Springer Theses
Recognizing Outstanding Ph.D. Research

Toni Mäkelä

Towards Global Interpretation of LHC Data

SM and EFT Couplings from Jet and Top-Quark Measurements at CMS

Springer

Springer Theses

Recognizing Outstanding Ph.D. Research

Aims and Scope

The series "Springer Theses" brings together a selection of the very best Ph.D. theses from around the world and across the physical sciences. Nominated and endorsed by two recognized specialists, each published volume has been selected for its scientific excellence and the high impact of its contents for the pertinent field of research. For greater accessibility to non-specialists, the published versions include an extended introduction, as well as a foreword by the student's supervisor explaining the special relevance of the work for the field. As a whole, the series will provide a valuable resource both for newcomers to the research fields described, and for other scientists seeking detailed background information on special questions. Finally, it provides an accredited documentation of the valuable contributions made by today's younger generation of scientists.

Theses may be nominated for publication in this series by heads of department at internationally leading universities or institutes and should fulfill all of the following criteria

- They must be written in good English.
- The topic should fall within the confines of Chemistry, Physics, Earth Sciences, Engineering and related interdisciplinary fields such as Materials, Nanoscience, Chemical Engineering, Complex Systems and Biophysics.
- The work reported in the thesis must represent a significant scientific advance.
- If the thesis includes previously published material, permission to reproduce this must be gained from the respective copyright holder (a maximum 30% of the thesis should be a verbatim reproduction from the author's previous publications).
- They must have been examined and passed during the 12 months prior to nomination.
- Each thesis should include a foreword by the supervisor outlining the significance of its content.
- The theses should have a clearly defined structure including an introduction accessible to new PhD students and scientists not expert in the relevant field.

Indexed by zbMATH.

Toni Mäkelä

Towards Global Interpretation of LHC Data

SM and EFT Couplings from Jet and Top-Quark Measurements at CMS

Doctoral Thesis accepted by
The University of Hamburg, Hamburg, Germany

Springer

Author
Dr. Toni Mäkelä
BP2 Theoretical Physics Division
National Centre for Nuclear Research
Warsaw, Poland

Supervisor
Prof. Dr. Katerina Lipka
DESY
Hamburg, Germany

University of Wuppertal
Wuppertal, Germany

ISSN 2190-5053 ISSN 2190-5061 (electronic)
Springer Theses
ISBN 978-3-031-29778-6 ISBN 978-3-031-29779-3 (eBook)
https://doi.org/10.1007/978-3-031-29779-3

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Thus ought man to be an impersonation of the divine process of nature, and to show forth the union of the infinite with the finite.

—James Clerk Maxwell

Supervisor's Foreword

The elementary structure and interactions of matter are described by the standard model of particle physics, a quantum field theory based on gauge symmetries and operating with masses of particles and the interaction couplings as fundamental parameters. Similar to the periodic system in chemistry, the standard model takes the central role in the description of the subatomic world. In spite of the great success of this model to describe the observed phenomena, some aspects of its structure and relations of its parameters hint at the presence of new physics—yet unknown fundamental interactions, setting up at very large energy scales. For example, the standard model sorts the fundamental building blocks of matter, the quarks, into three families. However, the stable nuclear matter—everything we can see and touch—is built from the members of the first family only. This fact is not explained yet. One hypothesis suggests that the known quarks could have a structure themselves and be subject to a yet unknown interaction setting in at the very high energy scale, inaccessible by direct measurements in existing experiments. To date, the most promising indirect searches for such new physics are performed by studying the quark interactions at the Large Hadron Collider, LHC, at CERN. There, the strongly interacting particles called hadrons, in particular protons, are brought to collisions at very high energies.

The main challenge of studying quark interactions in hadron collisions is the fact that quarks are confined inside hadrons by the strong force and cannot be observed as free particles. Once kicked out of a proton, quarks form collimated sprays of hadrons, called jets. By measuring the energy and the direction of a jet, one can conclude on the properties of the original quark. Each deviation of the experimentally measured jet rates from the standard model predictions could be a signal of new physics. The new physics contributions are usually described by the effective field theory.

The observed jets may emerge from a new interaction, as well as from the strong interaction of the standard model. Therefore, the interpretation of the measurements relies to large extent on the precision of the standard model parameters and of the distributions of quarks and gluons in the proton. This problem persisted since the first searches for new physics using jets in proton-antiproton collisions, performed in the 1990s at the Tevatron in Fermilab, demanding a novel approach to extract the couplings of the new physics, unbiased from the strong interaction of the standard model. This long-standing issue is resolved in the thesis by Toni Mäkelä, who has developed a new analysis method for the simultaneous extraction of the standard model parameters and the couplings of a possible new interaction. In this thesis, the standard model prediction of jet rates in proton-proton collisions at the LHC is extended to account for new physics contributions from so-called quark contact interactions, provided by effective field theory. The resulting prediction is confronted with the most recent measurement of inclusive jet and top quark-antiquark pair production at the LHC at the center of mass energy of 13 TeV, collected by the CMS experiment. Together with the CMS measurements, the data of the electron-proton collisions at HERA accelerator at DESY are used, which serve as a basis for any determination of the quark and gluon distributions in the proton.

In the thesis, for the first time, the data of HERA and LHC are used to simultaneously constrain the minimal energy at which the novel interaction can emerge, together with the quark and gluon distributions in the proton, and the values of the top quark mass and strong coupling constant. As a result, the most stringent constraints on the couplings of the new interaction are imposed, and the best precision is achieved for the values of the top quark pole mass and of the strong coupling constant, to date. The latter is determined with an accuracy of about 1.5%, which is currently the most precise single measurement at a hadron collider.

Hamburg, Germany
December 2022

Prof. Dr. Katerina Lipka

Abstract

This thesis presents the first global interpretation of the measurements of double-differential cross sections of inclusive jet and top quark-antiquark pair ($t\bar{t}$) production in proton-proton collisions at the Large Hadron Collider (LHC) at the center of mass energy of $\sqrt{s} = 13$ TeV, together with inclusive measurements in Deep Inelastic electron-proton Scattering (DIS) at HERA. The LHC data have been recorded by the Compact Muon Solenoid (CMS) experiment in 2016 and correspond to the integrated luminosity of 33.5 fb^{-1} for inclusive jet measurement or 35.9 fb^{-1} for the $t\bar{t}$ measurement.

This novel analysis is performed at next-to-leading order (NLO) of perturbative quantum chromodynamics (QCD), where the inclusive neutral and charged current DIS measurements at HERA, the CMS jet measurements and the normalized triple differential cross sections of $t\bar{t}$ production are used together to constrain the parameters of QCD and impose constraints on new physics. In particular, the Standard Model Lagrangian is extended to incorporate effective contributions from quark contact interactions (CI). In this analysis, the constraints on the parton distributions, the strong coupling constant $\alpha_S\,(m_Z)$ and the top quark pole mass m_t^{pole} are obtained, and simultaneously, limits on the scale Λ of the CI are imposed. This results in $\alpha_S(m_Z) = 0.1188 \pm 0.0031$, $m_t^{\text{pole}} = 170.4 \pm 0.7$ GeV and the 95% confidence level limits of $\Lambda > 24$ TeV for purely left-handed, $\Lambda > 32$ TeV for vector-like and $\Lambda > 31$ TeV for axial vector-like exchanges. For the first time, the LHC data are used for extracting the Wilson coefficient of the CI, unbiased from the assumptions on the Standard Model parameters.

The jet and top quark production at the LHC are further investigated individually, to determine the QCD parameters at the highest precision. A Standard Model analysis is performed at next-to-next-to-leading order (NNLO) of perturbative QCD using the CMS jet cross section measurements together with the HERA DIS data. The parton distribution functions in the proton are extracted simultaneously with the strong coupling constant $\alpha_S\,(m_Z) = 0.1166 \pm 0.0017$. This is the most precise single measurement of the strong coupling constant at the LHC, to date.

Further, the CMS measurements of the $t\bar{t}$ cross sections are used to extract the top quark mass in the MSR renormalisation scheme. In this scheme, the mass dependence in the difference between the pole and modified minimal subtraction ($\overline{\mathrm{MS}}$) masses is replaced by a scale parameter *R*. The first study of the behaviour of the *R* scale is presented. The extracted top quark MSR mass obtained using CMS measurements results in $m_t^{\mathrm{MSRn}}(R = 80\ \mathrm{GeV}) = 169.3^{+0.6}_{-0.7}$ GeV, which is up to now the most precise value of the top quark MSRn mass extracted from experimental data.

The presented investigations of the fundamental couplings of the Standard Model, and of new physics, are at the edge of precision in experimental physics and high energy phenomenology, paving the road towards a global interpretation of LHC data.

Publications Related to this Thesis

- T. Mäkelä, A. Hoang, K. Lipka, S.O. Moch. "Investigation of the scale dependence in the MSR and $\overline{\text{MS}}$ top quark mass schemes for the $\text{t}\bar{\text{t}}$ invariant mass differential cross section using LHC data". January 2023. arXiv: 2301.03546 [hep-ph]
- CMS Collaboration. "Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at $\sqrt{s}$ 13 TeV", *JHEP* 02 (2022), p. 142. http://dx.doi.org/10.1007/JHEP02(2022)142
- T. Mäkelä, A. Hoang, K. Lipka, S.O. Moch. "Single-differential top quark pair production cross sections with running mass schemes at NLO", arXiv: 2211.13607 [hep-ph]
- H. Abdolmaleki et al. "xFitter: An Open Source QCD Analysis Framework. A resource and reference document for the Snowmass study". arXiv: 2206.12465 [hep-ph]
- D. d'Enterria et al. "The strong coupling constant: State of the art and the decade ahead". arXiv: 2203.08271 [hep-ph]
- T. Mäkelä [for the CMS Collaboration]. "Precision QCD measurements from CMS". arXiv: 2206.11624 [hep-ph]

Acknowledgements

Knowing now in hindsight the effort that a Ph.D. project requires, I first and foremost wish to express my gratitude to my supervisor Prof. Dr. Katerina Lipka, and to Prof. Dr. Sven-Olaf Moch, for the work we carried out together, and for everything I learnt about high energy physics as well as research and life in academia. PD. Dr. Alexander Glazov deserves special thanks for sharing his priceless insight on using and extending the xFITTER framework. I also thank Prof. Dr. Dieter Horns for chairing the defense, and Prof. Dr. Gregor Kasieczka for being a part of the committee.

All this wouldn't have been possible without the helpful discussions I had with many colleagues. In particular, I wish to thank Dr. Patrick Connor, Dr. Matteo Defranchis, Dr. Engin Eren, Dr. Jun Gao, Prof. Dr. André Hoang, Dr. Svenja Pflitsch, Prof. Dr. Harrison Prosper, PD. Dr. Klaus Rabbertz and Dr. Oleksandr Zenaiev.

My fellow group members, office inhabitants, other students, etc. at DESY and Universität Hamburg deserve to be thanked for the good times in- and outside of the campus. Of those not already mentioned, thanks to Ali, Ana, Bakar, Beatriz, Evan, Fang-Ying, Federico, Josry, Qun, Sam, Sebastian, Simone, Valentina and Vitaly, and apologies to anyone I might have forgotten. However, since a large part of my time as a Ph.D. student coincided with the COVID-19 pandemic, I'd like to thank Antti, Caro, Frauke, Tony, Ville and the progressive minds at Retrogramofoni + Ilkka and The Boys for keeping in touch virtually. I remain especially indebted to Ted for being there with me in Hamburg through it all. 梁先輩の有意義な教訓を忘れません。友達になってくれてありがとう!

I wish to thank my family for always being there for me, regardless of the spatial distance. Kiitokset tuestanne ja ymmärryksestänne sekä siitä, että piditte mielessäni mikä elämässä on tärkeää.

Contents

Chapter 1
Introduction

The current understanding of the dynamics between elementary particles is encoded into a quantum field theory (QFT) known as the Standard Model (SM) of particle physics. So far, the SM has described the observations in particle physics experiments with great success. However, it is not anticipated to remain the final theory of particle physics due to its well-known shortcomings. While the SM predicts the relations of its parameters, which include the couplings of the interactions and the masses of elementary particles, their values need to be determined experimentally. Furthermore, it is unexplained why the SM operates with 3 fermion families, although only one is needed for building stable matter, or why there are large differences in the masses of the particles within a family. It is also possible that the electroweak vacuum is not stable in the SM, but resolving whether or not this is the case requires increasing precision in the measurements of the top quark mass m_t and the value of the strong coupling constant α_S. Both α_S and m_t are measured with extreme precision in the context of this thesis, along with investigating the phenomenological aspects of their determination—individually, using cutting-edge phenomenology for jet and $\mathrm{t\bar{t}}$ production, but also together, exploring the sensitivity of both processes to the gluon distribution and α_S in proton-proton collisions.

There is an abundance of theories beyond the Standard Model (BSM), but so far none have been confirmed experimentally. The extensive searches carried out at collider experiments have lead to increasing exclusion limits for the scale of new physics, particularly in proton-proton collisions at the Large Hadron Collider (LHC) at CERN. With the most recently accomplished Run 2 at a center-of mass-energy of 13 TeV, LHC is the facility at the very frontier of collision energy.

The aim of a *direct search* for new physics is the observation of a resonance peak in the spectrum of an observable, such as a cross section. However, there are hints that a direct observation of new physics, a yet unknown very massive elementary particle or a fundamentally new interaction, is out of reach even at the LHC. Moreover, the SM assumes elementary particles to be pointlike, and experimental limits of $\mathcal{O}(10^{-17})$ cm have been set for the quark radii. If the radii are fundamentally finite, resulting

T. Mäkelä, *Towards Global Interpretation of LHC Data*, Springer Theses,
https://doi.org/10.1007/978-3-031-29779-3_1

for instance from more elementary particles constituting the quarks as composite particles, it would appear as a deviation from the SM cross section spectra [1, 2]. If the BSM effects are not expected to be seen as resonance peaks, but as a continuously increasing deviation from the SM prediction as a function of energy, the search is referred to as *indirect*. A historical example of a successful indirect search is provided by the observation of the continuous energy spectrum in beta decay, which led to the formulation of the electroweak theory.

The contributions of new physics are mostly approximated by the Standard Model effective field theory (SMEFT), which extends the SM by operators that describe new interactions between SM particles. The cause of the interaction, e.g. the exchange of new particles or substructure of SM particles, is integrated out, and predictions at energies below the BSM scale are obtained without the need to resolve the exact mechanism. However, achieving sensitivity to minute differences in the spectra of physical observables requires both the measurements and the predictions to be extremely precise.

In the SM, the dominant processes in the proton-proton collisions at the LHC are described by quantum chromodynamics (QCD), and the most fundamental process for studying QCD is the production of *jets*: since quarks or gluons (together called partons) cannot be observed as free particles due to the confinement in QCD, the partons kicked out of a proton in a pp collision form collimated sprays of color-neutral hadrons, so-called jets, which preserve the momentum and direction of the original parton. Jet production thus sheds light on the parton distribution functions (PDFs) in the proton, is directly sensitive to the value of the strong coupling constant α_S, and at the same time probes the scale of new physics at high transverse momenta. In particular, the PDFs and the value of α_S play a central role in the interpretation of the LHC measurements.

The PDFs are functions of the energy scale at which the proton structure is probed and of the partonic fraction x of the proton momentum. While the energy scale dependence of the PDFs is known in perturbative QCD, the x-dependence cannot yet be calculated. Therefore, the PDFs need to be extracted using experimental data. The PDF determination assumes the validity of the SM also at high transverse momenta where new physics can emerge, possibly absorbing the new physics inside the proton structure and therefore causing a bias in the interpretation of the measurements in an indirect search. This problem is assessed in this thesis by following an unbiased analysis strategy in a search for 4-quark contact interactions (CI).

The main focus of the thesis is the comprehensive QCD analysis of the double-differential inclusive jet data measured by the CMS Collaboration in pp collisions at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 33.5 fb^{-1}. The jets are reconstructed using the anti-k_T algorithm with the distance parameter 0.7. In an analysis at next-to-next-to-leading order (NNLO) in SM QCD, these data are used together with the HERA inclusive deep inelastic scattering (DIS) cross sections [3], resulting in a simultaneous extraction of the PDFs and $\alpha_S(m_\text{Z})$. The obtained NNLO value of $\alpha_S(m_\text{Z})$ is the most precise value extracted in a single experiment at the LHC, to date. In an analysis performed at next-to-leading order (NLO), also the CMS triple-differential top quark-antiquark production cross section measurements