNRAS, 458, 2836
Plavin, A. V., Burenin, R. A., Kovalev, Y. Y., et al. 2024, JCAP, 2024, 133
Ricci, C., Chang, C.-S., Kawamuro, T., et al. 2023, ApJL, 952, L28
Sahakyan, N., Harutyunyan, G., Gasparyan, S., & Israyelyan, D. 2024, MNRAS, 528, 5990
Wilms, J., Allen, A., & McCray, R. 2000, ApJ, 542, 914