Run-III Dileptonic Top Quark Pair Production Cross-section Measurement Introduction

credits: TTXS members and FNAL staffs

TTXS Facilitators

Alice Bean

Beren Ozek

Lingqiang He

Daniel Li

Yibo Zhong

Roy Candelaria

Aidan Grummer

Introduction

" $\sigma_{t\bar{t}}$ rises to $\mathbf{921\ pb}$ at $\sqrt{s} = 13.6$ TeV from 834 pb in Run 2."

"We expect sensitivity on the order of $\pm 5\%$ ($\pm$ lumi)."

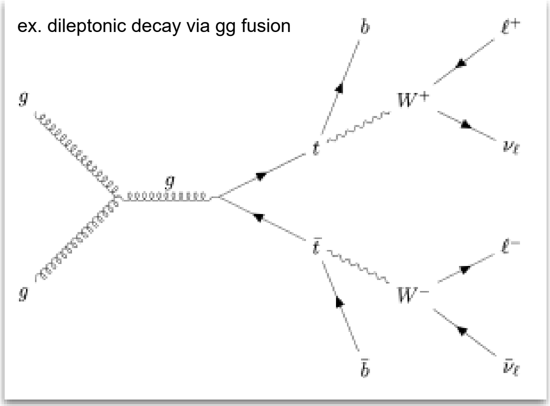
Top-antitop quark production cross-section is an important attribute to validate the Standard Model (SM) physics. It can also be used as a probe to access new physics among many observables provided by the LHC.

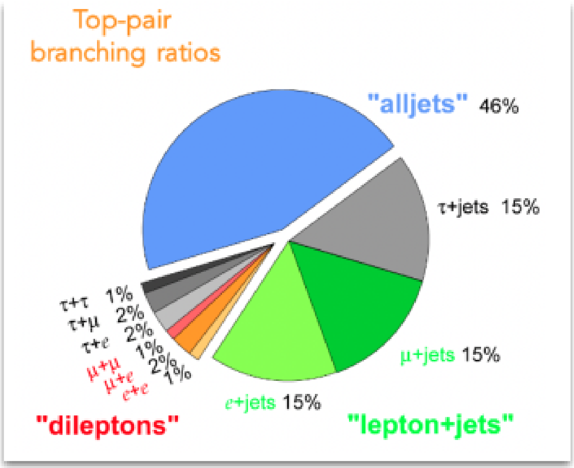
The CMS collaboration expects to gain 10% increase in the x-sec from 13 TeV to 13.6 TeV with PDF4LHC21 predictions. (TOP-22-012)

Top Pair Decay Channels

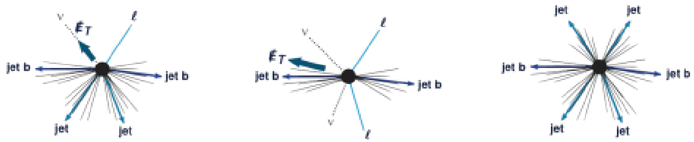
$t\bar{t}$ dileptonic production cross-section

$$\sigma \propto \left|\mathcal{M}(gg/qq \to t\bar{t} \to \ell^+\nu b; \ell^-\bar{\nu}\bar{b})\right|^2$$





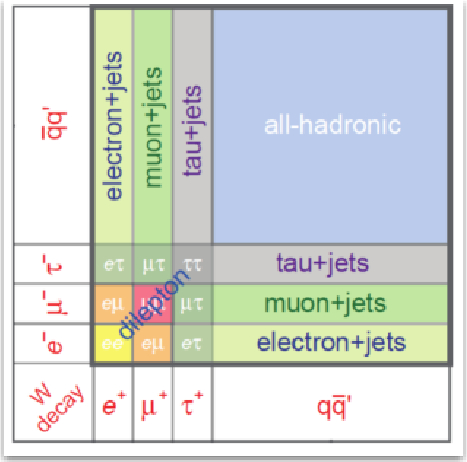
$t\bar{t}$ semi-leptonic | $t\bar{t}$ dileptonic | $t\bar{t}$ hadronic



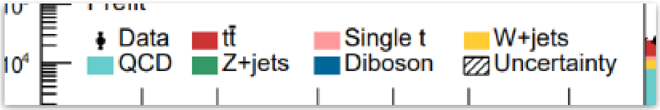
$BR \sim 45\%$ | $BR \sim 9\%$ | $BR \sim 46\%$

Signal Selection

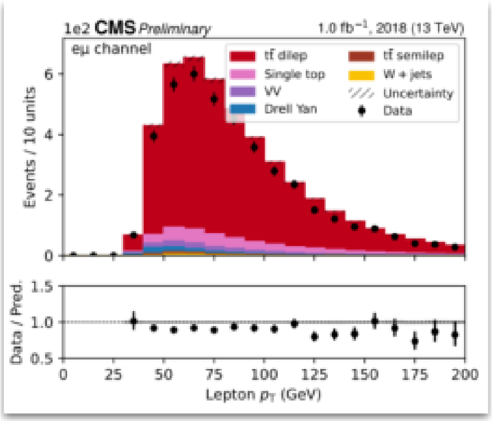
1. Lepton pair with opposite signs at high p_T
    a. ee, eμ, μμ, e+jets, μ+jets
2. Not including leptonic tau decays
3. At least 1 b-tagged jet



Background Events



1. V+Jets
2. SingleTop
3. QCD
4. Diboson
5. …



Data & MC Samples

Luminosity of 1.23 fb^{-1} of certified data from Era C (golden JSON from 16.08.)

MuonEG, Egamma, Muon DoubleMuon data stream [Look up the definitions here]

Single & Dilepton high level triggers (HLT) applied

The ABCD Method

Region A: Signal Region (SR)

- Meets both selection criteria (e.g., tight isolation and a certain jet multiplicity).
- Contains contributions from both signal and background.

Region B: Background Region 1

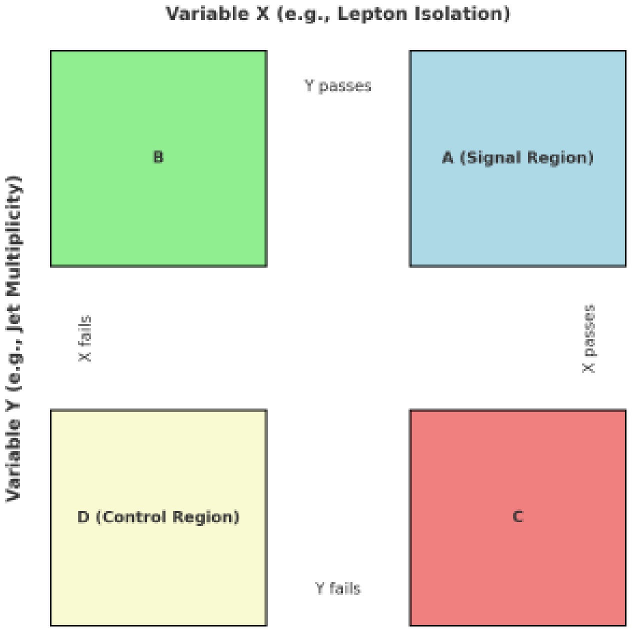
- Meets one selection criterion but fails the other (e.g., loose isolation).

Region C: Background Region 2

- Fails the first criterion but meets the second.

Region D: Control Region

- Fails both selection criteria.



Scale Factors

1. Lumi SF $SF_{lumi} = \frac{\sigma \times \mathcal{L}}{N_{Gen}}$

1. Trigger SF $SF_{Trigger} = \frac{\epsilon_{\text{trigger}}^{\text{data}}}{\epsilon_{\text{trigger}}^{\text{MC}}}$

1. Non-prompt SF $N_{\text{NP}}^{\text{SR}} = \left(N_{\text{data}}^{\text{CR}} - N_{\text{MC}}^{\text{CR}}\right) \times \frac{\left(N_{\text{data}}^{\text{SR,1j}} - N_{\text{MC}}^{\text{SR,1j}}\right)}{\left(N_{\text{data}}^{\text{CR,1j}} - N_{\text{MC}}^{\text{CR,1j}}\right)}$

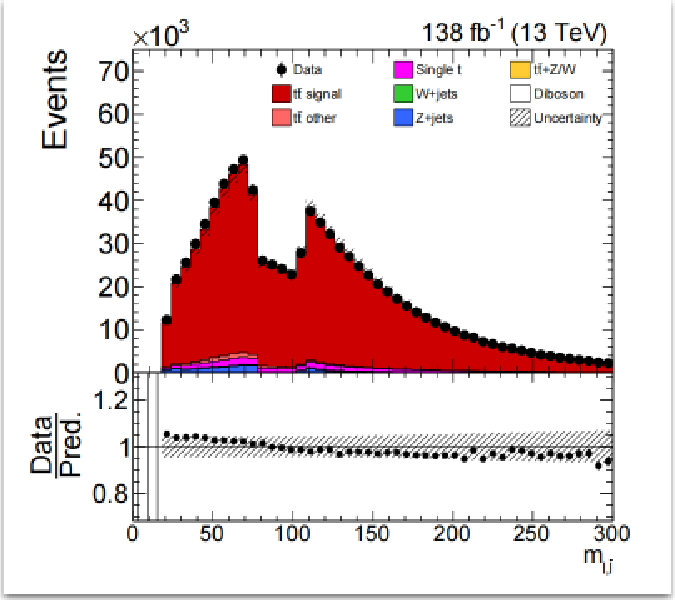
1. Drell-Yan SF $SF_{DY} = \frac{\left(N_{\text{out}}^{\geq 1b}\right)_{\text{data}}}{\left(N_{\text{out}}^{\geq 1b}\right)_{\text{MC}}} = \frac{\left(N_{\text{in}}^{\geq 1b}\right)_{\text{data}}}{\left(N_{\text{in}}^{\geq 1b}\right)_{\text{MC}}} \cdot \frac{\left(R_{\text{in/out}}^{0b}\right)_{\text{MC}}}{\left(R_{\text{in/out}}^{0b}\right)_{\text{data}}}$

5. BTagging SF

Event Selections

- ➔ Electrons: p_T > 35 GeV, $|\eta|$ < 2.4
  - ◆ tight cut-based ID
  - ◆ $|\eta_{SC}|$ not in [1.444; 1.566]
- ➔ Muons: p_T > 35 GeV, $|\eta|$ < 2.4
  - ◆ tight cut-based ID, tight ISO
- ➔ $m_{\ell\ell}$ > 20 GeV
- ➔ Z peak: remove 76 GeV < $m_{\ell\ell}$ < 106 GeV (ee and μμ only)
  - ◆ No MET Cuts

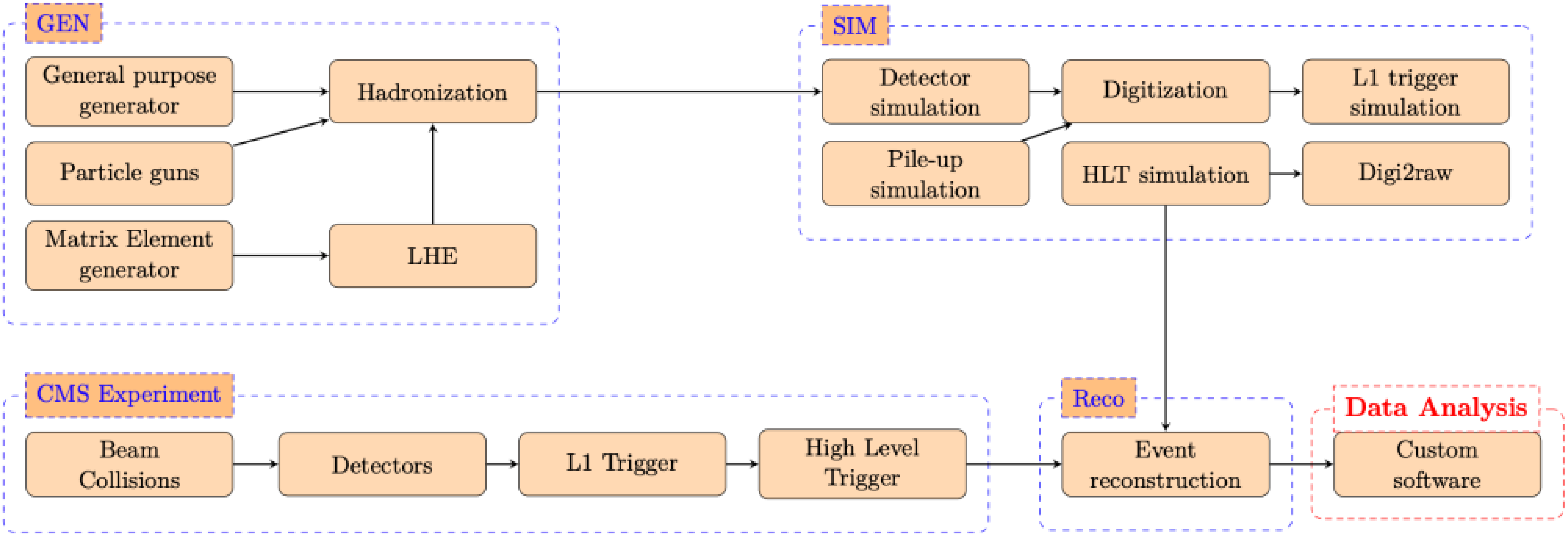
ex. excluded Z peak in mutual-flavor lepton pairs



[TOP-23-006]

Jet Selections

- ➔ Charged Hadron Subtraction (CHS), Tight jet ID
- ➔ $|\eta| < 2.4$
- ➔ $p_T > 30$ GeV


GEN
General purpose generator
Particle guns
Matrix Element generator
Hadronization
LHE
SIM
Detector simulation
Pile-up simulation
Digitization
HLT simulation
L1 trigger simulation
Digi2raw
CMS Experiment
Beam Collisions
Detectors
L1 Trigger
High Level Trigger
Reco
Event reconstruction
Data Analysis
Custom software

Common Analysis Tools

1. Columnar Analysis in Python
    a. Pepper
    b. Coffea
    c. Uproot
    d. NanoAODSchema
    e. Awkward Array
2. Anaconda & Jupyter-lab
    a. A convenient UI
    b. Support many formats, as competitive as MS VisualStudio

[OpenAI image]

Statistical Method

Likelihood fit via Combine Tool

$$\mathcal{L}_{\text{bin}} = \Gamma\left[n_{\text{obs}}^{\text{bin}} \middle| r\, s^{\text{bin}}(\{\theta_i\}) + b^{\text{bin}}(\{\theta_i\})\right] \times \prod_i p_i(\theta_i)$$

Overall Likelihood Function

$$\mathcal{L} = \prod_{\text{bin}} \mathcal{L}_{\text{bin}}$$

Poisson Distribution

$$\Gamma[n|\lambda] = \frac{\lambda^n e^{-\lambda}}{n!}$$

s - signal

b - background

$\{\theta_i\}$ - nuisance parameters

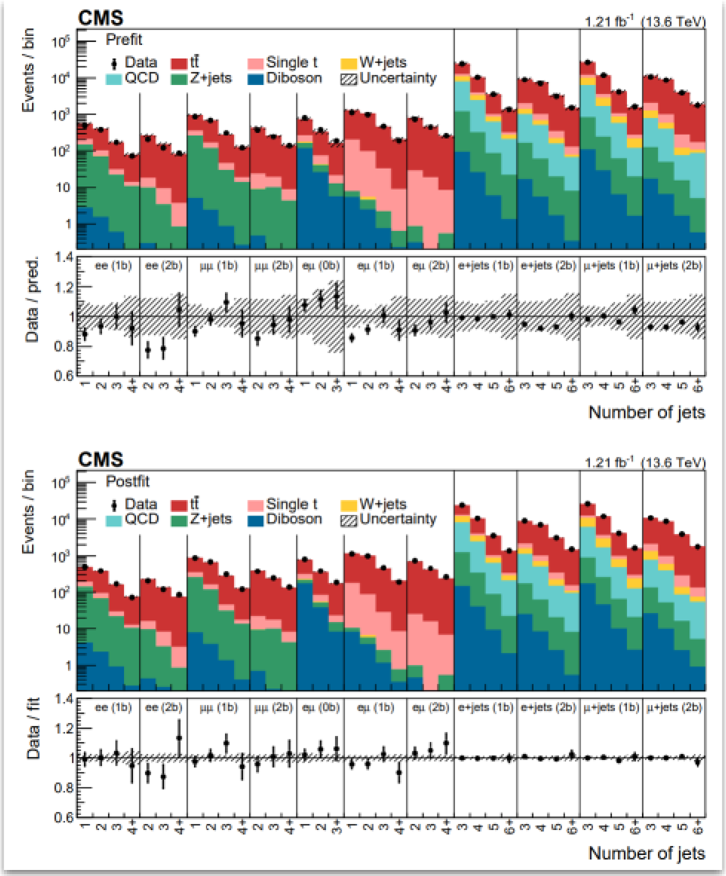
$p_i(\theta_i)$ - penalties

Statistical Method

prefit (top) vs. postfit (bottom)

pay attention to

1. normalization changes
2. changes in uncertainties



[TOP-22-012]

Statistical Method

Externalized: Lumi

Included in LL fit:

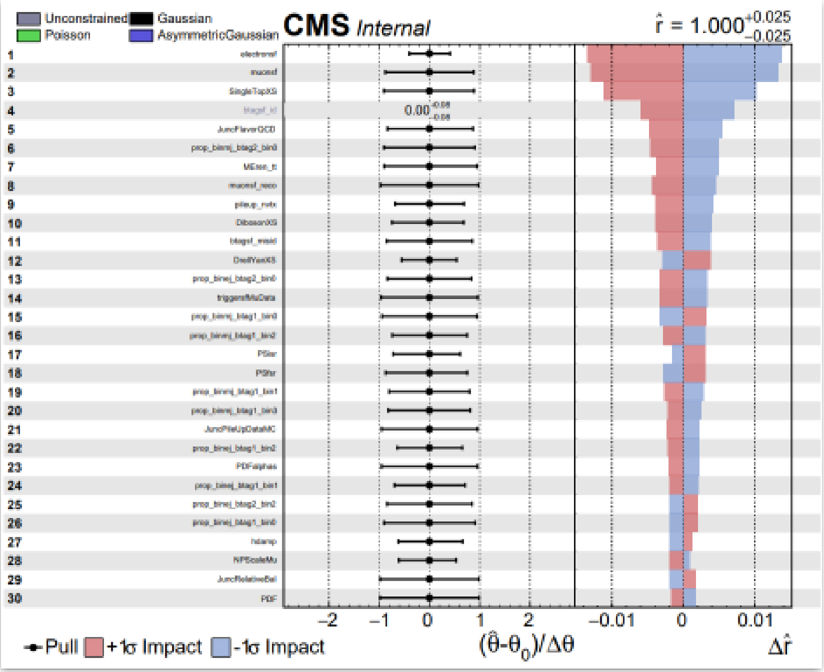
check out do_syst_table.py
for the systematic uncertainty names

Category	Parameter	Implementation	Treatment	tt	Single t	Drell Yan	Diboson	W+jets
Dominant Experimental	Lepton ID	Unconstrained	Correlated	✓	✓	✓	✓	✓
	JES	Shape	Correlated	✓	✓	✓	✓	✓
	b-tag SF	Shape	Correlated	✓	✓	✓	✓	✓
	Light Mistag SF	Shape	Correlated	✓	✓	✓	✓	✓
	Pileup	Shape	Correlated	✓	✓	✓	✓	✓
	Trigger SF	Shape	Correlated	✓	✓	✓	✓	✓
Dominant Theory	PDF $(+\alpha_s)$	Normalized Shape	Correlated	✓	✓	✓	✓	✓
	ME Scale	Normalized Shape	Uncorrelated	✓	✓	✓	✓	✓
	PS Scale	Normalized Shape	Correlated	✓	✓	✓	✓	✓
	h_{damp}	Dedicated Sample	Signal Only	✓	–	–	–	–
	BG Cross Section	Normalization	Uncorrelated	✓	15%	20%	30%	30%

Statistical Method

Summarize the effect of nuisance parameters on the best-fit value of the signal strength parameter **r.**

A well-behaved fit typically has most pulls within ±1σ, indicating that the data reasonably constrain the nuisance parameters



[TOP-22-012]

“It’s time to stop!”

Now we take a break and divide into two groups.

Group 1: Fill the histograms with lumi SF. Consider what samples to include in the control plots. Consider parton shower variation. Plot electron ID SF vs. nominal.

Group 2: Also fill the histograms with lumiSF. Consider pile-up reweighting. Plot muon ID SF vs. nominal.

Back Up

Actual Stack Plots in AN