Rare Single Top-Quark production at CMS

Seddigheh Tizchang

Department of Physics, Faculty of Science, Arak University, Arak 384817758, Iran.
School of particles and accelerators Institute for Research in Fundamental Science

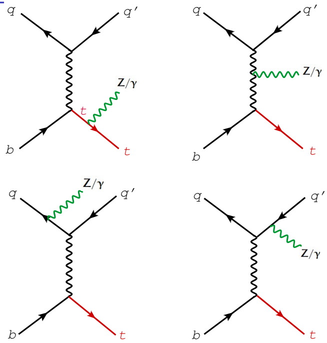
for the CMS collaboration

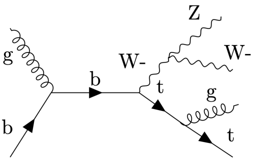
Conference on High Energy Physics
A.I. Alikhanyan National Science Laboratory, Yerevan, Armenia

11-14 September 2023

Rare single top production at LHC

- Single top quark production with $\gamma/Z/WZ$ are remarkably rare processes predicted by the SM.
- Measuring X-sections of these processes are crucial as:
  - Sesitive to top quark EW coupling
  - Sensitive to many BSM models and EFT operators.
- Production of $t\gamma/tZ$ at the LHC: t-channel, s-channel, and tW-channel (t-channel mode is the dominant one).





Results covered in this talk:

$t\gamma q$ first evidence: PRL 122(2019)132003

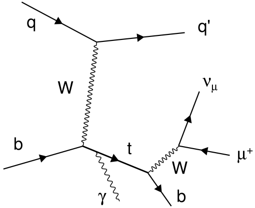
tZq Observation: JHEP 02 (2022)107

tWZ first evidence: CMS-PAS-TOP-22-008

CMS Result for tγq measurement

X-section measurement using 2016 Run II data

tqγ production at t-channel



- The X-section is sensitive to the top quark electric and magnetic dipole moments.
- μ channel, $35.9\,\mathrm{fb}^{-1}$

Object selections:

- Well isolated photon with $E_T > 25$ GeV in barrel region
- Single isolated muon with $p_T > 26$ GeV and $|\eta| < 2.4$ veto extra loose muon or electron.
- AK4PF Jet with $p_T > 40$GeV and $|\eta| < 4.7$
- One b-tagged jet with WP medium
- MET> 30GeV

Object reconstruction:

- W boson is reconstructed with μ+MET.
- Analytic neutrino solution(s).
- Top mass is reconstructed with W boson and a b-jet candidate.

Additional requirements:

- $\Delta R(\gamma, X) > 0.5$ with X=jets,b-jet,μ
- $\Delta R(\mu, X) > 0.5$ with X=jets,b-jet

Background estimation strategy

- Backgrounds with prompt photon:
  - Dominant BGs: $t\bar{t}\gamma$, Wγ+jets,Zγ+jets.
  - Estimated with both MC and data-driven method.
- Backgrounds with fake photon
  - Jets can be misreconstructed as photon.
  - $t\bar{t}$,W+jets, DY+jets.
  - Data driven method is applied to estimate fake photon.

Fake photon estimation

- Shape and normalization of events with fake photon is estimated by using data.
- $t\bar{t}$, W+jets,DY+jets, Single top, diboson have contribution.
- $j \rightarrow \gamma$**:**
  - **First**: fake photon fraction in the SR is estimated by using the ABCD method.
  - **Second**: A control region dominated by misidentified photon, known as Photon Like Jet (PLJ) is defined.
  - Correction factor depends on the p_T of photon.
  - Obtain normalization and shape of fake events in SR by weighting each event in PLJ sample with correction factor.

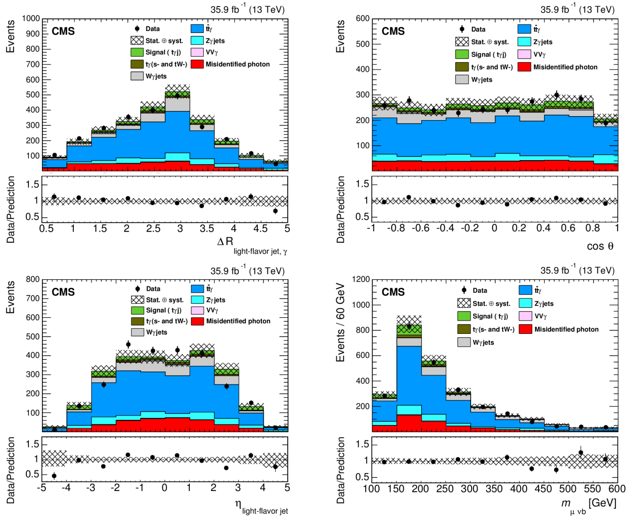
$t\bar{t}\gamma$ Control Region

- Main BG is $t\bar{t}\gamma$.
- Its shape is obtained from MC and normalization is estimated by using data with **relevant control region:**
  - Exactly 2 b-tagged jet is required.
  - Other selection criteria are the same as signal region.
- Normalization of $t\bar{t}\gamma$ is left floating in both SR & CR.
- Simultaneous fit into SR and CR is applied to extract the $t\bar{t}\gamma$ normalization in SR.

Process	Event yield
$t\bar{t}+\gamma$	1401 ± 131
$W\gamma$+jets	329 ± 78
$Z\gamma$+jets	232 ± 55
Misidentified photon	374 ± 74
$t\gamma$ (s- and tW-channel)	57 ± 8
$VV\gamma$	8 ± 3
Total background	2401 ± 178
Expected signal	154 ± 24
Total SM prediction	2555 ± 180
Data	2535

Analysis strategy

To enhance the signal and background separation, signal against the main background ($t\bar{t}\gamma$) is trained.

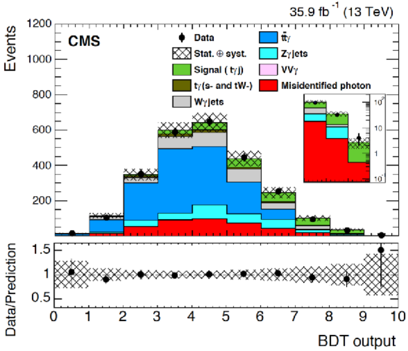


The most discriminant variables used as BDT input:

- η of the light-flavor jet
- Cosine of the angle between muon & light-flavor jet in top quark rest frame
- η of the muon
- ΔR(light-flavor jet, γ)
- Reconstructed top quark mass
- Jet multiplicity
- Transverse mass of the reconstructed W boson
- Muon charge

Signal extraction

- Binned likelihood fit on BDT output of 1-btagged and 2b-tagged regions.
- Normalization of $t\bar{t}\gamma$ determined from specific control region with 2 b-tagged jet.
- Main source of uncertainties:
  JEC (12%), signal modeling(9%) and b-tagging and mis-tagging rates (7%).



The fiducial X-section:

$$\sigma(t\gamma j)\mathcal{B}(t \to \mu\nu b) = 115 \pm 17(\text{stat.}) \pm 30(\text{syst.})\text{fb}$$

Observed (expected) significance: 4.4 (3.0) σ

Expected SM prediction: $\sigma^{\text{SM}}(t\gamma q)B(t \to \mu\nu b) = 81 \pm 4\text{fb}$

Recent CMS Result for tZq measurement

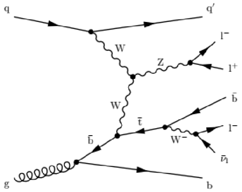
Inclusive and differential X-section measurements using full Run II data

tZq signal selection

- 3 leptons (e/μ) with $p_T > 25/15/10$ GeV.
  -Loose selection is used to estimate non-prompt background.
  -Tight selection (based on new lepton MVA) is used for prompt leptons.
- 2 leptons consistent with Z-boson mass within 15 GeV.
- at least 2 jets ($p_T > 25\text{GeV} \,\&|\eta| < 5$)
- at least one tagged as b-jet (differential measurement: 4 central jets)

Categorized based on number of jets and b-jets(inclusive):

- 2-3 jets with one b-tagged
- $\geq$ 4 jets with one b-tagged
- $\geq$ 2 b-tagged jets

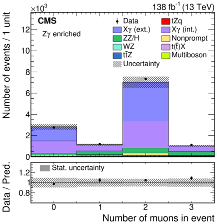


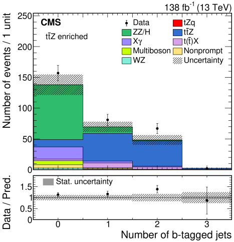
Inclusive vs. differential measurement:
Mostly similar strategy with some small differences:

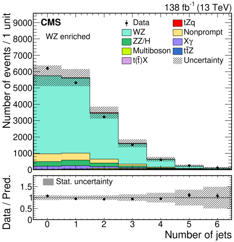
- Inclusive:
  - 3 signal categories based on number of jets and bjets.
  - Binary one vs. all classification with BDT.
- Differential:
  - Signal region inclusive in jets and b-jets.
  - Multiclass NN to regain discrimination between different BGs.

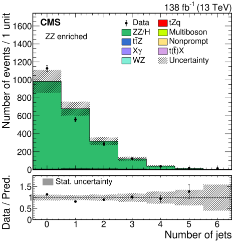
Prompt background estimation

- Most important BGs: WZ and $t\bar{t}Z$.
- Control region enriched with main backgrounds: WZ, ZZ, $t\bar{t}Z(4\ell)$, and $Z\gamma$.
- The agreement between the data and predictions is checked.
- The control regions are included in the final fit.









$Z\gamma$ CR

- 3 leptons
- no lepton pair within the Z mass window
- 3 leptons invariant mass within the Z mass window

$t\bar{t}$Z CR

- 4 leptons
- $\geq$ 2 jet
- only one lepton pair within the Z mass window

WZ CR

- 3 leptons
- At least one bjet veto
- Z boson candidate

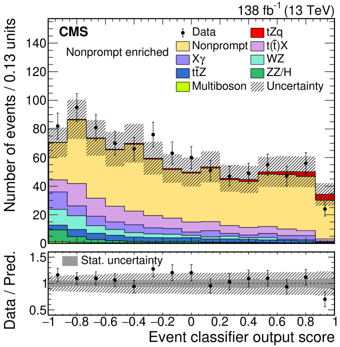
ZZ CR

- 4 leptons
- 2 lepton pairs within the Z mass window

Nonprompt background estimation

Jets can be misidentified as leptons.

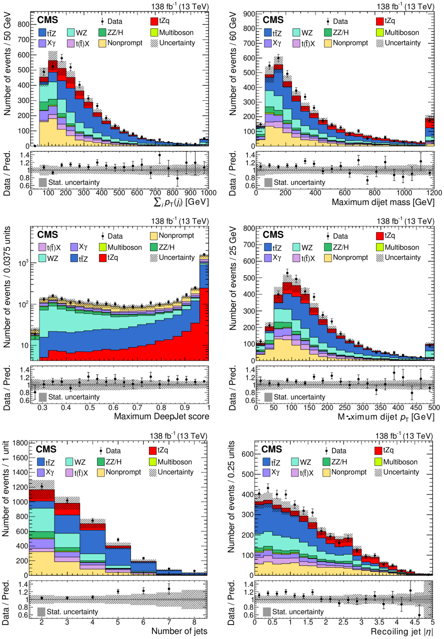
- Mainly DY+jet and $t\bar{t}$.
- BG with ≥ 1 nonprompt lepton(s).
- Strongly suppressed by lepton MVA, but still considerable.
- Estimated using data-driven technique naming fake rate method.



Fake rate:

- A CR in data enriched with multi-jet events is defined:
  - Only one loose lepton
  - At least one jet($p_T > 30$GeV and $|\eta| < 2.4$)
  - $\Delta R(\text{jet},\ell) > 0.7$
  - $p_T^{\text{miss}} < 20$ GeV
- Fake rate: probability that a nonprompt loose lepton also passes the tight identification in defined CR.
- The fake rate is measured as a function of the p_T and η of the nonprompt lepton.
- Apply in signal region side-bands:
  - with 3 looser leptons instead of 3 tight
  - 3 tight leptons are vetoed.

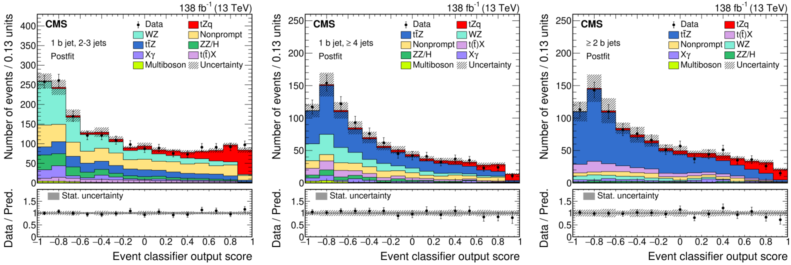
MVA input variables



BDT is trained to separate signal from background.
Some of the most discriminating features included in BDT training:

- $|\eta|$ of most forward jet
- Maximum deepFlavour b-tagging score
- Scalar sum of transverse momenta of all leptons + missing transverse momentum
- Maximum di-jet transverse momentum among all di-jet combinations
- Maximum deepFlavour b-tagging score among all jets
- Multiplicity of jets and bjets

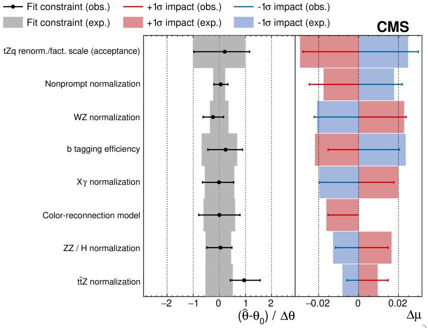
Results: inclusive measurement



Signal extraction:
Maximum-likelihood fit in the all SRs and CRs is performed on output distribution of BDT.

$$\sigma_{\text{tZq}}^{\text{Obs.}} = 87.9^{+7.5}_{-7.3}(\text{stat.})^{+7.3}_{-6.0}(\text{syst.})\text{fb}$$

for $m_{\ell\ell}$ >30 GeV.**Total uncertainty is** 11%. 3 – 4% lower uncertainty comparing to previous CMS and ATLAS results.



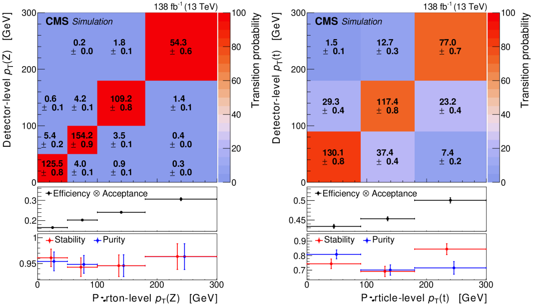
The SM predictions for $m_{\ell\ell}$ >30 GeV at NLO:

$$\sigma_{\text{tZq}}^{\text{SM}} = 94.2^{+1.9}_{-1.8}(\text{scale.}) \pm 2.5(\text{PDF})\text{fb}$$

Differential measurement

- Differential tZq cross section measurements as functions of several observables at the parton and particle levels are performed.
- A likelihood-based **unfolding procedure** is performed to measure the X-section in each kinematic region defined by each generator-level (parton or particle) bin.

- **Parton level**
  - Generator level objects after ISR/FSR, before hadronization.
  - Whole phase space
  - Strong dependency to used generator.
- **Particle level**
  - Stable particles after parton showering.
  - Collected into jets, MET and leptons.
  - Less dependency on used generator.
  - Object cannot be directly mapped to matrix element objects.
  - Unfolding in fiducial phase space.



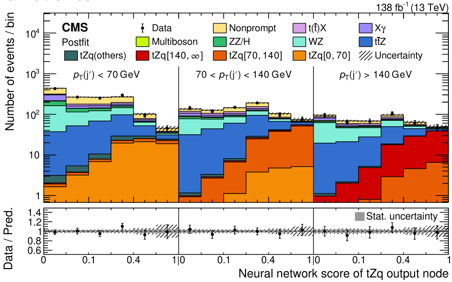
Results: differential measurement

Multi-class NN with several input variables and five separate class (tZq,ttZ, WZ,$t(\bar{t})X$ and other BGs) is performed to classify the signal events.

Signal extraction:

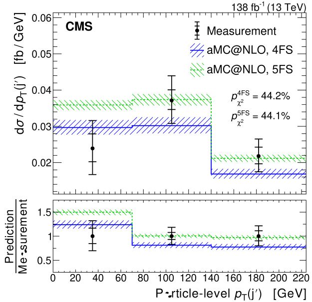
- σ^k_{tqz}: X-section at generator level with generator level bin k as a function of physical quantities(x):

$$\sigma^k_{tqz} = \int_{x_k^{low}}^{x_k^{high}} \frac{d\sigma_{tqz}}{dx} dx$$

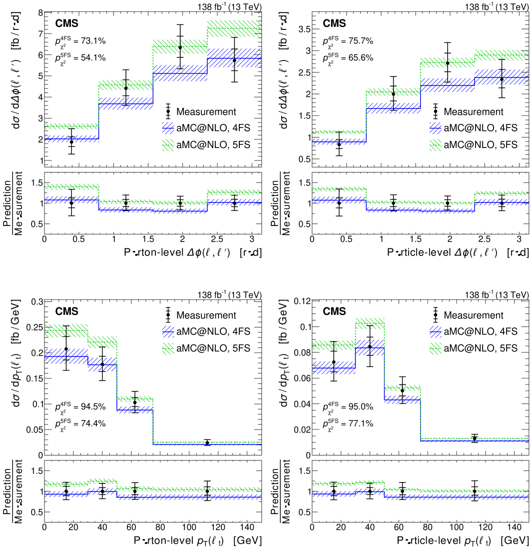


Signal region: $N_{\text{jets}} < 4 \rightarrow$ to get rid of the contribution of $t\bar{t}Z$ backgrounds.

- Maximum likelihood based unfolding:
  - SR is break into subregions based on the observable at detector level.
  - each subregion dominated by signals from its generator bin.
  - each subregion fit to classifier output to separate signal from BGs.
  - all CRs are into the fit.
  - all systematic are considered as nuisance parameter.



Results: absolute differential X-sections



- Good agreement between prediction and measurement.
- 4FS and 5FS predictions compared.
- Down to 25% uncertainties in including jets observables.
- Down to 15% uncertainties in purely leptonic observables.
- **Dominant uncertainties:** Statistical, b-tag identification, BG modeling.

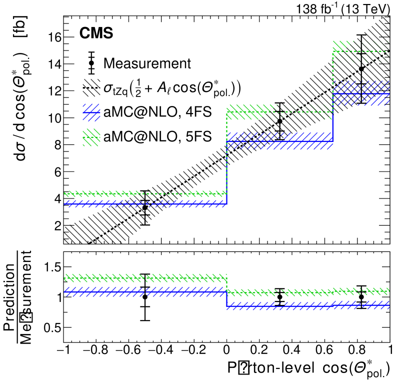
Results: spin asymmetry

- In tZq production, due to **V-A electroweak coupling**, generated top quark is polarized.
- Deviations could potentially indicate **anomalous coupling** structure.
- At the parton level, spin asymmetry is connected to the differential cross-section as a function of the polarization angle:

$$\frac{d\sigma}{dcos(\theta^{\star}_{\text{pol}})} = \sigma_{tZq}.(\frac{1}{2} + A_{\ell}cos(\theta^{\star}_{\text{pol}}))$$

$$cos(\theta^{\star}_{\text{pol}}) = \frac{\overrightarrow{p}(q'^{\star}).\overrightarrow{p}(\ell^{\star}_{t})}{|\overrightarrow{p}(q'^{\star})||\overrightarrow{p}(\ell^{\star}_{t})|}$$

A_{ℓ} is directly extracted from a re-parameterized fit using the full likelihood and uncertainties. Dominant uncertainty is statistical Uncert.



Measurement result:

$$A_{\ell} = 0.54 \pm 0.16(\text{stat}) \pm 0.06(\text{syst})$$

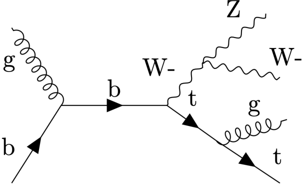
The SM prediction:

$$A_{\ell}^{4FS} = 0.44, A_{\ell}^{5FS} = 0.45$$

CMS Result for tWZ measurement

X-section measurement using full Run II data

tWZ production



Low energy region:

Search Semi-leptonic final states

- One OSSF lepton pair compatible with Z mass.
- **One of the W** decay **leptonically**.
- jets multiplicity ($p_T > 25$GeV,$|\eta| < 2.5$):
  - **SR1:** at least 3j & at least 1b-jet.
  - **SR2:** exactly 2j & at least one b-jets.

Search in fully leptonic final states

SR3:

- One OSSF lepton pair compatible with Z mass.
- **Both W bosons** decay **leptonically**.
- at least one jet & at least one b-jet.

Boosted (high energy) region:

Hadronic top (two jets from W decay are reconstructed as a fat jet)

- Hadronic top
  - SR1 selection
  - at least one fat jet ($p_T > 300$GeV & 105 GeV$< m_{SD} <$210 GeV)
  - ΔR(fat jet , b-jet)<0.8
- Leptonic top
  - SR1 selection
  - veto hadronic selection
  - $p_T^{\ell_3} > 30$GeV
  - at least one b-jet ($p_T > 200$GeV)
  - Δ R(b-jet , ℓ_3)<2

Background estimation

Main BG: ttZ

- Very similar kinematically to signal.
- ttZ differ via one b-jet in the final state with signal.
- ttZ BG is estimated from MC then constrained from DNN output score.

Multi-lepton Control Regions:

Are estimated from MC and normalizations are constrained in dedicated control regions.

ZZ CR

- 4 leptons
- 2 pairs OSSF compatible with Z mass
- no jet is required

WZ+jet CR

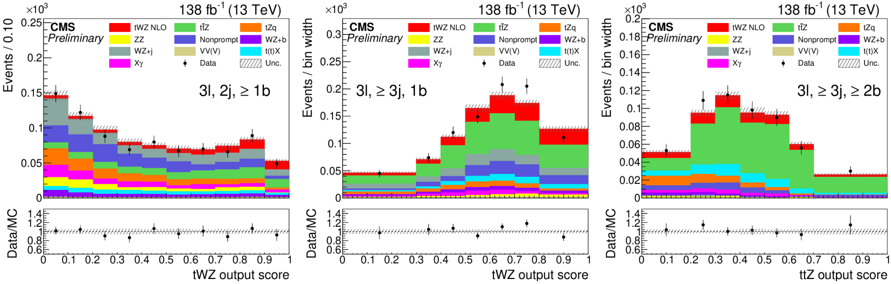
- Exactly 3 leptons
- one OSSF lepton pair from Z mass
- any b-jet veto
- MET >50 GeV

Non-prompt backgrounds: Fake rate method is used to estimate this BG.

- Estimated from data

Other BGs are estimated from MC

Signal extraction



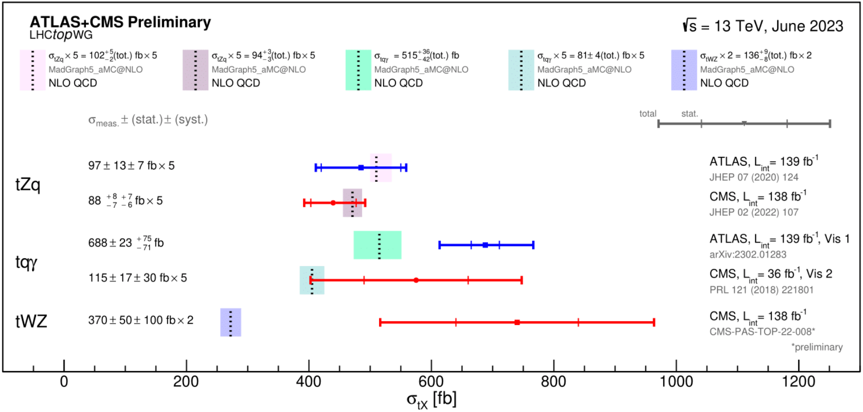
- MVA techniques employed to separate signal from ttZ and other backgrounds.
- Binned maximum likelihood fit is used to extract the tWZ signal significance and cross section.
- Evidence of tWZ with an **observed(expected) significance of 3.5 (1.4) standard deviations**.
- Cross section (2.1 standard deviations away from the SM prediction):

$$\sigma_{\mathrm{tWZ}} = 0.37 \pm 0.05(\mathrm{stat}) \pm 0.10(\mathrm{sys})\ \mathrm{pb}$$

  Expected SM prediction: $\sigma^{\mathrm{SM}} = 136$ fb
- **Most impactful source of uncertainties:** ttZ normalization, b-tagging efficiency corrections and estimation of nonprompt leptons.

Summary of tX measurements

- $t\gamma q$ analysis
  - First evidence by CMS (2016 data).
  - Observation by ATLAS(full Run II).
  - Upcoming: full Run II results for CMS.
- tZq analysis
  - Observed by CMS and ATLAS.
  - Good agreement with the SM.
- tWZ analysis
  - First evidence by CMS.
  - 2.1 standard deviation from the SM.



Summary

- First evidence of tWZ production with 3.5 standard deviation, performed in multilepton final states using the full Run II data at CMS experiment.

- Inclusive and differential measurements of tZq production have been done with full Run II data at CMS experiment.
- CMS collaborations has observed tZq production with significance greater than 5σ.
- The measured inclusive X-section is in agreement with the SM prediction.
- Compared to the SM differential cross section, all observables are in good agreement within the uncertainties with one exception.

- CMS has reported first evidence of $t\gamma$q process using $35.9\,\text{fb}^{-1}$ of data at 13 TeV with 4.4 observed significance.
- Measurement of $t\gamma$q X-section via full Run II data is in progress.

Backup

Previous tZq search by CMS

First evidence (2016 data):

$$\sigma(PP \to tZq \to t\ell^+\ell^- q) = 123^{+33}_{-31}(\text{stat.})^{+29}_{-23}(\text{syst.})\text{fb}$$

Significance =3.7(obs.)/3.1(exp.)

Phys.Lett. B 779 (2018)358

Observation at CMS (2016&2017) data

$$\sigma(PP \to tZq \to t\ell^+\ell^- q) = 111^{+13}_{-13}(\text{stat.})^{+11}_{-9}(\text{syst.})\text{fb}$$

Significance = 8.2(obs.)/7.7(exp.)

Phys. Rev. Lett. 122, 132003

Observation at ATLAS (full Run II):

$$\sigma(PP \to tZq \to t\ell^+\ell^- q) = 97 \pm 13(\text{stat.})7(\text{syst.})\text{fb}$$

Significance $\gg 5\sigma$ JHEP 07 (2020) 124