1SC4710 – Data Analysis in Particle Physics and Cosmology

Instructors: Jean-Christophe Hamilton, Guillaume Mention
Department: DÉPARTEMENT PHYSIQUE
Language of instruction: ANGLAIS
Campus: CAMPUS DE PARIS - SACLAY
Workload (HEE): 60
On-site hours (HPE): 34,50

Description

Particle Physics and Observational Cosmology are two fields which experienced considerable advances over the last few years, improving in a significant manner our understanding of the Universe and its content. These topics require considerable human, scientific and technological efforts in order to be able to detect rare events or very weak signals, therefore hard to detect which we call "Black Swans". In order to be scientifically meaningful, these measurements need to be reported with an accurate consciousness of their limitation s(statistical and systematic uncertainties).
This course will cover the fundamental physics questions as well as a wide variety of algorithms used within the scientific community in order to analyse data from the most recent instruments (LHC, Planck, SDSS, ...). It aims at giving the students both a synthetic and a technical view of the scientific activities in these research fields. It will give an overview and basic knowledge of the challenges in particle physics and cosmology data analysis.

Quarter number

ST4

Prerequisites (in terms of CS courses)

Cours dispensé en anglais : une bonne maîtrise de la langue est donc nécessaire.
CIP PED course

Syllabus

1 Particle Physics (Neutrinos, Dark Matter, Higgs, Supersymmetry) - 20 HEE
Scientific Introduction - 10 HEE
Standard model of particle physics

Basic knowledge of particle physics (interaction probability, cross section, event identification)
Overview of experimental detection methods in particle physics
Simulation tools developed by the community (Monte Carlo, ...)
Current research in particle physics

Specific Data Analysis Techniques - 10 HEE
Counting rate, detection efficiency, event classification, uncertainties. Signal / Noise Ratio, Statistical Significance.
Signal, background noise, classification, performance.
Systematic, nuisance, ...
Experimental uncertainties, theoretical. Balance sheets of uncertainties. Propagation of uncertainties.
Statistical inference, adjustments with χ2, Likelihood, Likelihood profile, confidence contours, p-value.
Combination of results and suitability tests
2 Cosmology- 20 HEE
General introduction 10 HEE
Pillars of cosmology (General Relativity, Expansion of the Universe, Cosmological Principle)
Big-Bang model : cosmological parameters
Cosmological Probes : Type Ia Supernovae, Acoustic Baryon Oscillations, Cosmic Microwave Background
Opening to the primordial Universe, inflation
Specific Analysis Techniques - 10 HEE
In astronomy: Imaging, photometry, spectroscopy, CMB map, spherical harmonics, pixelation
In statistics: Adjustments with χ2, Likelihood, marginalization, Bayes theorem, posterior, Monte-Carlo-Markov-Chain
3 Automatic Learning- 20 HEE
General Introduction
Classifiers: principles
Nearest Neighbor
Decision trees
OCR curves, learning curve
Overtraining, cross validation
Neural network, principles
General principles of optimization
Building tips and optimization of neural networks
Some architectures and their applications : dense network, convolutional network, recurrent network, adversary network

Class components (lecture, labs, etc.)

The "séquence thématique" starts with 4 half day of introduction with seminars from external people and round tables.

The courses proper are 8 half day (3 hours) of courses, 3 half days (3 hours) of exercices.
The séquence thématique end with the week long "enseignement d'intégration" . The students are split in 4 groups, 2 in cosmology and 2 in particle physics.
Each group has to put in place and optimise a data analysis pipe line, resulting in a scientific measurement (including uncertainties), each group divided in teams focussing on a particular aspect.

Grading

Combination of several methods. Synthetic notes on the introduction seminars. 1H30 final exam for the specific course. Enseignement d'integration : grade combining the notebooks written and the final presentation.

Course support, bibliography

- F. James, Statistical Methods in Experimental Physics: 2nd Edition
- G. Cowan, Statistical Data Analysis
- I. Narsky et F. C. Porter, Statistical analysis techniques in particle physics: fits, density estimation and supervised learning
- G. D'Agostini, Bayesian reasoning in data analysis: a critical introduction
- J. Rich, Fundamentals of Cosmology, Springer
- S. Serjeant, Observational Cosmology, The Open University
- C. Giunti, C. W. Kim, Fundamentals of Neutrino Physics and Astrophysics
- D. Griffiths, Introduction to Elementary Particle Physics.

Resources

- Teaching team : Jean-Christophe Hamilton (CNRS-APC), Guillaume Mention (CEA-IRFU)
- Practice work: 100 students on their laptops
- Software and licenses : iPython Notebooks (Jupyter, JupyterHub on a dedicated serve). No specific licence.

Learning outcomes covered on the course

Applying knowledge of statistics and machine learning to scientific issues