First studies towards top-quark pair differential cross section measurement in the dilepton channel at $\sqrt{s}$ = 13 TeV with the CMS detector

Mykola Savitskyi, Maria Aldaya, Carmen Diez Pardos et al.

DESY-CMS Top Group, Hamburg

The top quark is special

Heaviest elementary particle known to date: $m_t \approx 173$ GeV

Mass close to scale of electroweak symmetry breaking (EWSB):
$\rightarrow y_t \approx 1$: important role in EWSB?

Decays before hadronizing: $\tau \approx 5 \cdot 10^{-25}$ s << $1/\Lambda_{QCD}$ – unique window on "bare" quark

Sensitive to physics phenomena beyond the Standard Model (BSM):
$\rightarrow$ new physics may preferentially couple/decay to top

Major source of background for many Higgs and BSM searches

Why measure differentially?

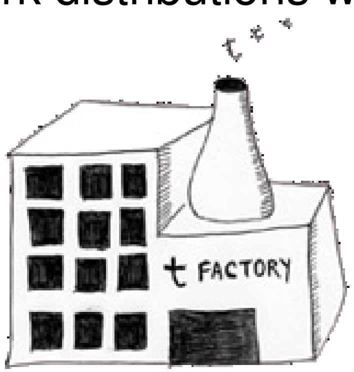
Precise understanding of top quark distributions is crucial:

- Precision tests of perturbative QCD for top-quark production at different phase space regions
- Tune and test theory predictions and models:
  $\rightarrow$ potential to reduce signal modelling systematics
- Enhance sensitivity to BSM physics

LHC is a "top factory"

- Several millions of top-quark pair ($t\bar{t}$) events produced already in Run-I ($\sqrt{s}$ = 7, 8 TeV)
- Run-II: expect much larger data sets at $\sqrt{s}$ = 13(14) TeV

$\rightarrow$ important to measure top quark distributions with very high precision

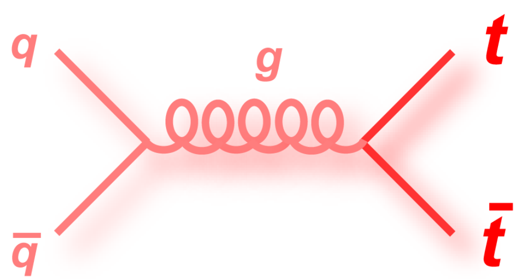


Top Pair Production and Decay

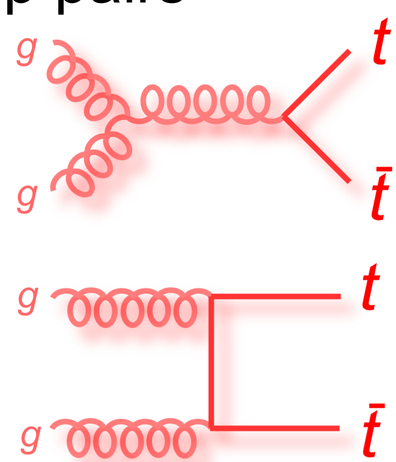
Production cross sections:

Energy, TeV	8	13
$\sigma_{incl}^{NNLO+NNLL}(t\bar{t})$, pb	252.89	831.76

Strong interaction: Top pairs



$q\bar{q}$ annihilation



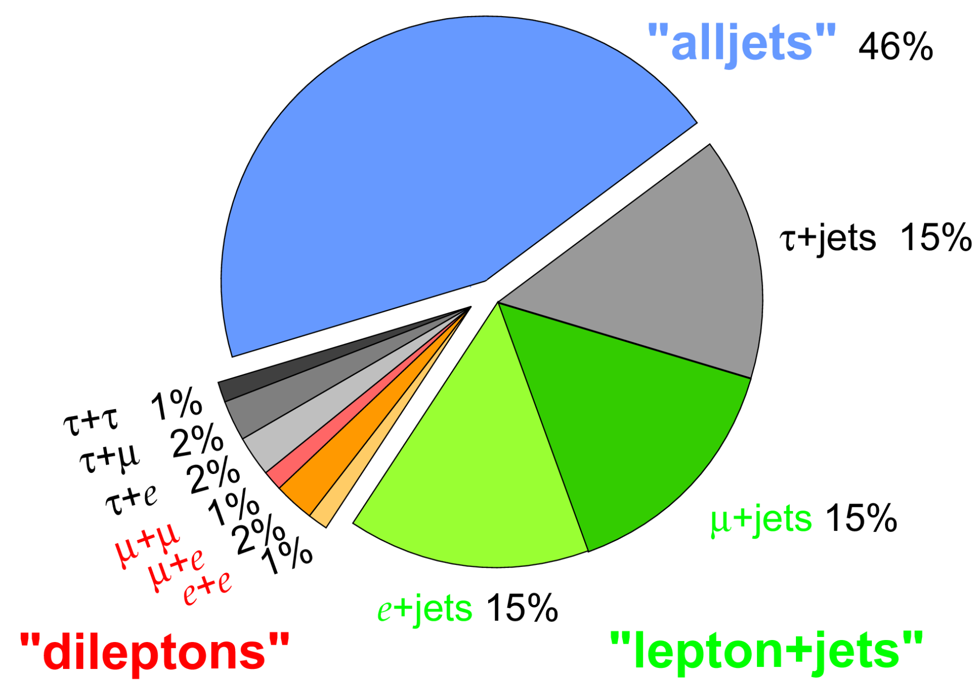
gg fusion

LHC Energy, TeV	8	13
$gg \rightarrow t\bar{t}$	~82%	~90%
$q\bar{q} \rightarrow t\bar{t}$	~18%	~10%

Weak interaction: Top decay

$t \rightarrow W+b$ ~ 100 %

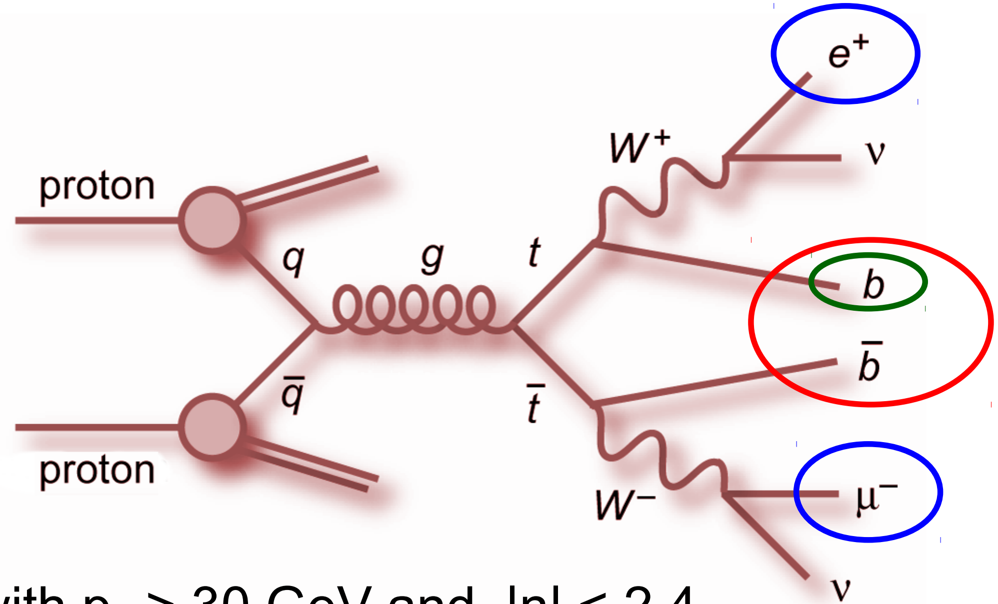
Top Pair Branching Fractions



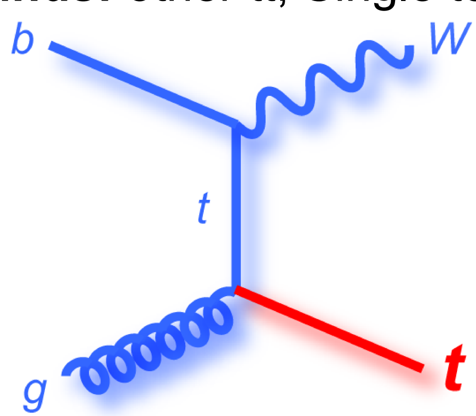
Samples are classified according to W-decay: dilepton ($\ell\ell$), lepton+jets (ℓ+jets), all jets

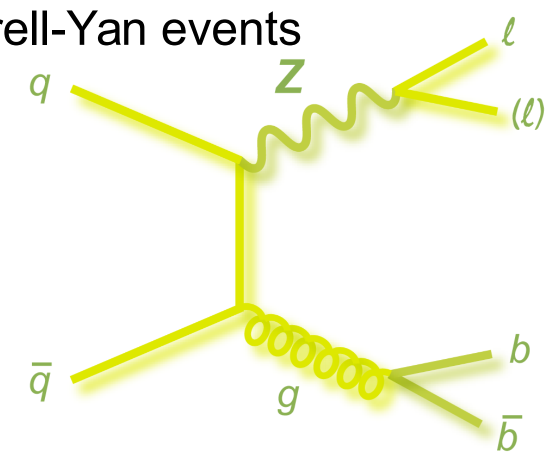
Selection in Dilepton Channel

1. Trigger conditions
2. **Lepton pair** selection:
   - opposite charge
   - e and μ isolation criteria
   - p_T > 20 GeV, $|\eta|$ < 2.4
   - invariant mass: m_{ll} > 20 GeV
3. Exclusion of Z -region:
   - in ee and μμ: 76 GeV < m_{ll} < 106 GeV
4. Presence of **two jets** (anti-k_t R=0.4) with p_T > 30 GeV and $|\eta|$ < 2.4
5. Missing E_T > 40 GeV in ee and μμ
6. At least one **b−tagged jet**
7. Meaningful solution for kinematic event reconstruction



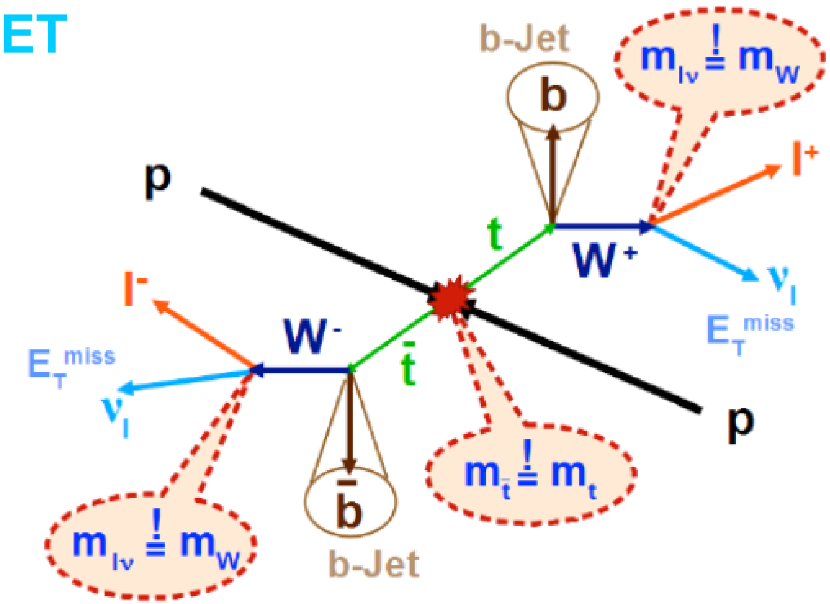
Largest backgrounds: other $t\bar{t}$, Single top, Drell-Yan events





Kinematic Reconstruction

- **Measured input: 2 jets, 2 leptons, MET**
- **Unknowns:** $\bar{p}_\nu, \bar{p}_{\bar{\nu}} \rightarrow 6$
- **Constraints:**
  - > $m_t, m_{\bar{t}} \rightarrow 2$
  - > $m_{W(+)}, m_{W(-)} \rightarrow 2$
  - > $(\bar{p}_\nu + \bar{p}_{\bar{\nu}})_T$ = MET $\rightarrow 2$



- Reconstructing each event 100 times and smearing inputs by their resolution:
  - > top mass fixed to 172.5 GeV
  - > W mass at RECO level smeared accordingly to W mass distribution
  - > Jet and lepton energies are corrected for detector effects
- Consider weighted average of solutions for all smeared points:

$$p_{x,y,z}^{top} = \frac{1}{w} \sum_{i=0}^{100} w_i \cdot \left(p_{x,y,z}^{top} \right)_i$$

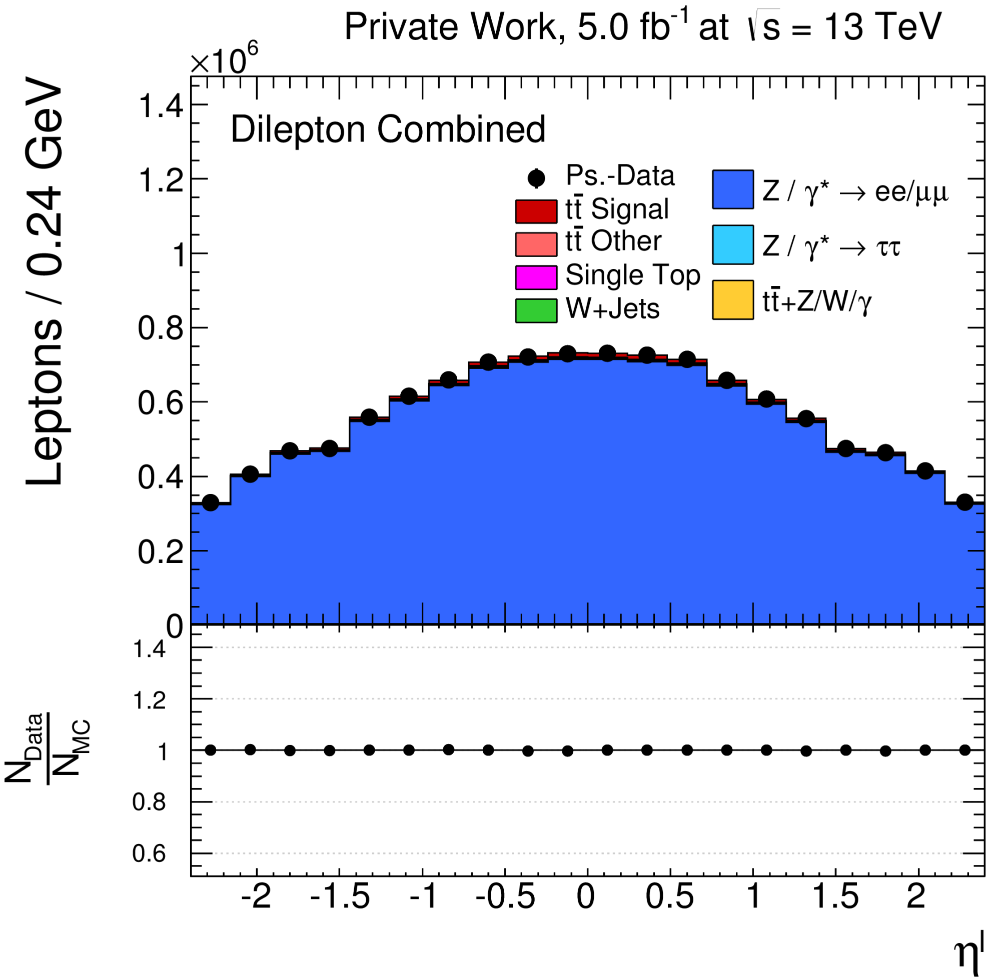
Selection in Dilepton Channel

Channel: combined (ee+eμ+μμ)
Pseudo-data: poisson-smeared sum of all signal and background simulation samples
Normalized to: L_{int} = 5.0 fb^{-1}

Observable: Lepton η
Signal: MadGraph+Pythia8

Selection steps:

1. Trigger conditions
2. Lepton pair selection:
   - opposite charge
   - e and μ isolation criteria
   - p_T > 20 GeV, |η| < 2.4
   - invariant mass: m_{ll} > 20 GeV



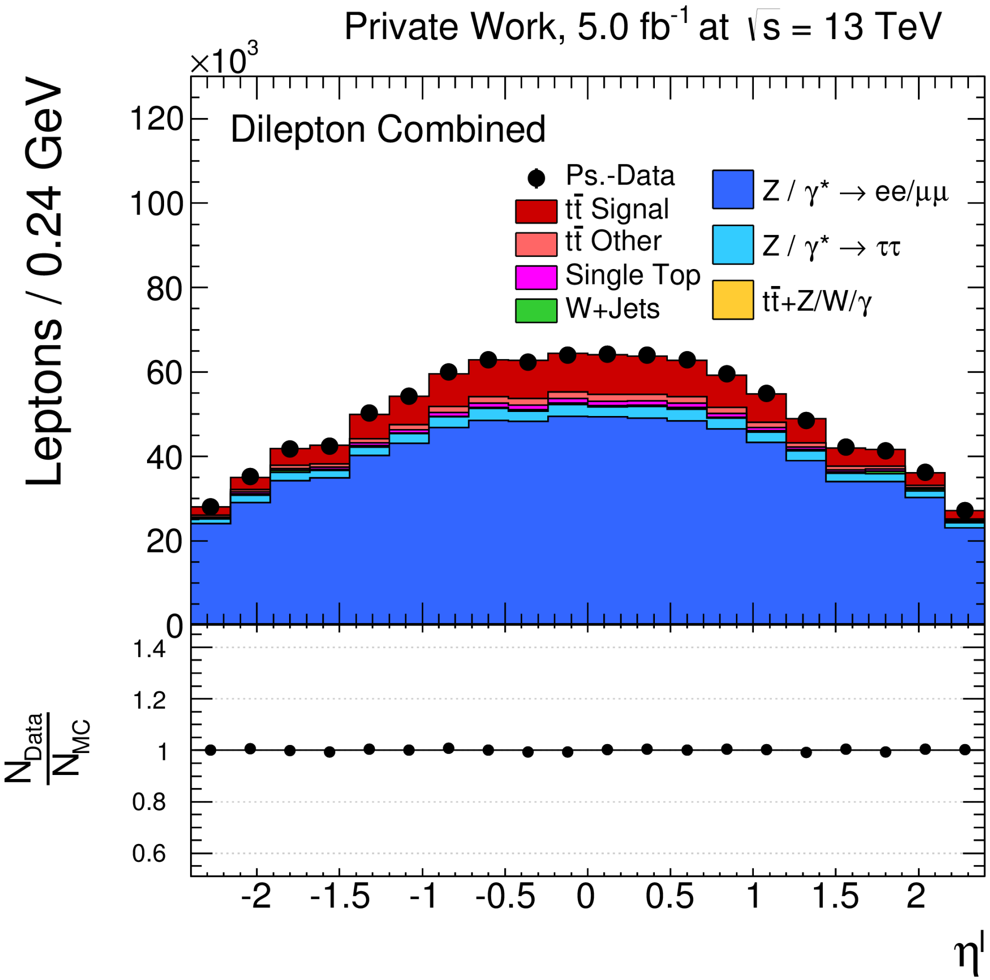
Selection in Dilepton Channel

Channel: combined (ee+eμ+μμ)
Pseudo-data: poisson-smeared sum of all signal and background simulation samples
Normalized to: L_{int} = 5.0 fb^{-1}

Observable: Lepton η
Signal: MadGraph+Pythia8

Selection steps:

1. Trigger conditions
2. Lepton pair selection:
   - opposite charge
   - e and μ isolation criteria
   - p_T > 20 GeV, |η| < 2.4
   - invariant mass: m_{ll} > 20 GeV
3. Exclusion of Z -region:
   - in ee and μμ: 76 GeV < m_{ll} < 106 GeV



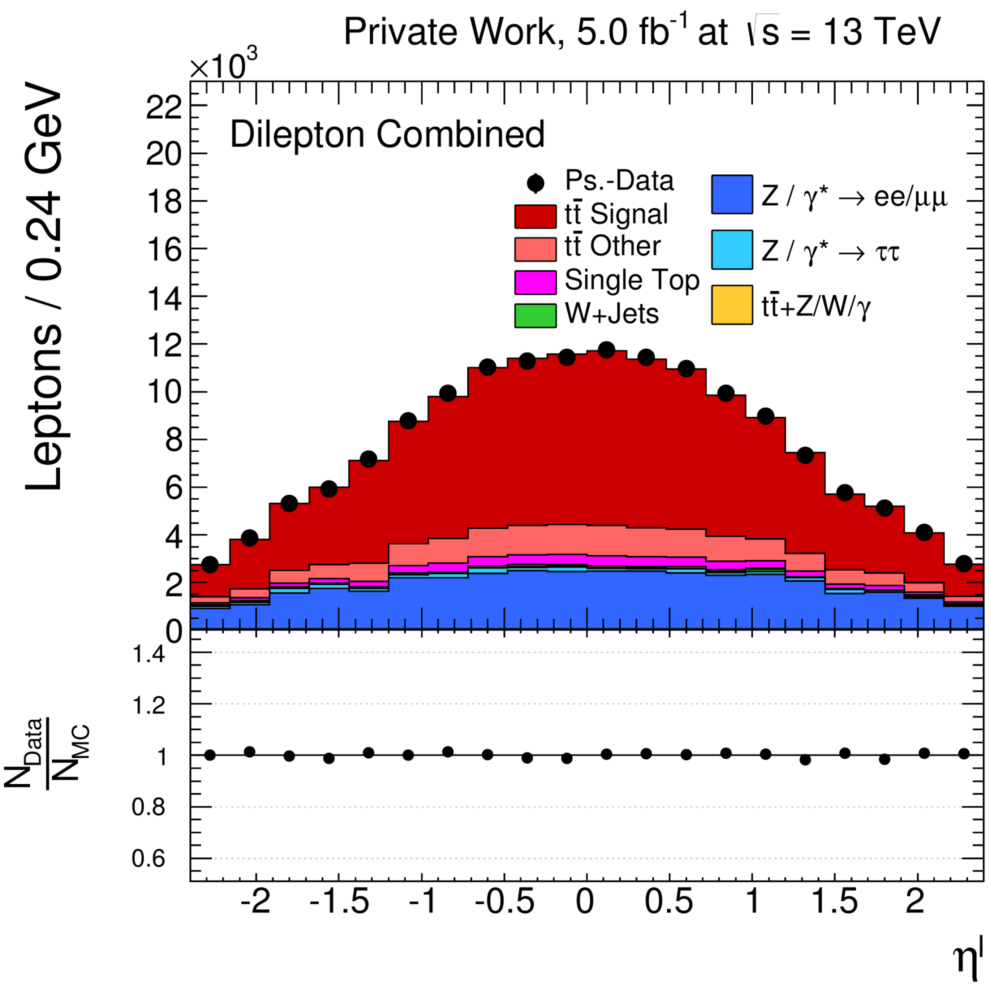
Selection in Dilepton Channel

Channel: combined (ee+eμ+μμ)
Pseudo-data: poisson-smeared sum of all signal and background simulation samples
Normalized to: L_{int} = 5.0 fb^{-1}

Observable: Lepton η
Signal: MadGraph+Pythia8

Selection steps:

1. Trigger conditions
2. Lepton pair selection:
   - opposite charge
   - e and μ isolation criteria
   - p_T > 20 GeV, |η| < 2.4
   - invariant mass: m_{ll} > 20 GeV
3. Exclusion of Z -region:
   - in ee and μμ: 76 GeV < m_{ll} < 106 GeV
4. Presence of two jets:
   - with p_T > 30 GeV and within |η| < 2.4



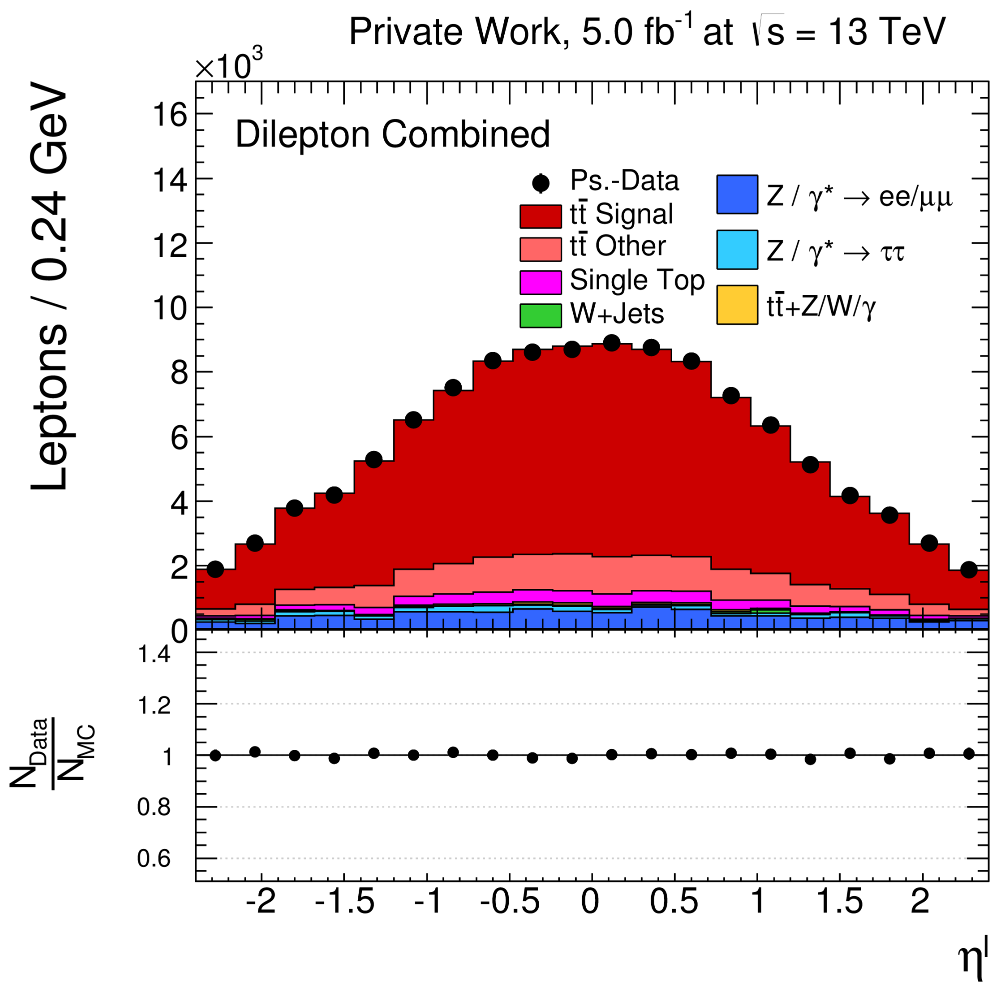
Selection in Dilepton Channel

Channel: combined (ee+eμ+μμ)
Pseudo-data: poisson-smeared sum of all signal and background simulation samples
Normalized to: L_{int} = 5.0 fb^{-1}

Observable: Lepton η
Signal: MadGraph+Pythia8

Selection steps:

1. Trigger conditions
2. Lepton pair selection:
   - opposite charge
   - e and μ isolation criteria
   - p_T > 20 GeV, |η| < 2.4
   - invariant mass: m_{ll} > 20 GeV
3. Exclusion of Z -region:
   - in ee and μμ: 76 GeV < m_{ll} < 106 GeV
4. Presence of two jets:
   - with p_T > 30 GeV and within |η| < 2.4
5. Missing E_T > 40 GeV in ee and μμ



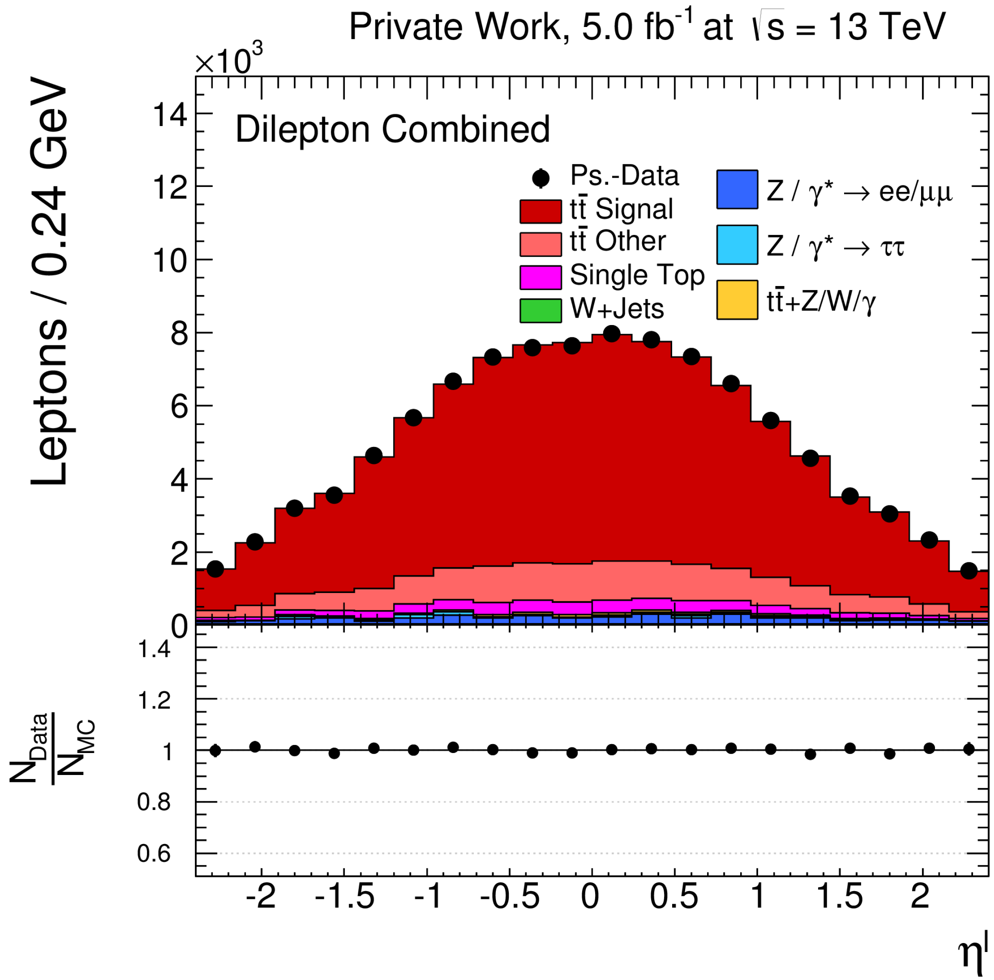
Selection in Dilepton Channel

Channel: combined (ee+eμ+μμ)
Pseudo-data: poisson-smeared sum of all signal and background simulation samples
Normalized to: $L_{int} = 5.0\ fb^{-1}$

Observable: Lepton η
Signal: MadGraph+Pythia8

Selection steps:

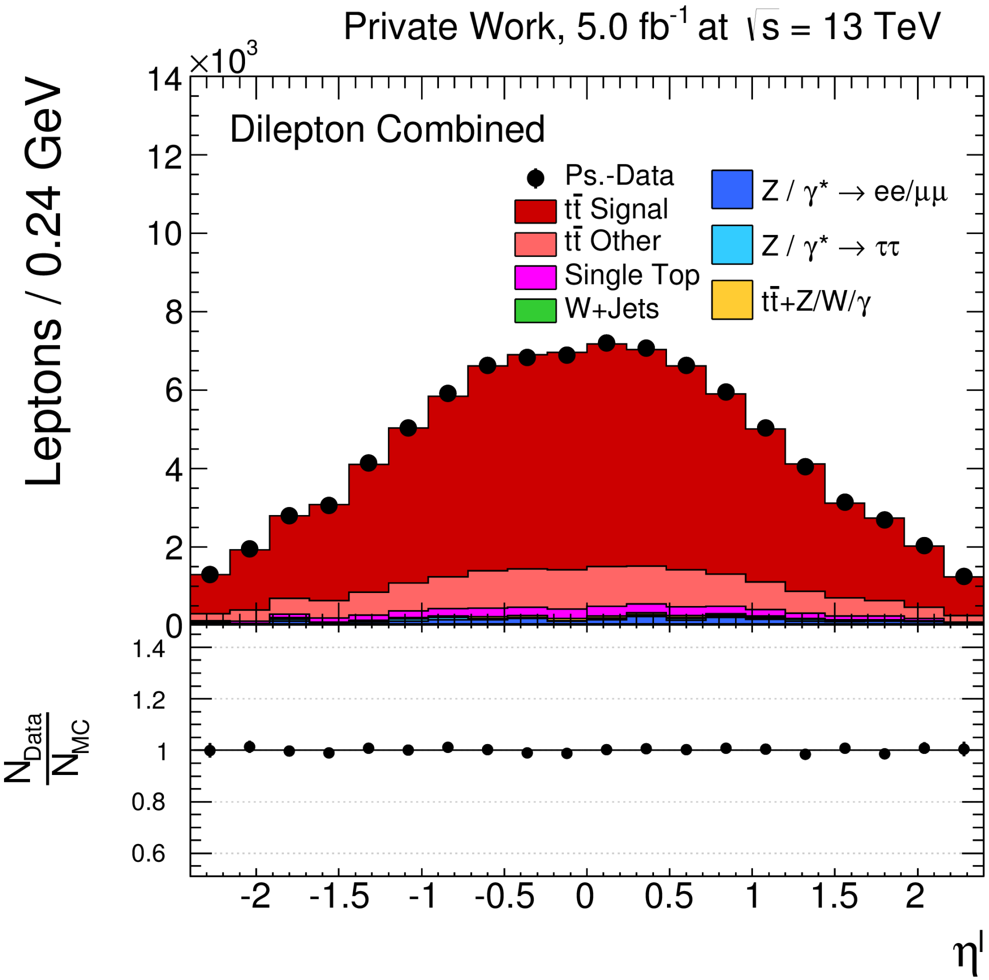
1. Trigger conditions
2. Lepton pair selection:
   - opposite charge
   - e and μ isolation criteria
   - $p_T > 20$ GeV, $|\eta| < 2.4$
   - invariant mass: $m_{ll} > 20$ GeV
3. Exclusion of Z -region:
   - in ee and μμ: 76 GeV < m_{ll} < 106 GeV
4. Presence of two jets:
   - with $p_T > 30$ GeV and within $|\eta| < 2.4$
5. Missing $E_T > 40$ GeV in ee and μμ
6. At least one b−tagged jet



Selection in Dilepton Channel

Channel: combined (ee+eμ+μμ)
Pseudo-data: poisson-smeared sum of all signal and background simulation samples
Normalized to: $L_{int} = 5.0\ fb^{-1}$

Observable: Lepton η
Signal: MadGraph+Pythia8



Selection steps:

1. Trigger conditions
2. Lepton pair selection:
   - opposite charge
   - e and μ isolation criteria
   - p_T > 20 GeV, |η| < 2.4
   - invariant mass: m_{ll} > 20 GeV
3. Exclusion of Z -region:
   - in ee and μμ: 76 GeV < m_{ll} < 106 GeV
4. Presence of two jets:
   - with p_T > 30 GeV and within |η| < 2.4
5. Missing E_T > 40 GeV in ee and μμ
6. At least one b−tagged jet
7. Meaningful solution for kinematic event reconstruction

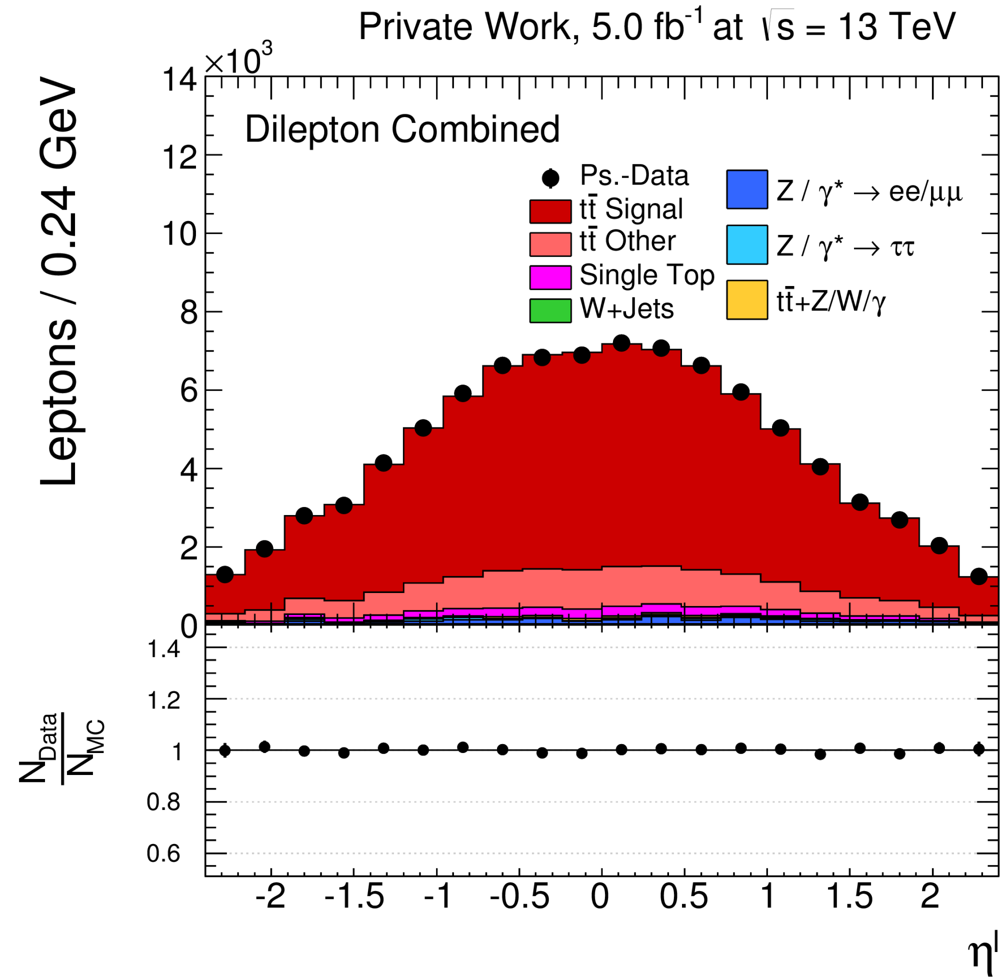
Selection in Dilepton Channel

Channel: combined (ee+eμ+μμ)
Pseudo-data: poisson-smeared sum of all signal and background simulation samples
Normalized to: L_{int} = 5.0 fb^{-1}

Observable: Lepton η
Signal: MadGraph+Pythia8

Table given after all cuts:

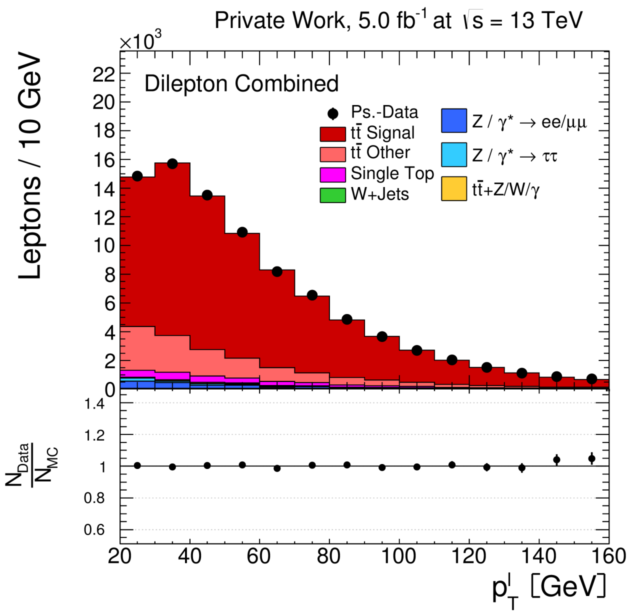
Process	Fraction, %
$t\bar{t}$ Signal	78.7
$t\bar{t}$ Other	14.1
tW	3.3
W+Jets	0.2
DY → ee/μμ	2.4
DY → ττ	0.6
$t\bar{t}$+Z/W/γ	0.7

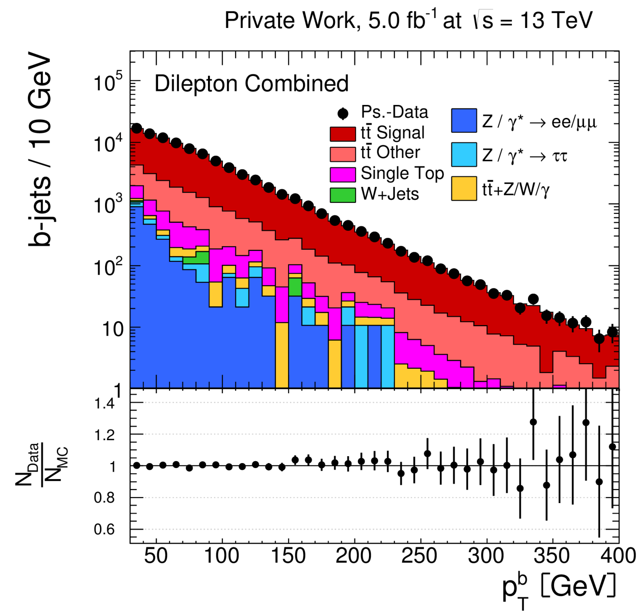


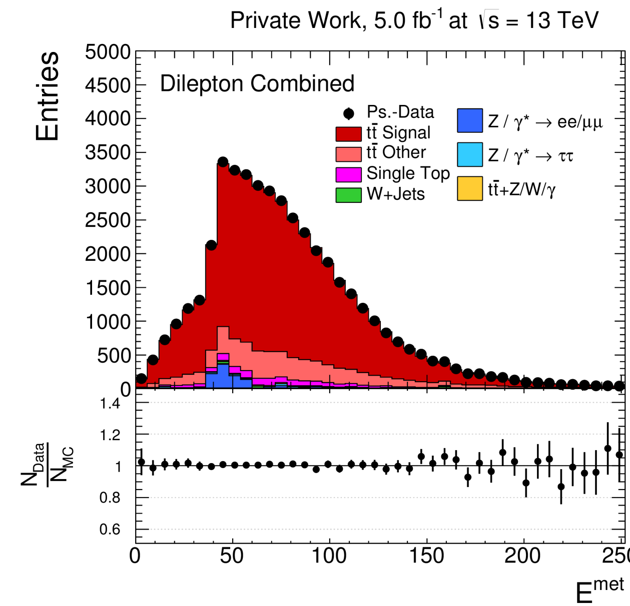
Control Distributions

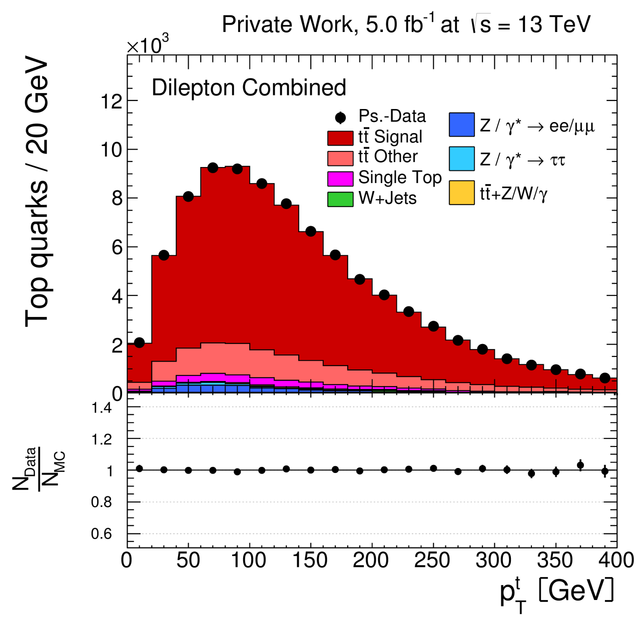
Channel: combined (ee+eμ+μμ) , $L_{int} = 5.0\ fb^{-1}$

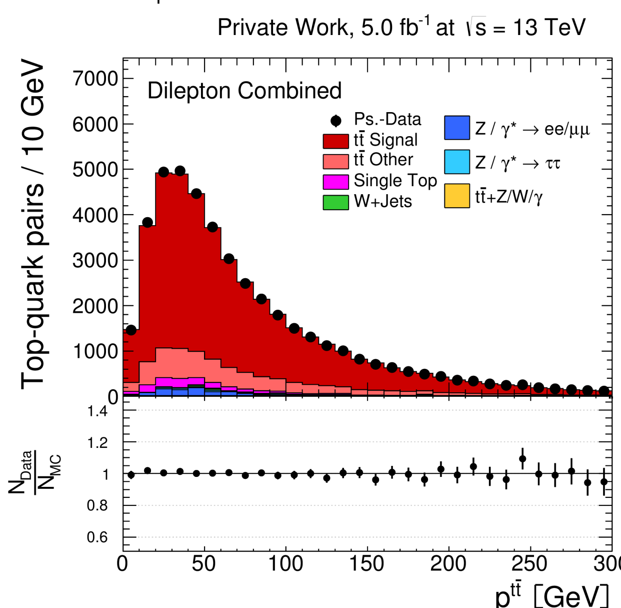
Plots given after full selection











Pseudo-data: poisson-smeared sum of all signal and background simulation samples

Differential Cross Section

- For each observable X the **normalized differential cross section** in the *i*-th bin is defined as:

$$\left(\frac{1}{\sigma}\frac{d\sigma}{dX}\right)^i = \frac{1}{\sigma}\frac{N^i_{events}}{\Delta^i_X L}$$

N^i_{events} - number of events after background subtraction, efficiency, acceptance and bin-to-bin migration correction

σ - total $t\bar{t}$ cross section in same phase space

L - integrated luminosity

Δ^i_X - bin width

- **Phase space** definition:

  Top quarks, $t\bar{t}$ system (obtained via **kinematic reconstruction of event**) – extrapolated to full phase space after corrections for detector and hadronization effects

  Leptons or b-jets – measured in visible phase space (*leptons*: $p_T > 20$ GeV, $|\eta| < 2.4$; *jets*: $p_T > 30$ GeV, $|\eta| < 2.4$) after correction for detector effects

- Bin-to-bin migrations are reduced by **binning optimization** and corrected by **unfolding**

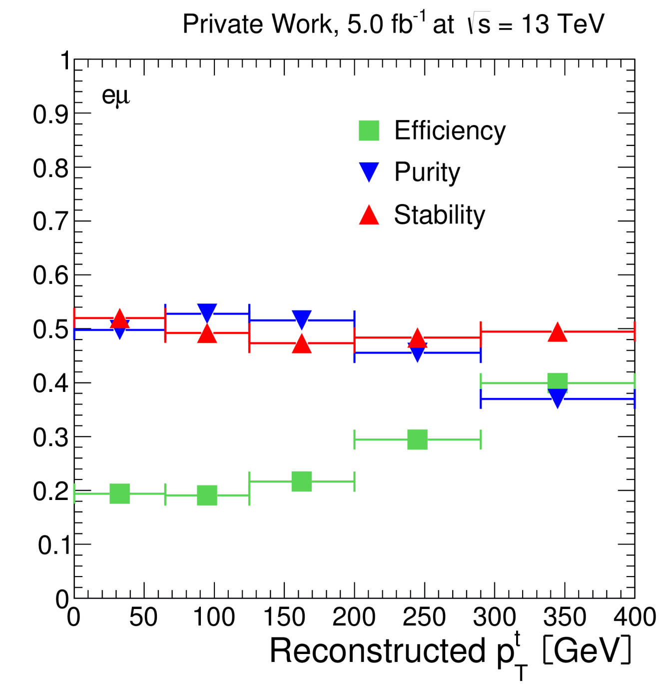
Binning and Migrations

- **Migration effects** studied by:

$$p_i = \frac{N_i^{rec\,\&\,gen}}{N_i^{rec}}$$ - **purity**: sensitive to migrations to *i*-th bin

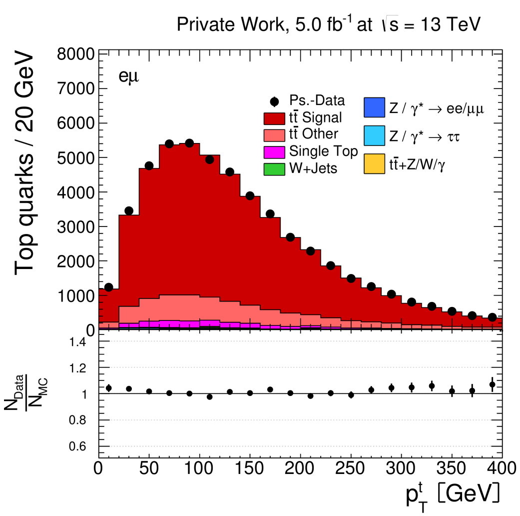
$$s_i = \frac{N_i^{rec\,\&\,gen}}{N_i^{gen}}$$ - **stability**: sensitive to migrations out of *i*-th bin

$$\varepsilon_i = \frac{N_i^{rec\,\&\,sel}}{N_i^{all\ generated}}$$ - **efficiency** in *i*-th bin

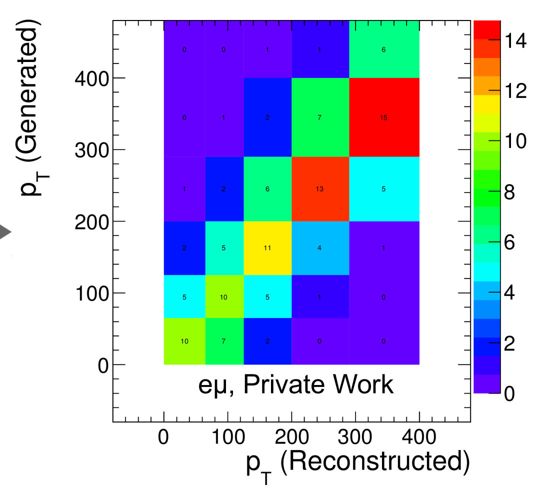


- **Binning criteria:** stability or purity ≥ ~0.5

- **Response matrices** are constructed from signal MC

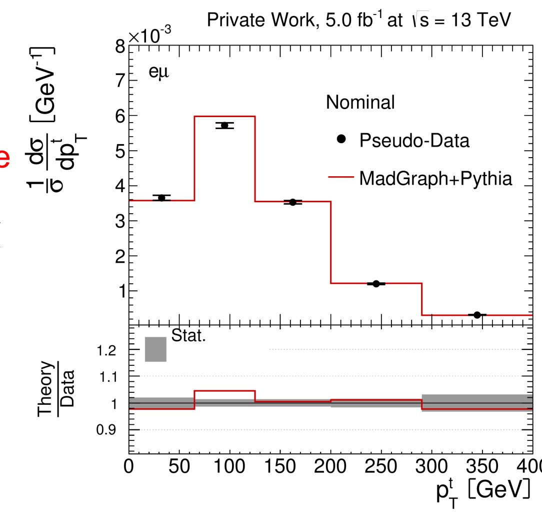


rebin

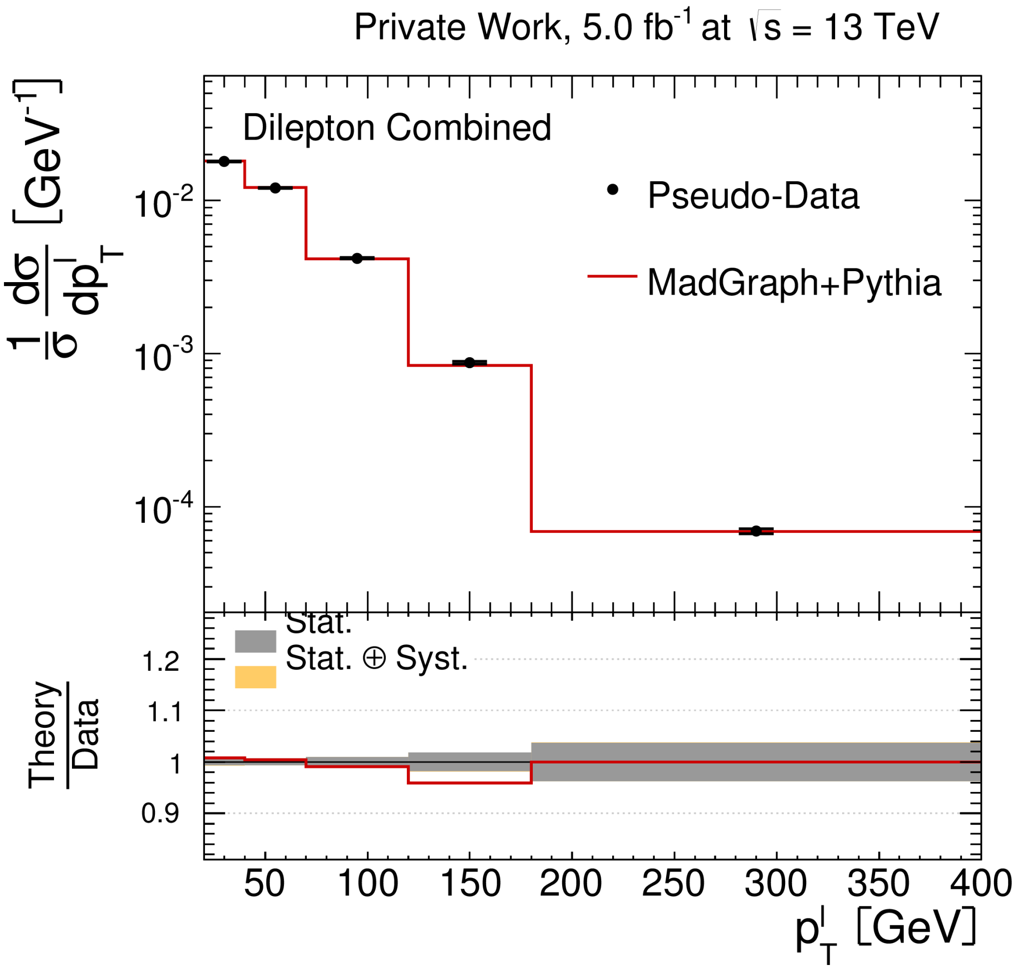


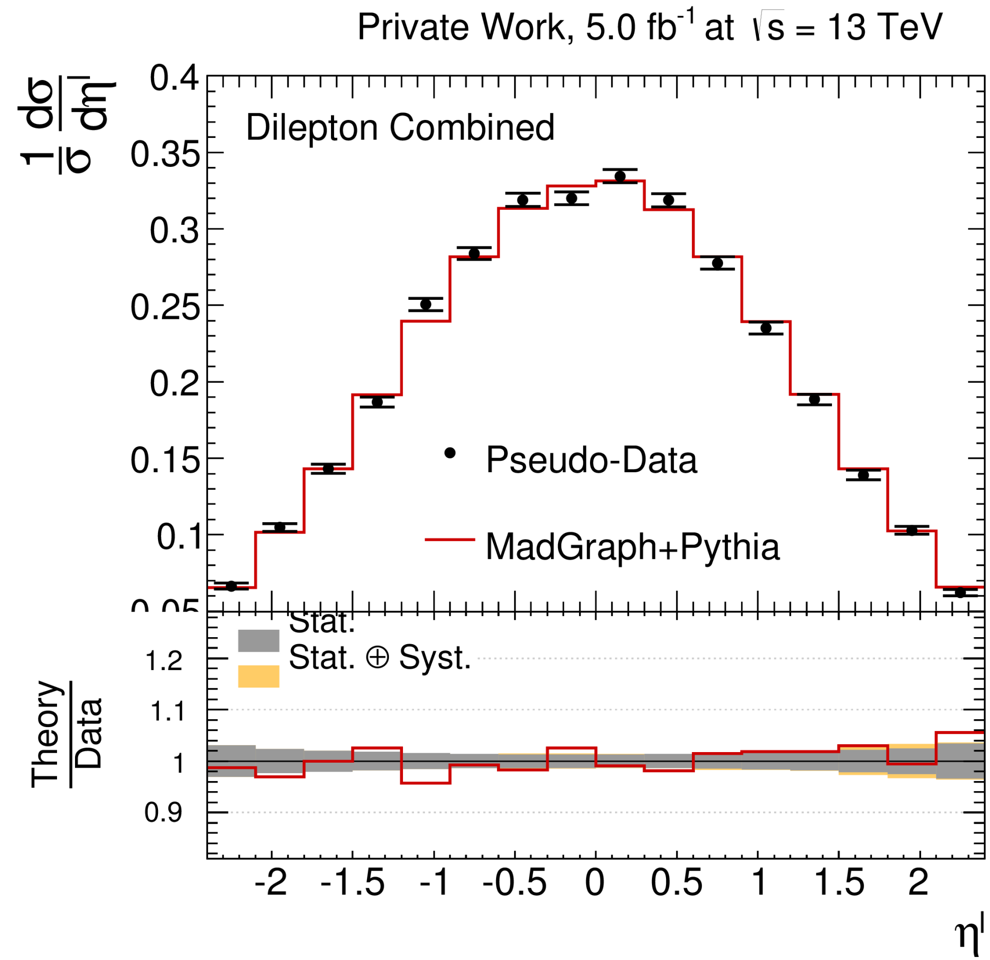
phase space

unfolding



Measured Cross Sections: leptons



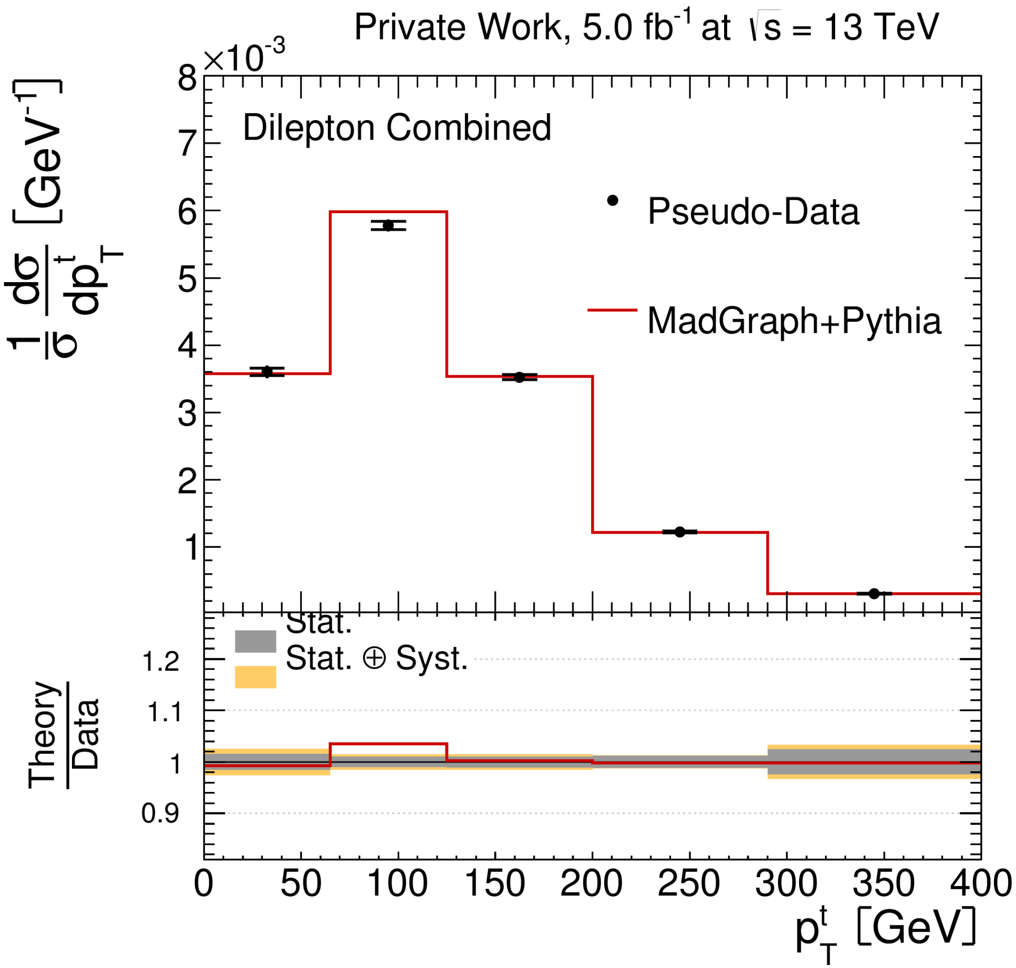


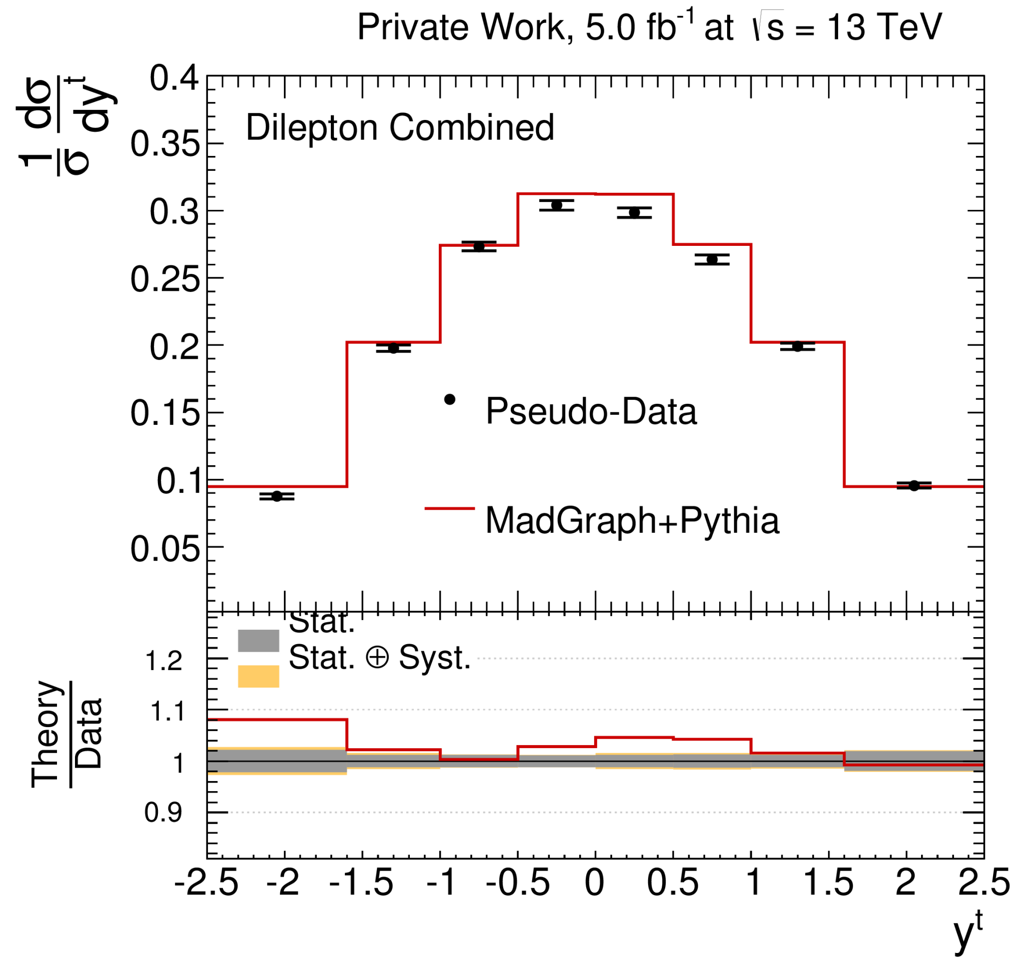
- **Normalization allows to:**

  - reduce systematic uncertainties
  - perform shape comparisons of different theory models to data

- **Systematic uncertainty:** include JES and JER uncertainties for now

Measured Cross Sections: tops



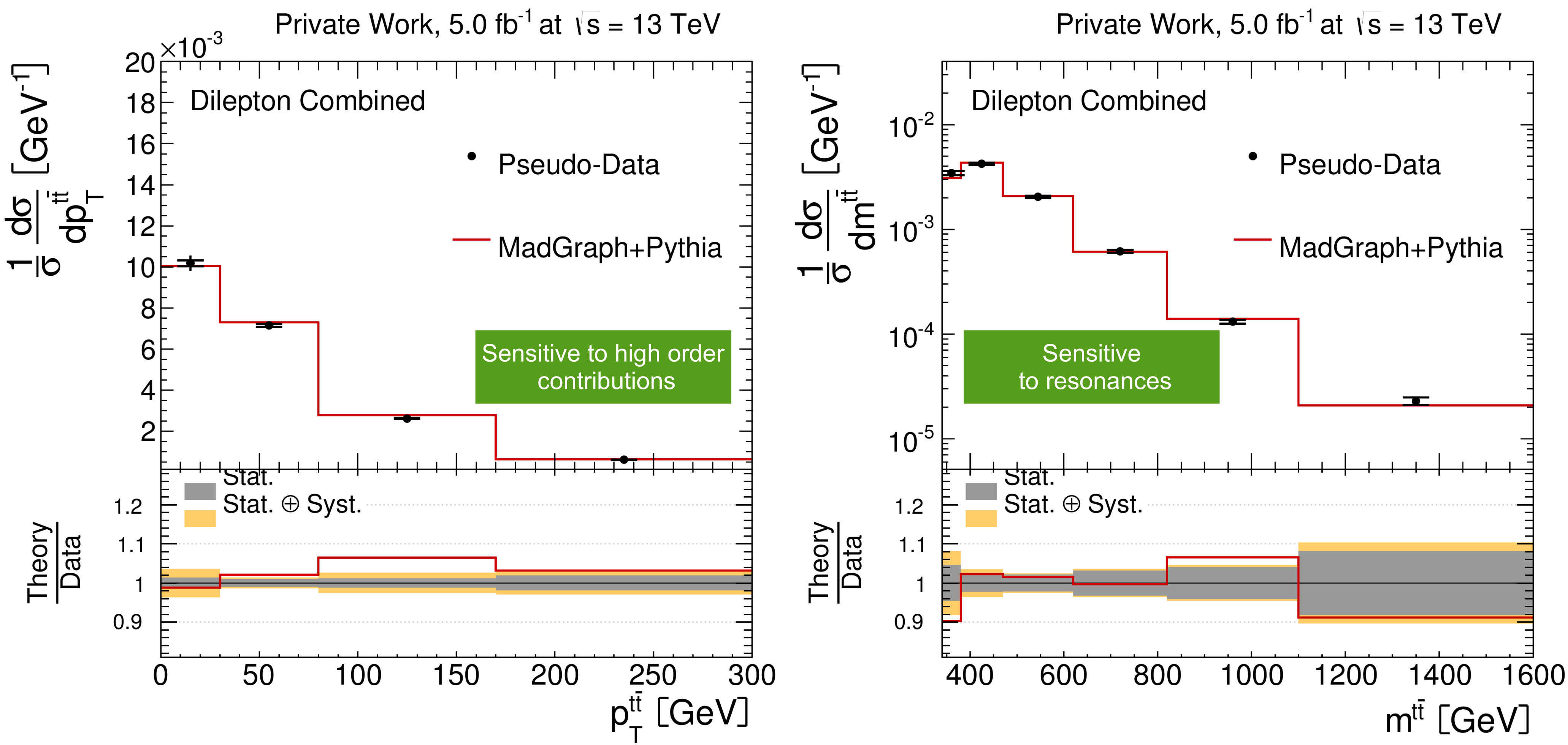


- **Normalization allows to:**

  - reduce systematic uncertainties
  - perform shape comparisons of different theory models to data

- **Systematic uncertainty:** include JES and JER uncertainties for now

Measured Cross Sections: $t\bar{t}$-pair



- **Normalization allows to:**

  - reduce systematic uncertainties
  - perform shape comparisons of different theory models to data

- **Systematic uncertainty:** include JES and JER uncertainties for now

Conclusions & Outlook

- **Top-quark pair differential cross section measurements:**

  - Essential for constraining the SM

  - Ideal probe for looking for new physics beyond the SM

  - Needed for tuning of PDF sets and modern art Monte-Carlo event generators

- **First studies towards measurements at 13 TeV were presented:**

  - Not latest status & results → currently **Top** secret at CMS and work in progress!

  - Optimize binning with final configuration of the data analysis

  - Evaluate systematic uncertainties

  - Compare to different theory predictions

Conclusions & Outlook

- **Top-quark pair differential cross section measurements:**

  - Essential for constraining the SM

  - Ideal probe for looking for new physics beyond the SM

  - Needed for tuning of PDF sets and modern art Monte-Carlo event generators

- **First studies towards measurements at 13 TeV were presented:**

  - Not latest status & results → currently **Top** secret at CMS and work in progress!

  - Optimize binning with final configuration of the data analysis

  - Evaluate systematic uncertainties

  - Compare to different theory predictions

Thank you for your attention! :-)

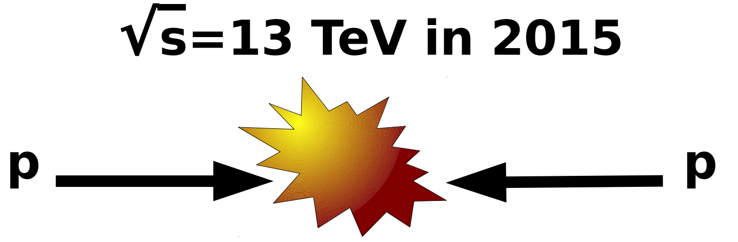
Backup

LHC Collider

Run 1: $L_{int} \approx 25\ fb^{-1}$ in 2010-2012 at 7/8 TeV

Run 2: Phase 0 in 2015-2017
- design energy: $\sqrt{s}$=13 TeV~14 TeV
- nominal luminosity: $L \sim 1 \cdot 10^{34} cm^{-2} s^{-1}$
- bunch spacing: 25 ns
- pile up: ~49
- L_{int} per year ~ 45 fb^{-1}

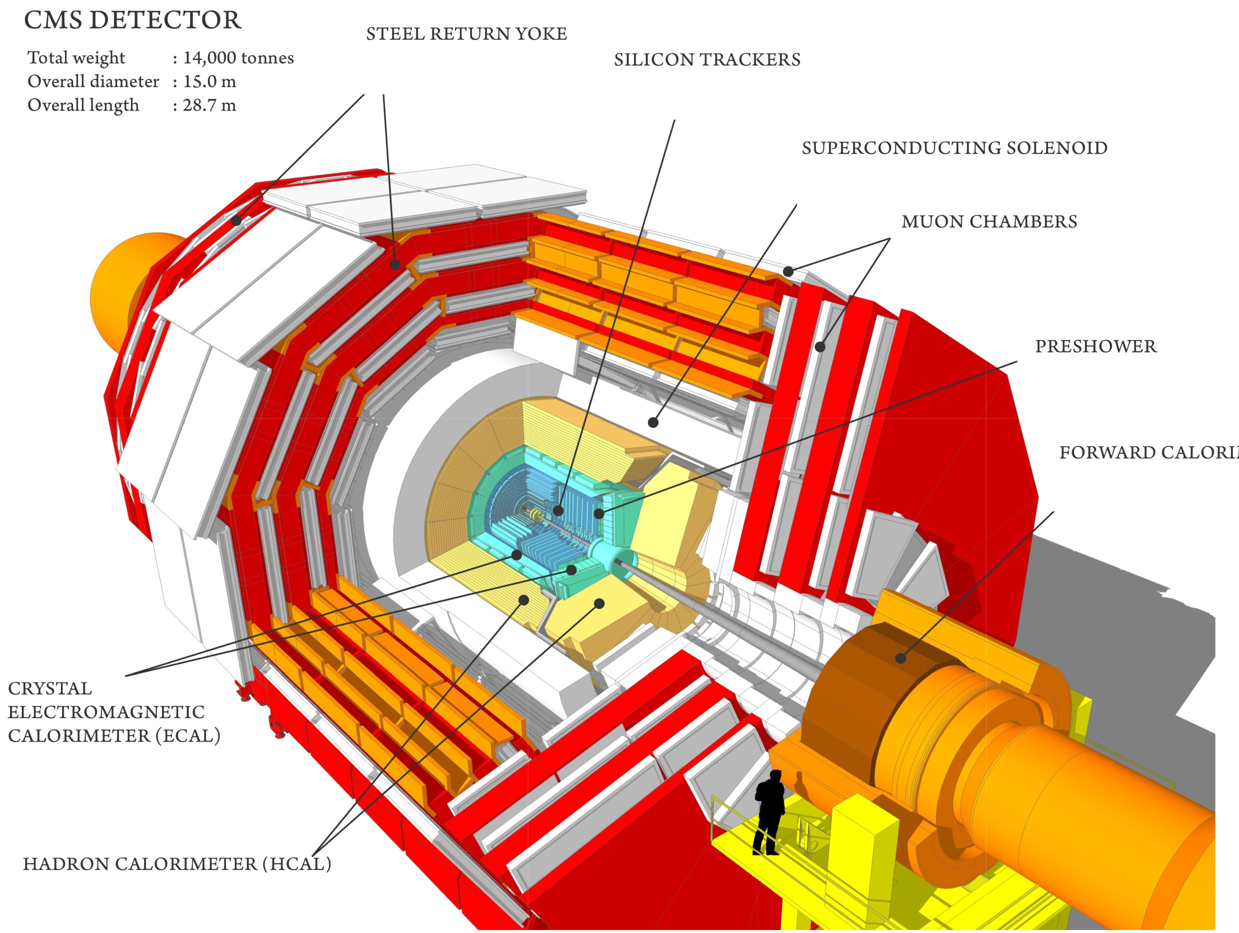
Planned to collect: $L_{int} \sim 75 – 100\ fb^{-1}$



The CMS Detector

General purpose 4π detectors:

Tracker: Detection and momentum measurement for charged particles
Calorimeter: Identification and energy measurement of jets and electrons
Muon system: Identification and momentum measurement of muons



Phase 0 (2015-2017) upgrade:

- Complete muon coverage
- Colder tracker
- Photodetectors in HCAL
- New beampipe and infrastructure updates

Visible Phase Space Definition

- **Object definition at generator level:**
  - > particles after radiation and hadronization
  - > leptons: from W-decay
  - > jets: anti-kT (with cone of ΔR=0.5) algorithm
  - > b-jets: identified by B-hadrons
- **Directly measured quantities:** leptons and b-jets

Visible particle LVL Phase Space

↓

Correct for:
Detector Effects

↓

Measure at visible Phase Space:
Leptons: p_T > 20 GeV, $|\eta|$ < 2.4
Jets: p_T > 30 GeV, $|\eta|$ < 2.4

Unfolding

- **Unfolding** techniques correct migrations between bins
- **Response matrix (A):** represents bin-by-bin correlations
- Unfolding problem is transformed to χ^2 - minimization problem:

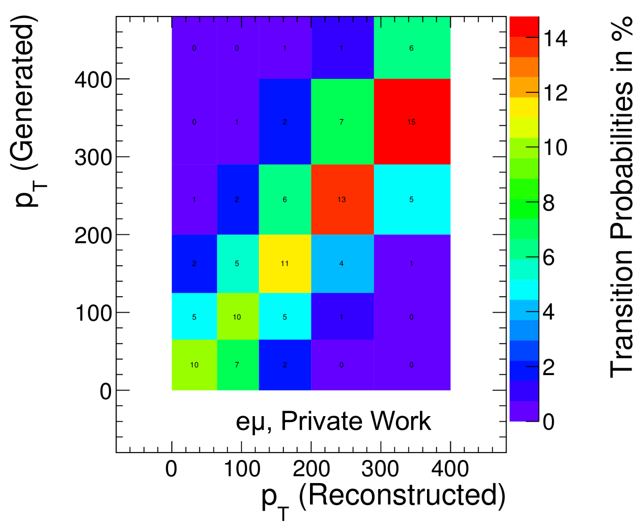
$$\chi^2 = \overbrace{(\vec{N} - A\cdot\vec{x})^T \, COV_{\vec{N}}^{-1} (\vec{N} - A\cdot\vec{x})}^{\text{unfolding}} - \overbrace{\tau^2 \cdot K(\vec{x})}^{\text{regularization}}$$

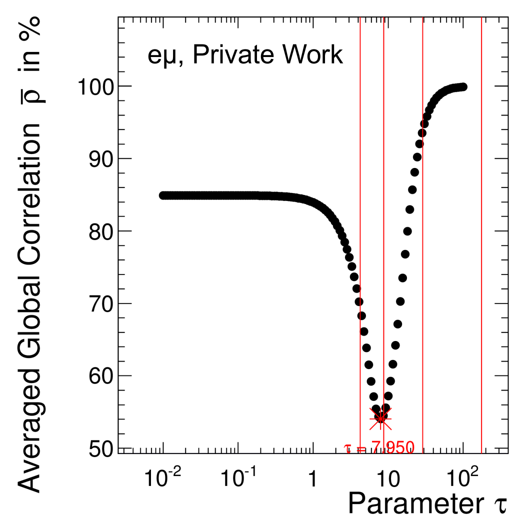
- **N:** BG corrected data
- **x:** unfolded result

- Non-physical fluctuations removed by means of the regularization:

  → τ – continuous regularization parameter

  → selected at minimum of average global correlation





- In this measurement regularized unfolding is used