

# Pablo de Castro Manzano

## Curriculum Vitae

+34 699 08 48 86  
pablodecm@gmail.com  
Santander Spain

### Profile

Young Physics PhD student about to graduate at the University of Padua.  
Domain knowledge about particle physics at collider experiments.  
Technical research expertise in statistics and machine learning for scientific data analysis and simulation.  
Competent C++ and Python programmer and technological tinkerer.  
Broadly interested in the role of computational tools in science and society.  
Strong advocate for open and reproducible science.

### Education

- 2015-2019 **PhD in Physics**, *University of Padua, Italy*  
Completed, waiting for PhD thesis review and defence (end of March 2019)  
*PhD Thesis*: "Statistical Learning and Inference in Particle Collider Experiments"
- 2014-2015 **Master's Degree in Physics, Instrumentation and the Environment**, *University of Cantabria, Spain*  
Average grade: 9.7/10.0 - Specialty in Advanced Physics  
*Master's thesis*: "Measurement of CMS b-tagging efficiencies using the Flavour-tag Consistency Method at a center-of-mass-energy of 13 TeV "
- 2010-2014 **Bachelor's Degree in Physics (4 years)**, *University of Cantabria, Spain*  
Average grade: 8.6/10.0 - Mention in Fundamental Physics  
*Final Year Project*: "Measurement of the W+W- production cross section in pp collisions at a center-of-mass energy of 8 TeV "
- 2012-2013 **Physics Exchange Student**, *Imperial College London, UK*  
1st Class (70% GPA)
- 2008-2010 **Spanish Baccalaureate in Science and International Baccalaureate**, *I.E.S. Santa Clara, Spain*  
University Access Qualification: 12.1/14.0

### Experience

- 2015 - 2018 **Early Stage Researcher**, *INFN - Sezione di Padova, Italy*  
Within the AMVA4NewPhysics H2020 project, whose aim is to develop and apply state of the art machine learning techniques for High Energy Physics data analyses. Main projects:  
- New machine learning technique to construct inference-aware summary statistics.  
- Non-resonant Higgs pair production analysis (bbbb channel) at the LHC with the CMS detector.  
- Integration of TensorFlow-based multi-class jet tagger model DeepJet in CMS experiment software.
- Winter 2016 **Academic Secondment**, *University California Irvine, US*  
Collaboration with researches at the UCI Center of Machine Learning on differentiable approximations of histograms to build inference-aware losses for neural networks and the role of new deep learning techniques on jet quark-gluon tagging.
- Autumn 2016 **Industrial Secondment**, *SDG Consulting Milan, Italy*  
Worked on possible applications topological data analysis and developed a open-source package re-implementing the MAPPER algorithm with a scikit-learn-like API.
- 2014 - 2015 **Research Project Associate**, *University of Cantabria, Spain*  
Collaborating in data analyses within the CMS Collaboration, mainly related with b-tagging and top quark pair production.
- Summer 2015 **Research Internship**, *Brown University, US*  
Carry out part of Master's thesis with the Experimental Particle Physics research group.
- Summer 2014 **CERN Summer Student**, *CERN, Switzerland*  
Working with an experimental research team on characterising silicon detectors using lasers.  
Developed an open-source simulator of the drift dynamics of carrier distributions in complex semiconductor detectors.

Spring 2014 **Research Internship**, *Instituto de Física de Cantabria (IFCA)*, Spain  
Focused on the use of ontologies, knowledge bases and semantic web technologies to design a system for data preservation in High Energy Physics.

---

## Skills

Languages **Spanish**: Native speaker  
**English**: Proficient user (> **C1 level**) with the following certifications:  
- Cambridge Advance English (**CAE**): Grade B (June 2013)  
- Test of English as a Foreign Language (**TOEFL**): 101 / 120 Score (December 2013)  
Experienced technical writer.  
**Italian**: Intermediate

Computing **Advanced Linux and Unix system administrator** (>5 years)  
**Control version, continuous integration and other open-source software practises**  
**Programming Languages**: projects carried out using Python, C++ and JS among others.  
**Data Analysis**: numpy, pandas, TensorFlow, scikit-learn, scipy, ROOT, R and many more.  
**Visualization**: matplotlib, ggplot and D3JS libraries.  
**Used to complex data analysis algorithms and data structures**  
**Scientific/technical document creation with Latex/Markdown**

---

## Awards and Grants

2015-2018 **Marie Skłodowska-Curie ESR fellowship**, *AMVA4NewPhysics ITN*, EU

2015 **Brown University Exchange scholarship**, *University of Cantabria*, Spain

2014 **CERN Summer Student**, *CERN*, Switzerland

2013-2014 **Undergraduate Research Scholarship**, *Spanish Government*, Spain

2012-2013 **Erasmus Scholarship with Excellence Mention**, *Spanish Government*, Spain

---

## Publications

Author of 200+ publications as a member of the CMS Collaboration. See [🔗 Google Scholar Profile](#) for full list.  
Selected subset of CMS publications with substantial personal contribution and non-CMS publications:

stat-ml preprint "INFERNO: Inference-Aware Neural Optimisation". A. de Castro and T. Dorigo. June 2018.  
[🔗 arxiv:1806.04743](#) (submitted to CPC)

hep-ex preprint "Search for nonresonant Higgs boson pair production in the bbbb final state at 13 TeV". CMS Collaboration. October 2018. [🔗 arxiv:1810.11854](#) (submitted to JHEP).

hep-ex preprint "Combination of searches for Higgs boson pair production in proton-proton collisions at 13 TeV". CMS Collaboration. November 2018. [🔗 arxiv:1811.09689](#) (submitted to PRL).

Journal Publication "TRACS: A multi-thread transient current simulator for micro strips and pad detectors". J. Calvo, P. de Castro et al. Nucl. Instrum. Methods Phys. Res. February 2019. [🔗 doi:10.1016/j.nima.2018.11.132](#).

DSPS workshop at NIPS "DeepJet: Generic physics object based jet multi-class classification for LHC experiments". Markus Stoye on behalf of CMS Collaboration. December 2017. [🔗 Workshop Paper](#).

CMS PAS "Search for non-resonant pair production of Higgs bosons in the bbbb final state with 13 TeV CMS data", CMS Collaboration, August 2016, [🔗 cds:2209572](#)

hep-ph preprint "Analytical parametrization and shape classification of anomalous HH production in the EFT approach", LHC Higgs Cross Section Working Group, July 2016, [🔗 arxiv:1608.06578](#)

---

## Presentations and Posters

Conference Presentation "INFERNO: Inference-Aware Neural Optimisation"  
**Advanced Computing and Analysis Techniques in Physics Research**, March 2019, Sans-Fee, Switzerland

- Invited Talk "INFERNO: Inference-Aware Neural Optimisation"  
**Dark Machines Monthly Meeting**, October 2018, Remote Contribution
- Workshop Poster "INFERNO: Inference-Aware Neural Optimisation"  
**Advanced Statistics for Physics Discovery**, September 2018, Padova, Italy
- Invited Talk "Direct Learning of Systematics-Aware Summary Statistics"  
**CMS Machine Learning Forum**, August 2018, Remote Contribution
- Presentation and Poster "Direct Learning of Systematics-Aware Summary Statistics" (*awarded best poster prize*)  
**XIIIth Quark Confinement and the Hadron Spectrum**, August 2018, Maynooth, Ireland
- Workshop Presentation "Direct Learning of Systematics-Aware Summary Statistics". **2nd Inter-experimental Machine Learning Working Group Workshop**, April 2018, CERN, Switzerland
- Conference Presentation "QCD multijet background modelling by hemisphere mixing" **XIIth Quark Confinement and the Hadron Spectrum**, September 2016, Thessaloniki, Greece
- Workshop Presentation "Non-resonant HH to bbbb analyses"  
**HH searches with CMS workshop**, January 2016, Lyon, France
- Workshop Presentations "TRACS: Transient Current Simulator" (main author and presenter)  
Also co-author of the following contributions: "Two Photon Absorption and carrier generation in semiconductors"  
"TPA-TCT: A novel Transient Current Technique based on the Two Photon Absorption (TPA) process"  
**25th RD50 Workshop on Radiation hard semiconductor devices for very high luminosity colliders**, CERN, Switzerland
- Conf. Presentation (co-author) "Increasing the capacitance beyond the classical limits in capacitors with free-electron like electrodes"  
**APS March Meeting 2015 (Volume 60 - Number 1)**, San Antonio, Texas
- Conference Poster (co-author) "Implementation of the IFCA CMS Open Data Portal using EGI FedCloud resources"  
**EGI Conference 2015**, Lisbon, Portugal

## Seminars, Certifications, Courses and other Events

- Conference Attended **35th International Conference on Machine Learning 2018**  
Stockholm, Sweden
- School Attended **European School of High Energy Physics 2018** Evora, Portugal
- Public Seminar **Adapting Machine Learning for Scientific Discovery** (in Spanish), March 2017, Oviedo, Spain
- Conference Attended **5th International Conference on New Frontiers in Physics 2016**  
Crete, Greece
- School Attended **Second Machine Learning in High Energy Physics Summer School 2016**, Yandex School of Data Analysis, and National Research University Higher School of Economics, Lund, Sweden
- Workshop Attended **Data Science @ LHC 2015 Workshop**  
November 2015, CERN, Switzerland
- School Attended **CMS Data Analysis School 2015**, CMS Collaboration, Bari, Italy
- Workshop Attended **TALLER DE ALTAS ENERGÍAS 2014**  
CPAN, Benasque, Spain

• [🐦 pablodecm](#) • [🌐 pablodecm](#) • [🔗 http://pablodecm.com/](http://pablodecm.com/)

January 16, 2019