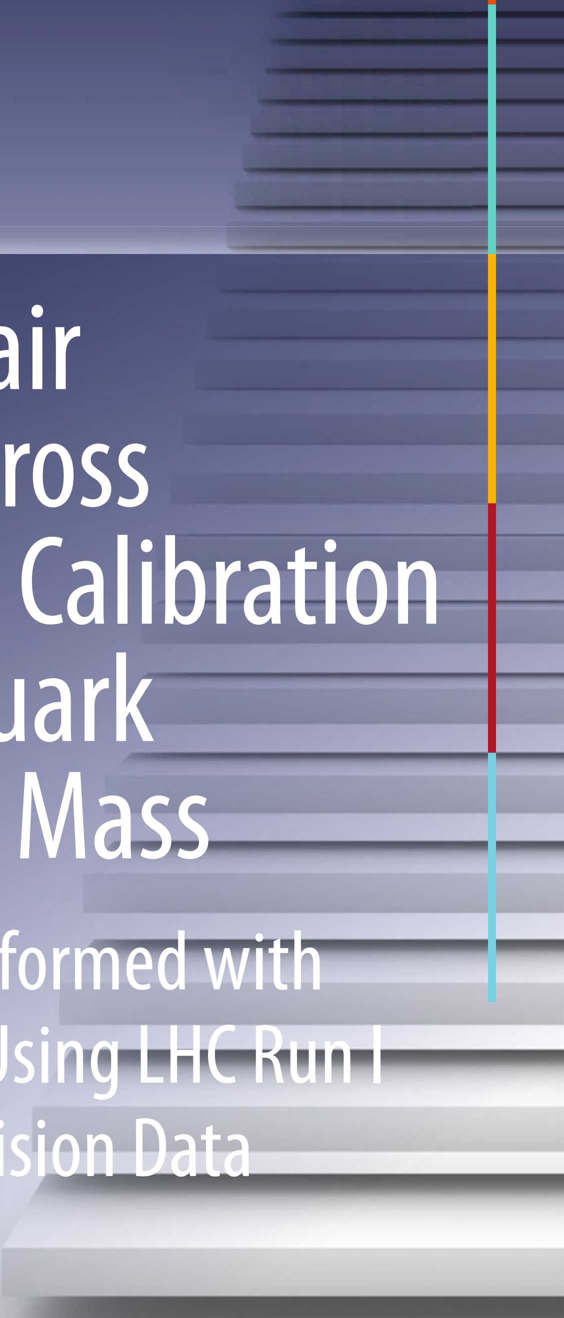
Springer Theses
Recognizing Outstanding Ph.D. Research

Jan Kieseler

Top-Quark Pair Production Cross Sections and Calibration of the Top-Quark Monte-Carlo Mass

Measurements Performed with the CMS Detector Using LHC Run I Proton-Proton Collision Data



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 are accepted into the series by invited nomination only and must 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.
- 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 scientists not expert in that particular field.

More information about this series at http://www.springer.com/series/8790

Jan Kieseler

Top-Quark Pair Production Cross Sections and Calibration of the Top-Quark Monte-Carlo Mass

Measurements Performed with the CMS Detector Using LHC Run I Proton-Proton Collision Data

Doctoral Thesis accepted by
the University of Hamburg, Germany

Springer

Author
Dr. Jan Kieseler
DESY-CMS
Hamburg
Germany

Supervisor
Dr. Katerina Lipka
DESY Hamburg
Hamburg
Germany

ISSN 2190-5053 ISSN 2190-5061 (electronic)
Springer Theses
ISBN 978-3-319-40004-4 ISBN 978-3-319-40005-1 (eBook)
DOI 10.1007/978-3-319-40005-1

Library of Congress Control Number: 2016941284

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Supervisor's Foreword

Understanding of the matter structure is the fundamental task of particle physics. The mass of the proton and therefore of the baryonic matter in the universe is ascribed to the strong force, carried by massless gluons, acting between the quarks, the fundamental constituents of nucleons. The underlying theory of the strong force, Quantum Chromodynamics (QCD), describes the interactions of the quarks and the gluons. Quark masses have an unambiguous definition in a renormalization scheme of QCD and are its fundamental parameters.

The heaviest elementary particle known, the top quark, is an object of particular interest in particle physics, being the only quark decaying before hadronizsation, thus allowing the study of bare quark properties and tests of QCD with ultimate precision. The value and the precision of the top-quark mass have far-reaching implications affecting conclusions about the stability of the vacuum state of our universe.

The top quark was discovered in the 1990s at the Tevatron proton-antiproton collider. Since 2011, the Large Hadron Collider (LHC) at CERN has served almost as a factory for top-quark pair production in the proton–proton collisions at centers-of-mass of 7, 8, and 13 TeV. In recent precise measurements at hadron colliders, a top-quark mass parameter used in simulations based on heuristic models is obtained. Its interpretation in terms of a theoretically well-defined top-quark mass has been a long-standing problem. The thesis of Jan Kieseler addresses this problem for the first time employing novel techniques of experimental analysis. The top-quark pair production cross section and the top-quark mass are determined simultaneously in proton–proton collisions at the LHC using the data, collected by the CMS detector at a center-of-mass energy of 7 and 8 TeV.

Dr. Kieseler has developed an analysis technique which employs a combination of a template fit of multi-differential distributions of observed kinematics of top-quark production and a parameterization of the top-quark signal based on the expected event topology. The fitted distributions include the invariant mass of the lepton and the b-quark originating in the top-quark decay. This distribution and the sensitivity of its shape to the top-quark mass is explicitly investigated in the thesis.

A simultaneous fit to the cross section and the top-quark mass, as used in the simulation, is performed, with the dependence of the measured cross section on the top-quark mass parameter being diminished. The resulting top-quark pair production cross sections are the most precise measurements at the LHC. By using the theoretical dependence of the top-quark pair production cross section on the pole or running mass of the top-quark at the highest perturbative order available, the resulting top-quark mass is calibrated to all orders and the direct connection of the measurement to the well-defined QCD parameter is established. This method can be extended to other experimental observables. By means of this fully consistent comparison with the simultaneously obtained top-quark mass parameter in the simulation, the long-standing problem of its interpretation is solved. The top-quark mass, unambiguously defined in various renormalization schemes of quantum chromodynamics, is determined directly and with unprecedented precision.

The analysis strategy developed by Jan Kieseler opens new paths for top-quark measurements and their interpretation. One prominent application would be the first determination of the top-quark running mass, which may become possible with more precise data to be collected by the LHC experiments at a center-of-mass energy of 13 TeV in the coming years.

Hamburg
April 2016

Dr. Katerina Lipka

Abstract

In this thesis, measurements of the production cross sections for top-quark pairs and the determination of the top-quark mass are presented. Dileptonic decays of top-quark pairs ($\mathrm{t\bar{t}}$) with two opposite-charged lepton (electron and muon) candidates in the final state are considered. The studied data samples are collected in proton–proton collisions at the CERN Large Hadron Collider with the CMS detector and correspond to integrated luminosities of 5.0 fb^{-1} and 19.7 fb^{-1} at center-of-mass energies of $\sqrt{s}$ = 7 TeV and $\sqrt{s}$ = 8 TeV, respectively. The cross sections $\sigma_{\mathrm{t\bar{t}}}$, are measured in the fiducial detector volume (visible phase space), defined by the kinematics of the top-quark decay products, and are extrapolated to the full phase space. The visible cross sections are extracted in a simultaneous binned-likelihood fit to multi-differential distributions of final state observables, categorized according to the multiplicity of jets associated to b quarks (b jets) and other jets in each event. The fit is performed with emphasis on a consistent treatment of correlations between systematic uncertainties and taking into account features of the $\mathrm{t\bar{t}}$ event topology. By comparison with predictions from the Standard Model at next-to-next-to leading order (NNLO) accuracy, the top-quark pole mass, m_t^{pole} is extracted from the measured cross sections for different state-of-the-art PDF sets.

Furthermore, the top-quark mass parameter used in Monte Carlo simulations m_t^{MC}, is determined using the distribution of the invariant mass of a lepton candidate and the leading b in the event, $\mathrm{m_{lb}}$. Being defined by the kinematics of the top-quark decay, this observable is unaffected by the description of the top-quark production mechanism. Events are selected from the data collected at $\sqrt{s}$ = 8 TeV that contain at least two jets and one b jet in addition to the lepton candidate pair. A novel technique is presented, in which fixed-order calculations in quantum chromodynamics (QCD) are employed to determine the top-quark mass from the shape of the measured $\mathrm{m_{lb}}$ distribution.

The analysis is extended to a simultaneous fit of the $\mathrm{t\bar{t}}$ production cross sections and m_t^{MC}, including the $\mathrm{m_{lb}}$ distribution to increase the sensitivity to m_t^{MC}. The resulting $\mathrm{t\bar{t}}$ production cross sections at $\sqrt{s}$ = 7 and 8 TeV do not depend on assumptions on m_t^{MC} and are the most precise ones obtained with the CMS experiment. The extracted m_t^{MC} reaches an unprecedented precision for a single measurement of m_t^{MC} in the dileptonic decay channel. The measured $\sigma_{\mathrm{t\bar{t}}}$ are further used to determine m_t^{pole} and the $\overline{\mathrm{MS}}$ mass, $m_t^{\overline{\mathrm{MS}}}$, at up to NNLO QCD. The extracted $m_t^{\overline{\mathrm{MS}}}$ exhibits a better perturbative convergence and is converted to the pole mass $m_t^{\mathrm{p,conv}}$, using recent calculations at 4-loop QCD. For the first time, the direct relation of m_t^{MC} to $m_t^{\overline{\mathrm{MS}}}$, m_t^{pole}, and $m_t^{\mathrm{p,conv}}$ is quantified experimentally at the highest available precision.

Acknowledgments

I would like to thank my supervisor, Dr. Katerina Lipka, for offering me the opportunity to work within her group already during my diploma thesis and for giving me a lot of freedom to work independently—but also for giving clear directions at the same time and saying clear words every once in a while—also beyond the scope of this thesis.

My gratitude to Prof. Dr. Sven Moch, for very useful discussions and support, and also for evaluating my thesis efficiently and accurately. I would also like to thank my third referee, Prof. Dr. Zhiqing Zhang, for his very thorough and extraordinarily fast evaluation.

Thanks also to Dr. Olaf Behnke, Prof. Dr. Elisabetta Gallo, and Prof. Dr. Caren Hagner for taking part in my disputation and for agreeing on a date on short notice. I would like to thank Olaf additionally for many inspirational discussions, in particular concerning statistics.

Many thanks to Maria for her invaluable and intense support and her constant willingness to help in every department—starting from my first steps with programming and the computing facilities in “the” office several years ago, up to discussions about almost everything as friends, nowadays.

Many thanks to Sebastian for push-starting the efforts and his help as convener.

I would like to thank everyone I shared an office with for making it pleasent to come there every morning: Monica with her extraordinary collection of tea, Marco for discussions about theoretical physics and guitar equipment, and Roberval for his calmness while I was writing.

My gratitude to my other former and present colleagues, it was a pleasure working (and having lunch) with you: Johannes and Benjamin being interested in even the most technical programming aspects, Jasone always having a piece of chocolate and a smile in almost every situation, Eleni providing the complementary cookies and spotting even the tiniest missing sub-subscript, Ivan enjoying class-to-class conversion complications, Wolf probably solving them, Ali and Simon building yet another piece of hardware, Anya and Jenya being in either the one or the other office, and Gregor as the expert on administrative procedures. But it was also a pleasure discussing and working with the other more or less recent additions to the top group: Mykola, Christian C., Till, Alexander, Christian S., Kelly, and Gerrit.

Thanks to Samantha, Paolo, Francesco, and Ringaile for several conversations about physics and more.

Auch möchte ich meiner gesamten Familie danken—dafür, dass ihr mich unterstützt habt, wann immer ich es brauchte.

Besonders unterstützt während der Promotion—aber nicht nur in Bezug darauf—hast du mich, Carmen, so dass diese Arbeit im positiven Sinne eng mit dir verbunden ist. Danke!

Contents

Chapter 1
Preamble

Particle physics studies the fundamental components of matter and their interactions. Within the last decades impressive advancements in this field have been achieved. The variety of physics phenomena are explained in terms of fundamental interactions between elementary particles. All of them, except the gravitational force, are combined into the Standard Model (SM) of particle physics, describing the building blocks of matter. It comprises leptons and quarks which build up matter, and gauge bosons, which mediate the exchange forces between them. Combinations of two or three quarks, held together by gluons, create hadrons, thereby defining their properties.

The SM has been successfully tested over the past 30 years. Very recently, at the Large Hadron Collider (LHC) experiments at CERN, experimental evidence has been found for a Higgs boson with a mass of around 125 GeV [1, 2], which was the last missing building block of the SM. However, the SM cannot explain particular features of our universe, such as the origin of dark matter and the predominance of matter over antimatter. Therefore, it is believed that the SM is only an effective description of the structure of matter up to a certain energy scale and that there must be a truly fundamental underlying theory. Most of the extensions of the theory that have been proposed to solve the shortcomings of the SM have a common feature: they predict the existence of new physics phenomena not considered by the SM at the TeV scale. Nevertheless, no evidence of such phenomena has yet been observed.

Within the SM, the top quark is the heaviest fundamental particle. Due to its large mass, the top quark decays very rapidly (within 5×10^{-25} s), before hadronizing, and is thus the only quark that gives direct access to its properties such as spin and charge. With its large mass, the top quark has a uniquely strong coupling to the Higgs boson. Thus, the top quark is believed to play a special role in the electroweak symmetry breaking. Various scenarios of physics beyond the SM expect the top quark to couple to new particles. Furthermore, SM top quark processes are a dominant background to many searches for physics beyond the SM. It is crucial to understand the production mechanisms and properties of the top quark to the highest possible precision. Apart

J. Kieseler, *Top-Quark Pair Production Cross Sections and Calibration of the Top-Quark Monte-Carlo Mass*, Springer Theses,
DOI 10.1007/978-3-319-40005-1_1

from this, the description of final states of all processes at hadron colliders such as the LHC relies on Quantum Chromodynamics (QCD), describing the interaction of colliding partons and subsequently produced hadrons. The exceptional properties of the top quark offer unique possibilities to test QCD as well as predictions from the electroweak theory.

The top quark was discovered at the Tevatron proton-antiproton collider in 1995 at a center-of-mass energy ($\sqrt{s}$) of 1.8 TeV [3, 4]. Although some of its properties and interactions have been measured very precisely as reviewed in [5], others suffer from the relatively low rate of top quarks produced at the Tevatron. At the LHC, which is in operation since 2009, protons collide with protons at $\sqrt{s} = 7$ TeV (2010–2011) and 8 TeV (2012), and since June 2015, also 13 TeV. These high collision energies allow for a large production rate for top quarks. In consequence, several million of top-quark pairs have been produced at the LHC, around 100 times more than at the Tevatron. This allows to perform precise measurements of the top-quark production and properties, challenging the accuracy of the theoretical predictions, potentially constraining QCD parameters, and opening the possibility to search for new physics by studying deviations of the top-quark properties from the SM expectation.

The top-quark mass, m_t, is one of the fundamental parameters of the SM. Its value significantly affects predictions for many observables either directly or via radiative corrections. As a consequence, the measured top-quark mass is one of the crucial inputs to electroweak precision fits, which enable comparisons between experimental results and predictions within and beyond the SM. Furthermore, together with the Higgs-boson mass, it has critical implications on the stability of the electroweak vacuum when extrapolating the SM to high energy scales [6, 7]. The top-quark mass has been determined with remarkable precision: the current world average is 173.34 $\pm$ 0.76 GeV[1] [8], determined by combining results from the Tevatron and the LHC. However, these *direct* measurements rely on the relation between the top-quark mass and the respective experimental observable, e.g. the reconstructed invariant mass of the top-quark decay products. This relation is derived from Monte Carlo (MC) simulations. Hence, the direct measurements determine the top-quark mass parameter implemented in this simulation, m_t^{MC}, that is most compatible with the data. It is important to understand how to interpret the experimental result in terms of well-defined theory parameters used in QCD and electroweak calculations. In calculations beyond leading order (*LO*), the top-quark mass depends on the renormalization scheme [9], e.g. the *pole* or $\overline{MS}$ scheme. The relation between m_t^{MC} and these mass parameters is of particular relevance since the uncertainty on the measured m_t^{MC} parameter has become smaller than the uncertainty on its theoretical interpretation, which is of the order of 1 GeV [10]. Therefore, a calibration of m_t^{MC} to a theoretically well-defined top-quark mass is necessary and an important aspect of this thesis.

The top-quark mass can be extracted by confronting a measured observable with its prediction, e.g. the inclusive top-quark pair ($t\bar{t}$) production cross section ($\sigma_{t\bar{t}}$) [11–13], calculated beyond LO QCD in a well-defined top-quark mass scheme. Studies of $\mathrm{t\bar{t}}$ production have been performed on a variety of production and decay channels

---

[1] Throughout this thesis $c = \hbar = 1$ applies.

in the recent years by the LHC and Tevatron experiments. So far, all these results are consistent with predictions from the SM at next-to-next-to leading order (*NNLO*) with a precision of about 4 % [14], depending on the dominant production mechanism and the center-of-mass energy. Reaching a similar or lower uncertainty in measurements at hadron colliders is experimentally challenging.

Compared to previous precision analyses of $\sigma_{\mathrm{t\bar{t}}}$ [12, 15, 16], the measurements presented in this thesis bring the following improvements. The full data sample recorded by the Compact Muon Solenoid (CMS) experiment at the LHC in the years 2011–2012 is analyzed, corresponding to two center-of-mass energies, 7 and 8 TeV. An innovative cross section extraction method is employed: a simultaneous fit of the $\mathrm{t\bar{t}}$ production cross sections at $\sqrt{s} = 7$ and 8 TeV is performed, with emphasis on a consistent treatment of correlations regarding simulation modeling and detector uncertainties. The cross sections are measured in the visible phase space, defined by the detector fiducial volume, and extrapolated to the full phase space. With total uncertainties of 3.6 % (7 TeV) and 3.8 % (8 TeV), the measurements presented in this thesis constitute the most precise determinations of $\sigma_{\mathrm{t\bar{t}}}$ with the CMS detector and are competitive with recent results published by the ATLAS collaboration [12]. These precise results are used to determine the top-quark pole mass through comparison with recent calculations.

Alternatively, differential $\mathrm{t\bar{t}}$ production cross sections are studied, aiming for the determination of m_t in a well-defined scheme. These measurements can improve the precision compared to the results from $\sigma_{\mathrm{t\bar{t}}}$ if the chosen observable is particular sensitive to m_t or not affected by certain systematic uncertainties. Already long before the first start-up of the LHC, it was proposed to employ the invariant mass of lepton and b jet (m_{lb}) in dileptonic $\mathrm{t\bar{t}}$ events as such an observable [17]. It has been noted that the distribution is under good theoretical control up to next-to-leading order (NLO) QCD over the entire range that is relevant for measurements of m_t [9, 18, 19]. In this thesis, a novel technique is presented to relate such fixed-order calculations to the distribution observed in data and extract the top-quark mass. The method is applied to predictions of the $\mathrm{t\bar{t}}$ production cross section as a function of $\mathrm{m_{lb}}$ calculated with MCFM [20] and the resulting distribution is compared to the observation in data at $\sqrt{s} = 8$ TeV.

Finally, a simultaneous fit of m_t^{MC} and $\sigma_{\mathrm{t\bar{t}}}$ at $\sqrt{s} = 7$ and 8 TeV is performed, profiting from the individual studies of these parameters. In consequence, the production cross sections are determined for the first time without prior assumptions on m_t^{MC}. The extracted $\sigma_{\mathrm{t\bar{t}}}$ are employed to determine the top-quark mass in well-defined renormalization schemes with unprecedented precision. The simultaneously determined m_t^{MC} parameter not only represents the by far most precise single measurement of m_t^{MC} in dileptonic $\mathrm{t\bar{t}}$ events, but also provides a consistent treatment of correlations between the extracted $\sigma_{\mathrm{t\bar{t}}}$ and m_t^{MC}, and as a consequence between the extracted well-defined top-quark mass parameters and m_t^{MC}. Thus, this measurement represents the first experimental calibration of the m_t^{MC} parameter to these mass parameters.

This thesis is organized as follows: The SM is introduced in Chap. 2 with particular focus on the top quark, its production, decay and the interpretation of its mass.

Furthermore, the MC generators and detector modeling used in the analyses are discussed. In Chap. 3, the LHC machine, the CMS detector, and the data employed in this thesis are described. The event reconstruction and selection are reviewed in Chap. 4, which also comprises a description of data-driven corrections applied to the simulation. The simultaneous fitting technique to extract $\sigma_{\mathrm{t\bar{t}}}$ at 7 and 8 TeV is presented and validated in Chap. 5, together with a description of the sources of systematic uncertainties. The resulting visible and total cross sections are also discussed there. Chapter 6 is dedicated to the extraction of the top-quark mass from the measured $\sigma_{\mathrm{t\bar{t}}}$ and the $\mathrm{m_{lb}}$ distribution. The simultaneous fit of $\sigma_{\mathrm{t\bar{t}}}$ at 7 and 8 TeV and m_t^{MC} is described in Chap. 7, as well as the determination of the theoretically well-defined top-quark masses and the calibration of m_t^{MC}. The summary and conclusions are discussed in Chap. 8.

1.1 Frequently Used Terms and Expressions

Throughout this thesis, the term "uncertainties on A due to variations of x and y are added in quadrature" is defined in the following way: let the variation increasing A due to a variation of κ be $\Delta_{\kappa,+}$ and the term decreasing A $\Delta_{\kappa,-}$. Then, both uncertainties are added to the total asymmetric uncertainty on A, $\Delta_{A,\pm}$, as:

$$\Delta^2_{A,+} = \Delta^2_{x,+} + \Delta^2_{y,+} \tag{1.1}$$

$$\Delta^2_{A,-} = \Delta^2_{x,-} + \Delta^2_{y,-} \tag{1.2}$$

The term "the difference in quadrature between A and B is C" refers to the following operation:

$$C^2 = |A^2 - B^2|. \tag{1.3}$$

References

1. ATLAS Collaboration: Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC. Phys. Lett. B **716**, 1–29 (2012)
2. CMS Collaboration: Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. Phys. Lett. B **716**, 30–61 (2012)
3. CDF Collaboration: Observation of top quark production in $\bar{p}p$ collisions with the collider detector at fermilab. Phys. Rev. Lett. **74**, 2626–2631 (1995)
4. DO Collaboration: Observation of the top quark. Phys. Rev. Lett. **74**, 2632–2637 (1995)
5. Galtieri, A.B., Margaroli, F., Volobouev, I.: Precision measurements of the top quark mass from the Tevatron in the pre-LHC era. Rept. Prog. Phys. **75**, 056201 (2012)
6. Alekhin, S., Djouadi, A., Moch, S.: The top quark and Higgs boson masses and the stability of the electroweak vacuum. Phys. Lett. B **716**, 214–219 (2012)
7. Degrassi, G., Di Vita, S., Elias-Miro, J., Espinosa, J.R., Giudice, G.F., et al.: Higgs mass and vacuum stability in the standard model at NNLO. JHEP **1208**, 098 (2012)