Quantum Interference Measurement with Physics Aware Machine Learning at the Large Hadron Collider at CERN

Aishik Ghosh, David Rousseau
Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France

Introduction

Particle Physics has gone through a few waves of machine learning innovations and is giving back ideas to the machine learning research. Moving from the regime of shoehorning physics problems into forms where existing state-of-the-art machine learning solutions can be applied, particle physics is starting to marry machine learning tools with physics insight to create a new family of "physics-aware machine learning" algorithms where the objective of the tools more closely matches the actual objective of the physicist. This allows leveraging extra information from existing physics tools that can boost performances beyond the use of off-the-shelf machine learning algorithms.

We will compare a neural network model aware of the flexibility of a theoretical physics model (developed by [5]) and a traditional approach optimised at a single point in the phase space being probed with no explicit knowledge of the physics model, for a particular particle physics study very important for physics at the Large Hadron Collider (LHC), CERN. We will show that the former is better even at the particular point at which the traditional approach was optimised, simply because it "understands" the physics being studied better.

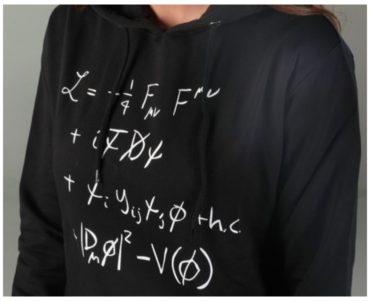
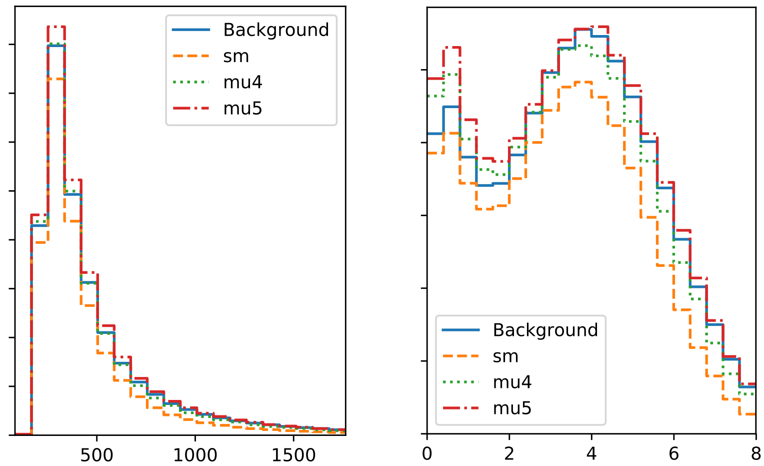


Figure 1: The Standard Model Lagrangian

The Problem of Quantum Interference of the Offshell Higgs

The Lagrangian of the Standard Model (SM) of particle physics is a mathematical formula (Figure 1) that condenses our current understanding of the universe from a quantum perspective, and is known to be incomplete (it does not explain gravity, neutrino mass or matter-antimatter dis-balance). There are several proposed mathematical extensions to the SM (Lagrangians with extra terms) but the most promising ones are already being excluded by data. The SM is continuously being tested at the LHC where the ATLAS experiment [2] collects a huge amount of data to perform precision measurements to find hints of a direction in which to expect new physics. The data is too complex to interpret without involved statistical techniques and a deep understanding of precisely what the SM predicts.

The predictions of a model (Lagrangian) using Quantum Field Theory (QFT) calculations is too expensive to compute analytically so an entire sub-field of particle physics develops Monte-Carlo based simulation techniques for it. At the end, physicists are interested in the inverse problem of going from the measured data to the value of the theory parameters (parameters in the Lagrangian) that best describe it. Re-doing the data analysis for each hypothesis (each new proposed Lagrangian) is impossible given the limited resources, and hence studies are selected that can have an impact on the assessment of multiple promising proposals. One such study is the interference between the



(a) High level feature constructed based on physics motivations

(b) A different feature

Figure 2: Distributions of (a) a physics motivated feature that is usually used for a "four lepton" analysis, but cannot differentiate between $\mu = 0$ and $\mu = 4$ (b) another physics motivated feature which can break the degeneracy between $\mu = 0$ and $\mu = 4$.

"offshell" Higgs boson processes and other "background" processes where four leptons are observed by the detector at the very end ("final state").

The "offshell" Higgs boson particle is "virtual", with a mass far away from the one described by special relativity's $E = MC^2$ (Heisenberg allowed particles to disobey Einstein as long as they do so for a very short period of time, through his famous uncertainty principle, $\sigma_E \sigma_t \geq \frac{\hbar}{2}$). Quantum mechanics also prescribes that given an initial and final state, all possible intermediate states can and will occur, and they might interfere with one another. For the ATLAS "Higgs to four leptons" study this implies that the observed physics could look different for small changes in the "Higgs Couplings", i.e. parameters in the theory that determine how strongly the Higgs interacts with other fields. In this project, these parameters are assumed to scale in similar ways and are represented together by the "signal strength" μ.

Quantum Interference is Problematic to Traditional Algorithms

Usually the signal and background quantum processes come from disjoint phase spaces and can thus be simulated separately in a particle physics simulation. However, in the presence of quantum inference between the signal and background processes, they need to be simulated together to model the probability distributions correctly. As a simplified example, the probability of having one particular sample X, denoted $P(X)$ (with $0 \leq P(X) \leq 1$) is a function of two complex functions from Quantum Field Theory, $M_s(X)$, $M_b(X)$ (with $M_s, M_b \in \mathbb{C}$), for the signal and background process respectively, is given by Eq. 1. If the third term was insignificant and could be ignored, the signal and background contributions could be simulated separately and combined when needed. However in this case, the contribution from the mixed term cannot be ignored.

$$P(X) = |M_s(X) + M_b(X)|^2 = |M_s(X)|^2 + |M_b(X)|^2 + 2Re(\overline{M_s(X)} M_b(X)) \tag{1}$$

This renders the notion of "true class labels" ill-defined, and thus the task cannot translate into a classification problem. Further, since the inference describes very different kinds of physics depending on the value of μ, any algorithm will have to be aware of the physics going on at various values of μ, not just the one at the SM (where $\mu = 1$). Figure 2 demonstrates how a high level, physics motivated feature can fail to distinguish between two very different kinds of physics, which happens due to the added complications of quantum interference.

A new family of machine learning algorithms [3, 4, 5, 6] have recently been in development that are at the intersection of machine learning, probabilistic programming, statistics and particle physics phenomenology. The techniques rely on the ability to simulate accurate samples and "cheat", i.e. extract additional information about the physics model from the simulator that would be unavailable in real data recorded at the LHC. The additional information allows to train neural networks that are not just aware of "Signal" and "Background" classes but rather learn the flexibility of the Largrainians themselves. The authors are able to show on a toy particle physics data-sets that training on such augmented data allows to use neural networks as a tool for calculus of variations and arrive at the likelihood ratio between any two physics models.

The actual ATLAS Higgs to four leptons analysis [1] is more complicated, and the family of physics models is confined by extra assumptions from the inference strategy and also prior knowledge from other measurements. We investigate the possibility to adapt the technology presented in [5] to this problem. Further, particle physicists have found a way to study almost all possible alternatives to the SM that might be measurable at the LHC using an "Effective Field Theory Framework" (EFT Framework) [7]. This is possible because of some mathematical and physical properties any Lagrangian must satisfy, making the number of terms of the Lagrangian to study finite. A successful use of these new algorithms within the ATLAS experiment will pave the way for further investigation into them for ongoing studies within ATLAS in the context of EFT.

Results

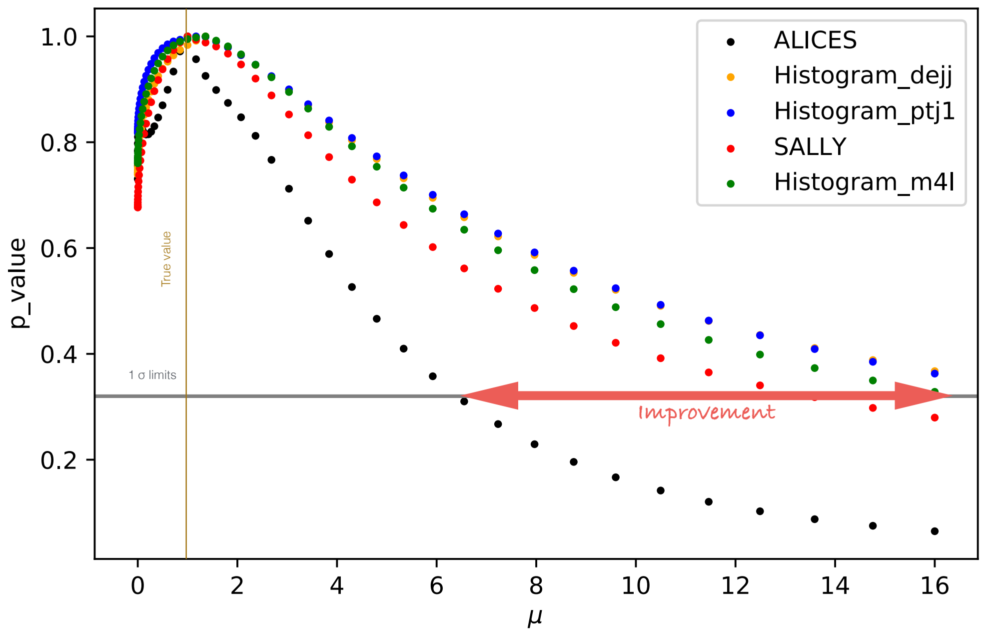


Figure 3: P-value scan using various Histogram techniques compared to SALLY and ALICES for a true value of $\mu = 1$ (sharper is better). The horizontal grey line indicates the p-value corresponding to a 1σ confidence interval

Some preliminary results (Figure 3) are shown to compare a traditional "histogram" approach of particle physics with more and more "physics-aware" algorithms to infer the true parameters of the Lagrangian. At inference time, the inputs of the neural network, for a given sample, are the features measured by the detector, as well as the hypothesis being tested (i.e. one particular value of μ). The output of the network is the likelihood ratio between the test hypothesis and the null hypothesis ($\mu = 1$). The output for all samples in the test dataset for a given test hypothesis is converted into a single p-value, as in standard statistics, and the entire process is redone for the same test dataset with a new test hypothesis (new value of μ). The p-values for the histogram techniques is calculated using multi-binned Poisson likelihood with the normalised histogram of particular physics motivated features. The "SALLY" (**S**core **A**pproximates **L**ikelihood **L**ocall**y**) model is aware of physics in the neighbourhood of the SM ($\mu = 1$) whereas "ALICES" (**A**pproximate **L**ikelihood with **I**mproved **C**ross-entropy **E**stimator and **S**core) is aware of physics in the entire range of μ, and shows better results (narrower peaks in a p-value scan, smaller 1σ margin of uncertainty for measuring μ), thus demonstrating the usefulness of a physics-aware model.

Conclusion

There is a long history of cross-pollination between particle physics and machine learning. A first study is performed to try to adapt a family of novel "physics-aware" machine learning algorithms to a realistic Higgs to four leptons analysis for the ATLAS experiment at CERN. Other efforts along similar lines such as probabilistic programming, graph networks, physics-aware generative models, adversarial networks, and so on also indicate the impending shift in the particle physics community from shoehorning physics problems into state-of-the-art machine learning algorithms to developing physics-aware algorithms that can leverage available physics insight as well as inject inductive biases to algorithms in a way that was not possible before.

Our initial studies indicate that a neural network aware of the theoretical physics model performs better inference than traditional physics-agnostic techniques, in the presence of severe quantum interference. Further studies need to be done taking into account all signal and background processes as well as simulating within the ATLAS software infrastructure to take into account the true detector effects. These machine learning models for the first time could be extended to also be aware of systematic uncertainties (when there is a known systematic difference between the simulated training data, and the real unlabeled data to which we will apply the model, but the amount and nature of the difference is unknown) that were difficult to incorporate in traditional machine learning techniques. Success with these techniques encourages the idea of extending this philosophy to other fields, such as "Maths-Aware Machine Learning".

References

[1] Morad Aaboud et al. Constraints on off-shell Higgs boson production and the Higgs boson total width in $ZZ \to 4\ell$ and $ZZ \to 2\ell 2\nu$ final states with the ATLAS detector. *Phys. Lett.*, B786:223–244, 2018.

[2] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[3] Johann Brehmer, Kyle Cranmer, Gilles Louppe, and Juan Pavez. A Guide to Constraining Effective Field Theories with Machine Learning. *Phys. Rev.*, D98(5):052004, 2018.

[4] Johann Brehmer, Kyle Cranmer, Gilles Louppe, and Juan Pavez. Constraining Effective Field Theories with Machine Learning. *Phys. Rev. Lett.*, 121(11):111801, 2018.

[5] Johann Brehmer, Felix Kling, Irina Espejo, and Kyle Cranmer. MadMiner: Machine learning-based inference for particle physics. 2019.

[6] Johann Brehmer, Gilles Louppe, Juan Pavez, and Kyle Cranmer. Mining gold from implicit models to improve likelihood-free inference. 2018.

[7] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2016.