CMS Physics Analysis Summary

Contact: cms-pag-conveners-top@cern.ch 2023/03/22

Observation of four top quark production in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$

The CMS Collaboration

Abstract

The first observation of the production of four top quarks in proton-proton collisions is reported, based on a data sample collected by the CMS experiment at a center-of-mass energy of 13 TeV in 2016 to 2018 at the CERN LHC and corresponding to an integrated luminosity of 138 fb^{-1}. Events with two same-sign, three, and four charged leptons (electrons and muons) and additional jets are analyzed with multivariate discriminants to distinguish the signal process from the main backgrounds. The signal cross section is measured with a profile likelihood fit to be $17.9\,^{+3.7}_{-3.5}\,(\text{stat})\,^{+2.4}_{-2.1}\,(\text{syst})$ fb, in agreement with the best available theoretical predictions. The observed (expected) significance of the signal is 5.5 (4.9) standard deviations above the background-only hypothesis.

1 Introduction

The production of four top quarks ($\mathrm{t\bar{t}t\bar{t}}$) in proton-proton (pp) collisions is among the rarest processes of the standard model (SM) of particle physics searched for at hadron colliders. The SM cross section at $\sqrt{s} = 13\,\text{TeV}$ is calculated at next-to-leading order (NLO) in quantum chromodynamics (QCD) and electroweak theory as $12.0\,^{+2.2}_{-2.5}\,\text{fb}$ [1], and the prediction has recently been updated with soft gluon emission corrections at next-to-leading logarithmic accuracy to $13.4\,^{+1.0}_{-1.8}\,\text{fb}$ [2]. In the pp collision data sets recorded with the ATLAS [3] and CMS [4] experiments between 2015 and 2018 at the CERN LHC, corresponding to about $140\,\text{fb}^{-1}$ of integrated luminosity each, less than 2000 $\mathrm{t\bar{t}t\bar{t}}$ events are expected to have been produced per experiment. The measurement of the $\mathrm{t\bar{t}t\bar{t}}$ production cross section provides sensitivity to the top quark Yukawa coupling [5, 6], allows constraining two Higgs doublet models [7–10], supersymmetry [11–17], and other beyond-the-SM models [18–20], and is an important input to effective field theory interpretations [21–28].

The ATLAS and CMS experiments have presented searches for $\mathrm{t\bar{t}t\bar{t}}$ production in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ targeting events with zero to four charged leptons (electrons and muons) and additional jets [29–38]. In the latest result by the CMS Collaboration [37], combining searches in events with zero, one, two (opposite- and same-sign), and at least three charged leptons from a data set corresponding to $138\,\text{fb}^{-1}$, the observed (expected) significance of $\mathrm{t\bar{t}t\bar{t}}$ production was 4.0 (3.2) standard deviations (s.d.).

This note presents a search for $\mathrm{t\bar{t}t\bar{t}}$ production using pp collision data recorded by the CMS experiment in 2016–2018, corresponding to an integrated luminosity of $138\,\text{fb}^{-1}$ [39–41], selecting events with two same-sign, three, and four charged leptons. This search supersedes the results from Ref. [34], which used events with two same-sign or at least three leptons selected from the same data set and found 2.6 (2.7) s.d. of observed (expected) significance. Notable improvements include the use of state-of-the-art machine learning techniques in the lepton identification, in the tagging of jets originating from the hadronization of a b quark, and in the discrimination between signal and background processes. The $\mathrm{t\bar{t}t\bar{t}}$ production cross section $\sigma(\mathrm{pp} \to \mathrm{t\bar{t}t\bar{t}})$ is extracted with a profile likelihood fit to optimized distributions that provide good signal-to-background discrimination.

2 The CMS detector & event reconstruction

The CMS apparatus [4] is a multipurpose, nearly hermetic detector, designed to trigger on [42, 43] and identify electrons, muons, photons, and (charged and neutral) hadrons [44–46]. A global particle-flow (PF) algorithm [47] aims to reconstruct all individual particles in an event, combining information provided by the all-silicon inner tracker, the crystal electromagnetic calorimeter (ECAL), and the brass-scintillator hadron calorimeter (HCAL), operating inside a 3.8 T superconducting solenoid, with data from the gas-ionization muon detectors embedded in the flux-return yoke outside the solenoid. The reconstructed particles are used to build τ leptons, jets, and missing transverse momentum [48–50].

Hadronic jets are clustered from the PF objects using the anti-k_T algorithm [51, 52] with a distance parameter of 0.4. Jet energy corrections are derived from simulation studies and corrected with in situ measurements to match the energy scale in data and simulation [53]. We retain jets for further analysis if they have transverse momentum $p_\mathrm{T} > 25\,\text{GeV}$ and pseudorapidity $|\eta| < 2.4$, and are separated by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.4$ from any identified lepton, where $\Delta\eta$ and $\Delta\phi$ are the η and azimuthal angle differences between the jet and lepton directions. Jets

originating from b quarks are identified with the DEEPJET algorithm [54–56], using a working point with a selection efficiency for b quark jets of more than 90%, and a misidentification rate for c quark jets (light quark and gluon jets) of 50 (20)%. Jets selected with the DEEPJET algorithm are referred to as "b jets". The missing transverse momentum vector $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is computed as the negative vector sum of the transverse momenta of all the PF candidates in an event, and its magnitude is denoted as $p_{\mathrm{T}}^{\mathrm{miss}}$ [50].

The primary vertex (PV) is taken to be the vertex corresponding to the hardest scattering in the event, as described in Section 9.4.1 of Ref. [57]. Electrons are reconstructed via a combination of tracker and ECAL measurements [44, 58] in the range $|\eta| < 2.5$, and electrons in the barrel–endcap transition region $1.44 < |\eta| < 1.57$ are removed. Muons are reconstructed in the range $|\eta| < 2.4$ by combining information from the tracker, the muon spectrometers, and the calorimeters in a global fit [45]. We select electrons and muons with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and require that they are compatible with originating from the PV and fulfill a loose set of identification (ID) criteria [44, 45]. The relative isolation I_{rel} of a lepton is defined as the scalar p_{T} sum of all particles within a certain distance ΔR around the lepton, divided by the lepton p_{T}, with $\Delta R < 0.2$ for leptons with $p_{\mathrm{T}} < 50\,\mathrm{GeV}$, $\Delta R < 1\,\mathrm{GeV}/p_{\mathrm{T}}$ for $50 < p_{\mathrm{T}} < 200\,\mathrm{GeV}$, and $\Delta R < 0.05$ for $p_{\mathrm{T}} > 200\,\mathrm{GeV}$ [59]. We require all electrons and muons to have $I_{\mathrm{rel}} < 0.4$.

To distinguish leptons originating directly from the prompt decay of top quarks or massive bosons (prompt leptons) from both genuine leptons produced in hadron decays and photon conversions or jet constituents misidentified as leptons (nonprompt leptons), we define two additional sets of ID criteria labeled "loose" and "tight". For the tight ID, we train gradient boosted decision trees (BDTs) with the XGBOOST program [60] to distinguish between prompt and nonprompt leptons, following the methods developed for various CMS measurements and searches with multilepton signatures [61–66]. The input variables for the BDT are kinematic properties, isolation variables, and impact parameter of the leptons, properties of the closest jet, and additional standard ID variables specific for electrons [44] and muons [45]. The full list of input variables is given in Table 1. Tight electrons and muons are required to be above a threshold of the BDT discriminant, and the efficiency of these requirements are compared in Fig. 1. The loose ID is defined by requiring electrons or muons to either pass the tight ID, or selection requirements on some properties of the closest jet that are also used as inputs to the discriminant.

For electrons, a nonneglibile background contribution arises from charge mismeasurements. The sign of this charge is evaluated with three different observables that measure the electron curvature using different methods. We require all three charge evaluations to agree for each selected electron, which reduces this background contribution by a factor of five with an efficiency of about 97% [67]. The charge mismeasurement of muons is negligible [68, 69].

3 Simulated event samples

Events are simulated for signal and background processes with Monte Carlo generators to determine the $\mathrm{t\bar{t}t\bar{t}}$ signal acceptance, to estimate most background contributions, and to provide training data for machine learning discriminants. The $\mathrm{t\bar{t}t\bar{t}}$ signal process is simulated at NLO in QCD with the MADGRAPH5_aMC@NLO v2.6.0 program [70, 71]. Background samples for top quark pair ($\mathrm{t\bar{t}}$) production in association with a boson ($\mathrm{t\bar{t}W}$, $\mathrm{t\bar{t}Z}$, $\mathrm{t\bar{t}H}$, and $\mathrm{t\bar{t}}\gamma$), single top quark production (s channel, Z boson associated, and WZ diboson associated), Z+jets production, and WZ, $\mathrm{Z}\gamma$, WH, and ZH diboson production are simulated at NLO in QCD with MADGRAPH5_aMC@NLO as well. A leading-order (LO) simulation with MADGRAPH5_aMC@NLO is performed for $\mathrm{t\bar{t}}$ production in association with two bosons, triple top quark production, single

Table 1: List of the input variables to the prompt-lepton ID BDTs. The nearest jet $\mathrm{j}_{\mathrm{near}}$ is defined as the jet that includes the PF particle corresponding to the reconstructed lepton, and its momentum is recalibrated after subtracting the contribution from the lepton. The last two rows list input variables only used in the electron or muon ID BDTs, respectively, and are defined in Refs. [44, 45].

Symbol	Definition
p_{T}	Lepton p_{T}
$\lvert\eta\rvert$	Absolute value of the lepton η
$I_{\mathrm{rel}}^{\mathrm{fixed}}$	Relative isolation using a fixed distance $\Delta R < 0.4$
$I_{\mathrm{rel}}^{\mathrm{ch}}$	Relative isolation using a p_{T}-dependent distance and including only charged particles
$I_{\mathrm{rel}}^{\mathrm{neu}}$	Relative isolation using a p_{T}-dependent distance and including only neutral particles
$N_{\mathrm{ch}}(\mathrm{j}_{\mathrm{near}})$	Number of charged particles associated with the jet
$p_{\mathrm{T}}^{\mathrm{ratio}}$	Ratio of the lepton p_{T} to the nearest jet p_{T}, $p_{\mathrm{T}}(\ell)/p_{\mathrm{T}}(\mathrm{j}_{\mathrm{near}})$, or $1/(1+I_{\mathrm{rel}}^{\mathrm{fixed}})$ if no nearest jet is found
$p_{\mathrm{T}}^{\mathrm{rel}}$	Component of the lepton momentum in direction transverse to the nearest jet, $p(\ell)\sin\theta\big(\vec{p}(\ell),\vec{p}(\mathrm{j}_{\mathrm{near}})\big)$
$\mathrm{DJ}(\mathrm{j}_{\mathrm{near}})$	DEEPJET score of the nearest jet
$\log\lvert d_{xy}\rvert$	Distance of closest approach of the lepton track to the PV in the transverse plane
$\log\lvert d_z\rvert$	Distance of closest approach of the lepton track to the PV in the longitudinal plane
$d/\delta d$	Significance of the distance of closest approach of the lepton track to the PV
$P_{\mathrm{ID}}^{\mathrm{e}}$	Electron ID discriminant
P_{seg}^{μ}	Muon segment compatibility

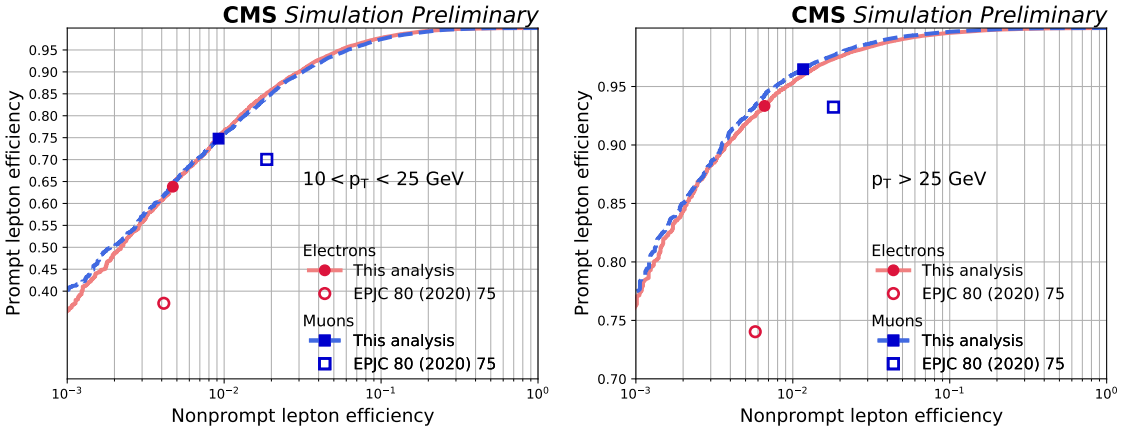

Figure 1: Efficiency of selecting prompt leptons as a function of the misidentification probability for nonprompt leptons evaluated in simulated $\mathrm{t\bar{t}}$ events for the electron (red solid line) and muon (blue dashed line) ID BDT, shown for leptons with $10 < p_{\mathrm{T}} < 25\,\mathrm{GeV}$ (left) and $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ (right). Indicated with filled markers are the efficiencies for the ID criteria applied in this measurement and with empty markers for the ID criteria applied in Ref. [34], where red circles and blue squares are used for electron and muon criteria, respectively.

top quark production (Higgs boson associated), W+jets production, and same-sign WW diboson production. The POWHEG v2 program [72–80] is used for the simulation of $t\bar{t}$ production, single top quark production (t channel and W boson associated), $q\bar{q}$-initiated ZZ and opposite-sign WW diboson production, and Higgs boson production through vector boson and gluon fusion at NLO in QCD. The gg-initiated ZZ diboson production process is simulated with the MCFM v7.0.1 generator [81–83] at LO in QCD.

The NNPDF3.1 [84] parton distribution functions (PDFs) are used in the matrix-element calculation of all processes. The simulation of parton showering, hadronization, and underlying event is performed with the PYTHIA v8.230 program [85], using the CP5 tune [86]. Double counting of partons generated with MADGRAPH5_aMC@NLO and PYTHIA is eliminated with the FxFx [87] (MLM [88]) matching scheme for NLO (LO) samples. In the POWHEG samples of Higgs boson production, the decay to four leptons is simulated with the JHUGEN v5.2.5 program [89]. Additional pp collisions in the same or nearby bunch crossings (pileup) are overlayed to each simulated event, and the generated distribution of the number of events per bunch crossing is matched to that observed in data. All simulated events are processed with a full simulation of the CMS detector based on the GEANT4 toolkit [90], and are reconstructed with the same software as used for the data.

4 Event selection & search strategy

The analyzed event sample is collected with a combination of triggers that require the presence of one, two, or three charged leptons (electrons and/or muons). To be considered in the analysis, events must contain between two and four loose leptons, with $p_T > 25$ and 20 GeV for the highest p_T (leading) and second-highest p_T (subleading) lepton, and at least two jets, of which at least one is a b jet. Events with two (more than two) leptons are removed if any lepton pair has an invariant mass below 20 (12) GeV, to reduce background events with leptons from the decay of low-mass resonances. Signal and control regions are defined for the events where all leptons also pass the tight ID criteria, whereas events with at least one loose-not-tight lepton are used as a sideband for the nonprompt-lepton background estimation. In events with two leptons, the signal and control regions additionally require that both leptons have the same charge, and events with oppositely charged leptons are used as a sideband for the estimation of backgrounds with charge-misidentified leptons. In events with four leptons, we additionally require that the sum of the lepton charges is zero. For the definition of the signal regions, the number of jets and b jets (N_j and N_b), the scalar p_T sum of all jets (H_T), and the invariant mass $m(\ell\ell)$ of opposite-sign same-flavor (OSSF) lepton pairs are used. A schematic representation of the signal and control region definitions is shown in Fig. 2.

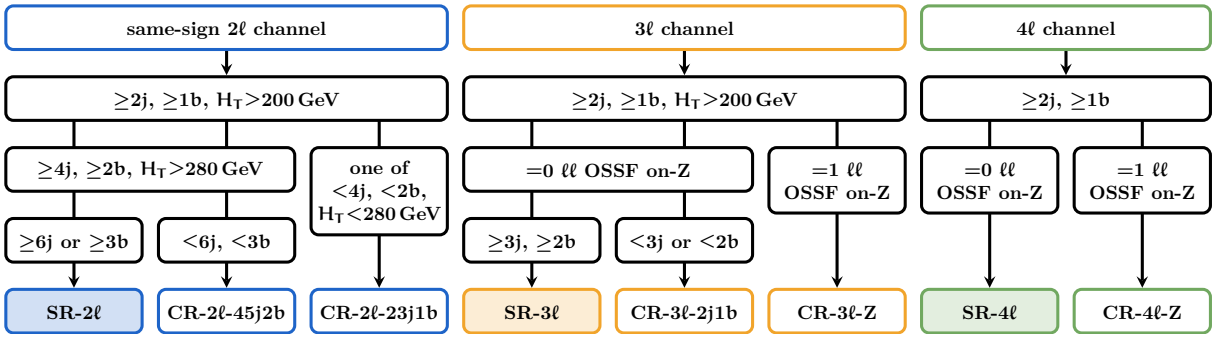

Figure 2: Schematic representation of the event selection and categorization.

For the events with two same-sign leptons (2ℓ channel), the signal region SR-2ℓ is defined by $N_j \geq 4$, $N_b \geq 2$, and $H_T > 280\,\text{GeV}$, which additionally have either $N_j \geq 6$ or $N_b \geq 3$. The control region CR-2ℓ-45j2b, which is enriched in $t\bar{t}W$ production, comprises all events with $N_j \geq 4$, $N_b \geq 2$, and $H_T > 280\,\text{GeV}$ that do not pass the tighter SR-2ℓ requirements. The control region CR-2ℓ-23j1b, which is used to constrain both $t\bar{t}W$ production and nonprompt-lepton backgrounds, is defined by all events with $H_T > 200\,\text{GeV}$ that fail exactly one of the requirements $N_j \geq 4$, $N_b \geq 2$, or $H_T > 280\,\text{GeV}$.

The signal region SR-3ℓ for events with three leptons (3ℓ channel) is defined by $N_j \geq 3$, $N_b \geq 2$, $H_T > 200\,\text{GeV}$, and the requirement that there is no OSSF lepton pair with $|m(\ell\ell) - m_Z| < 15\,\text{GeV}$. The last requirement, which uses the world-average Z boson mass [91], rejects events consistent with leptonically decaying Z bosons. The control region CR-3ℓ-2j1b, used to constrain nonprompt-lepton backgrounds, consists of events that pass the SR-3ℓ requirements except that they have $N_j = 2$ or $N_b = 1$. To constrain $t\bar{t}Z$ production, the control region CR-3ℓ-Z is defined by requiring $H_T > 200\,\text{GeV}$ and the presence of an OSSF lepton pair with $|m(\ell\ell) - m_Z| < 15\,\text{GeV}$.

No additional jet requirements are imposed on events with four leptons (4ℓ channel). Events that have no OSSF lepton pair with $|m(\ell\ell) - m_Z| < 15\,\text{GeV}$ form the signal region SR-4ℓ, while all events that have exactly one such pair form the control region CR-4ℓ-Z enriched in $t\bar{t}Z$ production.

To enhance the separation of signal events and those from different background processes in the signal regions, we employ two multiclassification BDTs trained with the TMVA program [92], where one BDT is trained for the 2ℓ channel and the other for the combined 3ℓ+4ℓ channel. The BDTs provide output scores for three classes of events (labeled "nodes"): the $t\bar{t}t\bar{t}$ signal, $t\bar{t}$ production in association with a heavy boson ($t\bar{t}Z$, $t\bar{t}W$, and $t\bar{t}H$, referred to as $t\bar{t}X$), and backgrounds with nonprompt and charge-misidentified leptons (referred to as $t\bar{t}$). The output scores can be interpreted as measures of how likely an event originates from the corresponding classes.

The variables used in the BDT training are related to the kinematic properties of the jets, b jets, and leptons, the jet and b jet multiplicities, the DEEPJET scores of the jets, the separations between leptons and/or jets, H_T, p_T^{miss}, and invariant masses and transverse masses [93, 94] of combinations of leptons and/or jets consistent with top quark decays. The full list of input variables is shown in Table 2. Simulated event samples of $t\bar{t}t\bar{t}$, $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, and $t\bar{t}$ production are used in the training, with 20% of events randomly excluded and used for testing. The training and testing samples are not used in the other parts of the analysis. The modeling of each input variable in simulation was validated in data, and good agreement between simulation and data was found for the shapes of the output score distributions in the control regions.

Events selected in the signal regions are further split according to the node that yields the highest BDT output score, resulting in "$t\bar{t}t\bar{t}$-like", "$t\bar{t}X$-like", and "$t\bar{t}$-like" signal region classes. In the cross section extraction, the signal region classes for SR-2ℓ and SR-3ℓ are fitted separately. Due to the limited number of events expected in the SR-4ℓ, only the $t\bar{t}t\bar{t}$ and $t\bar{t}X$ classes are fitted separately, and the events that have the highest output score for the $t\bar{t}$ class are added to the $t\bar{t}X$ class.

The approach followed here is different from the BDT approach of Ref. [34] in several aspects. We train separate BDTs for the 2ℓ and 3ℓ+4ℓ channels, with an extended set of input variables that have been separately optimized for the two channels. The performance also profits from the looser requirements in the initial event selection, allowing the BDTs to improve with re-

Table 2: List of the input variables to the event multiclassification BDTs, sorted by importance in the 2ℓ BDT. Observables not (only) used in the 3ℓ+4ℓ BDT are marked with † (‡). The m_{T2} variable is defined in Ref. [93, 94] and constructed from $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ and two four-momenta of particle (systems) specified in the table.

Symbol		Definition
$\max_2 \mathrm{DJ}$		Second-highest DEEPJET score of any jet
$\Delta R(\ell_1, \ell_2)$		ΔR between leading and subleading lepton
$\min \Delta R(\mathrm{b}, \mathrm{b})$		Smallest ΔR between any two b jets
$\min_2 \Delta R(\ell, \mathrm{b})$		Second smallest ΔR between any lepton and b jet
$\Delta\phi(\ell_1, \ell_2)$	†	$\Delta\phi$ between leading and subleading lepton
$\min_1 \Delta R(\ell, \mathrm{b})$	†	Smallest ΔR between any lepton and b jet
$m(\mathrm{t}_1)$		Invariant mass of any three jets, of which one is a b jet, that is closest to the top quark mass
$\max_3 \mathrm{DJ}$		Third-highest DEEPJET score of any jet
$\mathrm{DJ}(\mathrm{j}_1)$		DEEPJET score of the leading jet
$m(\mathrm{W}_1)$		Invariant mass of any two jets used for $m(\mathrm{t}_1)$ that is closest to the W boson mass
$p_{\mathrm{T}}(\mathrm{j}_4)$	†	Fourth-highest p_{T} of any jet
$\mathrm{DJ}(\mathrm{j}_2)$		DEEPJET score of the subleading jet
$p_{\mathrm{T}}(\mathrm{j}_5)$		Fifth-highest p_{T} of any jet
H_{T}		Scalar sum of p_{T} of all jets
$p_{\mathrm{T}}(\ell_2)$	†	Second-highest p_{T} of any lepton
$\mathrm{DJ}(\mathrm{j}_3)$	†	DEEPJET score of the jet with the third-highest p_{T}
$m_{\mathrm{T}}(\ell_1)$		Transverse mass of the leading lepton and $p_{\mathrm{T}}^{\mathrm{miss}}$
$p_{\mathrm{T}}(\mathrm{j}_1)$		Highest p_{T} of any jet
$p_{\mathrm{T}}(\ell_1)$		Highest p_{T} of any lepton
$p_{\mathrm{T}}^{\mathrm{miss}}$		Missing transverse momentum
$m_{\mathrm{T}}(\ell_2)$	†	Transverse mass of the subleading lepton and $p_{\mathrm{T}}^{\mathrm{miss}}$
$p_{\mathrm{T}}(\mathrm{j}_2)$	†	Second-highest p_{T} of any jet
$m(\mathrm{t}_2)$	†	Invariant mass of any three jets, of which one is a b jet and none of which is used for $m(\mathrm{t}_1)$, that is closest to the top quark mass
N_{j}		Number of jets
$m_{\mathrm{T2}}(\mathrm{b})$	†	m_{T2} variable constructed from the leading and subleading b jet
$m_{\mathrm{T2}}(\ell{+}\mathrm{b})$	†	m_{T2} variable constructed from two lepton+jet systems built with the leading two leptons and the leading two b jets
$N_{\mathrm{b}}^{\mathrm{tight}}$		Number of jets passing the "tight" DEEPJET working point (tighter than in the event selection)
$m(\mathrm{W}_2)$	†	Invariant mass of any two jets used for $m(\mathrm{t}_2)$ that is closest to the W boson mass
$\max_4 \mathrm{DJ}$	†	Fourth-highest DEEPJET score of any jet
$N_{\mathrm{b}}^{\mathrm{medium}}$	†	Number of jets passing the "medium" DEEPJET working point (tighter than in the event selection, but looser than "tight")
$p_{\mathrm{T}}(\ell_3)$	‡	Third-highest p_{T} of any lepton
$p_{\mathrm{T}}(\mathrm{j}_3)$	‡	Third-highest p_{T} of any jet
$m_{\mathrm{T2}}(\ell)$	‡	m_{T2} variable constructed from the leading and subleading lepton
$\mathrm{DJ}(\mathrm{j}_4)$	‡	DEEPJET score of the jet with the fourth-highest p_{T}

spect to the sensitivity achieved by a series of tight one-dimensional selection requirements. Finally, the setup as a multiclassifier that simultaneously separates two characteristically different background classes facilitates a good separation of processes, and the use of all output nodes in the fit improves the constraints on the main background contributions.

5 Background estimation

Background contributions to the signal regions arise from various processes, and are categorized into processes where all selected leptons are prompt (prompt background) and processes with at least one nonprompt lepton (nonprompt background). In the 2ℓ channel, we additionally separate prompt-background contributions where one electron is reconstructed with the wrong charge (labeled "charge misID"). Prompt backgrounds are estimated from simulation, while nonprompt and charge-misID backgrounds are estimated from sideband samples in data. All predictions from simulation are normalized according to the measured integrated luminosity and state-of-the-art cross section predictions.

The dominant prompt background contributions are from $t\bar{t}X$ production: $t\bar{t}W$ in the 2ℓ and 3ℓ channels, $t\bar{t}Z$ in the 3ℓ and 4ℓ channels with smaller contributions in the 2ℓ channel, and $t\bar{t}H$ in all channels. For the $t\bar{t}Z$ and $t\bar{t}H$ background predictions, cross sections of 859 ± 80 fb and 504 ± 39 fb as provided by Ref. [95] are used, respectively, for the normalization. The $t\bar{t}W$ prediction is normalized with a cross section of 722 ± 74 fb [96]. The quoted uncertainties are from scale variations and PDF uncertainties. Other contributions from top quark production processes are small, with the largest remaining contribution being single top quark production in association with a Z boson, normalized with a cross section of 94.2 ± 3.1 fb [97]. These processes are grouped together and labeled as "Other t".

The production of two or three heavy vector bosons contributes mainly via WZ production in the 2ℓ and 3ℓ channels, and ZZ production in the 4ℓ channel. Due to the low overall event yields, we group all contributions together as "VV(V)". Contributions with photon conversions, both in processes with top quarks and only vector bosons, are grouped together as "Xγ". The contributions from the dominant prompt background processes are treated separately, and get constrained effectively by the inclusion of the signal region $t\bar{t}X$ classes and the control regions with a reconstructed Z candidate (CR-3ℓ-Z and CR-4ℓ-Z) in the cross section extraction.

The modeling of additional jets in the predicted WZ and ZZ samples is validated in data. A sample of WZ candidate events in data is selected by requiring three leptons, one OSSF lepton pair with $|m(\ell\ell) - m_{\mathrm{Z}}| < 15\,\mathrm{GeV}$, a three-lepton invariant mass $m(\ell\ell\ell)$ with $|m(\ell\ell\ell) - m_{\mathrm{Z}}| > 15\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$, and no b jets. More than 70% of the selected events originate from WZ production, with smaller contributions from nonprompt leptons and ZZ production. A sample of ZZ candidate events in data is selected by requiring four leptons that form two OSSF lepton pairs with $|m(\ell\ell) - m_{\mathrm{Z}}| < 15\,\mathrm{GeV}$, and contributions from other processes to this selection are found to be negligible. A disagreement is observed in the jet multiplicity distribution in both data samples, and scale factors are derived per N_{j} bin that increase the contribution of WZ (ZZ) production at higher jet multiplicities by factors of up to 2.7 (1.8).

Nonprompt-background contributions are estimated with a "tight-to-loose" ratio method [98]. The probability for a loose lepton to also pass the tight ID is measured, separately for electrons and muons, as a function of p_{T} and $|\eta|$ in a data sample enriched in events composed uniquely of jets produced through the strong interaction, which is rich in nonprompt leptons. The measured probability is then applied to data events in a sideband of the signal and control regions where one or more of the selected leptons fail the tight selection while passing

the loose selection. The method is validated in simulation both with $\mathrm{t\bar{t}}$ and Z+jets production events, separately for nonprompt electrons and muons, as well as in comparison of data and predicted background yields in the control regions with lower hadronic activity (CR-2ℓ-23j1b, CR-2ℓ-45j2b, and CR-3ℓ-2j1b), with an agreement of better than 30% in the most relevant kinematic distributions.

To determine the charge-misID background, the electron charge-misID probability is determined in simulated samples of Z+jets production and parameterized as a function of p_{T} and $|\eta|$. This probability is applied as a scale factor to events with two opposite-sign electrons in the data sideband to obtain a prediction of the charge-misID background in the SR-2ℓ. To evaluate possible discrepancies between the charge-misID probability in simulation and data, two data samples are selected with exactly two tight same-sign or opposite-sign electrons with $|m(\mathrm{ee} - m_{\mathrm{Z}})| < 15\,\mathrm{GeV}$, and the difference between the same-sign event yields and the scaled opposite-sign event yields is used to derive correction factors to the integral normalization, separately for each year of data taking.

6 Systematic uncertainties

Several sources of systematic uncertainty affect the prediction of the signal acceptance and reconstruction efficiency, the background event yields, and the distributions of the observables used for the signal extraction. Experimental uncertainties in the integrated luminosity, lepton selection efficiency, trigger efficiency, jet energy scale, b tagging efficiency, and nonprompt background estimation are considered with several sources each, of which some are correlated and uncorrelated between the years of data-taking. Only the experimental uncertainties in the jet energy resolution and in unclustered $p_{\mathrm{T}}^{\mathrm{miss}}$ contributions are fully uncorrelated between the data-taking years. Uncertainties in the simulated pileup distribution, background normalization, and the shapes of the fitted observables estimated from simulation are fully correlated between the data-taking years.

The integrated luminosities for the three data-taking years have uncertainties between 1.2 and 2.5% [39–41], while the overall uncertainty for the combined data set is 1.6%, affecting both the background predictions from simulation and the measurement of the $\mathrm{t\bar{t}t\bar{t}}$ cross section. The uncertainty in the distribution of pileup events in simulation is estimated by varying the total pp inelastic cross section by $\pm$4.6%.

The efficiency of the tight ID selection of electrons and muons is measured in data and simulation with a "tag-and-probe" method in $\mathrm{Z} \to \ell^{+}\ell^{-}$ events [99]. Per-lepton correction factors are derived and applied to simulated events, and statistical and systematic uncertainties are considered. The trigger efficiency is measured by selecting events with two or three leptons in an unbiased data sample, collected with triggers on $p_{\mathrm{T}}^{\mathrm{miss}}$ or hadronic activity, and corrections are derived to match the efficiency in simulation to that observed in data. A systematic uncertainty of 3% is assigned to simulated event yields to cover residual differences, and additionally the statistical uncertainties in the measurement are considered. Special correction factors are applied for a gradual shift in the timing of the ECAL inputs of the first-level trigger, which caused a specific trigger inefficiency during the 2016 and 2017 data-taking periods [42].

Uncertainties in the jet energy scale and resolution are applied by p_{T} variations of the reconstructed jets in simulated events, considering various uncertainty sources split between detector regions and data-taking years [53]. The variations in the jet energy scale, as well as an additional variation to account for uncertainty contributions from unclustered PF particles [50], are also propagated to $p_{\mathrm{T}}^{\mathrm{miss}}$. Corrections are applied to simulated events to match the shape

of the DEEPJET discriminant in simulation to that in data, and uncertainties in the correction factors are evaluated separately for heavy- and light-flavor jets [54].

Uncertainties in the theoretical cross section predictions of the background contributions estimated from simulation are taken into account. For the $\mathrm{t\bar{t}H}$ background contribution, the uncertainty in the SM prediction [95] is considered. The normalization of the $\mathrm{t\bar{t}Z}$ and $\mathrm{t\bar{t}W}$ background contributions are not constrained a priori but treated as free parameters in the fit, and get constrained by the data in the control regions as well as in the signal region $\mathrm{t\bar{t}X}$ classes. A normalization uncertainty of 6.2% is applied for the WZ background contribution, based on the precision of the measured cross section [100], and additionally the statistical uncertainties in the N_j scale factors are propagated. For the other background classes, normalization uncertainties are assigned based on the experimental precision of the cross section measurement of the dominant contributions, resulting in 20% for the Other t group [66], 6% for the $\mathrm{VV(V)}$ group [101, 102], and 5% for the $\mathrm{X}\gamma$ group [103].

For the nonprompt-background prediction, three sources of uncertainty are considered: the statistical uncertainty in the measured tight-to-loose ratio, the statistical uncertainty of the event yields in the sideband regions, and an overall normalization uncertainty. The normalization uncertainty is applied separately for nonprompt electrons and nonprompt muons, and split into two components of 20% each that are either fully correlated or uncorrelated between the data-taking years, to account both for possible systematic effects in the method itself and for potential year-to-year differences. For the charge-misID background, the statistical uncertainty of the event yields in the sideband regions and a normalization uncertainty of 15% are considered.

Various uncertainty sources in the modeling of the simulated signal and background processes are considered. The choice of the renormalization and factorization scales in the matrix-element calculation are assessed by individual and simultaneous variations of these scales up and down by a factor of 2 (excluding the two extreme variations), and evaluating the envelope of the six obtained variations in the fitted distributions separately for each process. The limited knowledge of the proton PDFs is taken into account using NNPDF replicas [84] and applying the procedure described in Ref. [104], resulting in 100 individual variations that are evaluated simultaneously for all processes. For the scale choice of the initial-state (final-state) radiation in the parton shower simulation, individual variations by a factor of 2 up and down are considered, separately (simultaneously) for the considered processes. In all cases, any impact on the normalization of each process before any selection is removed, and only the acceptance change and the modification to the shapes of the fitted distributions are taken into account.

Additional uncertainties are considered to account for a possible mismodeling in the number of additional jets and b jets in $\mathrm{t\bar{t}X}$ production processes. Following dedicated studies of the cross section ratio of $\mathrm{t\bar{t}}$ production in association with additional b jets or additional light jets [105, 106], we assign a conservative uncertainty of 40% to the contribution of $\mathrm{t\bar{t}W}$, $\mathrm{t\bar{t}Z}$, and $\mathrm{t\bar{t}H}$ production with at least one additional b jet at particle level. For $\mathrm{t\bar{t}W}$ production, an additional uncertainty is introduced corresponding to a reweighting of the number of reconstructed jets distribution according to the theory evaluation from Ref. [96], resulting in a variation of up to 55% in 2ℓ events with seven or more jets.

7 Results

The measured cross section of the $\mathrm{t\bar{t}t\bar{t}}$ signal process is extracted from a simultaneous binned profile likelihood fit to the data in the signal and control regions, as described in Sec. 3.2 of

Ref. [107]. All sources of systematic uncertainty are taken into account as nuisance parameters in the fit, using appropriate correlations between the data-taking years and the processes. Statistical uncertainties in the predicted yields per bin are taken into account through a single nuisance parameter for all processes [108, 109]. The fitted distributions are the BDT output scores in the signal region classes, the BDT output score for the $\mathrm{t\bar{t}}$ class in the CR-2ℓ-23j1b and CR-2ℓ-45j2b, the jet multiplicity in the CR-3ℓ-Z and CR-4ℓ-Z, and the event yields with positive or negative sum of lepton charges of the CR-3ℓ-2j1b. In the SR-2ℓ, the $\mathrm{t\bar{t}t\bar{t}}$ and $\mathrm{t\bar{t}X}$ classes are also split according to the lepton flavors, to achieve a better separation of the nonprompt, charge-misID, and Xγ backgrounds.

The fitted distributions in the signal and control regions are shown in Figs. 3–7. A summary of the signal purity in the distributions in the $\mathrm{t\bar{t}t\bar{t}}$ classes of the signal regions is shown in Fig. 8.

The observed (expected) statistical significance of the $\mathrm{t\bar{t}t\bar{t}}$ signal is determined using the asymptotic approximation of the distribution of the profile likelihood test statistic [110, 111] and found to be 5.5 (4.9) s.d. from the background-only hypothesis. The results when fitting each channel on its own are summarized in Table 3, with the largest sensitivity provided by the 2ℓ channel. This constitutes the first observation of $\mathrm{t\bar{t}t\bar{t}}$ production.

The cross sections of $\mathrm{t\bar{t}t\bar{t}}$, $\mathrm{t\bar{t}W}$, and $\mathrm{t\bar{t}Z}$ production are measured to be

$$\begin{aligned}\sigma(\mathrm{pp} \to \mathrm{t\bar{t}t\bar{t}}) &= 17.9\,^{+3.7}_{-3.5}\,(\mathrm{stat})\,^{+2.4}_{-2.1}\,(\mathrm{syst})\,\mathrm{fb},\\ \sigma(\mathrm{pp} \to \mathrm{t\bar{t}W}) &= 997 \pm 58\,(\mathrm{stat})\,^{+79}_{-72}\,(\mathrm{syst})\,\mathrm{fb},\\ \sigma(\mathrm{pp} \to \mathrm{t\bar{t}Z}) &= 1134\,^{+52}_{-43}\,(\mathrm{stat}) \pm 86\,(\mathrm{syst})\,\mathrm{fb}.\end{aligned}$$

For $\mathrm{t\bar{t}t\bar{t}}$ production, the measured cross section is in agreement with the SM prediction of $13.4\,^{+1.0}_{-1.8}$ fb [1] at the level of 1.1 s.d., when taking uncertainties of both prediction and measurement into account. The measured $\mathrm{t\bar{t}W}$ and $\mathrm{t\bar{t}Z}$ cross sections are larger than the SM predictions [95, 96] with an agreement at the level of 2.3 and 2.2 s.d., respectively.

The measurement of $\sigma(\mathrm{pp} \to \mathrm{t\bar{t}t\bar{t}})$ is limited by the statistical uncertainty in the signal region event yields. From the systematic uncertainty sources related to modeling, the dominant contributions are due to additional jets and b jets in $\mathrm{t\bar{t}X}$ production, and the matrix-element scale variations for $\mathrm{t\bar{t}t\bar{t}}$ production. The nuisance parameter associated with additional b jets in $\mathrm{t\bar{t}X}$ production is increased in the fit by about 40%, resulting in an increase of the $\mathrm{t\bar{t}X}$ yield in the signal regions of about 16%. The additional nuisance parameter associated with additional jets in $\mathrm{t\bar{t}W}$ production is not increased in the fit. The largest experimental contributions to the systematic uncertainty arise from the DEEPJET correction factors and the jet energy calibration.

Table 3: Comparison of fit results in the channels individually and in their combination.

Channel	Obs. (exp.) significance	$\sigma(\mathrm{pp} \to \mathrm{t\bar{t}t\bar{t}})$
2ℓ	4.1 (4.1) s.d.	$17.6\,^{+4.7}_{-4.3}\,(\mathrm{stat})\,^{+2.8}_{-2.7}\,(\mathrm{syst})$ fb
3ℓ	3.5 (3.0) s.d.	$19.4\,^{+7.1}_{-6.4}\,(\mathrm{stat})\,^{+2.9}_{-2.3}\,(\mathrm{syst})$ fb
4ℓ	0.0 (0.8) s.d.	—
Combined	5.5 (4.9) s.d.	$17.9\,^{+3.7}_{-3.5}\,(\mathrm{stat})\,^{+2.4}_{-2.1}\,(\mathrm{syst})$ fb

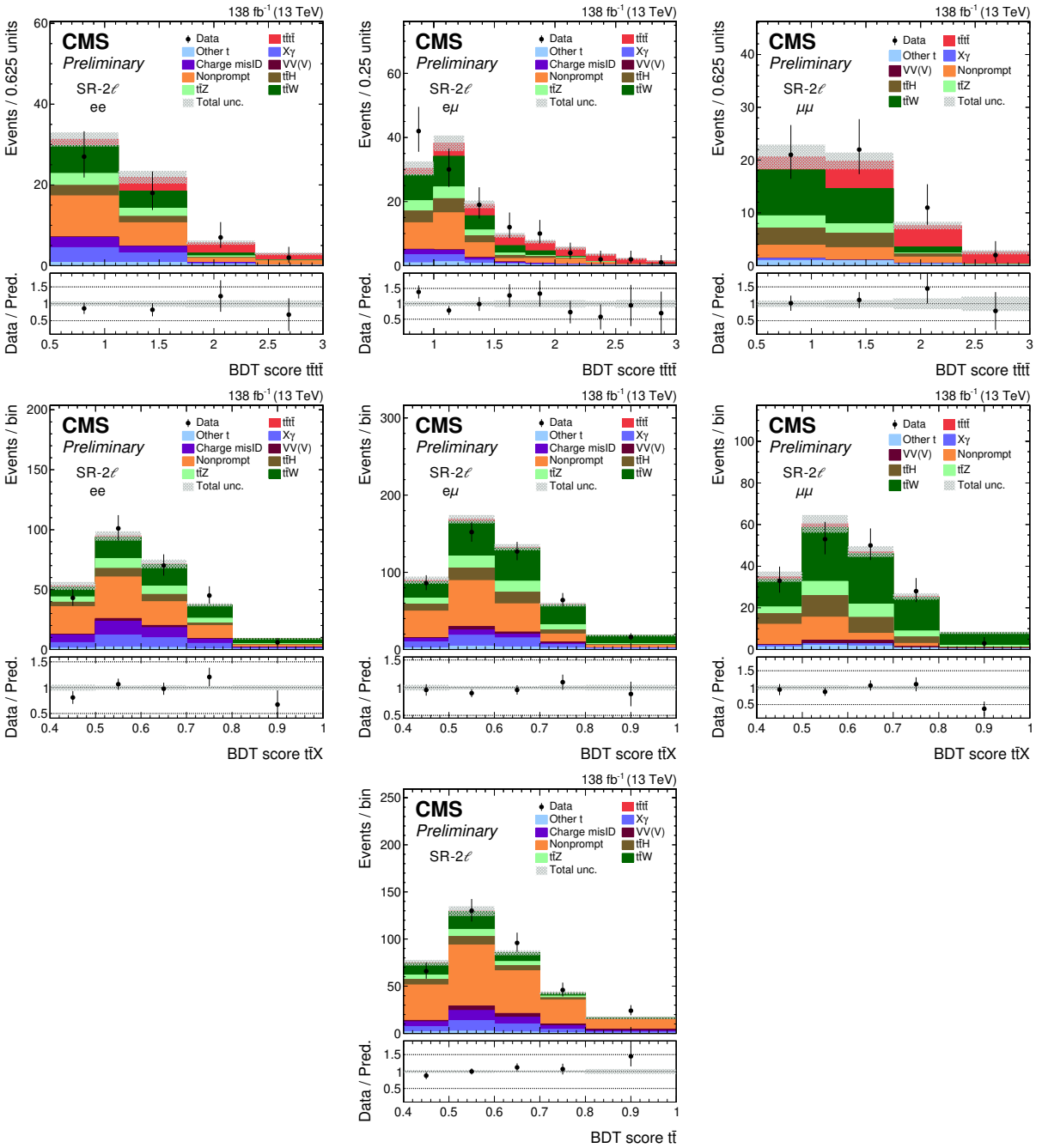

Figure 3: Comparison of the number of observed (points) and predicted (colored histograms) events in the BDT output distributions for the $t\bar{t}t\bar{t}$ signal in the $t\bar{t}t\bar{t}$ class (upper row), for the $t\bar{t}X$ background in the $t\bar{t}X$ class (middle row), and for the $t\bar{t}$ background in the $t\bar{t}$ class (lower row) of SR-2ℓ. The $t\bar{t}t\bar{t}$ and $t\bar{t}X$ classes are shown separately in the ee (left), eμ (middle) and $\mu\mu$ (right) categories. The vertical bars on the points represent the statistical uncertainties in the data, and the hatched bands the systematic uncertainty in the predictions. The predictions are shown "postfit", i.e., with the values of the signal and background normalizations and nuisance parameters obtained in the fit to the data applied.

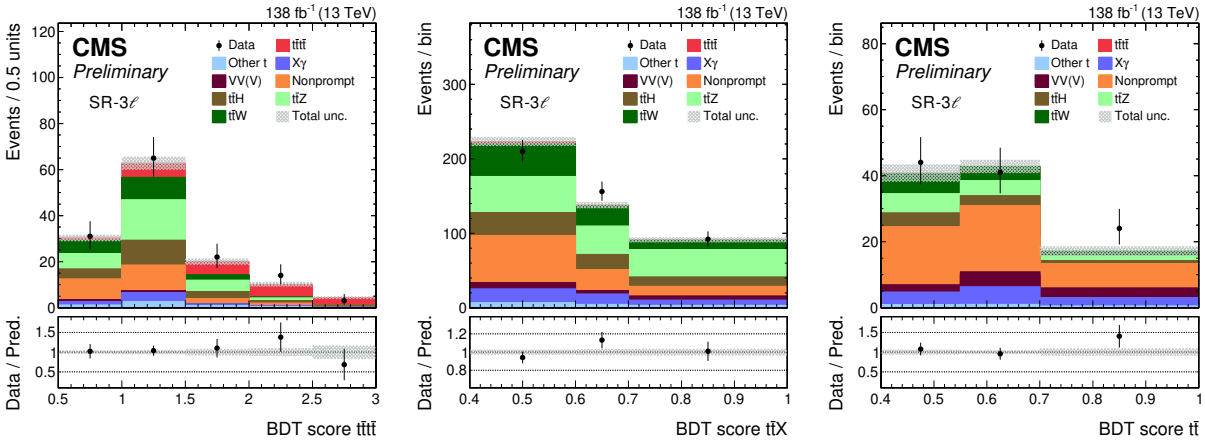

Figure 4: Comparison of the number of observed (points) and predicted (colored histograms) events in the BDT output distributions for the $t\bar{t}t\bar{t}$ signal in the $t\bar{t}t\bar{t}$ class (left), for the $t\bar{t}X$ background in the $t\bar{t}X$ class (middle), and for the $t\bar{t}$ background in the $t\bar{t}$ class (right) of SR-3ℓ. The vertical bars on the points represent the statistical uncertainties in the data, and the hatched bands the systematic uncertainty in the predictions. The predictions are shown "postfit", i.e., with the values of the signal and background normalizations and nuisance parameters obtained in the fit to the data applied.

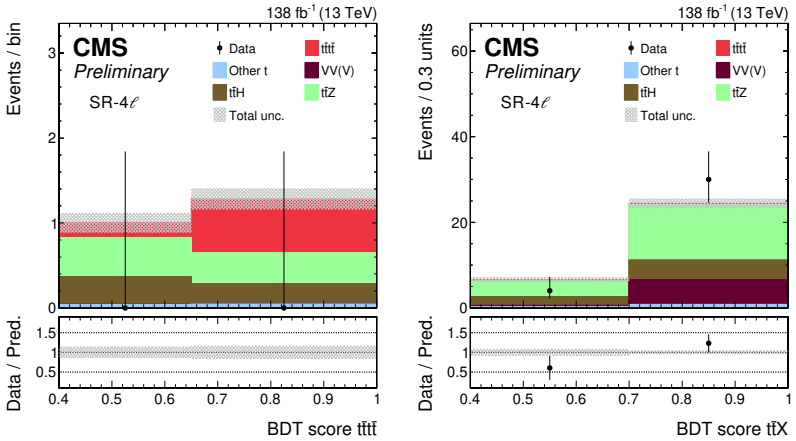

Figure 5: Comparison of the number of observed (points) and predicted (colored histograms) events in the BDT output distributions for the $t\bar{t}t\bar{t}$ signal in the $t\bar{t}t\bar{t}$ class (left) and for the $t\bar{t}X$ background in the combined $t\bar{t}X$ and $t\bar{t}$ classes (right) of SR-4ℓ. The vertical bars on the points represent the statistical uncertainties in the data, and the hatched bands the systematic uncertainty in the predictions. The predictions are shown "postfit", i.e., with the values of the signal and background normalizations and nuisance parameters obtained in the fit to the data applied. No data events are observed in the $t\bar{t}t\bar{t}$ class of SR-4ℓ.

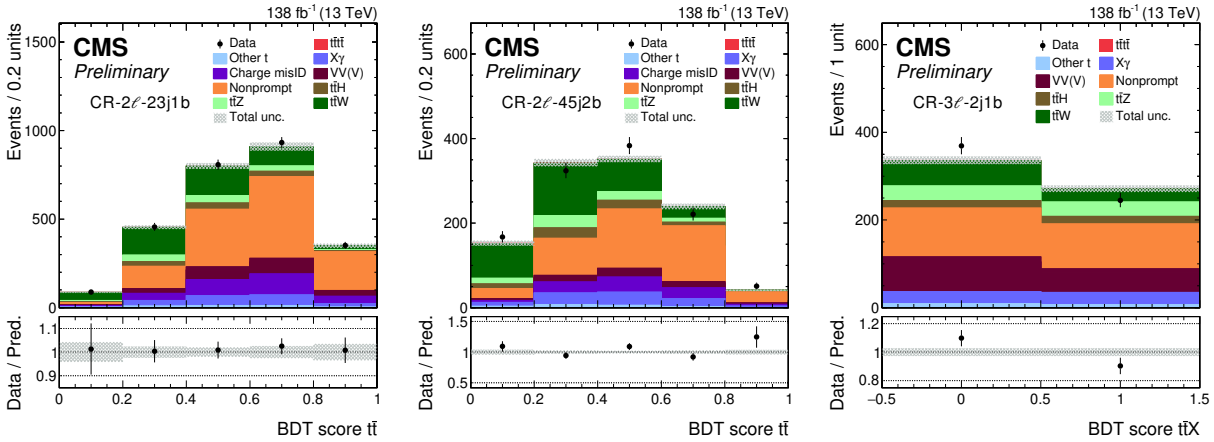

Figure 6: Comparison of the number of observed (points) and predicted (colored histograms) events in the BDT output distributions for the $t\bar{t}$ background, shown for CR-2ℓ-23j1b (left) and CR-2ℓ-45j2b (middle), and in the event yields with positive and negative sum of lepton charges in CR-3ℓ-2j1b (right). The vertical bars on the points represent the statistical uncertainties in the data, and the hatched bands the systematic uncertainty in the predictions. The predictions are shown "postfit", i.e., with the values of the signal and background normalizations and nuisance parameters obtained in the fit to the data applied.

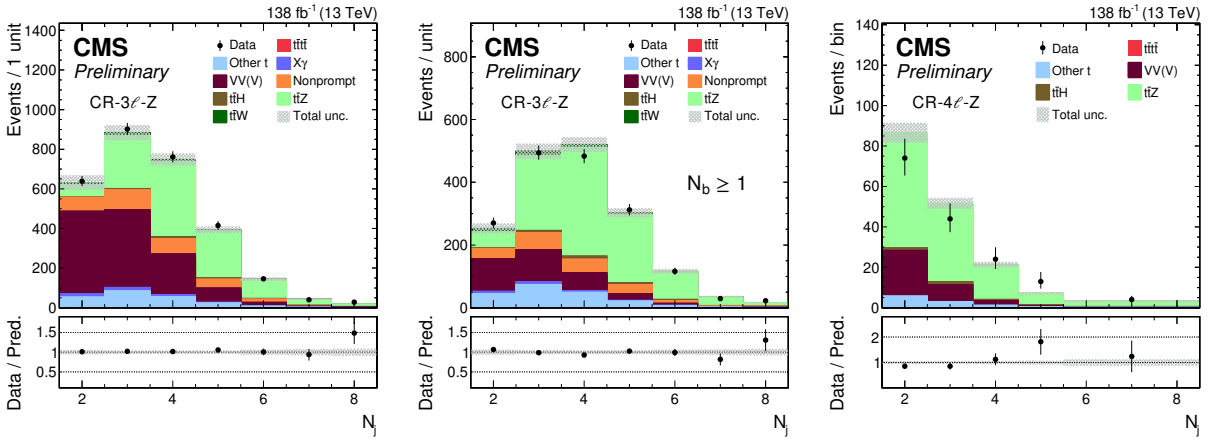

Figure 7: Comparison of the number of observed (points) and predicted (colored histograms) events in the jet multiplicity distribution, shown for CR-3ℓ-Z (left), CR-3ℓ-Z with the additional requirement of $N_b \geq 1$ (middle), and CR-4ℓ-Z (right). The vertical bars on the points represent the statistical uncertainties in the data, and the hatched bands the systematic uncertainty in the predictions. The predictions are shown "postfit", i.e., with the values of the signal and background normalizations and nuisance parameters obtained in the fit to the data applied.

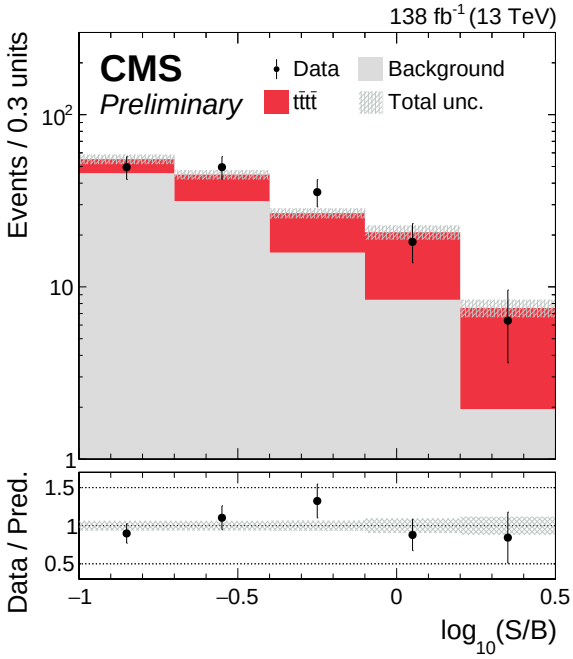

Figure 8: Comparison of the number of observed (points) and predicted (colored histograms) events as a function of $\log_{10}(S/B)$, where S and B are the predicted signal and background yields, respectively, evaluated for each bin in the $t\bar{t}t\bar{t}$ classes of the signal regions before the fit to data. Only bins with $\log_{10}(S/B) > -1$ are included, and bins with $\log_{10}(S/B) > 0.5$ are included in the last bin. The vertical bars on the points represent the statistical uncertainties in the data, and the hatched bands the systematic uncertainty in the predictions. The predictions are shown "postfit", i.e., with the values of the signal and background normalizations and nuisance parameters obtained in the fit to the data applied.

8 Summary

In summary, we have reported the first observation for the production of four top quarks ($t\bar{t}t\bar{t}$) in proton-proton (pp) collisions, using events with two same-sign, three, and four charged leptons (electrons and muons) and additional jets. The observed (expected) significance of the $t\bar{t}t\bar{t}$ signal above the background-only hypothesis is 5.5 (4.9) standard deviations. The signal cross section is measured to be $\sigma(pp \to t\bar{t}t\bar{t}) = 17.9\,^{+3.7}_{-3.5}\,(\text{stat})\,^{+2.4}_{-2.1}\,(\text{syst})$ fb, in agreement with the standard model prediction.

References

[1] R. Frederix, D. Pagani, and M. Zaro, "Large NLO corrections in $t\bar{t}W^{\pm}$ and $t\bar{t}t\bar{t}$ hadroproduction from supposedly subleading EW contributions", *JHEP* **02** (2018) 031, doi:10.1007/JHEP02(2018)031, arXiv:1711.02116.

[2] M. van Beekveld, A. Kulesza, and L. Moreno Valero, "Threshold resummation for the production of four top quarks at the LHC", 2022. arXiv:2212.03259.

[3] ATLAS Collaboration, "The ATLAS experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.

[4] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[5] Q.-H. Cao, S.-L. Chen, and Y. Liu, "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, doi:10.1103/PhysRevD.95.053004, arXiv:1602.01934.

[6] Q.-H. Cao et al., "Limiting top quark-Higgs boson interaction and Higgs-boson width from multitop productions", *Phys. Rev. D* **99** (2019) 113003, doi:10.1103/PhysRevD.99.113003, arXiv:1901.04567.

[7] D. Dicus, A. Stange, and S. Willenbrock, "Higgs decay to top quarks at hadron colliders", *Phys. Lett. B* **333** (1994) 126, doi:10.1016/0370-2693(94)91017-0, arXiv:hep-ph/9404359.

[8] N. Craig et al., "The hunt for the rest of the Higgs bosons", *JHEP* **06** (2015) 137, doi:10.1007/JHEP06(2015)137, arXiv:1504.04630.

[9] N. Craig et al., "Heavy Higgs bosons at low tan β: from the LHC to 100 TeV", *JHEP* **01** (2017) 018, doi:10.1007/JHEP01(2017)018, arXiv:1605.08744.

[10] Anisha et al., "On the BSM reach of four top production at the LHC", 2023. arXiv:2302.08281. Submitted to *JHEP*.

[11] H. Nilles, "Supersymmetry, supergravity and particle physics", *Phys. Rept.* **110** (1984) 1, doi:10.1016/0370-1573(84)90008-5.

[12] G. Farrar and P. Fayet, "Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry", *Phys. Lett. B* **76** (1978) 575, doi:10.1016/0370-2693(78)90858-4.

[13] M. Toharia and J. Wells, "Gluino decays with heavier scalar superpartners", *JHEP* **02** (2006) 015, doi:10.1088/1126-6708/2006/02/015, arXiv:hep-ph/0503175.

[14] T. Plehn and T. Tait, "Seeking sgluons", *J. Phys. G* **36** (2009) 075001, doi:10.1088/0954-3899/36/7/075001, arXiv:0810.3919.

[15] S. Calvet, B. Fuks, P. Gris, and L. Valery, "Searching for sgluons in multitop events at a center-of-mass energy of 8 TeV", *JHEP* **04** (2013) 043, doi:10.1007/JHEP04(2013)043, arXiv:1212.3360.

[16] L. Beck et al., "Probing top-philic sgluons with LHC Run I data", *Phys. Lett. B* **746** (2015) 48, doi:10.1016/j.physletb.2015.04.043, arXiv:1501.07580.

[17] L. Darmé, B. Fuks, and M. Goodsell, "Cornering sgluons with four-top-quark events", *Phys. Lett. B* **784** (2018) 223, doi:10.1016/j.physletb.2018.08.001, arXiv:1805.10835.

[18] K. Kumar, T. M. P. Tait, and R. Vega-Morales, "Manifestations of top compositeness at colliders", *JHEP* **05** (2009) 022, doi:10.1088/1126-6708/2009/05/022, arXiv:0901.3808.

[19] G. Cacciapaglia et al., "Composite scalars at the LHC: the Higgs, the sextet and the octet", *JHEP* **11** (2015) 201, doi:10.1007/JHEP11(2015)201, arXiv:1507.02283.

[20] O. Ducu, L. Heurtier, and J. Maurer, "LHC signatures of a Z' mediator between dark matter and the SU(3) sector", *JHEP* **03** (2016) 006, doi:10.1007/JHEP03(2016)006, arXiv:1509.05615.

[21] C. Degrande et al., "Non-resonant new physics in top pair production at hadron colliders", *JHEP* **03** (2011) 125, doi:10.1007/JHEP03(2011)125, arXiv:1010.6304.

[22] C. Zhang, "Constraining qqtt operators from four-top production: a case for enhanced EFT sensitivity", *Chin. Phys. C* **42** (2018) 023104, doi:10.1088/1674-1137/42/2/023104, arXiv:1708.05928.

[23] N. Hartland et al., "A Monte Carlo global analysis of the standard model effective field theory: the top quark sector", *JHEP* **04** (2019) 100, doi:10.1007/JHEP04(2019)100, arXiv:1901.05965.

[24] C. Englert, G. F. Giudice, A. Greljo, and M. Mccullough, "The $\widehat{\mathrm{H}}$-parameter: an oblique Higgs view", *JHEP* **09** (2019) 041, doi:10.1007/JHEP09(2019)041, arXiv:1903.07725.

[25] G. Banelli et al., "The present and future of four top operators", *JHEP* **02** (2021) 043, doi:10.1007/JHEP02(2021)043, arXiv:2010.05915.

[26] L. Darmé, B. Fuks, and F. Maltoni, "Top-philic heavy resonances in four-top final states and their EFT interpretation", *JHEP* **09** (2021) 143, doi:10.1007/JHEP09(2021)143, arXiv:2104.09512.

[27] SMEFiT Collaboration, "Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC", *JHEP* **11** (2021) 089, doi:10.1007/JHEP11(2021)089, arXiv:2105.00006.

[28] R. Aoude, H. El Faham, F. Maltoni, and E. Vryonidou, "Complete SMEFT predictions for four top quark production at hadron colliders", *JHEP* **10** (2022) 163, doi:10.1007/JHEP10(2022)163, arXiv:2208.04962.

[29] CMS Collaboration, "Search for physics beyond the standard model in events with two leptons of same sign, missing transverse momentum, and jets in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Eur. Phys. J. C* **77** (2017) 578, doi:10.1140/epjc/s10052-017-5079-z, arXiv:1704.07323.

[30] CMS Collaboration, "Search for standard model production of four top quarks with same-sign and multilepton final states in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Eur. Phys. J. C* **78** (2018) 140, doi:10.1140/epjc/s10052-018-5607-5, arXiv:1710.10614.

[31] ATLAS Collaboration, "Search for new phenomena in events with same-charge leptons and b jets in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector", *JHEP* **12** (2018) 039, doi:10.1007/JHEP12(2018)039, arXiv:1807.11883.

[32] ATLAS Collaboration, "Search for four-top-quark production in the single-lepton and opposite-sign dilepton final states in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector", *Phys. Rev. D* **99** (2019) 052009, doi:10.1103/PhysRevD.99.052009, arXiv:1811.02305.

[33] CMS Collaboration, "Search for the production of four top quarks in the single-lepton and opposite-sign dilepton final states in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **11** (2019) 082, doi:10.1007/JHEP11(2019)082, arXiv:1906.02805.

[34] CMS Collaboration, "Search for production of four top quarks in final states with same-sign or multiple leptons in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Eur. Phys. J. C* **80** (2020) 75, doi:10.1140/epjc/s10052-019-7593-7, arXiv:1908.06463.

[35] ATLAS Collaboration, "Evidence for $\mathrm{t\bar{t}t\bar{t}}$ production in the multilepton final state in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector", *Eur. Phys. J. C* **80** (2020) 1085, doi:10.1140/epjc/s10052-020-08509-3, arXiv:2007.14858.

[36] ATLAS Collaboration, "Measurement of the $\mathrm{t\bar{t}t\bar{t}}$ production cross section in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector", *JHEP* **11** (2021) 118, doi:10.1007/JHEP11(2021)118, arXiv:2106.11683.

[37] CMS Collaboration, "Evidence for four-top quark production in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", 2023. arXiv:2303.03864. Submitted to *Phys. Lett. B*.

[38] F. Blekman, F. Déliot, V. Dutta, and E. Usai, "Four-top quark physics at the LHC", *Universe* **8** (2022) 638, doi:10.3390/universe8120638, arXiv:2208.04085.

[39] CMS Collaboration, "Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ in 2015 and 2016 at CMS", *Eur. Phys. J. C* **81** (2021) 800, doi:10.1140/epjc/s10052-021-09538-2, arXiv:2104.01927.

[40] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13\,\text{TeV}$", CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018.

[41] CMS Collaboration, "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13\,\text{TeV}$", CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019.

[42] CMS Collaboration, "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JINST* **15** (2020) P10017, doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165.

[43] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[44] CMS Collaboration, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC", *JINST* **16** (2021) P05014, doi:10.1088/1748-0221/16/05/P05014, arXiv:2012.06888.

[45] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.

[46] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[47] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[48] CMS Collaboration, "Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13\,\text{TeV}$", *JINST* **13** (2018) P10005, doi:10.1088/1748-0221/13/10/P10005, arXiv:1809.02816.

[49] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015) P02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.

[50] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector", *JINST* **14** (2019) P07004, doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.

[51] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_{T} jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[52] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[53] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[54] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[55] CMS Collaboration, "Performance of the DeepJet b tagging algorithm using 41.9 fb^{-1} of data from proton-proton collisions at 13 TeV with Phase 1 CMS detector", CMS Detector Performance Note CMS-DP-2018-058, 2018.

[56] E. Bols et al., "Jet flavour classification using DeepJet", *JINST* **15** (2020) P12012, doi:10.1088/1748-0221/15/12/P12012, arXiv:2008.10519.

[57] CMS Collaboration, "Technical proposal for the Phase-II upgrade of the Compact Muon Solenoid", CMS Technical Proposal CERN-LHCC-2015-010, CMS-TDR-15-02, 2015.

[58] CMS Collaboration, "ECAL 2016 refined calibration and Run2 summary plots", CMS Detector Performance Note CMS-DP-2020-021, 2020.

[59] K. Rehermann and B. Tweedie, "Efficient identification of boosted semileptonic top quarks at the LHC", *JHEP* **03** (2011) 059, doi:10.1007/JHEP03(2011)059, arXiv:1007.2221.

[60] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system", in *Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining: San Francisco CA, USA, August 13–17, 2016*. 2016. arXiv:1603.02754. doi:10.1145/2939672.2939785.

[61] CMS Collaboration, "Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying τ leptons at $\sqrt{s} = 13$ TeV", *JHEP* **08** (2018) 066, doi:10.1007/JHEP08(2018)066, arXiv:1803.05485.

[62] CMS Collaboration, "Observation of single top quark production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. Lett.* **122** (2019) 132003, doi:10.1103/PhysRevLett.122.132003, arXiv:1812.05900.

[63] CMS Collaboration, "Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **81** (2021) 378, doi:10.1140/epjc/s10052-021-09014-x, arXiv:2011.03652.

[64] CMS Collaboration, "Search for electroweak production of charginos and neutralinos in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **04** (2022) 147, doi:10.1007/JHEP04(2022)147, arXiv:2106.14246.

[65] CMS Collaboration, "Measurements of the electroweak diboson production cross sections in proton-proton collisions at $\sqrt{s} = 5.02\,\text{TeV}$ using leptonic decays", *Phys. Rev. Lett.* **127** (2021) 191801, doi:10.1103/PhysRevLett.127.191801, arXiv:2107.01137.

[66] CMS Collaboration, "Inclusive and differential cross section measurements of single top quark production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **02** (2022) 107, doi:10.1007/JHEP02(2022)107, arXiv:2111.02860.

[67] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8\,\text{TeV}$", *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[68] CMS Collaboration, "Performance of CMS muon reconstruction in cosmic-ray events", *JINST* **5** (2010) T03022, doi:10.1088/1748-0221/5/03/T03022, arXiv:0911.4994.

[69] CMS Collaboration, "Performance of the reconstruction and identification of high-momentum muons in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JINST* **15** (2020) P02027, doi:10.1088/1748-0221/15/02/P02027, arXiv:1912.03516.

[70] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[71] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations", *JHEP* **03** (2013) 015, doi:10.1007/JHEP03(2013)015, arXiv:1212.3460.

[72] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[73] S. Frixione, G. Ridolfi, and P. Nason, "A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction", *JHEP* **09** (2007) 126, doi:10.1088/1126-6708/2007/09/126, arXiv:0707.3088.

[74] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[75] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **09** (2009) 111, doi:10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: doi:10.1007/JHEP02(2010)011].

[76] P. Nason and C. Oleari, "NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG", *JHEP* **02** (2010) 037, doi:10.1007/JHEP02(2010)037, arXiv:0911.5299.

[77] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[78] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[79] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini, "Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM", *JHEP* **02** (2012) 088, doi:10.1007/JHEP02(2012)088, arXiv:1111.2854.

[80] P. Nason and G. Zanderighi, "W^+W^-, WZ and ZZ production in the POWHEG-BOX-V2", *Eur. Phys. J. C* **74** (2014) 2702, doi:10.1140/epjc/s10052-013-2702-5, arXiv:1311.1365.

[81] J. M. Campbell and R. K. Ellis, "An update on vector boson pair production at hadron colliders", *Phys. Rev. D* **60** (1999) 113006, doi:10.1103/PhysRevD.60.113006, arXiv:hep-ph/9905386.

[82] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", *JHEP* **07** (2011) 018, doi:10.1007/JHEP07(2011)018, arXiv:1105.0020.

[83] J. M. Campbell, R. K. Ellis, and W. T. Giele, "A multi-threaded version of MCFM", *Eur. Phys. J. C* **75** (2015) 246, doi:10.1140/epjc/s10052-015-3461-2, arXiv:1503.06182.

[84] NNPDF Collaboration, "Parton distributions from high-precision collider data", *Eur. Phys. J. C* **77** (2017) 663, doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.

[85] T. Sjöstrand et al., "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[86] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA 8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020) 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.

[87] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[88] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[89] S. Bolognesi et al., "On the spin and parity of a single-produced resonance at the LHC", *Phys. Rev. D* **86** (2012) 095031, doi:10.1103/PhysRevD.86.095031, arXiv:1208.4018.

[90] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[91] Particle Data Group, R. L. Workman et al., "Review of particle physics", *Prog. Theor. Exp. Phys.* **2022** (2022) 083C01, doi:10.1093/ptep/ptac097.

[92] H. Voss, A. Höcker, J. Stelzer, and F. Tegenfeldt, "TMVA, the toolkit for multivariate data analysis with ROOT", in *Proc. 11th Int. Workshop on Advanced Computing and Analysis Techniques in Phys. Research (ACAT 2017): Amsterdam, The Netherlands, April 23–27, 2007*. 2007. arXiv:physics/0703039. [PoS (ACAT2007) 040]. doi:10.22323/1.050.0040.

[93] C. G. Lester and D. J. Summers, "Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders", *Phys. Lett. B* **463** (1999) 99, doi:10.1016/S0370-2693(99)00945-4, arXiv:hep-ph/9906349.

[94] C. G. Lester, "The stransverse mass, m_{T2}, in special cases", *JHEP* **05** (2011) 076, doi:10.1007/JHEP05(2011)076, arXiv:1103.5682.

[95] A. Kulesza et al., "Associated top quark pair production with a heavy boson: differential cross sections at NLO+NNLL accuracy", *Eur. Phys. J. C* **80** (2020) 428, doi:10.1140/epjc/s10052-020-7987-6, arXiv:2001.03031.

[96] R. Frederix and I. Tsinikos, "On improving NLO merging for $\mathrm{t\bar{t}W}$ production", *JHEP* **11** (2021) 029, doi:10.1007/JHEP11(2021)029, arXiv:2108.07826.

[97] CMS Collaboration, "Measurement of the associated production of a single top quark and a Z boson in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$", *Phys. Lett. B* **779** (2018) 358, doi:10.1016/j.physletb.2018.02.025, arXiv:1712.02825.

[98] CMS Collaboration, "Search for new physics in same-sign dilepton events in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$", *Eur. Phys. J. C* **76** (2016) 439, doi:10.1140/epjc/s10052-016-4261-z, arXiv:1605.03171.

[99] CMS Collaboration, "Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$", *JHEP* **01** (2011) 080, doi:10.1007/JHEP01(2011)080, arXiv:1012.2466.

[100] CMS Collaboration, "Measurements of the $\mathrm{pp} \to \mathrm{WZ}$ inclusive and differential production cross section and constraints on charged anomalous triple gauge couplings at $\sqrt{s} = 13\,\mathrm{TeV}$", *JHEP* **04** (2019) 122, doi:10.1007/JHEP04(2019)122, arXiv:1901.03428.

[101] CMS Collaboration, "Measurements of the $\mathrm{pp} \to \mathrm{ZZ}$ production cross section and the $\mathrm{Z} \to 4\ell$ branching fraction, and constraints on anomalous triple gauge couplings at $\sqrt{s} = 13\,\mathrm{TeV}$", *Eur. Phys. J. C* **78** (2018) 165, doi:10.1140/epjc/s10052-018-5567-9, arXiv:1709.08601.

[102] CMS Collaboration, "$\mathrm{W^+W^-}$ boson pair production in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$", *Phys. Rev. D* **102** (2020) 092001, doi:10.1103/PhysRevD.102.092001, arXiv:2009.00119.

[103] CMS Collaboration, "Measurement of the inclusive and differential $\mathrm{t\bar{t}}\gamma$ cross sections in the dilepton channel and effective field theory interpretation in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$", *JHEP* **05** (2022) 091, doi:10.1007/JHEP05(2022)091, arXiv:2201.07301.

[104] J. Butterworth et al., "PDF4LHC recommendations for LHC Run II", *J. Phys. G* **43** (2016) 023001, doi:10.1088/0954-3899/43/2/023001, arXiv:1510.03865.

[105] CMS Collaboration, "Measurement of the cross section for $\mathrm{t\bar{t}}$ production with additional jets and b jets in pp collisions at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **07** (2020) 125, doi:10.1007/JHEP07(2020)125, arXiv:2003.06467.

[106] CMS Collaboration, "Inclusive and differential cross section measurements of $\mathrm{t\bar{t}b\bar{b}}$ production in the lepton+jets channel at $\sqrt{s} = 13\,\text{TeV}$ with the cms detector", CMS Physics Analysis Summary CMS-PAS-TOP-22-009, 2023.

[107] CMS Collaboration, "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV", *Eur. Phys. J. C* **75** (2015) 212, doi:10.1140/epjc/s10052-015-3351-7, arXiv:1412.8662.

[108] R. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples", *Comput. Phys. Commun.* **77** (1993) 219, doi:10.1016/0010-4655(93)90005-W.

[109] J. S. Conway, "Incorporating nuisance parameters in likelihoods for multisource spectra", in *Proc. 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding (PHYSTAT 2011): Geneva, Switzerland, January 17–20, 2011*. 2011. arXiv:1103.0354. doi:10.5170/CERN-2011-006.115.

[110] ATLAS and CMS Collaborations, and LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in Summer 2011", Technical Report CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, 2011.

[111] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: doi:10.1140/epjc/s10052-013-2501-z].