Samuele Grossi

Third Year PhD Report

samgrs1997@gmail.com

Tutors

Riccardo Torre and Simone Marzani

Research Activity

Over the past year, my research focused on two distinct projects. The first was a natural continuation of my work from the second year of my PhD on proposing metrics to evaluate two-sample testing, the second was within the SMEFT framework, to obtain constraints over some dimension-6 operators looking at a variable called $A_0 - A_2$ in Drell-Yan process. In the following I will briefly discuss the former, given that the basic have been already presented in the Second Year Report and I will focus more on the latter, as it represents a new line of research. I will refer to the two projects as "NPLM for evaluating two-sample test" and "Lam-Tung relation violation in the SMEFT" respectively.

NPLM for evaluating two-sample test:

In this project, we extended our work from last year, which has in the meantime been published at `https://iopscience.iop.org/article/10.1088/2632-2153/adb3ee`. In that work, we proposed a set of statistical tools, called *metrics*, which have the purpose to compare independent datasets in order to assess if they are drawn from the same distribution. This is useful in high energy particle physics to validate datasets provided by machine learning (ML) generators trained on more reliable traditional tools such as Powheg. Alongside the proposal of specific *metrics*, our paper also established a clear benchmarking methodology for evaluating and comparing their performance. This year, we introduced a new *metric* based on a ML classifier called the New Physics Learning Machine (NPLM), to compare its performance against the others. We found that in general, as expected, NPLM is more powerful and flexible, achieving among the highest discriminating power across a wide range of scenarios. However, it also turned out to be significantly slower, making it not the a-priori best choice in every situation. Our conclusion was to rely on the simple and fast metric proposed in our paper when fine discriminating power is not essential (such as when tuning the hyper-parameters of a generator) and to adopt ML-based approaches only when higher precision is required.

Lam-Tung relation violation in the SMEFT:

The Standard Model Effective Field Theory (SMEFT) is a tool to explore new physics (NP) scenarios with an indirect and model independent approach. In the context of the Large Hadron Collider (LHC), the high p_T precision program is a line of research developed to target those operators whose effects grow with the energy. Selecting a specific process, a SM prediction on a given observable can be compared to the respective SMEFT prediction, and using statistical tools (such as χ^2 minimization or the likelihood ratio test) bounds on NP operators can be calculated. This is a standard and well known approach in the scientific literature.

During this project, we focused on the so-called Drell-Yan (DY) process, namely the production of a lepton and an anti-lepton in the final state. This is one of the most studied processes in high energy physics (HEP) and many experimental measurements are available. In particular, when focusing on the multi-differential cross section, the angular part can be parametrized as follows:

$$\begin{aligned}\frac{d\sigma}{dm_{\ell\ell}dp_{T,\ell\ell}dy_{\ell\ell}d\cos\theta d\phi} =& \frac{3}{16\pi}\frac{d\sigma}{dm_{\ell\ell}dp_{T,\ell\ell}dy_{\ell\ell}} \\ & \Big\{(1+\cos^2\theta) + \frac{1}{2}A_0(1-3\cos^2\theta) + A_1\sin 2\theta\cos\phi \\ & + \frac{1}{2}A_2\sin^2\theta\cos 2\phi + A_3\sin\theta\cos\phi + A_4\cos\theta \\ & + A_5\sin^2\theta\sin 2\phi + A_6\sin 2\theta\sin\phi + A_7\sin\theta\sin\phi\Big\}\ .\end{aligned} \tag{1}$$

The SM predicts that, up to $\mathcal{O}(\alpha_S)$, $A_0 - A_2 = 0$, which is the so called Lam-Tung relation.

This observable can be analysed in the SMEFT context, including within theoretical calculations the relevant dimension-6 operators. Given its particular features, $A_0 - A_2$ is insensitive to the class of operators interfering with the SM (the vector 4-fermions we studied during my first year of PhD), providing an interesting case study to observe the impact of the class of operators referred to as dipoles ($\psi^2 X \varphi$) and scalar and tensor 4-fermions (ψ^4 with a different chirality with respect to the SM one), which appear in the squared amplitude just as the squares of themselves. The former are expected to be better constrained around the Z pole, looking for example at the p_T distribution, while the latter when looking at high invariant mass values.
We extracted the bounds on the Wilson coefficients at 95% confidence level (CL) considering a center of mass (c.o.m.) energy of 13 TeV, looking both at p_T distribution (with $m_{\ell\ell}$ around the Z pole) to constrain the dipoles operators and at a $m_{\ell\ell}$ distribution extending to high values to constrain the scalar and tensor 4-fermions operators. We generated fake data to simulate a SM measurement (given that no experimental data are available with the same configuration) and used the likelihood ratio test to obtain the bounds. The results obtained suggest that this observable provides a nice insight into possible NP, providing tight constraints and a configuration where non-interfering dimension-6 operators are not overshadowed by the interfering ones that naturally vanish.

Courses

Exam passed:
- Advanced Statistics for Data Analysis (F. Badaracco, F. Di Bello, F. Parodi, 3CFU)

Exam in program:
- Introduction to the Foundations of Quantum Mechanics and Applications (P. Solinas, N. Zanghì, 3 CFU) (programmed around the end of September)

Publications

- *Refereeing the referees: evaluating two-sample tests for validating generators in precision sciences*, published the 27/02/2025 in: Machine Learning Science and Technology, (2025).
- (e-Print) *Comparing Generative Models with the New Physics Learning Machine*, `https://arxiv.org/pdf/2508.02275`
- (In progress) *The contribution of dipoles and 4-fermions operators to the Lam-Tung violation*

Schools, Seminars and Conferences

- Seminars: Weekly seminars organized by INFN in Genoa throughout the whole year.
- Seminar: Online talk at the LHC EFT WG Meeting, CERN, Geneve, 18 April 2025.
- (Soon to be attended) Conference: "23rd International Workshop on Advanced Computing and Analysis Techniques in Physics Research", in Hamburg, from 8 to 12 September 2025. Speaker. I will give a talk about the paper *Refereeing the referees: evaluating two-sample tests for validating generators in precision sciences*.
- Conference: "European AI for fundamental physics conference, EuCAIFCon25", in Cagliari, from 16 to 20 June 2025. I presented a poster about the paper *Refereeing the referees: evaluating two-sample tests for validating generators in precision sciences*.
- Conference: "Fifth MODE Workshop on Differentiable Programming for Experiment Design", in Chaniá, from 9 to 13 June 2025. Speaker. I gave a talk about the paper *Refereeing the referees: evaluating two-sample tests for validating generators in precision sciences*.
- Conference: "7th Inter-Experimental LHC Machine Learning Workshop", in CERN, Geneva from 19 to 23 May 2025. Speaker. I gave a talk about the paper *Refereeing the referees: evaluating two-sample tests for validating generators in precision sciences*.

Teaching activities

- High School Internship Tutor, Università degli studi di Genova, from 27/01 to 06/02