

Using artificial neural networks in searches for Lorentz invariance violation

Tomislav Terzić^{1,*}

¹University of Rijeka, Faculty of Physics, Rijeka 51000, Croatia

Abstract. Lorentz invariance violation (LIV) in gamma rays can have multiple consequences, such as energy-dependent photon group velocity, photon instability, vacuum birefringence, and modified electromagnetic interaction. Depending on how LIV is introduced, several of these effects can occur simultaneously. Nevertheless, in experimental tests of LIV, each effect is tested separately and independently. For the first time, we are attempting to test for two effects in a single analysis: modified gamma-ray absorption and energy-dependent photon group velocity. In doing so, we are using artificial neural networks. In this contribution, we discuss our experiences with using machine learning for this purpose and present our very first results.

1 Introduction and motivation

Some theoretical approaches to quantum gravity (QG) propose that at extremely high energies space-time may no longer be smooth and continuous, and small Lorentz invariance violation (LIV) might arise. Measuring such deviations could steer the search for the theory of QG. We focus on searching for effects of LIV using very high energy gamma rays (VHE, $E > 100$ GeV). LIV is usually modelled as a modified photon dispersion relation (MDR)

$$E^2 \simeq p^2 \times \left[1 + \sum_{n=1}^{\infty} S_n \left(\frac{E}{E_{\text{QG},n}} \right)^n \right], \quad (1)$$

where E and p are gamma-ray energy and momentum, respectively, c is the standard speed of light, S_n is a prefactor which can take values ± 1 , and E_{QG} is the QG energy scale, typically expected to be on the order of Planck energy $E_{\text{P}} \approx 1.2 \times 10^{19}$ GeV. While MDR is not a direct consequence of any particular QG model, it is a simple way to parametrize non-standard effects, and the usual starting point in searches for LIV effects with gamma rays.

Consequences of MDR include energy dependent photon group velocity, modified reaction thresholds and cross sections, vacuum birefringence, etc. All of these have been tested, but so far, no trace of LIV has been found. Depending on the underlying theory, some of these effects can coexist. For example, quantum electrodynamics (QED) with LIV operators of dimension up to six in the Lagrangian, assuming preserved gauge invariance and CPT and P invariance, will lead to energy-dependent photon group velocity and anomalous QED interactions [1, 2]. Therefore, VHE gamma rays will experience a time of flight from source to detector to be energy dependent, while their emission, interaction with the background

*e-mail: tterzic@uniri.hr

radiation, and detection will be affected by LIV. Nevertheless, all of the experimental studies test only one of the effects independently. To date, only one study tested two effects on the same data set, alas again treating them mutually independent [3].

In this work, we are reporting on the first attempt to test two coexisting effects of LIV in a single analysis. Namely, we consider the gamma-ray arrival time delay and the probability of survival to the detector without being absorbed on the extragalactic background light (EBL). The LIV effect of the former is given as [4]

$$\Delta t_n \cong -S_n \frac{n+1}{2} \frac{E^n}{E_{\text{QG},n}^n} \kappa_n(z_s), \quad \kappa_n(z_s) = \frac{1}{H_0} \int_0^{z_s} \frac{(1+z)^n}{\sqrt{\Omega_\Lambda + \Omega_m(1+z)^3}} dz. \quad (2)$$

For the standard cosmological parameters of the Λ CDM cosmology model, we use $H_0 = 70 \text{ km Mpc}^{-1} \text{ s}^{-1}$, $(\Omega_\Lambda, \Omega_m) = (0.7, 0.3)$. The survival probability $P(E, z_s) = \exp(-\tau(E, z_s))$, depends on the opacity

$$\tau(E, z_s) = \int_0^{z_s} dz \frac{dl}{dz} \int_{-1}^1 d \cos \theta \left(\frac{1 - \cos \theta}{2} \right) \int_{\omega_{\text{th}}(E, \theta)}^{\infty} d\omega n(\omega, z) \sigma(E(1+z), \omega, \theta), \quad (3)$$

where E is the gamma-ray energy at detection, z_s the source redshift, n the EBL spectral density, and ω_{th} the reaction threshold of the soft photon. In this case, LIV affects both ω_{th} and the cross section $\sigma(E(1+z), \omega, \theta)$. It is important to note that the expressions for cross section usually used in the literature are approximations, and not good ones. We follow the recommendation from [2], which offers the most complete calculation up to date. For simplicity, we restrict ourselves to the case $n = 2$, i.e. only second order modification in Eq. (1) is considered, and $S_2 = -1$.

While such analysis should be possible using standard analysis tools, such as likelihood maximisation (e.g. [3, 5–7]), adding more stages of gamma-ray life in the analysis, such as inverse-Compton scattering or pion decay at the emission, or extensive air shower (EAS) development as a part of the detection process, these approaches might prove unmanageable. Therefore, we are attempting to use artificial neural networks (ANN) to search for LIV.

2 Artificial neural network training

We use the Mrk 501 flare from 2014, observed by the H.E.S.S. Collaboration [3] as a sandbox. Both the spectrum and light curve have rather simple functional forms. Moreover, since this is the only study so far in which the two LIV effects were tested on the same sample (although independently), it should allow us to compare the sensitivity of our approach. We generated a training sample modelled on the Mrk 501 data set through the following steps:

1. We generated 10^6 ordered pairs (t_e, E_e) , where t_e represents the emission time, and E_e the gamma-ray energy at emission. Temporal distribution of gamma rays was a double Gaussian with the following parameters $(A_1, \mu_1, \sigma_1) = (80.5, 2361, 2153)$, $(A_2, \mu_2, \sigma_2) = (60.5, 6564, 676)$, where A_i, μ_i, σ_i are the prefactor, mean, and width of each Gaussian, respectively. The time interval is $[0, 7200]$ s. The spectral distribution of events is a power-law with slope $\alpha = -2.03$ for energies between 35 GeV and 50 TeV.
2. We calculate the detection time $t_d = t_e + \Delta t_2$, the detected energy $E_d = E_e/(1+z_s)$, and the survival probability for every event. The time delay and survival probability are calculated for $E_{\text{QG}} \rightarrow \infty$, which corresponds to Lorentz invariant case, and 209 different values of E_{QG} , logarithmically distributed on the $[10^9, 10^{30}]$ eV interval.

3. The training sample is created by randomly picking from the pool of 10^6 events. For each event, the survival probability is compared to a random number $p_i \in [0, 1]$. All the events are considered to be emitted, but only the ones with $P(E, z_s, E_{\text{QG}}) > p_i$ ¹ are detected. The size of the detected sample is 2000. The distributions of events in both emitted and detected samples are fitted in time with a double Gaussian, and in energy with a power law. The process is repeated 100 times for each E_{QG} value, generating in total 21,000 training samples.

The training sample prepared in such a way assumes no source-intrinsic correlation between gamma-ray energy and emission time. In addition, we assumed a perfect detector.

The ANN consists of the input layer with 8 nodes, one for each parameter describing the detected sample: 6 light curve parameters, and 2 spectral parameters. There are three hidden layers of 1000 nodes each. The output layer consists of 9 nodes: 8 for the light curve and spectral parameters of the emitted sample, and one representing LIV. All input and output parameters are mapped to a $[0, 1]$ interval. The training is performed with 2500 iterations.

3 Results

To test the ANN, we feed it with parameters from the detected sample, and compare the parameters' values in the emitted sample with the values reconstructed by the ANN. The results are shown in Figure 1. There is a strong correlation between true and reconstructed values of parameters that appear both in the emitted and detected sample, and whose values are homogeneously distributed over the $[0, 1]$ interval. An example in the left plot of Figure 1, has the Pearson correlation coefficient $(r, p) = (0.94, 10^{-98})$. On the other hand, when the $[0, 1]$ interval is not well covered, the correlation is not as strong. This can be seen in the central plot of Figure 1, where, in addition to inhomogeneous distribution of the points, there are some outliers. In this case the correlation is $(r, p) = (0.63, 10^{-25})$. While $r = 0.63$ is considered to be a strong correlation, we aim for values much closer to unity for good reconstructive power. A similar situation happens with the LIV parameter (right plot of Figure 1), where $(r, p) = (0.64, 10^{-25})$. Just as in the central plot, the coverage can be optimised. This part is somewhat tricky because the null hypothesis corresponds to an infinite value of E_{QG} , and special attention should be paid to seamlessly connecting this case with finite values of E_{QG} .

4 Discussion and outlook

In this contribution we presented the first attempt at simultaneous testing of two coexisting LIV effects in the VHE gamma-ray data. We used a simple ANN structure, which successfully reconstructed some of the parameters which were present both in input and output layers of training samples. The parameter representing LIV was present in the output layer only, and in this case, the reconstruction power is substantially weaker. There is obviously room for improvement. Firstly, by more adequate mapping of the LIV parameter to the $[0, 1]$ interval and parameter space coverage, and secondly, through ANN hyperparameter optimization and different choices of training samples. These constitute our short-term plans for the continuation of this research.

Our mid-term plans include investigating different approaches; e.g. using individual events or bins in the input layer, instead of fit parameters. Furthermore, we will introduce realistic detectors with non-unit acceptance and finite energy resolution). That will be followed by an estimate of the sensitivity to systematic effects and comparison to the sensitivity to other analysis methods, such as likelihood.

¹Note that now the survival probability is also a function of E_{QG} .

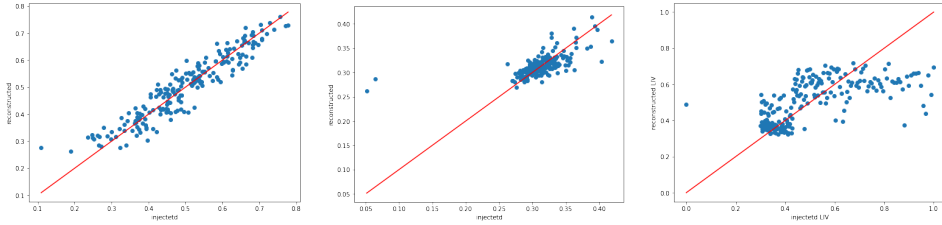


Figure 1. Comparison of injected and reconstructed parameter values. *Left:* Randomly chosen parameter from the light curve or spectral fit, with homogeneous distribution of values. *Centre:* Randomly chosen parameter, with inhomogeneous distribution. *Right:* LIV parameter. The red solid line shows the ideal reconstruction.

On the long-term, we intend to combine multiple observations, introduce additional LIV effects (e.g. EAS development), introduce additional free parameters representing source intrinsic correlation, cosmology, background models, etc.

Acknowledgements

The author acknowledges the networking support by COST actions CA18108 (QG-MM, <https://qg-mm.unizar.es/>) and CA23130 (BridgeQG, <https://web.infn.it/BridgeQG/>), and the financial support from the University of Rijeka through project uniri-iskusni-prirod-23-24 and the Croatian Science Foundation (HrZZ) Project IP-2022-10-4595. We thank the Institute for Fundamental Physics of the Universe (IFPU, <https://www.ifpu.it/>) for hosting the workshop “Astrophysical searches for quantum-gravity-induced time delays” where ideas important to this work were developed.

References

- [1] G. Rubtsov, P. Satunin, S. Sibiryakov, On calculation of cross sections in Lorentz violating theories, *Phys. Rev. D* **86**, 085012 (2012), 1204.5782. [10.1103/PhysRevD.86.085012](https://arxiv.org/abs/10.1103/PhysRevD.86.085012)
- [2] J.M. Carmona, J.L. Cortés, F. Rescic, M.A. Reyes, T. Terzić, F.I. Vrban, Approaches to photon absorption in a lorentz invariance violation scenario, *Phys. Rev. D* **110**, 063035 (2024). [10.1103/PhysRevD.110.063035](https://arxiv.org/abs/10.1103/PhysRevD.110.063035)
- [3] H. Abdalla et al. (H.E.S.S.), The 2014 TeV γ -Ray Flare of Mrk 501 Seen with H.E.S.S.: Temporal and Spectral Constraints on Lorentz Invariance Violation, *Astrophys. J.* **870**, 93 (2019), 1901.05209. [10.3847/1538-4357/aaf1c4](https://arxiv.org/abs/10.3847/1538-4357/aaf1c4)
- [4] U. Jacob, T. Piran, Lorentz-violation-induced arrival delays of cosmological particles, *J. Cosmology Astropart. Phys.* **1**, 031 (2008), 0712.2170. [10.1088/1475-7516/2008/01/031](https://arxiv.org/abs/10.1088/1475-7516/2008/01/031)
- [5] M. Martinez, M. Errando, A new approach to study energy-dependent arrival delays on photons from astrophysical sources, *Astropart. Phys.* **31**, 226 (2009), 0803.2120. [10.1016/j.astropartphys.2009.01.005](https://arxiv.org/abs/10.1016/j.astropartphys.2009.01.005)
- [6] S. Abe et al., Constraints on Lorentz invariance violation from the extraordinary Mrk 421 flare of 2014 using a novel analysis method, *Journal of Cosmology and Astroparticle Physics* **2024**, 044 (2024), 2406.07140. [10.1088/1475-7516/2024/07/044](https://arxiv.org/abs/10.1088/1475-7516/2024/07/044)
- [7] T. Terzić, D. Kerszberg, J. Strišković, Probing Quantum Gravity with Imaging Atmospheric Cherenkov Telescopes, *Universe* **7**, 345 (2021), 2109.09072. [10.3390/universe7090345](https://arxiv.org/abs/10.3390/universe7090345)