

Application of an artificial neural network for constraining masses in particle decays

Uygar Şaşmaz

Biruni University

Received: 9 February 2022, Accepted: 1 July 2022

Published online: 28 March 2023.

Abstract: The application of artificial neural networks (ANN) is a new and important trend in experimental particle physics. In this study a novel approach of using an ANN for constraining masses in particle decays is described and the performance compared to a traditional method. The ANN approach is found to perform slightly better than the traditional method and, importantly, is significantly faster. Reducing the CPU footprint of algorithms is critical for “big data” environments such as experimental particle physics experiments on the Large Hadron Collider.

Keywords: Experimental Particle Physics; Mass constraint; Artificial Neural Network

1 Introduction

In particle physics, mother particles are reconstructed from their decay products (daughter particles) that are measured in particle detectors. The reconstructed momentum resolution of the mother is limited by the momentum and angular resolutions of the particle detectors. Optimization of momentum resolution is the concern of detector design, reconstruction algorithms and calibration.

In this study, which focusses on the decay of the neutral pion to two photons ($\pi^0 \rightarrow \gamma\gamma$), a mass-constraint method for the further improvement of momentum resolution is presented. The purpose of the study is compare a well-known standard implementation of the mass-constraint to a new approach using an artificial neural network (ANN). In this comparison, two measures are important; the first is the improvement in momentum resolution for each approach, and the second is the processing time taken to achieve the improvement. The later case is critical in the “big data” environments of modern particle physics experiments.

The layout of this paper is as follows. The data used for the study is described in Section 2, the principle of the mass-constraint method is explained in Section 3, two methods for applying a mass-constraint are described in Sections 4 and 5. The performance of these methods is presented and compared in Section 6, and finally a conclusion is given in Section 7.

2 Data

This study is based on simulated data representing hadronic Z decays at the LEP collider. Events are passed through the full simulation and the reconstruction program for the ALEPH detector [1] and so provide a realistic simulation of particle momentum resolutions. Totally, 4,923,816 reconstructed events are used in study. The momentum of reconstructed neutral pions provided from this data range from a few GeV/c up to 25 GeV/c beyond which statistics become very low.

* Corresponding author e-mail: usasmaz@biruni.edu.tr

3 Neutral pion mass-constraint

While the natural (nominal) mass of the neutral pion is $0.135 \text{ GeV}/c^2$ [2] with a very small natural width, the effect of detector resolution is to smear the reconstructed invariant mass giving a broad peak around the nominal mass. Figure 1 shows a reconstructed invariant mass spectra formed from pairs of photons; pairs of photons that originate from the same mother form a π^0 signal peak while other combinations form a combinatorial background.

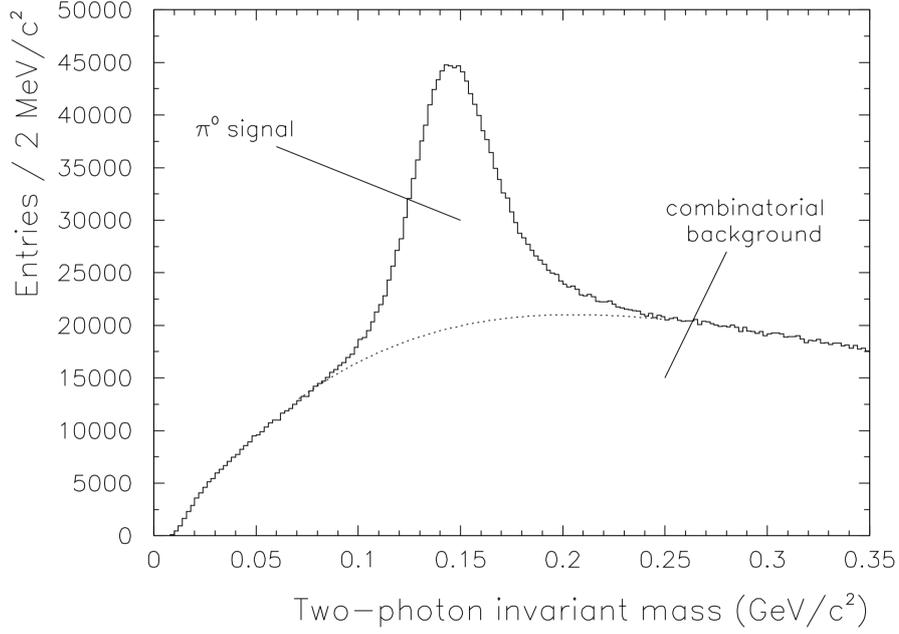


Fig. 1: Mass spectra of photon pairs. A broad signal for the decay $\pi^0 \rightarrow \gamma\gamma$ is seen clearly superimposed on a combinatorial background.

The construction of invariant mass, and its subsequent constraint is described next. Consider a mother particle of mass M denoted by 0 decaying to n daughters:

$$0 \rightarrow 1 + 2 + \dots + n \quad (1)$$

The reconstructed invariant mass of the daughters, W , can be evaluated from the relativistic relationship between energy, mass and momentum: $E^2 = P^2 + M^2$ which can be rearranged and rewritten as:

$$W^2 = \left(\sum_{i=1}^n [|\mathbf{p}_i|^2 + m_i^2]^{1/2} \right)^2 - \left(\sum_{i=1}^n \mathbf{p}_i \right)^2 \quad (2)$$

where \mathbf{p}_i and m_i are the reconstructed momentum vector and the mass of the i th daughter respectively.

For the specific decay $\pi^0 \rightarrow \gamma\gamma$, $n = 2$ and $m_1 = m_2 = 0$ (since photons are massless), Equation 2 reduces to:

$$W^2 = (|\mathbf{p}_1| + |\mathbf{p}_2|)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2 \quad (3)$$

Equation 3 can be rewritten in the equivalent but more convenient form:

$$W^2 = 2p_1p_2(1 - \cos \theta) = 2p_1p_2K \quad (4)$$

where $p_i = |\mathbf{p}_i|$, θ is the opening angle between the two photons, and $K = 1 - \cos \theta$.

A knowledge of the nominal mass, $M = 0.135\text{GeV}/c^2$, of the mother particle can be exploited to improve the reconstructed momentum of the mother particle by applying a mass-constraint. In this technique the magnitude of the photon momenta are adjusted such that their invariant mass is constrained to the nominal mass of the mother. In this refit, Equation 4 becomes:

$$M^2 = 2f_1f_2K \quad (5)$$

where f_1 and f_2 are the refitted momenta of photon 1 and 2 respectively. Note that the opening angle is not refitted and so remains constant.

In this study, two methods for the refitting process are compared. The first is a traditional Newtonian iterative method for minimizing a χ^2 [3], and the second is in the form of an artificial neural network (ANN) that performs a non-iterative refit. These two refitting solutions are explained in Section 4 and Section 5 respectively.

4 Newtonian mass-constrainer

This refit of the daughter momenta is traditionally an iterative process that takes into consideration the uncertainties in the measured values of p_1 and p_2 which are calibrated as a function of momentum. A common approach is to minimize a chi-square function of the following form under the constraint of Equation 5:

$$\chi^2 = \frac{(f_1 - p_1)^2}{\sigma_1^2} + \frac{(f_2 - p_2)^2}{\sigma_2^2} \quad (6)$$

Where σ_1 and σ_2 are the uncertainties on the measured momenta p_1 and p_2 respectively. To minimize the above χ^2 a Newtonian method of the following form is employed:

$$\mathbf{f}_{i+1} = \mathbf{f}_i - H_i^{-1} \nabla \chi_i^2, \quad (7)$$

where $\mathbf{f} = \{f_1, f_2\}$ and i is the iteration count. If the Hessian matrix H is non-singular, and \mathbf{f} is near the solution, then each iteration quadratically converges on the solution.

5 ANN mass-constrainer

Applications of sophisticated machine learning algorithms have become increasingly popular in particle physics due to their superior regression capabilities and growth of available computing power in recent years. Multivariate regression methods based on Multi-Layer Perceptron analyzer provides further performance in high energy physics analyses. Under ROOT framework [4], TMLPAnalyzer is a toolkit that consists of training, testing, validation and application of MLP algorithms.

In this study an Artificial Neural Network (ANN) approach to the mass constraint is developed and applied. The solutions for the refitted values are obtained from the function produced by the Multi-Layer Perceptron (MLP) analyzer method. This novel method is described in detail below.

The multilayer perceptron is a simple feed-forward network. It is made of neurons characterized by a bias and weighted links between them. MLP analysis requires hidden layers and neurons on the hidden layer(s). It was understood that multiple hidden layers or more neurons do not improve performance significantly after testing various number of hidden layers and neurons, but it increases the training time. Hence, it is found to be optimum for the network to contain a single hidden layer with 50 neurons.

Th MLP training procedure has various learning methods and the BFGS method (Broyden, Fletcher, Goldfarb, Shanno) is chosen for this study after testing other learning methods performances. In the training procedure, 500 epochs is found to be optimal as it is observed that improvement in performance with a higher number of epochs is insignificant while time consumption increases. 500 number of epochs is found to be optimum after testing with various epochs and observed that improvement in performance with higher number of epochs is insignificant while time consumption gets bigger.

In total, about 50,000 photons are selected to train and test neural network. Half of the data are used in the training and the other half for testing. Table 1 lists input and target variables used for training the neural network. Note that the true momentum of the photons (generated momentum) is available from the generator-level of the simulated event.

Table 1: Input and target variables for the training. Input layer of MLP consists of the first three variables in the list. Last two variables belong to the output layer.

Input layer variables	
p_{1reco}	Reconstructed momentum of the first daughter photon.
p_{2reco}	Reconstructed momentum of the second daughter photon.
W	Reconstructed invariant mass of the daughters.
Output (target) layer variables	
p_{1gen}	Generator-level momentum of the first daughter photon.
p_{2gen}	Generator-level momentum of the second daughter photon.

After training, the MLPAnalyzer toolkit provides mapping functions that directly evaluates the refitted photon momenta without iteration.

6 Performance

This section presents a comparison of the performance of the Newtonian χ^2 refit and the ANN refit methods. An important figure-of-merit is the momentum resolution of the reconstructed neutral pion; this is defined as the standard deviation of the relative momentum residual, $(p_{reco} - p_{gen})/p_{gen}$, where p_{reco} and p_{gen} are the reconstructed and generated pion momenta respectively.

Distributions of the relative momentum residual for unconstrained, Newtonian and ANN methods are given in Figure 2 for the test data. It can be seen clearly that the distributions for the Newtonian and ANN refitted data have reduced widths indicating significant improvements in the momentum resolution. Both methods appear to perform similarly. Also both methods correct for a systematic over-estimation of the momentum seen in the unconstrained data.

It was understood that invariant mass of two daughters, W , has a crucial effect on training. Without W in the training, momentum resolution performance yields significantly worse results. This result demonstrates that the ANN is learning the relation between the reconstructed momentum of daughter particles and their invariant mass.

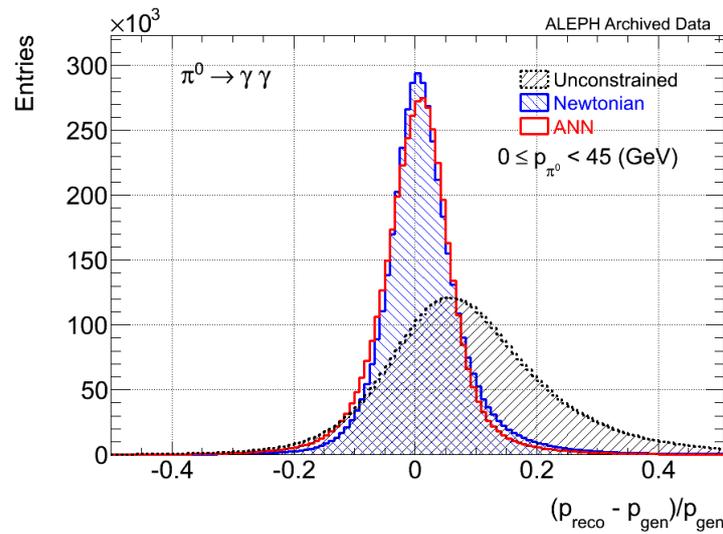


Fig. 2: Distributions of the relative momentum residual for Unconstrained, Newtonian and ANN methods.

Momentum resolution calculations are performed in intervals of pion momentum and presented in Figure 3. It is seen that the Newtonian and ANN methods perform equally well (ANN slightly out-performing the Newtonian method). The unconstrained resolution is the poorest at low momentum while the refitted resolution exhibits the best performance. Since the detector resolution is good at high momentum, both methods do not produce improvement compared to unconstrained resolution.

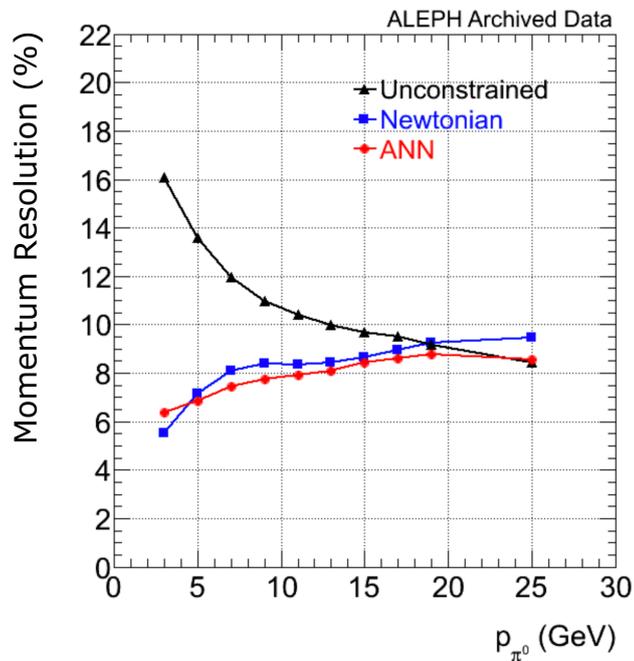


Fig. 3: Momentum resolutions for Unconstrained, Newtonian and ANN methods in momentum intervals. Errors bar are present, but smaller than the markers.

Table 2 shows the result of timing studies for the iterative χ^2 Newtonian and ANN methods. The speed of each method is considered by comparing the average time taken to refit a $\pi^0 \rightarrow \gamma\gamma$ decay. Each method is implemented in exactly the same environment and overheads common to each constraint are subtracted. The speed of each method is also compared for momentum intervals. Finally, Figure 4 shows timing test results as a function of reconstructed momentum.

Table 2: Average calculation time taken to constrain decays after overheads common to each method have been subtracted. The speedup factor is calculated by dividing time required for the χ^2 method to ANN.

Newtonian time (μs)	ANN time (μs)	Speedup factor
5.500 +- 0.003	2.479 +- 0.002	2.219 +- 0.002

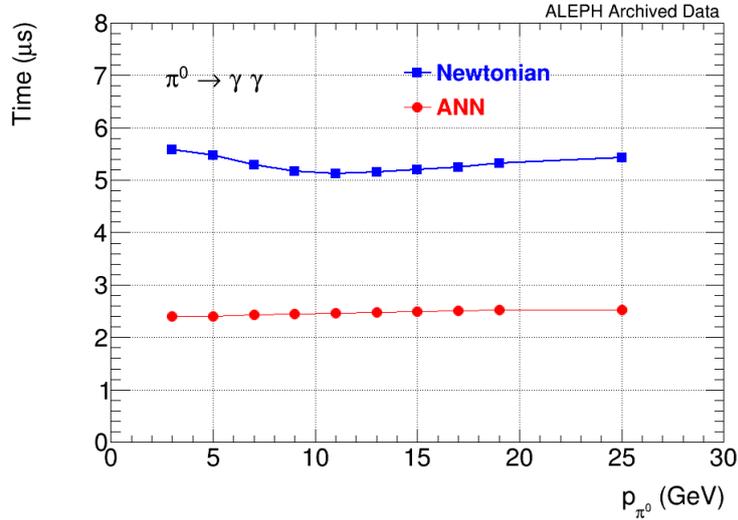


Fig. 4: Timing test results as a function of reconstructed momentum.

7 Conclusion

Using MLP provides a slightly better performance in terms of momentum resolution improvement compared to the Newtonian method. Comparing time consumptions for Newtonian and ANN method, it is understood that the ANN method is about two times faster than Newtonian. The ANN method provides an additional benefit of being easier to implement than the χ^2 Newtonian method that may also exhibit rare failures. The faster execution of the ANN provides an important advantage for big data environments such as that found in the Large Hadron Collider experiments.

The ANN method can be potentially adapted to many-body particle decays which cannot be treated by a more standard approach.

8 Acknowledgement

The author wishes to thank the ALEPH collaboration for access to the archived data [5] since the closure of the collaboration.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors have contributed to all parts of the article. All authors read and approved the final manuscript.

References

- [1] D. Buskulic et al., (ALEPH Collab.), *Nucl. Instr. Meth. A360 (1995) 481*.
- [2] P.A. Zyla et al. (Particle Data Group), *Prog. Theor. Exp. Phys. 2020, 083C01 (2020)*
- [3] K. Korzecka and T. Matulewicz, *Nucl. Instr. Meth. A453 (2000) 606*
- [4] Multi Layer Perceptron Library, algorithms and class for neural network usage.
ROOT Reference Guide (2022) <https://root.cern.ch/doc/master/classTMultiLayerPerceptron.html>
- [5] Statement on use of ALEPH data for long-term analyses
<https://dphep.web.cern.ch/content/aleph-preservation-policy>