Results on Top physics from CMS

Georgios Konstantinos Krintiras on behalf of the CMS Collaboration

Abstract

After the discovery of the top quark more than 20 years ago, top quark production cross sections have been meticulously studied. The rich variety of results from the LHC experiments are complemented with increasingly accurate theoretical predictions of heavy quark production and decay. Measurements of the top quark production provide a benchmark test of perturbative quantum chromodynamics and the standard model (SM), constraining at the same time the background in Higgs boson searches as well as extensions beyond the SM. Recent top quark measurements from CMS are reviewed, illustrating past and current experimental methods along with their attained precision. A perspective of top quark physics at the High-Luminosity LHC and at future colliders is briefly given.

Presented at LCWS2017 International Workshop on Future Linear Colliders
Results on TOP physics from CMS

Georgios Konstantinos Krintiras* on behalf of the CMS Collaboration
Université catholique de Louvain, Louvain-la-Neuve, Belgium
E-mail: gkrintir@cern.ch

After the discovery of the top quark more than 20 years ago, top quark production cross sections have been meticulously studied. The rich variety of results from the LHC experiments are complemented with increasingly accurate theoretical predictions of heavy quark production and decay. Measurements of the top quark production provide a benchmark test of perturbative quantum chromodynamics and the standard model (SM), constraining at the same time the background in Higgs boson searches as well as extensions beyond the SM. Recent top quark measurements from CMS are reviewed, illustrating past and current experimental methods along with their attained precision. A perspective of top quark physics at the High-Luminosity LHC and at future colliders is briefly given.

*Speaker.
Contents

1. Setting the scene 3
   1.1 A special role in the EWSB mechanism? 3
   1.2 Theory predates experiment; experiment keeps a tight rein on theory 5
   1.3 That could well be a top quark! 7

2. A multifaceted quark 9
   2.1 Most probably flies in pairs 9
      2.1.1 Measuring it inclusively with its partner 9
      2.1.2 Closing the gap between Tevatron and LHC energies 13
      2.1.3 Hot off the press; now also in nuclear collisions 13
      2.1.4 Just the scale has never been enough 16
      2.1.5 A factory of gluons over broad momenta 20
      2.1.6 Breaking the 1 TeV barrier 21
   2.2 “If we be doomed to remain single, we do” 24
   2.3 It pins down to two numbers; the mass and the width 28
   2.4 Rare and rarer signatures 30

3. “I like the dreams of the future better than the history of the past” 35
1. Setting the scene

With a mass close to the one of a gold atom, yet a size less than $10^{-3}$ fm, the top quark is the heaviest elementary particle known. When the top quark was independently discovered by the CDF [1] and DØ [2] Collaborations in 1995 at the Tevatron proton-antiproton ($p\bar{p}$) collider, a decades-long search for one of the last missing pieces of the standard model (SM) had come to an end. Despite an early blunder, that had resulted in claiming a “clear signal” compatible with a W boson decaying into a 40 GeV top quark in 1984 [3], indirect hints of a large top quark mass ($\gg 50$ GeV) later came from significant flavor oscillation frequencies in $B^0_d$ meson mass eigenstates and electroweak (EW) precision data. In Fig. 1, the limits on the top mass, $m_{\text{top}}$, as a function of time are compared to direct measurements at the Tevatron. The indirect determination within the framework of the standard model (SM) adheres to remarkably good agreement with the direct measurements. It is thus intriguing to test whether the top quark with such extremely heavy mass readily influences predictions regarding the stability of the Higgs field and its effects on the evolution of the Universe, as well as the sensitivity of SM consistency checks.

![Figure 1: Comparison of direct (at Tevatron) and indirect determinations of the top quark mass as a function of time [4]. The good agreement between the top quark mass measured directly and the predicted mass determined indirectly on the basis of measurements at the Z-pole is a convincing illustration of the validity of radiative corrections. Also shown is the 95% confidence level lower limit from the direct searches before the discovery of the top quark [1, 2].](image)

1.1 A special role in the EWSB mechanism?

In the SM, the quark and lepton masses have to be “inserted” by means of adjustable parameters that describe how feebly or strongly each type interacts with the Higgs boson. For a top quark the Yukawa coupling is almost exactly unity, unlike for any other fermion. This relatively strong coupling with the Higgs boson, and to some extent the mystique associated with a value of unity, suggests that the top quark may play a special role in the mechanism of electroweak symmetry

---

1 Alternative notations, like $M_{\text{top}}$, $M_t$, or $m_t$, exist and are typically used interchangeably.
breaking (EWSB). Quadratically divergent radiative corrections appear in the computation of the Higgs boson mass and are dominated by top quark loop contributions. Repeating the exercise of indirect determination of $m_{\text{top}}$ from global SM fits to EW precision data is thus essential for testing the overall SM consistency. A scan [5] of the confidence level profile of W boson mass versus $m_{\text{top}}$ is shown in Fig. 2a for the scenarios where the direct Higgs boson measurements [6, 7] are included (blue) or not (grey) in the fit. Both contours agree with the direct measurements (green bands and ellipse).

In addition, owing to its large value, top quark has a direct impact on extrapolations of the EWSB- to high-energy scales. Radiative corrections can drive the Higgs boson self-coupling ($\lambda$) towards zero or even negative values, potentially leading to an unstable vacuum. The determination of such energy scale ($\mu$), either requiring or not beyond the SM (BSM) physics at lower energies like the EWSB scale, is strongly influenced by the precision of the measurements in the top quark sector and by their unambiguous interpretation in a clear theoretical framework (in Fig. 2b $\mu$ varies by several orders of magnitude under a seemingly insignificant variation of $\pm 2$ GeV in $m_{\text{top}}$). Last but not least, the indication of a flat Higgs potential ($\lambda = 0$) at very high $\mu$ values could imply the possibility the Higgs boson to play the role of the scalar field in the early universe, as an appealing phenomenological consequence (e.g. [13]).

![Figure 2](image_url)

(a) Contours at 68% and 95% confidence level obtained from scans of fits [5] with fixed variable pairs of the W boson against the top quark mass. The corresponding direct measurements are excluded from the fit. The narrower blue and larger grey allowed regions are the results of the fit including and excluding the Higgs boson mass measurements [6, 7], respectively. The horizontal bands indicate the $\pm 1\sigma$ regions of the world averages used for the W boson [8] and top quark [9]. A theoretical uncertainty of 0.5 GeV–non-perturbative theoretical uncertainty of $\mathcal{O}(\Lambda_{\text{QCD}})$–is added to the direct top mass measurement. (b) Renormalisation-group evolution of the Higgs boson self-coupling varying the top mass, strong SU(3)$_c$ coupling and Higgs boson mass within a $\pm 3\sigma$ range [11]. Residual theoretical uncertainties [12] are safely smaller than the experimental uncertainty, dominated by the uncertainty on the top quark mass.
1.2 Theory predates experiment; experiment keeps a tight rein on theory

The basic tool for top quark physics is a high energy particle collider. For instance, the dominant production of top quark pairs (t¯t), which is a quantum chromodynamics (QCD) process, is accessible both at lepton and at hadron colliders, provided the √s value of the collisions is above the production threshold of twice m_{top}. The SM predicts the EW production of single top quarks in addition to the dominant t¯t process, while the associated production of top quarks with the EW gauge bosons or the Higgs boson is also possible (Fig. 3). A natural question can be then addressed: how theory and experiment have coped with the broad range of questions in top quark physics, from inclusive production cross sections and precise measurements of the top quark mass and of its couplings to a variety of searches for BSM physics with top quarks.

Figure 3: Compilation of CMS measurements and SM predictions in proton-proton (pp) collisions at √s = 7, 8, and 13 TeV for a variety of processes as indicated on the horizontal axis [10].

The probability (cross section), σ, for any process can be expressed as a weighted product of a high-energy (“hard”) parton-parton scattering multiplied with parton distribution functions (PDFs), integrated over all parton momenta and summed over all parton types (“phase space”):

\[ \sigma = \sum_{j,k} \int_0^1 dx_j dx_k f_k(x_k, \mu_F^2) \hat{\sigma}(sx_j, x_k, \mu_F, \alpha_s) \].

While the hard scattering cross section, \(\hat{\sigma}\), is process specific and can be computed in perturbative QCD at different levels of accuracy, PDFs, \(f_i(x_i, \mu_F^2)\), are deemed universal functions that describe the probability to find a parton \(i\) with a given longitudinal momentum fraction \(x_i\), when the hadron is probed at a momentum transfer scale of \(\mu_F\) (Fig. 4). The PDF parameterization absorbs all long-distance effects in the initial state, and ultraviolet divergences in \(\hat{\sigma}\) are treated with a renormalization procedure of the strong coupling constant, introducing an additional energy scale (\(\mu_R\)).
The “default” (but not unique) choice of energy scales are process specific, e.g., for t ¯t production a typical choice is $m_{\text{top}}$. Cross sections for t ¯t and single top quark production include higher-order corrections, and are currently available up to next-to-next-to-leading order (NNLO), including soft-gluon resummation at next-to-next-to-leading logarithm (NNLL), in perturbative QCD plus NLO EW and NNLO accuracy, respectively.

Figure 4: The proton gluon PDF against the longitudinal momentum $x$ as determined from the (a) ABMP16 [26], CT14 [27], MMHT14 [28], NNPDF3.0 [29], (b) HERAPDF2.0 [30], and JR14 [31] sets with their ±1σ uncertainties at the factorization scale equals to the Z boson mass. The distributions are normalized to the central values from ABMP16 for comparison [26].

QCD color confinement requires the final state particles to be color-neutral, but the process of turning partons into hadrons cannot be treated perturbatively, and it thus relies on phenomenological models. To compare calculations of hadron-hadron collisions to the data, software tools based on the Monte Carlo (MC) method are employed. “General-purpose” MC event generators (e.g. [16]-[18]) can automatically compute the full set of contributions to the hard process at NLO given the Feynman rules of the underlying theory (both SM and BSM). For almost all processes, the normalization of the MC event generation is corrected to match the most precise calculation, ignoring the effect that higher-order corrections may cause on the shape of kinematic observables. The final process of turning partons into hadrons is separated into two steps; a probabilistic method to model the fragmentation of partons that effectively resums soft and collinear radiation (parton shower), and hadron formation (hadronization). In order to avoid double-counting parton emissions due to higher order processes, specialized matching and merging techniques (e.g. [19, 20]) are available to consistently interface NLO MC event generators to the parton shower, typically to leading-logarithmic (LL) precision. Hadronization is described and validated (e.g. [21, 22]) with phenomenological approaches, the most popular being the Lund string [23, 24] and the cluster [25] models, albeit their unique implementation has been lately questioned in view of multiparticle production effects [32].
1.3 That could well be a top quark!

Until the end of Tevatron data-taking in 2011, the collider had delivered a total of 12 fb$^{-1}$ of integrated luminosity. The successful attempt from LHC at $\sqrt{s} = 7$ TeV in 2010 (“Run 1”) has marked the beginning of a new era in top quark physics. Enhanced cross sections for top quark production, on the one hand, and ever increasing data sets culminating at 25 fb$^{-1}$ by the end of 2012 (Fig. 5a), on the other hand, rendered LHC the first “top quark factory”. After a two-year shutdown for maintenance of the LHC machine and experiments, the LHC has restarted in early 2015 (“Run 2”). The $\sqrt{s}$ value has further increased to 13 TeV, boosting typical top quark cross sections by about a factor of three compared to Run 1 (Fig. 5b).

![Figure 5](image_url)

Figure 5: Cumulative (“integrated”) delivered and measured luminosity versus time for pp collisions during LHC Run 1 (2012) and Run 2 (2016) period at $\sqrt{s} = 8$ (a) and 13 (b) TeV, respectively. The measurements make use of the best available offline calibration conditions [33, 34]. With an availability for luminosity production of 50% an integrated luminosity of almost 50 fb$^{-1}$ has been further delivered in 2017 [35].

Unlike the up and down quarks, which are stable, the top quark has a mean lifetime of only about $10^{-24}$ sec. In the SM, the top quark will decay almost instantly and nearly all the time into a W boson and a bottom (b) quark. Neither the W boson nor the b quarks can be directly observed. When a quark emerges from a collision, it gets “dressed up” by a cloud of quarks and antiquarks. What is then experimentally observed is a jet, a directed spray of particles that in the vacuum have roughly the same energy and direction as the original quark. The W boson decays with a probability of almost 70% [8] into a quark and an antiquark of the same generation (“hadronically”). In this case, the quark and antiquark show up as two jets correlated in phase space. But the W boson also decays, though with a decreased probability of around 10% [8], “leptonically”, i.e., into a charged and a neutral lepton of the same generation, such as a muon and a neutrino. The neutrino traverses a detector completely unobserved. Its presence can be indirectly deduced due to a significant amount of missing momentum, assumed to have been carried away by the neutrino.

Among the most striking features of a top quark signature are therefore the jets containing b quarks. General-purpose detectors, like CMS [36], put great emphasis on the ability to accurately track the paths of individual particles in a magnetic field, and in parallel, rely on extremely precise...
segmented calorimeters. The b quark lies within a jet as part of a hadron, then decays roughly half a millimeter from where it was created (Fig. 6a). The particles in jets can be very precisely tracked using a silicon vertex detector. More specifically, the CMS silicon vertex detector can locate the path of a particle to within $\pm 15 \, \mu m$. By identifying most of the tracks in a jet and extrapolating them backwards towards the primary interaction vertex, the point where the bottom quark decayed could thereby be attained and the jet identified (“tagged”) as a b jet (Fig. 6b).

Figure 6: (a) Illustration [39] of a heavy-flavor jet with a secondary vertex (SV) from the decay of a b or c hadron resulting in charged-particle tracks that are displaced relative to the primary interaction vertex (PV), and hence with a large impact parameter (IP) value. (b) Misidentification probability for c and light-flavor jets against b jet identification efficiency for various b tagging algorithms applied to jets (with a minimum transverse momentum requirement) in $t\bar{t}$ events from a full simulation of the CMS detector response [37] based on POWHEG [17, 38] (v2) at NLO along with the NNPDF3.0 [29] PDF set, and using PYTHIA [24] (v8.205) for hadronization [39].
2. A multifaceted quark

2.1 Most probably flies in pairs

Figure 7: Representative diagrams [40] for \( t\bar{t} \) production at LO: quark-antiquark annihilation (a), gluon-gluon fusion in the \( t \)- (b), \( s \)- (c), and \( u \)-channels (d).

At the Born level two processes with cross sections proportional to \( \alpha_s^2 \) contribute leading to \( t\bar{t} \) final states: gluon-gluon (gg) fusion and quark-antiquark (q\bar{q}) annihilation (Fig. 7). The relative importance of q\bar{q} as opposed to gg-initiated processes is simply dictated by \( \sqrt{s} \) and the PDFs of the initial-state hadrons. The fraction of \( t\bar{t} \) events initiated by gg collisions grows monotonically with \( \sqrt{s} \). It is around 73% at 5.02 TeV, as calculated with POWHEG [17, 38] (v2) at NLO using the NNPDF3.0 [29] NLO PDFs, and increases to around 86% at 13 TeV (Fig. 8a).

At NLO, \( t\bar{t} \) production processes with cross sections proportional to \( \alpha_s^3 \) become relevant, while new production channels, like (qg) or (\bar{q}g), open up. Ultraviolet divergences occurring in NLO calculations are systematically treated by absorbing them into redefinitions of the couplings in the Lagrangian, introducing the \( \mu_R \) scale into the calculation. Infrared and collinear divergences in the initial state are systematically canceled by absorbing them into the PDFs, introducing the \( \mu_F \) scale into the calculation (Section 1.2). The precision of \( t\bar{t} \) cross section calculations is further improved at NNLO, i.e., including processes up to \( \alpha_s^4 \), and at NNLL order, i.e., resumming contributions that are proportional to \( \alpha_s^2 \ln(\ldots)^{2n} \) at the \( n \)-th order. The most precise prediction of the inclusive \( t\bar{t} \) production cross section to date (NNLO+NNLL) reaches an uncertainty of around 5% (at a fixed top quark mass of 172.5 GeV), dominated by the uncertainties associated with the choice of the \( \mu_R \) and \( \mu_F \) scales as well as with the PDF sets [41]. Comparisons of experimental measurements—already surpassed such a threshold—to the state-of-the-art theoretical predictions allow to test the perturbative QCD calculations.

2.1.1 Measuring it inclusively with its partner

Precise measurements of inclusive \( \sigma_t \) have been performed at all \( \sqrt{s} \) and in multiple collision systems delivered by the LHC. The most precise results in pp collisions are summarized in Fig. 9. All reported measurements are in excellent agreement with the theoretical predictions. Thanks to
the high integrated luminosity recorded, in particular for the higher \( \sqrt{s} \), those measurements are not limited by the statistical uncertainty in the available data samples, but they are rather dominated by the systematic variations. Less precise measurements in the all-hadronic channel [42, 43, 44] and in final states involving hadronically-decaying \( \tau \) leptons [45, 46] have also been reported. Despite the large branching fraction, the \( \sigma_{\ell \bar{\ell}} \) measurement in the all-hadronic channel poses the challenge of an overwhelming QCD multijet background. This type of background bears a major contribution to final states containing hadronically decaying \( \tau \) leptons, whose reconstruction is also susceptible to misidentification.

The decay channels with the best precision so far are the dilepton (\( e^{\pm} \mu^{\mp} \)) and \( \ell+\text{jets} (\ell = e, \mu) \) final states with either one electron (\( e \)) or muon (\( \mu \)). Although the branching ratio is small, the dilepton channel suffer less from backgrounds uncertainties; Drell-Yan contribution can be heavily suppressed by the requirement of opposite lepton flavor and in the presence of high jet multiplicity. While the initial \( \sigma_{\ell \bar{\ell}} \) measurements(e.g. [52, 53, 54]) have been based on simple counting experiments, the subsequent application of profile likelihood fits, simultaneously relying on the number of events counted in the different event categories, has considerably helped to reduce the impact of systematic variations. Indeed, by accounting for the main experimental uncertainties, such as the jet energy scale, in terms of auxiliary (“nuisance”) fit parameters, the choice of appropriate categories (e.g. based on the number of b-tagged jets and the number of additional non-b-tagged jets) and/or distributions (jet \( p_T \) spectrum of the non-b-tagged jet with the lowest \( p_T \)) led to a reduction of the sensitivity to those parameters [48].  

The \( \ell+\text{jets} \) channel is less hampered by the statistical limitation of the data sample, yet it brings additional backgrounds from the EW production of top quarks, W+jets and QCD multijet
Figure 9: Inclusive $\sigma_{t\bar{t}}$ in $pp$ collisions as a function of the center-of-mass energy; the most precise CMS measurements at $\sqrt{s} = 7, 8$ [48, 49], and 13 [50, 51] TeV in the separate dilepton and $\ell$+jets channels are displayed, along with the combined measurement at 5.02 TeV [47]. The NNLO+NNLL theoretical prediction [41] using the NNPDF3.0 [29] NNLO PDF set with $\alpha_s(M_Z) = 0.118$ and $m_{\text{top}} = 172.5$ GeV is shown in the main plot. In the inset, additional predictions at $\sqrt{s} = 5.02$ TeV using the MMHT14 [28], CT14 [27], and ABMP16 [26] PDF sets, the latter with $\alpha_s(M_Z) = 0.115$ and $m_{\text{top}} = 170.4$, are compared, along with the NNPDF3.0 prediction, to the individual and combined results from the same analysis. The vertical bars and bands represent the total uncertainties in the data and in the predictions, respectively [47].

production, that are typically estimated using data-driven techniques. The latest analysis at $\sqrt{s} = 13$ TeV [51] relied on a profile likelihood fit to all events with one isolated lepton, significant missing transverse momentum and at least one reconstructed jet, while the jet and b-tagged jet multiplicity was used to categorize the events. As shown in Fig. 10 (a), the low multiplicity categories are dominated by backgrounds, allowing both their normalization and shape (e.g. the invariant mass of the lepton and b-tagged jet) to be determined from the fit procedure. The large number of event categories with varying signal and background contributions allowed the nuisance parameters to be strongly constrained by the data, rendering the uncertainty in the integrated luminosity as the largest single source of systematic uncertainty (Fig. 10, right). The typical assumption that the model should correctly capture all the sources of systematic uncertainties, along with their correlations across different event categories and distributions, is implicitly made.

Theoretical predictions for ratios of $\sigma_{t\bar{t}}$ at different $\sqrt{s}$ values of a and b, $R(a/b)$, can benefit from cancellations in the total uncertainty, e.g., $R(8/7)$ is calculated to be 1.430 ± 0.013 at NNLO+NNLL [41] accuracy. The total relative uncertainty is below 1%, as the uncertainties in $\sigma_{t\bar{t}}$ at the two $\sqrt{s}$ values are highly correlated. In addition to the quoted PDF [55], $\alpha_s$, and $\mu_R$, $\mu_F$ scale uncertainties, $R(8/7)$ varies by less than 0.1% for a ±1 GeV variation in $m_{\text{top}}$. A series of systematic uncertainties also cancel in the experimentally-determined ratios, especially for measurements
Figure 10: (a) Event yields from data and the expected $\bar{t}t$ signal and backgrounds for each of the independent event categories of jet and b tagged jet multiplicity considered in the latest $\ell+{\text{jets}}$ analysis at $\sqrt{s} = 13$ TeV. Distributions are combined for the two lepton charges and flavors. The bottom panel shows the ratio between the data and the expectations with the relative uncertainty (shaded band) owing to the statistical uncertainty in the simulations and the systematic uncertainty in the total integrated luminosity. (b) Estimated change in the parametrized ratio of the measured cross section to the reference theoretical value, $\Delta \mu / \mu$, stemming from the listed experimental and theoretical sources of uncertainties in the same analysis. The various contributions are shown from the largest to the smallest observed impact on $\mu$ [51].

performed with the very same technique. The most precise CMS measurement of this kind is currently $R(8/7) = 1.41 \pm 0.06$ [48], obtained after a partial cancellation of systematic uncertainties, consistent with the predicted ratio.
2.1.2 Closing the gap between Tevatron and LHC energies

In November 2015, the LHC delivered pp collisions at $\sqrt{s} = 5.02$ TeV, producing an integrated luminosity of $27.4 \pm 0.6$ pb$^{-1}$ [56]. Measurements of $t\bar{t}$ production at various $\sqrt{s}$ probe different values of the fraction $x$ of the proton longitudinal momentum, and thus provide complementary information on the gluon PDF (Fig. 8b). The combined measurement using dilepton and $\ell$+jets channels (at the level of measured values instead of likelihood functions) achieved a total relative uncertainty of 12% [47], that represents a remarkable achievement and a significant improvement relative to the first observation using the $e^{\pm}\mu^{\mp}$ channel alone [57]. Indeed, the correlation in phase space of the light jets carries a distinctive hallmark with respect to the main backgrounds that are controlled by counting the number of $b$-tagged jets in the selected $\ell$+jets events. The signal extraction has been then performed by maximizing a profile likelihood fit to the distribution of a kinematic variable (Fig. 11), sensitive to the resonant behavior of the light jets, for different categories of lepton flavor and jet multiplicity. Analogously to the most recent LHC studies of the inclusive $\sigma_t$, the measurement is first performed in a fiducial phase space—a restricted region of the phase space that closely resembles the detector acceptance in $p_T$ and $\eta$ of leptons and jets—and it is then extrapolated to the full phase space based on MC simulation. The individual and combined results at $\sqrt{s} = 5.02$ TeV are separately compared to the predictions for the ABMP16 [26], CT14 [27], MMHT14 [28], and NNPDF3.0 [29] PDF sets. The nice agreement among them can be traced back to consistent values of $\alpha_s$ and $m_{top}$, that are associated with the respective PDF set. The limited-precision measurement can be complemented with the significantly larger data set recorded in 2017, equivalent to almost 0.3 fb$^{-1}$ [58].

![Figure 11: Distribution of the angular distance $R = \sqrt{\eta^2 + \phi^2}$ of the least separated jets, $\min\Delta R(j, j')$, for $\ell$+jets events in 0b- (left), 1b- (center) and $\geq$2b- (b) tagged jet categories. The distributions observed in the data are compared to the sum of the expectations for the signal and backgrounds prior to the fit. The shaded band represents the statistical and integrated luminosity uncertainties on the expected signal and background yields [47].](image)

2.1.3 Hot off the press; now also in nuclear collisions

Until recently, top quark measurements remained out of reach in nuclear collisions [59] due to the reduced amount of integrated luminosity produced during the first period at the LHC, and the relatively low nucleon-nucleon center-of-mass energies ($\sqrt{s_{NN}}$) available at the BNL RHIC. Novel studies of top quark cross sections have finally become feasible with the 2016 LHC proton-
lead (pPb) run at $\sqrt{s_{NN}} = 8.16$ TeV, surpassing in performance almost eight times the designed instantaneous luminosity and delivering to CMS 174 nb$^{-1}$ of pPb collision data.

Figure 12: (a) Accumulation of integrated luminosity in each LHC experiment in the pPb runs for years 2013 and 2016 [61]. (b) The EPPS16 nuclear modifications for the bound gluon PDF in Pb nucleus at the parametrization scale $Q^2 = 10^4$ GeV$^2$. The thick black curves correspond to the central fit values, the dotted curves to the individual error sets, and the total uncertainties are shown as blue bands [62].

Asymmetric collisions of $^{208}$Pb$^{82+}$ nuclei with protons had not been included in the initial LHC design. However, unexpected discoveries in small collision systems, reminiscent of flow-like collective phenomena, engaged further investigations [60]. After the short, yet remarkable, pilot physics run in 2012, and the first full one-month run in early 2013, the second full pPb run took place in late 2016, in order to deliver data at $\sqrt{s_{NN}} = 8.16$ TeV for each direction of the beams. Apart from complex bunch filling schemes due to the generation of the beams from two separate injection paths, the distinct feature of operation with asymmetric collisions at LHC is the difference in revolution and cavity frequencies. Given that the colliding bunches have significantly different size and charge, both beams are displaced transversely, onto opposite-sign off-momentum orbits, and longitudinally, to restore collisions at the proper interaction points, during a process colloquially known as “cogging” [61]. The long-term integrated luminosity goal of 100 nb$^{-1}$ has been clearly surpassed (Fig. 12a), rendering the 2016 pPb run the baseline for several years.

The binding energies of nucleons in the nucleus are several orders of magnitude smaller than the momentum transfers of deep-inelastic scattering (DIS). Therefore, such a ratio naively should be close to unity, except for small corrections for the Fermi motion of nucleons in the nucleus. Contrary to expectations, the EMC experiment discovered a declining slope to the ratio [63], a result later confirmed with high-precision electron- and muon-scattering DIS data. The conclusions from the combined experimental evidence demonstrates the consistency in modifying the free nucleon PDFs, $f_i(x, Q^2)$, at low $Q^2$ and letting the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi evolution to take care of the momentum transfer dependence, i.e., $f_i^{A}(x, Q^2) = R^A_i(x, Q^2)f_i(x, Q^2)$, where $R^A_i(x, Q^2)f_i$ is the scale-dependent nuclear modifications encoded in nPDFs (Fig. 12b).
The top pair production cross section has been measured for the very first time in proton-nucleus collisions [64], using pPb data at $\sqrt{s_{\text{NN}}} = 8.16$ TeV with a total integrated luminosity of 174 nb$^{-1}$. The measurement is performed by analyzing events with exactly one isolated lepton and at least four jets, and minimally relies on assumptions derived by simulating signal and background processes. The resonant nature of the invariant mass of $jj'$, $m_{jj'}$, provides a distinctive feature of the signal with respect to the main backgrounds, i.e., from QCD multijet and W+jets processes. The significance of the $t\bar{t}$ signal against the background-only hypothesis is above five standard deviations. The measured cross section is $45 \pm 8$ nb, consistent with the expectations from scaled pp data as well as perturbative quantum chromodynamics calculations (Fig. 13a). To further support the hypothesis that the selected data are consistent with the production of top quarks, a “proxy” of the top quark mass is constructed as the invariant mass of candidates formed by pairing the W candidate with a b-tagged jet (Fig. 13b). This first study clearly paves the way for further detailed investigations of top quark production in nuclear interactions (e.g., [65]), providing in particular a new tool for studies of the hot and dense matter created in nucleus-nucleus collisions.

Figure 13: (a) Top quark pair production cross section in scaled pp and pPb collisions at $\sqrt{s_{\text{NN}}} = 8.16$ TeV; the CMS measurements [48, 49, 64] are compared to the NNLO+NNLL theory predictions [41] using state-of-the-art nPDFs [62] with baseline proton PDFs [27]. (b) Invariant mass distributions of the $t \rightarrow jj'b$ candidates in the $\geq 2$ b-tagged jet category after all selections. All signal and background parameters are kept fixed to the outcome of the $m_{jj'}$ fit [64].
2.1.4 Just the scale has never been enough

Although inclusive $\sigma_{t\bar{t}}$ measurements already provide stringent tests of the theoretical dependence on the initial state and $\sqrt{s}$, the study of $\sigma_{t\bar{t}}$ as a function of various kinematic properties of analysis objects (differentially) allow for more detailed insights into the $t\bar{t}$ production mechanism. Measurements of the distribution of kinematic observables can be translated into differential cross sections using unfolding techniques (Fig. 14). Such procedures allegedly correct the observed distributions for efficiency, acceptance and resolution effects. After all detector-related effects are alleviated by unfolding, differential cross sections from different experiments, if performed in the same fiducial phase space, can be directly compared among each other and with predictions from MC event generators (at particle and parton level) or from theoretical calculations (at parton level). The directly measured observables, e.g., the kinematic distributions of leptons, are corrected back to the level of “stable” particles, which are accessible in MC generators in a fiducial region of the phase space. On the contrary, the kinematics of the top quarks and antiquarks or of the $t\bar{t}$ system are well defined on parton, i.e., before decay but after gluon and photon radiation, rather than on particle level.

Figure 14: Representative metrics for estimating the sensitivity to the underlying top quark physics [66]. Ideally the most pure, stable, and efficient approach (a) should reproduce as closely as possible the kinematic properties generated at parton level (b), while it should be resilient against large corrections for the experimentally reconstructed quantities (c). Exemplary distribution for the transverse momentum spectrum of the hadronically decaying top quark, $p_T(t_h)$, in the detector (smeared) phase space level in the $\ell$+jets channel, extracted from a POWHEG [17, 38] (v2) interfaced to PYTHIA [24] (v8.205) full simulation [37] of the CMS detector.

Differential cross section results can be also normalized to the measured inclusive or fiducial $\sigma_{t\bar{t}}$. In this way normalization uncertainties, e.g., the luminosity uncertainty, cancel and the sensitivity of the measurement to the shape of kinematic distributions is improved:

$$\frac{1}{\sigma_{t\bar{t}}} \left( \frac{d\sigma_{t\bar{t}}}{dX} \right)^i = \frac{1}{\sum_j (\Delta X^j / \Delta X^i \cdot BR_j \cdot L)} \cdot \frac{x^i}{\Delta X^i \cdot BR \cdot L},$$

i.e., the calculation proceeded by having divided the number of signal (background subtracted) events in the data corrected for efficiency and possible migration events, $x^i$, by an optimal bin width, $\Delta X^i$.

The study of differential cross-sections therefore allow to test and constrain the MC generator and theoretical predictions. Figure 15 shows differential cross-sections of the $t\bar{t}$ system at parton
level as function of $p_T(\bar{t}t)$, $|y(\bar{t}t)|$, and $M(\bar{t}t)$, and at particle level as a function of the number of additional jets (not arising from the decay products of the $\bar{t}t$ pair) compared to several MC generators. These studies [67] are crucial for BSM searches, where contribution from top quark production constitutes an important background, for it increases the confidence into the generation chain used or to define a domain of validity in case of discrepancies with the data. At the level of stable top quarks, differential cross section measurements can be compared with SM predictions directly. These predictions are available at various levels of QCD precision, that go beyond current NLO MC event generators, and even include electroweak corrections proportional to $\alpha_s^3$, $\alpha_s^2\alpha$, and $\alpha_s^2\alpha^2$ [68]. However, a visible trend at $M(\bar{t}t) \gtrsim 1$ TeV is revealed, a feature that has been initially covered by the limited statistical precision [21] and thus needs further investigation.

Another observed deviation concerns about the reconstructed top quark $p_T$, that displays a softer spectrum in the data than it is predicted by MC simulation. An improved description of the data is achieved by an ad-hoc reweighting of $\bar{t}t$ events, with the weight depending on the transverse top quark and antiquark momenta at parton level. Since the reweighting mitigates the observed deviation of the top quark $p_T$ spectrum, yet it could well introduce distortions of any correlated variables, it is treated as an additional systematic uncertainty. The recent full NNLO QCD+NLO EW calculation [68] of the top-quark $p_T$ shows improved agreement with the measured spectrum, as compared to calculations with decreased accuracy. As it has been documented in the literature, although EW effects are rather small at the level of inclusive $\sigma_{\bar{t}t}$, they can have a sizable impact on differential distributions. However, after separately inspecting the differential cross section at parton level as a function of the top quark with the higher (Fig. 16a) and lower (Fig. 16b) transverse momentum, $p_T(t_{\text{high}})$ and $p_T(t_{\text{low}})$, it has been observed that the former is very well modeled, while the later is predicted to be harder.

The study of the jet multiplicity allows to probe the whole simulation chain. The lower multiplicities are mainly sensitive to the matrix-element description, while the higher multiplicities are mainly sensitive to the parton shower description (Section 1.2). As seen in Fig. 15d, the predictions of $\bar{t}t + \geq 4$ jets could vary to up-to approximately 20% depending on the generation chain, while they agree at lower jet multiplicity. The differential measurements thus help to disambiguate which generators are able to correctly describe the dynamics of the production, though currently there seems to be no MC event generator to simultaneously grasp the production mechanism over the wide kinematic range.
Figure 15: Differential cross section in $\ell$+jets channel at parton level as a function of $p_T(\bar{t}t)$ (a), $|y(\bar{t}t)|$ (b), and $M(\bar{t}t)$ (c). The data are shown as points with light (dark) error bands indicating the statistical (statistical and systematic) uncertainties. The cross section is compared to the predictions of POWHEG [17, 38] (v2) combined with PYTHIA [24] (v8.205) or HERWIG++ [25], and the simulations from MG5_AMC@NLO [16] interfaced to PYTHIA [24] (v8.205) with FxFx merging [20], and separately from SHERPA [18]. The ratios of the predictions to the measured cross sections are shown at the bottom of each panel. In addition, the cross section is compared as a function of the additional jet multiplicity (d) [67].
Figure 16: Differential cross section in the $\ell$+jets channel at parton level as a function of the top quark with the higher (a) and lower (b) transverse momentum, $p_T^{(\text{t}_{\text{high}})}$ and $p_T^{(\text{t}_{\text{low}})}$ [51]. The cross sections is compared to the predictions of POWHEG [17, 38] (v2) combined with PYTHIA [24] (v8.205) or HERWIG++ [25], and the simulations from MG5_AMC@NLO [16] interfaced to PYTHIA [24] (v8.205) with FxFX merging [20], and separately from the NNLO QCD+NLO EW [68] calculations. The ratios of the various predictions to the measured cross section are shown at the bottom of each panel [67].
2.1.5 A factory of gluons over broad momenta

The impact of the $\sigma_{t\bar{t}}$ measurements on the knowledge of the proton PDFs can be well illustrated based on the example at $\sqrt{s} = 5.02$ TeV [47]. The evaluation represents the first QCD analysis at NNLO, together with the combined measurements of neutral- and charged-current cross sections for deep inelastic electron- and positron-proton scattering (DIS) at HERA [70], and the CMS measurement of the muon charge asymmetry in W boson production [71]. The precise HERA DIS data, obtained from the combination of the individual H1 and ZEUS results, are directly sensitive to the valence and sea quark distributions and probe the gluon distribution through scaling violations, rendering these data the core of all PDF fits. Parametrized PDFs are fitted to the data in a procedure that is referred to as the “PDF fit”. The experimental uncertainties in the measurement are propagated to the 14-parameter fit using the MC method. Indeed, a moderate reduction of the uncertainty in the gluon PDF at high $x \gtrsim 0.1$ is observed, once the measured values for $\sigma_{t\bar{t}}$ are included in the fit (Fig. 17a). All changes in the central values of the PDFs are well within the fit uncertainties, while the uncertainties in the valence quark distributions remain unaffected.

Due to the large $t\bar{t}$ data samples at the LHC, also the first double differential cross sections have been published, for example as a function of $M(t\bar{t})$ and $|y(t\bar{t})|$. Their study focuses on correlations between kinematic properties of the top quarks and provides insights into extreme regions of the phase space. Thereby the measurement of the double differential cross-section of $M(t\bar{t})$ and $|y(t\bar{t})|$ has been used to constrain the proton PDFs and especially to improve the gluon PDF [69]. The uncertainties in the gluon distribution at $x > 0.01$ are significantly reduced, once the $t\bar{t}$ data are included in the fit, with the largest improvement in gluon PDF uncertainty (by more than a factor of two) to be observed at $x \approx 0.3$ (Fig. 17b), as expected from the LO kinematic relation $x = (M(t\bar{t})/\sqrt{s})\exp[\pm y(t\bar{t})]$.

Figure 17: The relative fractional uncertainties in the gluon PDF as functions of $x$ at the scale of (a) [47] 100000 and (b) [69] 30000 GeV$^2$. The results of the fit including top quark measurements, and without those are compared.
2.1.6 Breaking the 1 TeV barrier

Top quarks produced in high-energy collisions may receive large momenta, either in regular SM processes or by hypothetical high-mass particles decaying to top quarks. To study the production of top quarks with large transverse momenta ("boosted" top quarks), differential cross section measurements using boosted-top reconstruction techniques have been devised. While at low $p_T$ the events appear in resolved topology (Fig. 18a), where all jets are visible, at high $p_T \gtrsim 400$ GeV the decay products (leptons and jets) start becoming collimated, such that they begin to overlap in the $\eta$-$\phi$ space (Fig. 18b). In the last decade, a large number of algorithms was conceived to analyze boosted-jet topologies. In these algorithms, jets are first reconstructed with a typical large radius parameter ("fat" jets), and their structure is subsequently either declustered (jet-declustering algorithms) to identify potential substructures in the jet related to the decay of the mother particle or assigned a probability of having stemmed from a given number of overlapping jets (jet-shape algorithms). Combinations of several techniques can also be used that also allow for b-tagging in the dense environment of fat jets (Fig. 18c), while the exact choice of algorithm depends on the expected event topology and momentum range of the boosted object.

![Figure 18: Illustration of resolved (a) and boosted (b) $t\bar{t}$ event topologies in the $\ell$+jets decay channel \cite{40}. The corrected mass (c) for radiation, underlying event, and pileup effects, $m_{SD}$, of the leading jet, before (solid line) and after (dashed line) the requirement of at least one b tagged substructure, in $t\bar{t}$ simulated events \cite{73}.

Boosted top quark events can effectively double the reach in transverse momentum relative to the resolved topology, allowing to cross the TeV boundary in the measurement of top quark $p_T$. Jet-shape algorithms have been proven to be robust for discriminating between boosted top quarks (a three-prong jet) and ordinary QCD multijet background that tends to have fewer substructures. Given that the three-prong jet mass should be similar to that of $m_{top}$, with a pair combination close to the W boson mass, it is crucial the effects of radiation, underlying event (hadronic activity not attributed to the hard scattering), and pileup contamination to be mitigated. Soft and collinear particles can be removed from the fat jet clustering in a recursive procedure regulated by two parameters; the first (typically denoted as "$z_{cut}$") controls the strength of the fractional $p_T$ selection, and the second (typically denoted as "$\beta$") suppresses the collinear radiation. For the present studies at $\sqrt{s} = 13$ TeV, the values of $z_{cut} = 0.1$ and $\beta = 0$ are used, which have been found to yield the best
performance for fat jets [72]. For the extraction of the cross section (inclusively and differentially as a function of the $p_T$ of the leading top quark) a template fit to the distribution of the corrected jet mass has been performed in the fully hadronic channel [73]. The signal and QCD multijet templates are taken from the MC simulation and a control sample in the data, respectively. Similar measurements have been performed at $\sqrt{s} = 8$ TeV in the fully hadronic [74] and $\ell$+jets [75] channels, respectively (Fig. 19).

![Figure 19: Differential cross section in the fully hadronic decay channel at $\sqrt{s} = 8$ [74] (a, boosted regime) and 13 [73] (b, resolved and boosted regimes) TeV, respectively. In addition, the unfolded result at $\sqrt{s} = 8$ TeV is compared to the measurement in the $\ell$+jets [75] channel, also in the boosted regime.](image)

The first differential $t\bar{t}$ cross section measurement [76] as a function of the mass of boosted top quarks has been also performed as a proof of principle to measure $m_{\text{top}}$ in the boosted final state (Fig. 20). More specifically, the peak position in the three-prong jet mass distribution is sensitive to $m_{\text{top}}$. The uncertainty in the top quark mass measurement is so far limited (± 9 GeV), dominated by the statistical and jet energy scale uncertainty in the boosted regime. Although, potential improvements and more constraints in systematic uncertainties are anyway possible in future application of this method on larger data samples, the prospect of reaching a more reliable correspondence between $m_{\text{top}}$ in any well-defined renormalization scheme and the top quark mass parameter in MC event generators should be faced with care. The first attempts to retain a universal description of the $m_{\text{top}}$ scheme with decay effects, and to remove soft contamination indicate significantly different values for the allowed region in both $z_{\text{cut}}$ and $\beta$ parameters [77]. Related to that, pileup mitigation could be also effectively handled experimentally with reverting to a per-particle-level suppression techniques, that were found to maintain top tagging performance up to very high number of simultaneous pp interactions [78].
Figure 20: (a) Differential cross section in the $\ell$+jets channel at particle level as a function of the mass of the leading jet. The cross sections from the combined electron and muon channels (points) are compared to LO predictions from MADGRAPH [14] (v.5.1.5.11) interfaced to PYTHIA [23] (v6.424) with MLM [19] merging, and NLO predictions separately using POWHEG [17, 38] (v1) combined with PYTHIA [23] (v6.424), and MC@NLO [15] (v3.41) combined with HERWIG++ [25]. The vertical bars represent the statistical (inner) and the total (outer) uncertainties. The horizontal bars show the bin widths. (b) The differential cross section is further normalized to unity and compared to LO predictions from MADGRAPH [14] (v.5.1.5.11)+PYTHIA[23] (v6.424) for three distinct values of $m_{\text{top}}$ [76].
2.2 “If we be doomed to remain single, we do”  

![Diagram for electroweak single top-quark production at LO: (a) t-channel, (b) associated tW, (c) s-channel production.]

Figure 21: Representative diagrams [40] for electroweak single top-quark production at LO: t-channel (a), associated tW (b), and s-channel production (c).

In addition to t¯t production, top quarks can also be produced singly via electroweak processes. Although the Born-level diagrams, classified by the virtuality of the exchanged W boson as shown in Fig. 21, exhibits a clear separation of three distinctive mechanisms, NLO corrections to tW diagram render these contributions to interfere with the t¯t production. In fixed-order calculations, the interference can be controlled by excluding the resonant region, where the invariant mass of the Wb system is close to the top quark mass, or by vetoing additional b quarks based on \( p_T \) spectrum. Due to initial- and final-state effects, such approaches are not directly applicable in MC generators. The overlap between tW and t¯t production is treated either by diagram removal (DR), where doubly-resonant diagrams are removed from the calculation, or diagram subtraction (DS), that cancels t¯t contributions locally by implementing a subtraction term. Both techniques should lead to comparable results, though in specific regions of the phase space discrepancies of \( O(30\%) \) exist for more complicated processes [79].

Typically, the single top-quark production measurements focus on the leptonic top quark decays, rendering a generic basis for signal scouting a relatively high-\( p_T \) lepton, a significant amount of missing momentum as well as a b-tagged jet from the top-quark decay. The additional signatures of single top-quark production are: a light flavor jet with large \( |\eta| \) in t-channel, an additional b-jet in s-channel, and two oppositely charged high-\( p_T \) leptons in associated tW production. For all single top-quark production channels, t¯t production is thus a major background.

The electroweak t-channel production of top quarks (Fig. 21a) is the most abundant single top production mode in the SM with a predicted cross section at NNLO that equals to \( \sigma_{t\text{-ch.}} = 214.5^{+1.0}_{-1.2} \) (scale) pb at \( \sqrt{s} = 13 \) TeV [80] (approximately 70% of the total single top quark cross section). It is also known fully differentially to NNLO accuracy, first calculated assuming stable top quarks [80], and recently for production and decay [81]. The NLO corrections to the LO t-channel cross section are accidentally small–of the order of a few percent–and the NNLO corrections are of the same order (Ref. [82] for an overview). Since single top-quark production in the t-channel has a moderately large cross section, it was established early in LHC Run 1 [83] (the first single top measurement that did not depend on multivariate techniques) and Run 2 [84].

In associated tW production (Fig. 21b) the W boson is real. This process is known to approximate NNLO QCD accuracy [85], and it was observed for the first time at the LHC [100], for its cross section was negligible at the Tevatron. In s-channel single top-quark production (Fig. 21c) a

\(^2\)T. Hardy, *Under the Greenwood Tree.*
Figure 22: (a) Yields observed in data compared with those expected from simulation as a function of the number of jets and b-tagged jets. The error band includes the statistical and all systematic uncertainties excluding the ones related to background normalization and integrated luminosity. Post-fit comparison between the observed data and simulated events in the classifier output (b) and \( p_T \) of the subleading jet (c) for events with exactly one jet which is tagged as a b jet and for events with exactly two jets which are tagged as b jets, respectively. The error band includes the statistical and systematic uncertainties. In all cases, the bottom of each panel shows the ratios of the data to the sum of the expected yields [89].

The process is also known to approximate NNLO QCD accuracy [86], and it is characterized by the smallest single top-quark production cross section at the LHC. So far it remains elusive [87], and it was only established as a separate single top-quark production channel at Tevatron [88].

Multivariate separation of signal and background, and profile likelihood fits to control the amount \( t\bar{t} \) background and to extract the cross section are adequate, if not mandatory. A characteristic example of measuring the associated \( tW \) production has been recently released [89], which also takes advantage of 36 fb\(^{-1} \) of pp collision data. Events have been categorized depending on the number of jets and b tagged jets as presented in Fig. 22a. The cross section has been then estimated through a profile likelihood fit to the distributions of the classifier output (Fig. 22b) and \( p_T \) of the...
subleading jet (Fig. 22c) simultaneously in signal- and background-enriched regions, respectively. A summary of inclusive single top-quark cross section measurements at the LHC in all production channels and for different center-of-mass energies is presented in Fig. 23. With the large data sets available at the LHC, also the first differential cross sections for $t$-channel single top-quark production as a function of the top quark transverse momentum and rapidity became feasible [90]. All measurements are in good agreement with the SM prediction.

Figure 23: Summary of LHC (and Tevatron as on the legend) measurements of the inclusive single top quark production cross section in $t$- [96, 97, 98] and $s$-channels [87], and in association with a $W$ [89, 99, 100] or $Z$ [94, 101] boson. The measurements are compared to theoretical calculations based on approximate NNLO QCD accuracy [85, 86, 95], expect for the $Z$-associated production that is based on MG5_aMC@NLO [16] (v.254). For all predictions the NNPDF3.0 [29] PDF set has been used and the uncertainties include variations due to $\mu_F$, $\mu_R$ scales, and PDF parameterization. The predictions for the $s$-channel production is given up to $\sqrt{s} = 3$ TeV for $p\bar{p}$ collisions and for $pp$ collisions starting from $\sqrt{s} = 5$ TeV, while for the rest of the production channels the curves corresponding to both $pp$ and $p\bar{p}$ collisions coincide at the considered accuracy [82, 102].

An interesting observable in $t$- and $s$-channel single top-quark production is the ratio of inclusive production rates for single top quark to top antiquark, $R_{t\text{-ch.}}$. More specifically, the top quark is produced with an up-type quark (or down-type antiquark) in the initial state, whereas the top antiquark is produced with a down-type quark (or up-type antiquark). Hence, in pp collisions $R_{t\text{-ch.}}$ is sensitive to the ratio of PDFs for up-type and down-type quarks and down-type and up-type antiquarks, with a naive expectation of $R_{t\text{-ch.}} = 2$ for up and down valence quarks only. The exact predictions of $R_{t\text{-ch.}}$ depend on $\sqrt{s}$ though; the higher the $\sqrt{s}$ the greater the impact of sea quarks, and thus the more the deviation from the naive expectation. Experimentally, $R_{t\text{-ch.}}$ is a robust observable in which many uncertainties cancel. Measurements of $R_{t\text{-ch.}}$ from LHC Run 1 (Fig. 24a) and Run 2 (Fig. 24a) are compatible with the SM prediction within the presently large uncertainties. The updated ABMP16 (NLO version retrieved from Ref. [91]) PDF sets predict ratios similar
to their ABM12 predecessor, while the latest NNPDF3.1 [92] version lies within $+1\sigma$ relative to NNPDF3.0. The values for the new ATLAS-epWZ16 [93] PDF set is compatible with the ones from the MMHT14 [28] PDFs. The relative importance of $R_{t\text{-ch.}}$ in PDF extractions is expected to grow with more integrated luminosity. Another complementary constraint on PDFs can be obtained from the ratio of $t$-channel cross sections at $\sqrt{s} = 7$ and 8 TeV [97].

Figure 24: Measurements of $R_{t\text{-ch.}}$, $\sqrt{s} = 8$ (a) and 13 (b) TeV, compared with theoretical expectations at NLO employing the CT14 [27], MMHT14 [28], NNPDF3.0 [29], and HERAPDF2.0 [30] PDF sets. Former versions of ABM16 [26] have been used. Error bars for the different PDF sets represent the quadratic sum of the 68% confidence level interval of the predictions of the eigenvectors in the respective set, and the uncertainty due to $\mu_F$, $\mu_R$ scales. An uncertainty of $\pm 2$ GeV in the top quark mass is also considered, while the statistical uncertainty due to the finite number of iterations employed for the integration is less significant (adapted from Ref. [82]).
2.3 It pins down to two numbers; the mass and the width

Moreover $\sigma_{t\bar{t}}$ measurements can be reinterpreted in terms of constraints on the top quark mass. The inclusive $t\bar{t}$ production cross section, as predicted by perturbative QCD, is a (steeply) falling function of $m_{\text{top}}$. In a given renormalization scheme, e.g., in the on-shell or the $\overline{\text{MS}}$ scheme, the mass parameter is defined unambiguously \cite{103}. The measured inclusive cross section also exhibits a (weak) dependence on the mass parameter used in the MC simulation \cite{104}; the higher the momenta transferred to the decay products, the higher the acceptance for $t\bar{t}$ events, hence the measured $\sigma_{t\bar{t}}$ decreases slightly with $m_{\text{top}}$. The top-quark mass can be thus determined from the intersection of the curves describing the $m_{\text{top}}$ dependence of the theoretical and the measured $\sigma_{t\bar{t}}$, as illustrated in Fig. 25a. This method has been inherited from Tevatron \cite{105,106}, and it has been subsequently implemented at LHC \cite{107} reaching a precision on the top quark pole mass of $\mathcal{O}(1\%)$ \cite{48}.

![Graph](image)

Figure 25: (a) Likelihood for the predicted dependence of the $t\bar{t}$ production cross section on the top quark pole mass for $\sqrt{s} = 7$ and 8 TeV, employing the NNPDF3.0 \cite{29} PDF set. The measured dependences on the mass are given by the dashed lines, along with their $\pm 1\sigma$ uncertainties represented by the dotted lines. The extracted mass at each value of $\sqrt{s}$ is indicated by a black point, with its $\pm 1\sigma$ uncertainty as the continuous contour \cite{48}. (b) Top quark pole mass at NNLO+NNLL \cite{41} accuracy extracted by comparing the combined measured $t\bar{t}$ production cross section at $\sqrt{s} = 7$ and 8 TeV \cite{48} with predictions employing the CT14 \cite{27}, MMHT14 \cite{28}, and NNPDF3.0 \cite{29} PDF sets \cite{48}.

The width of the top-quark, $\Gamma_t$, can be determined directly from the kinematic reconstruction of its decay products. Similar to measurements of $m_{\text{top}}$, an observable sensitive to the top quark width is built from reconstructed quantities. Such an observable is the invariant mass of charged lepton and $b$ jet pairs in dilepton $t\bar{t}$ events (Fig. 26a). In a series of binary hypotheses, the SM value of $\Gamma_t$, $\Gamma_t = \Gamma_{\text{SM}}$, has been tested \cite{108} against different BSM width hypotheses to extract an observed 95% confidence interval of $0.6 \leq \Gamma_t \leq 2.5$ GeV (Fig. 26b) to be compared with the NLO value of $\Gamma_{\text{SM}} = 1.330$ \cite{109} assuming $m_{\text{top}} = 172.5$ GeV and $\alpha_s = 0.118$. The sensitivity of this direct method \cite{108} (including single resonant $tW$ production) is lower than the sensitivity of indirect methods (e.g. \cite{110}), but both are still far away from challenging its $< 1\%$ theoretical
uncertainty.

Figure 26: (a) Expected sensitivity for the distribution of the invariant mass of charged lepton and b jet pairs, $M_{\ell b}$, in the dilepton $t\bar{t}$ channel for simulated POWHEG [17, 38] (v2) combined with PYTHIA [24] (v8.205) events, where the top quark mass is varied by $\pm 3$ GeV with respect to the default value of 172.5 GeV, and the width is varied by a factor of 4 with respect to the SM value. The plot describes boosted ($p_T \geq 100$) events, with at least two b tagged jets ($\geq 2b$ category). The top panel shows the distribution with the last bin displaying the overflow of the histograms, while the bottom plot shows the ratio with respect to the $m_{\text{top}} = 172.5$ GeV and $\Gamma_t = \Gamma_{\text{SM}}$ scenario. (b) Evolution of the confidence limits as a function of the top quark width. The derived limits at the 95% confidence level are derived from the intersection of the fits to the horizontal line at 0.05 [108].
2.4 Rare and rarer signatures

The production of top quarks in association with the electroweak gauge bosons, or even with the Higgs boson, is predicted to be rare in the SM, with inclusive production cross sections in pp collisions at $\sqrt{s} = 13$ TeV as predicted at NLO QCD below 10 pb. However, measuring these processes gives direct access to the couplings of the top quark, meaning that any deviation of the measured values from the SM prediction would be an indication of BSM physics. Although evidence for $t\bar{t}\gamma$ production had been already reported at Tevatron [111], the main challenge lies in separating signal photons from genuine (hadron decays into photon pairs) or misidentified (hadrons and electrons) photons. By simultaneously reconstructing a top quark in the event (Fig. 27a) and rejecting photons close to the tail of the photon isolation, a high fraction of signal events could be retrieved (Fig. 27b). The production cross section, defined in a fiducial region of phase space (exactly one charged lepton and the corresponding neutrino, at least three jets, and a photon), is compatible with the SM prediction at NLO [112].

![Figure 27](image_url)

Figure 27: (a) Distribution of the invariant mass of the three jet combination that gives the highest vectorial sum of individual jet transverse momentum, $M_3$, in data and simulation. The fraction of events passing the photon selection containing top quark pairs, referred to as the “top quark purity”, is used to distinguish $t\bar{t}$ pair events from other backgrounds. The lower panel shows the ratio of the data to the prediction from simulation. The uncertainty band is a combination of statistical and systematic uncertainties in the simulation. (b) Result of the fit to the photon charged-hadron isolation in the $\ell+\text{jets}$ decay channel. The uncertainty band shows the statistical uncertainties in the templates derived directly from the data. The measurement of the “photon purity”, defined as the fraction of reconstructed photons in the selection region, is used to discriminate between the genuine photons expected from $t\bar{t}\gamma$ and the nonprompt photons from the $t\bar{t}$ background. The lower panel shows the ratio of the distribution observed in data to the sum of the templates scaled to the fit result [112].

Measurements of $t\bar{t}Z$ and $t\bar{t}W$ production have only become feasible with the large LHC datasets [113, 114]. The processes feature very massive final states of $O(0.5)$ TeV, and the almost fourfold increase in production cross sections at LHC Run 2 compared to Run 1 resulted in promptly re-observing both $t\bar{t}Z$ and $t\bar{t}W$. The multilepton final states are studied in order to reduce the major background processes. The background in these events are generally low, though it is
challenging to precisely estimate the residual number of wrongly identified leptons. The fake lepton background is usually modeled from the data in control regions; the probability for a loosely identified nonprompt lepton to fulfill the full set of tight requirements is estimated and validated using MC simulation. While the requirement of two leptons compatible with the Z boson significantly improves the purity of the sample for the measurement of t\bar{t}Z production, the selection of t\bar{t}W relies on a multivariate approach based on kinematic variables such as the missing missing, the leading lepton or jet $p_T$. A categorization of the selected events based on jet and b tagged jet multiplicity helps to better control the background. The measurements [113, 114]–with uncertainties equally shared between statistic and systematic uncertainties–are compatible with SM predictions (Fig. 28), and have been used to constrain BSM physics contributions to the $Ztt$, $Wtt$, and $Htt$ couplings [114].

![Graph showing the simultaneous fit for $t\bar{t}Z$ and $t\bar{t}W$ cross sections](image)

Figure 28: Result of the simultaneous fit for $t\bar{t}Z$ and $t\bar{t}W$ cross sections (star), along with its 68 and 95% confidence level contours. The right panel presents the individual measured cross sections along with the 68 and 95% confidence level intervals and the theory prediction with their respective uncertainties for $t\bar{t}W$ (upper right) and $t\bar{t}Z$ (bottom right) [114].

In view of the null results of the searches for new heavy particles decaying into top quarks, a comprehensive effective field theory (EFT) approach [115] to study top quark couplings becomes attractive. In such an approach the indirect effects of BSM physics on the top quark couplings are treated in a consistent way, by constructing a full set of effective operators that mediate top quark couplings with mass-dimension six (the single dimension-five operator violates lepton number conservation). An extended Lagrangian is assumed based on a series expansion in the inverse of the energy scale of the BSM physics, $1/\Lambda$, hence operators are suppressed as long as $\Lambda$ is large.
compared with the experimentally-accessible energy:

\[ \mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \sum_i c_i^{(6)} \frac{O^{(6)}_i}{\Lambda^2} + \mathcal{O}(\Lambda^{-4}) . \]

In the above equation, the effective Lagrangian density \( \mathcal{L}_{\text{eff}} \) is given by the SM Lagrangian \( \mathcal{L}_{\text{SM}} \) and is complemented with a sum of dimension-six operators \( O^{(6)}_i \), each weighted with a Wilson coefficient \( c_i^{(6)} \), calculable in LO QCD perturbation theory. The EFT approach is a comprehensive description of the top couplings in a gauge invariant and renormalizable way that respects all SM symmetries, and the first NLO corrections in the top quark sector have been presented.

The \( t\bar{t}Z \) and \( t\bar{t}W \) results of Ref. \[114\] have been then used to set constraints on the Wilson coefficients of dimension-six operators. Eight operators have been identified which are of particular interest because they change the expected cross sections of \( t\bar{t}Z \), \( t\bar{t}W \), or \( t\bar{t}H \) without significantly impacting the expected background yields (Fig. 29a). The constraints, obtained by considering one operator at a time (Fig. 29b), are a useful first step towards more global approaches.

Figure 29: (a) Representative LO diagram with the most significant BSM effect for a given operator identified to affect the contribution from \( t\bar{t}Z \) and \( t\bar{t}W \). (b) The \( t\bar{t}Z \) and \( t\bar{t}W \) cross section corresponding to the best fit value for such an operator, along with the corresponding 68% (dashed) and 95% (dash-dotted) contours. The theory predictions (dot) with their uncertainties (bars), and the two-dimensional best fit to the \( t\bar{t}Z \) and \( t\bar{t}W \) cross sections (star) are superimposed [114].
An additional rare process in the SM is the production of single top quarks in association with a $Z$ boson, that has been already searched for during Run 1 [101]. Events containing three leptons, two or three jets, and significant missing momentum have been selected for the updated analysis using pp data corresponding to 36 fb$^{-1}$ [94]. A higher separation of signal from background events, e.g., from $t\bar{t}Z$ or $WZ+jets$, has been ensued by combining different types of multivariate discriminants that resulted in 20% increase in the expected significance. The search has exceeded the conventional $3\sigma$ threshold (“evidence” for the existence of the signal process) on the incompatibility with the background-only hypothesis, in contrast to the expected 1.8 standard deviations in Run 1.

![Figure 30](image-url)

Figure 30: Post-fit data-to-prediction comparisons of the employed variables, i.e, the classifier output (a, b) and transverse mass of the $W \rightarrow \ell\nu$ system (c), for events (a, signal enriched) with one b jet and one light flavor jet, (b, $t\bar{t}Z$ enriched) at least two jets with at least two of them b tagged, (c, $WZ+jets$ enriched) and at least one jet without any b-tagged. The hatched bands include the total uncertainty in the background and signal contributions. The pulls of the distributions, defined in each bin as the difference between the data and prediction, divided by the quadratic sum of total uncertainties in the predictions (systematic and statistical) and the data (statistical), are shown at the bottom of the plots [94].

Four top quark production ($t\bar{t}t\bar{t}$) is an even rarer process, and it is currently predicted at NLO precision [116] in the SM. The representative Born level Feynman diagrams of $t\bar{t}t\bar{t}$ production, which occurs either through the gluon, the electroweak gauge boson, or the Higgs boson mediation, are shown in Fig. 31. The Higgs-induced $t\bar{t}t\bar{t}$ production (Fig. 31b) exhibits no dependence on the Higgs boson width, with the cross section proportional to the top quark Yukawa coupling to the fourth power [117]. Although a high integrated luminosity is needed to reach a $5\sigma$ discovery of the rare production (outside the LHC Run 2 reach), null searching results in the low luminosity operation of the LHC are also useful because they can be used to constrain the top Yukawa coupling.

The same-sign dilepton and the three- (or more) lepton final states, though represent a very low branching fraction, are the most sensitive to $t\bar{t}t\bar{t}$ production in the regime with SM-like kinematic properties, due to the low level of backgrounds. By optimizing the signal selection for sensitivity to SM $t\bar{t}t\bar{t}$ production, by using an improved b jet identification algorithm, and by employing background estimation techniques that are adapted to take into account the higher jet and b jet mul-
Figure 31: Representative diagrams for $t\bar{t}t\bar{t}$ production at LO in the SM through (a) gluon and (b) the Higgs boson mediation [121].

toplicity requirements in the signal regions (Fig. 32a), the corresponding expected upper limits of have improved on previous (by 1.6 times relative to Ref. [118]), and most recent (by 3.5 relative to Ref. [119] and 1.4 times relative to Ref. [120]) searches. Reinterpreting the results to constrain the ratio of the top quark Yukawa coupling to its SM value, a lower bound of 1.9 at 95% confidence level (Fig. 32b) is attained [121].

Figure 32: (a) Observed yields in all (control and signal) regions, compared to the post-fit predictions for signal and background processes. The hatched areas represent the total uncertainties in the signal and background predictions. The upper panel shows the ratios of the observed event yield to the total prediction. (b) The predicted SM value for $t\bar{t}t\bar{t}$ production [117] (taking into account a NLO/LO factor at $\gamma_t^{\text{SM}}$ [116]) as a function of the ratio of the top quark Yukawa coupling to its SM value (dashed line), $|y_t/y_t^{\text{SM}}|$, compared with the observed value (solid line), and with the 95% confidence level upper limit (hatched line). The measured cross section dependence on $|y_t/y_t^{\text{SM}}|$ is induced due to the rescaling of the $t\bar{t}H$ background by a factor of $|y_t/y_t^{\text{SM}}|^2$ [121].
3. “I like the dreams of the future better than the history of the past” ³

Top physics has come a long way from discovery and first measurements to increasingly sophisticated analyses using the LHC data. All measurements so far are compatible with the expectations of the SM and contribute significantly to constraining possible BSM extensions. As data accumulate, uncertainties in signal modeling becomes a hindrance for further progress though. The rich variety of inclusive and differential cross sections can be extensively used to constrain theoretical predictions based on MC simulation and analytical calculations. More specifically, they serve as substantial input for MC tuning or performing dedicated ancillary measurements, such as the studies of color flow in $t\bar{t}$ events. The importance of higher-order corrections is also validated, albeit the gradual decrease in experimental uncertainties reveals discrepancies to be further scrutinized.

One unique property of top quark events is the high fraction of b-flavored quarks in the partonic final state. This is routinely exploited through “in-situ” measurements of the b-tagging efficiency in challenging experimental conditions characterized with several jets and charged leptons, resembling the signal region of many BSM physics searches. Meanwhile, the top quark properties are remarkably well understood, and their precise determination is essential for testing the overall consistency of the SM through precision electroweak fits. In parallel, top quark has an extraordinary impact on the Higgs sector and on the SM extrapolation to high-energies. Precision measurements and ever increasing in accuracy calculations for top quark production can be used to include $t\bar{t}$ and single top quark data into free and bound proton PDF fits. It is important though to test what is the bias in PDF-related uncertainties on the predictions given the potential circularity.

Top-quark physics will remain important after the scheduled upgrades of the LHC experiments (HL-LHC), and infrastructures at future colliders may take the quest for the top to the next level. Projections of the expected uncertainties from key results help to reveal the physics potential at the HL-LHC. Most prominently, the uncertainty in the top quark mass could reach a rather optimistic estimate of 200 MeV, with its simultaneous determination with the production cross section grazing the 1 GeV threshold (Fig. 33a). The former method rivals a prominent and long-sought way of precision measurement of the top quark mass; a threshold scan of the toponium state to be performed by a linear collider whose initial energy is controlled. The couplings of the top quark as well as rare processes, such as flavour-changing neutral currents (FCNCs), may reveal deviations from the SM. Future projections represent good prospects for pushing top-FCNC boundaries to even higher constraints (Fig. 33b).

³To J. Adams from T. Jefferson, 1 August 1816.
Figure 33: (a) Total uncertainty on top quark mass obtained with different methods and their projections to the HL-LHC conditions upgrade. The projections for $\sqrt{s} = 14$ TeV, with 0.3 ab$^{-1}$ or 3 ab$^{-1}$ of data, are based on measurements performed at Run 1, assuming that an upgraded detector will maintain the same physics performance despite of a severe increase in pileup [127]. (b) Compilation [129] of branching ratio limits for top quark FCNC decays (points), compared to the SM predictions (black lines), as well as to various projections for HL-LHC conditions and future experiments (colored lines).
References

[1] CDF Collaboration, Observation of Top Quark Production in p\overline{p} Collisions with the Collider Detector at Fermilab, Phys. Rev. Lett. 74 (1995) 2626


[6] ATLAS Collaboration, Measurement of the Higgs boson mass from the H \rightarrow \gamma\gamma and H \rightarrow ZZ \rightarrow 4\ell channels with the ATLAS detector using 25 fb^{-1} of pp collision data, Phys. Rev. D 90 (2014) 052004


[14] Johan Alwall et al., MadGraph 5 : Going Beyond, JHEP 1106 (2011) 128


Georgios Konstantinos Krintiras


[41] M. Czakon et al., *Total top-quark pair-production cross section at hadron colliders through $t\bar{t}(\alpha_s)$*, Phys. Rev. Lett. 110 (2013) 252004

38
[42] CMS Collaboration, *Measurement of the $t\bar{t}$ production cross section in the all-jet final state in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP 1305 (2013) 065


[47] CMS Collaboration, *Measurement of the inclusive $t\bar{t}$ cross section in proton-proton collisions at $\sqrt{s} = 5.02$ TeV using final states with at least one charged lepton*, arXiv:1711.03143 (submitted to JHEP, 2016)

[48] CMS Collaboration, *Measurement of the $t\bar{t}$ production cross section in the $e\mu$ channel in proton-proton collisions at $\sqrt{s} = 7$ and 8 TeV*, JHEP 08 (2016) 029


[51] CMS Collaboration, *Measurement of the $t\bar{t}$ production cross section using events with one lepton and at least one jet in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP 09 (2017) 051


[53] CMS Collaboration, *Measurement of the $t\bar{t}$ production cross section and the top quark mass in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP 07 (2011) 049

[54] CMS Collaboration, *Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP 11 (2012) 067


[57] CMS Collaboration, *First measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s} = 5.02$ TeV*, CMS Physics Analysis Summary CMS-PAS-TOP-16-015 (2016), https://cds.cern.ch/record/2161499


[87] CMS Collaboration, *Search for s channel single top quark production in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, JHEP 09 (2016) 027

[88] CDF and DØ Collaborations, *Observation of s-Channel Production of Single Top Quarks at the Tevatron*, PRL 112 (2014) 231803


[91] https://www-zeuthen.desy.de/~alekhin/PDFS/preliminaries/ABMP16_5_nlo_v6.zip


[95] N. Kidonakis, Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production, Phys. Rev. D 83 (2011) 091503


[100] CMS Collaboration, Observation of the Associated Production of a Single Top Quark and a W Boson in pp Collisions at $\sqrt{s} = 8$ TeV, Phys. Rev. Lett. 112 (2014) 231802

[101] CMS Collaboration, Search for associated production of a Z boson with a single top quark and for $tZ$ flavour-changing interactions in pp collisions at $\sqrt{s} = 8$ TeV, JHEP 07 (2017) 003


[105] DØ Collaboration, Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, Phys. Rev. Lett. 100 (2008) 192004

[106] DØ Collaboration, Measurement of the inclusive $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV and determination of the top quark pole mass, Phys. Rev. D 94 (2016) 092004


[110] CMS Collaboration, Measurement of the ratio $\mathcal{B}(t \rightarrow Wb)/\mathcal{B}(t \rightarrow Wq)$ in pp collisions at $\sqrt{s} = 8$ TeV, Phys. Lett. B 736 (2014) 33

[111] CDF Collaboration, Evidence for $t\bar{t}$ production and measurement of $\sigma_{t\bar{t}}/\sigma_{t\bar{t}}$, Phys. Rev. D 84 (2011), 031104

[112] CMS Collaboration, Measurement of the semileptonic $t\bar{t} + \gamma$ production cross section in pp collisions at $\sqrt{s} = 8$ TeV, JHEP 10 (2017) 006

42


[117] Q.-H. Cao et al., *Probing Higgs width and top quark Yukawa coupling from $t\bar{t}H$ and $t\bar{t}t\bar{t}$ productions*, Phys. Rev. D 95 (2017) 053004


[130] A.F. Zarneck, *TOP FCNC decays $t \rightarrow cH$ and $t \rightarrow c\gamma$ at CLIC*, Proceedings of the International Workshop on Future Linear Colliders (LCWS2017), 23-27 October 2017, Strasbourg, France