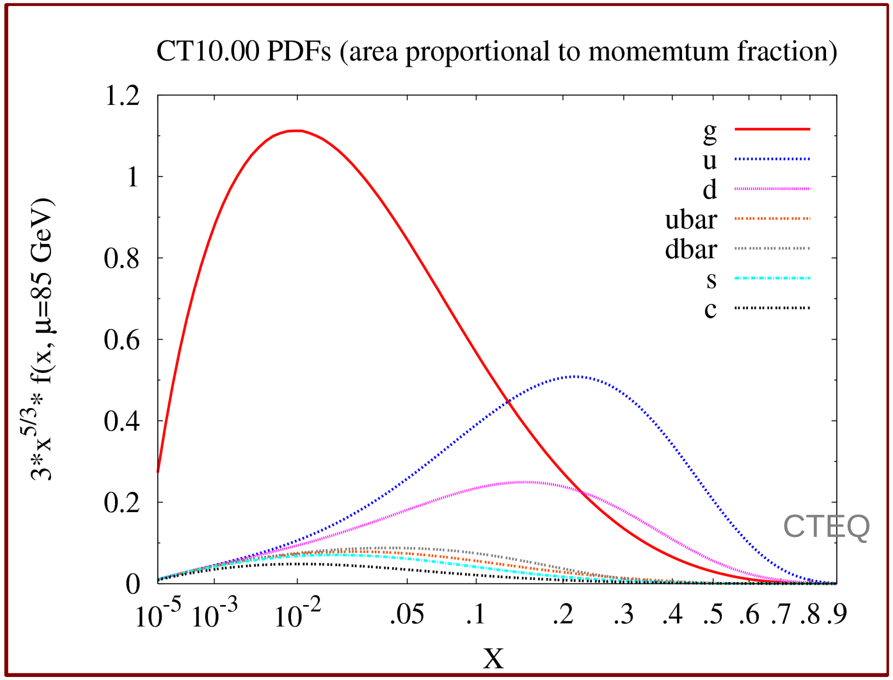
Off-Resonance Drell-Yan process at CMS/LHC

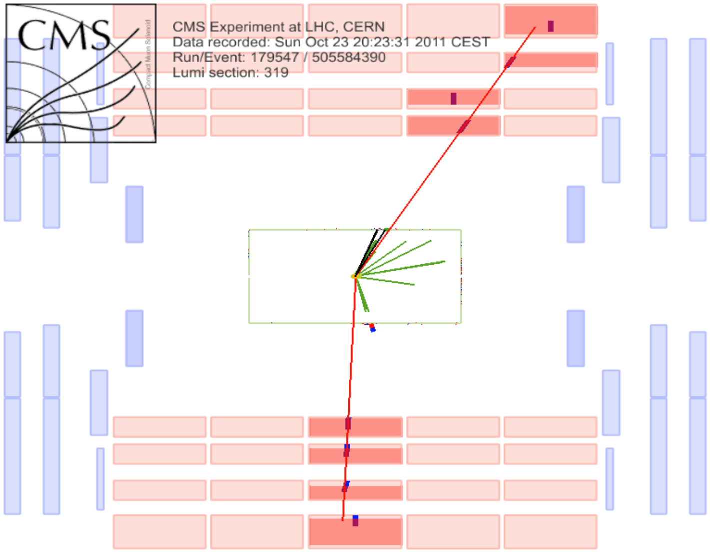
Stoyan Stoynev

Northwestern University

Outline

- Brief motivation for PDF studies
- Importance of DY in PDF fits
- Detector limitations at LHC
- (Off-Resonance DY) Physics program at CMS
- DY signal, backgrounds, resolution and "unfolding" in measurements
- Systematic sources to measurements
- Physics results (~4.5 fb^{-1}, preliminary)
- Sensitivity to PDFs
- Perspectives

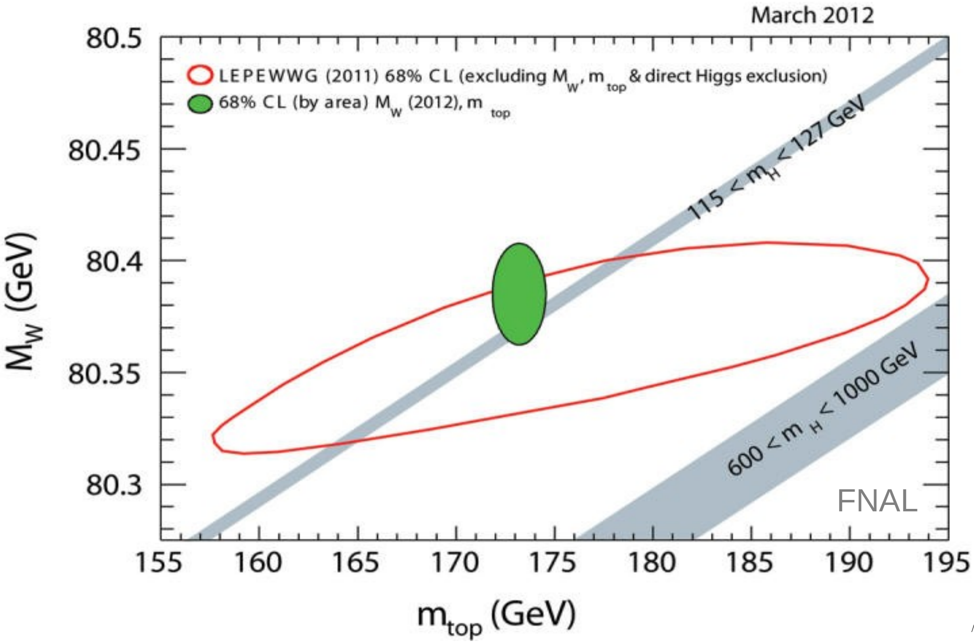




Motivation by example (for PDF studies)

- Precise determination of the Parton Distribution Functions (PDF) is crucial for understanding physics at hadron colliders
- Rigorous tests on the standard model rely heavily on PDF knowledge
- Constraints on New Physics are often driven by PDF uncertainties as well

Combined CDF/D0 result



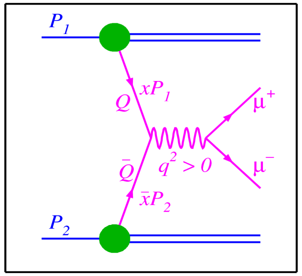
Transverse mass fit uncertainties
(CDF W mass measurement*, 2.2 fb-1)

Systematic (MeV)	Electrons	Muons	Common
Lepton Energy Scale	10	7	5
Lepton Energy Resolution	4	1	0
Recoil Energy Scale	5	5	5
Recoil Energy Resolution	7	7	7
$u_{\|\|}$ Efficiency	0	0	0
Lepton Removal	3	2	2
Backgrounds	4	3	0
$p_T(W)$ Model (g_2, g_3, α_s)	3	3	3
Parton Distributions	10	10	10
QED Radiation	4	4	4
Total	18	16	15

Dominated by PDF uncertainties

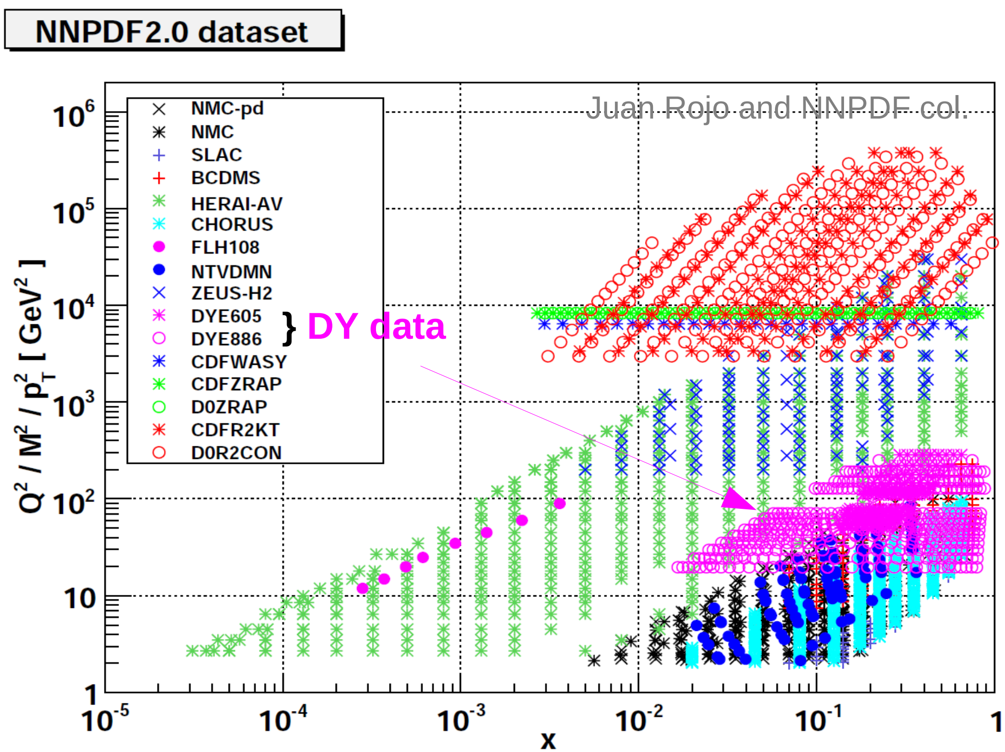
*http://www-cdf.fnal.gov/physics/ewk/2012/wmass/wmass_conf.pdf

Drell-Yan process and constraints on PDFs

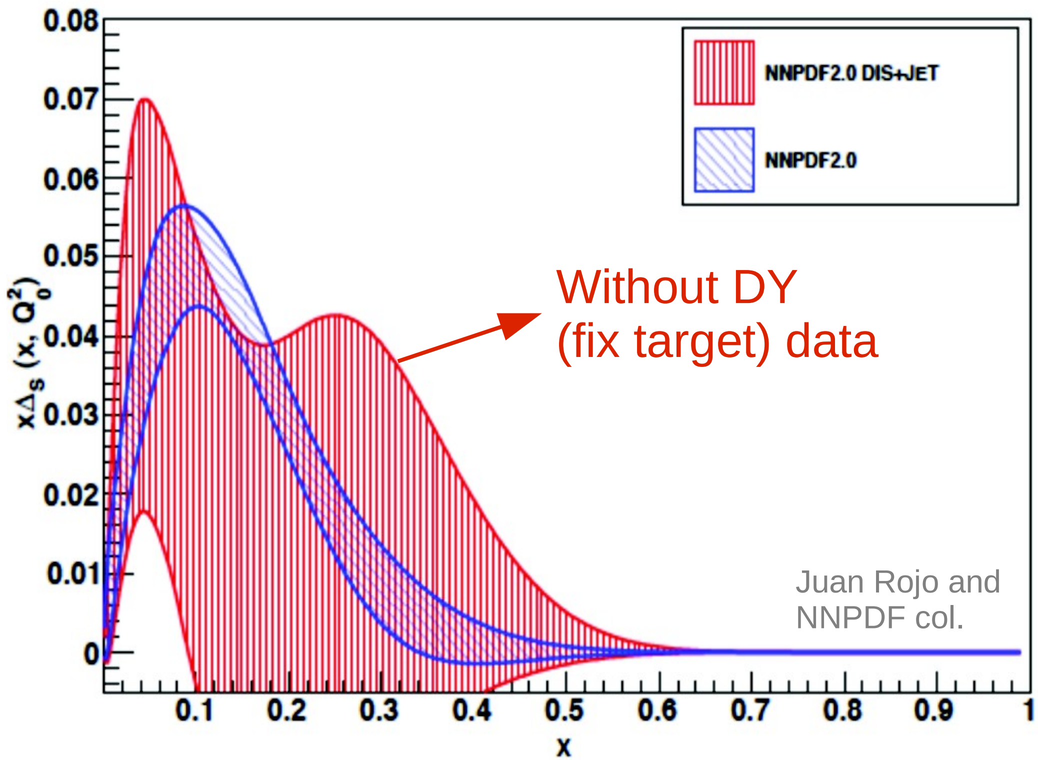


Ian C. Brock

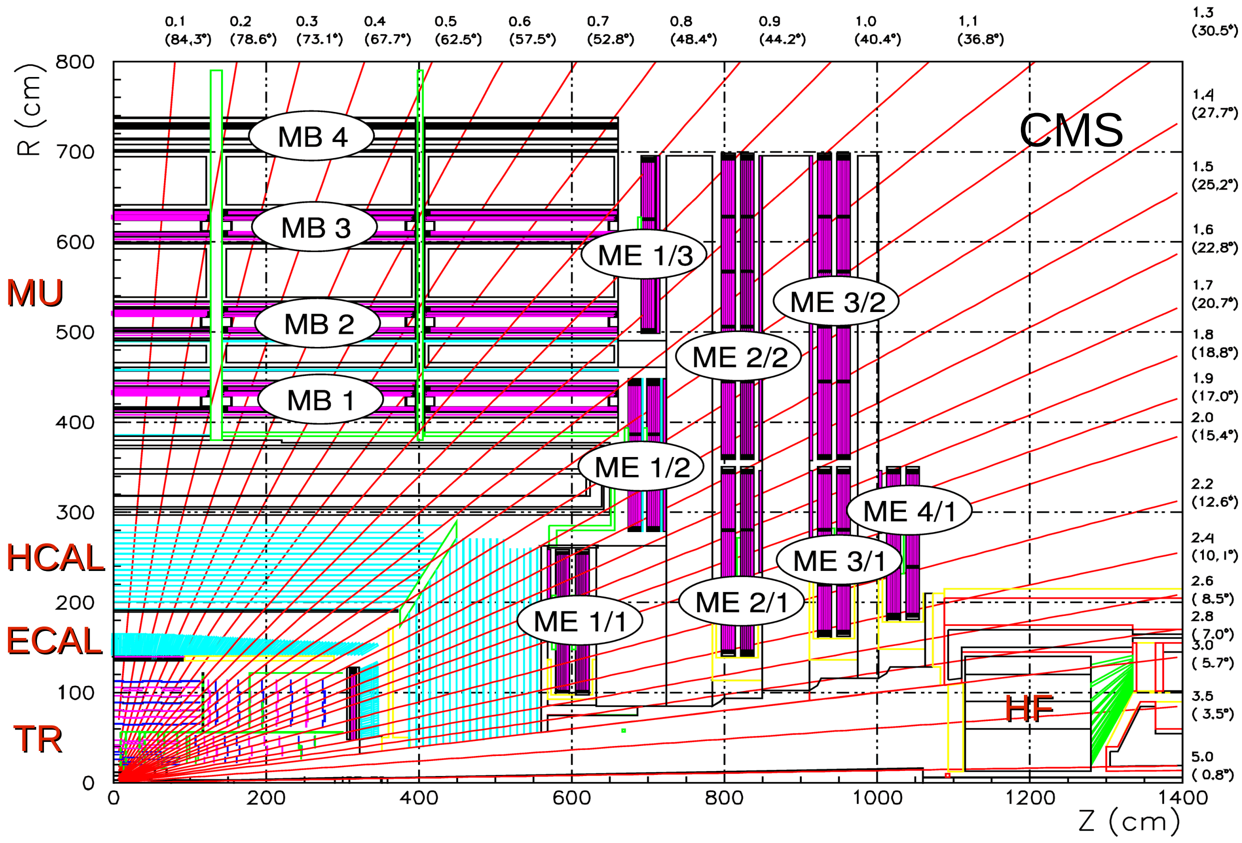
- DY fixed target experiments are a core ingredient in the current PDF fits
- Collider DY experiments are destined to play a major role partially substituting fix target results (avoiding the nuclear uncertainties)
- $x_{\pm}=\frac{m_{ll}}{\sqrt{s}}e^{(\pm y)}$: CMS/ATLAS can cover x ranges from 10^{-4} (or even 10^{-5}) to almost 1; LHCb can reach down to 10^{-6}

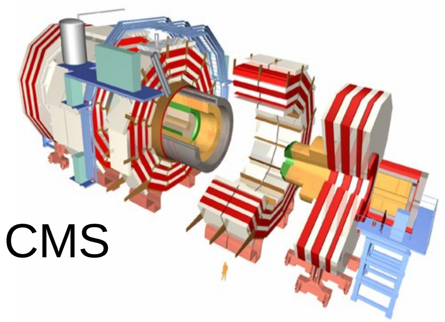


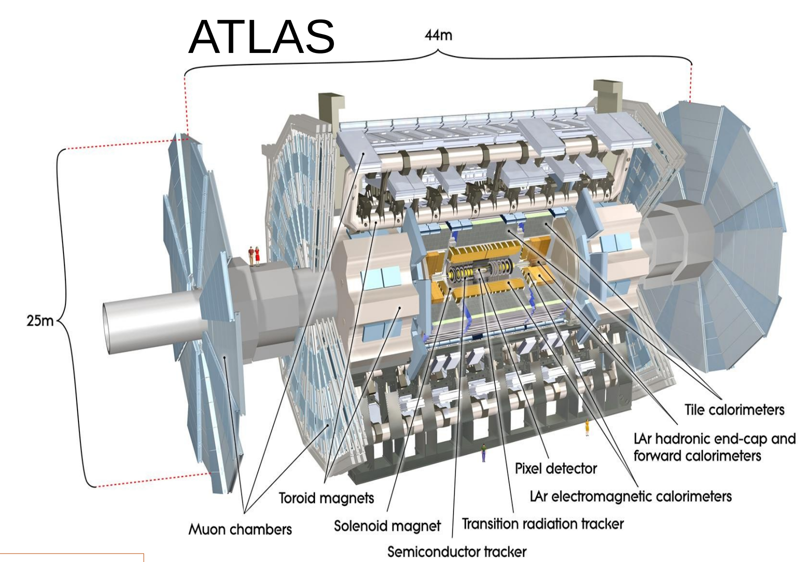
Sea asymmetry : $\Delta_s=\bar{d}-\bar{u}$

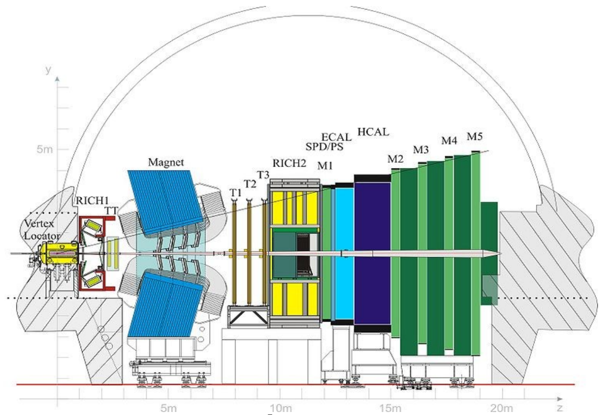


Detector reach









- CMS/ATLAS can detect leptons upto 2.4-2.5 in pseudorapidity (η=-log[tan(θ/2)]) and, effectively, in di-lepton rapidity (Y).
- CMS has the option to explore the HF calorimeter for electron reconstruction upto 5 in η (lower resolution than ECAL).
- LHCb operates in the region 1.9<η<4.9

Basic lepton selection in DY analyses (CMS)

MUONS

Main properties are driven by the tracker and muon sub-detectors.

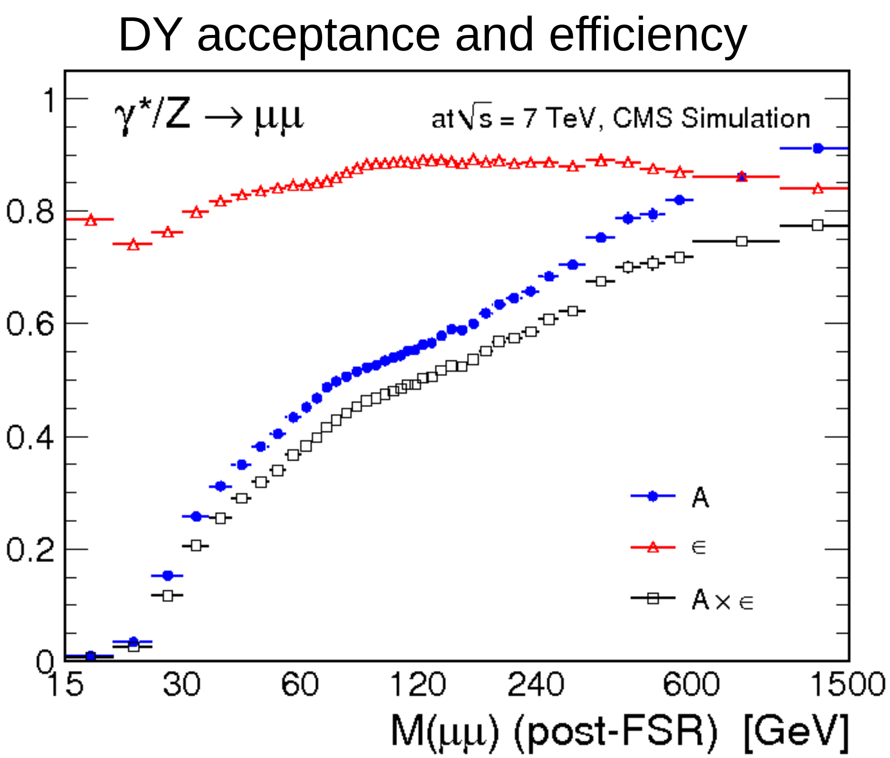
- Standard CMS reconstruction and selection providing quality muons and reducing background contributions - ~ 95% efficient
- Use Particle Flow algorithm in CMS to impose isolation requirements - ~ 99% efficient (momentum dependent)
- Double-muon trigger requirements (wrt previous selection) - >95% efficient
- Di-lepton requirements (vertex, collinearity)- ~ 98% efficient
- The muons are reconstructed upto 2.4 in η and transverse momentum (pT) requirements are chosen to avoid trigger inefficiency : 14 and 9 GeV for the two muons

Electrons

Main properties are driven by the electro-magnetic calorimeter (ECAL).

- Standard CMS reconstruction and selection providing quality electrons and reducing background contributions - ~80% (70% at low momenta) efficient
- Use Particle Flow algorithm in CMS to impose isolation requirements - ~ 98% efficient (momentum dependent)
- Double-electron trigger requirements (wrt to previous selection) - >98% efficient
- The electrons are reconstructed upto 2.5 in η and transverse momentum (pT) requirements are chosen to avoid trigger inefficiency : 20 and 10 GeV for the two electrons

DY Acceptance and efficiency (muons example)



At low invariant masses the acceptance is largely dominated by the lepton momentum requirements.

At high invariant masses it is solely the eta that restricts the acceptance.

Note the knee at invariant mass around the sum of the minimal pt of leptons : 14+9 GeV. At this point approximately the cross-section at leading order (LO) of the process vanishes.
Consequences:
lepton distributions drastically change, NLO is not enough to describe the full process with possible deviations from NNLO of tens of per-cents.

- Any acceptance correction (to the results) is highly dependent on theory at low invariant masses. The simulations should be typically re-weighted and this can not be done perfectly due to accuracy limitations (large log-corrections) and also MC statistical limitations. This introduces uncertainty due to “model dependency” .

- It makes sense to provide results in the detector acceptance and only IN ADDITION the results in the full phase space which should be considered harder to use for extracting constraints.

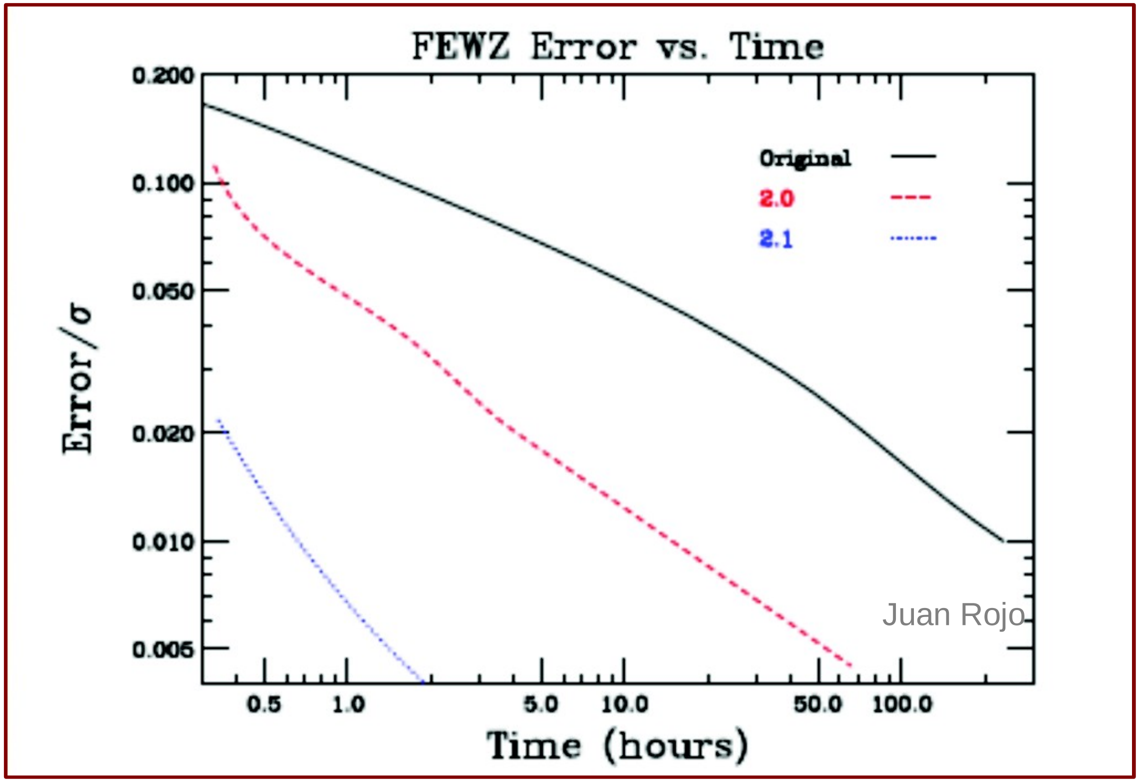
Theory code

- We use FEWZ* extensively in both NLO and NNLO
- We have verified (initially) that DYNNLO** gives compatible with FEWZ predictions and has similar to FEWZ capabilities (though using different theory methods)
- We found it difficult to perform all the calculations in NNLO due to the huge resources and time consumption needed. However this is seemingly about to change with the new 2.1 version available.

Currently theory calculations for DY (in CMS) are performed with FEWZ 2.0 and PDF errors are very easy to compute.

Calculations in the acceptance regions are straightforward.

I'd like to thank Frank Petriello (FEWZ author) for the fruitful discussions and advices.

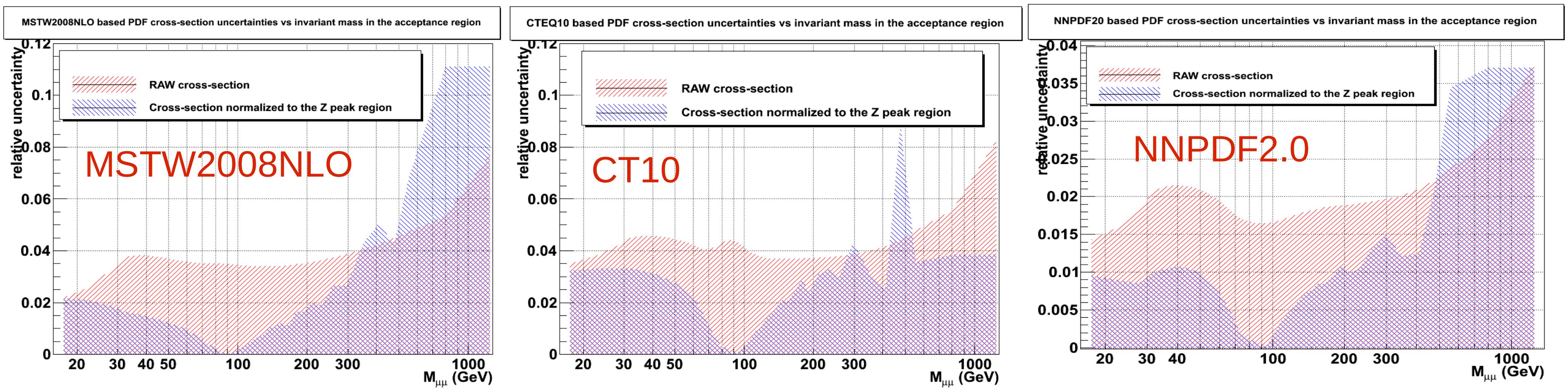


* Latest paper is: http://arxiv.org/abs/1201.5896
** See http://theory.fi.infn.it/grazzini/dy.html

(Off-Resonance DY) Physics program at CMS - 1

- Differential $d\sigma/dm$ cross-section measurements
  - start the measurement at invariant masses of 15 GeV and upto the highest reachable
  - for binning - take into account statistics, resolution, background, theory considerations

PDF uncertainties for different PDF sets vs invariant mass



For PDF constraints we benefit when we MAXIMIZE the above.

Apart from some integration error variations it is clear that the normalized to the Z peak region (60-120 GeV) cross-section is also a viable option to consider theoretically (except at the Z peak).

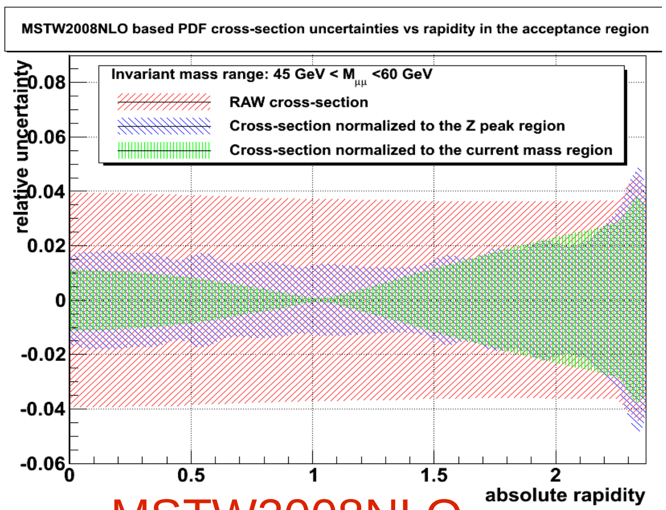
From experimental point of view working with absolute quantities is bad – the luminosity uncertainty alone is currently several percent (should drop to 2% in future).

We choose to report primary normalized results (shape analysis). We intend to provide full information about the normalization in future.

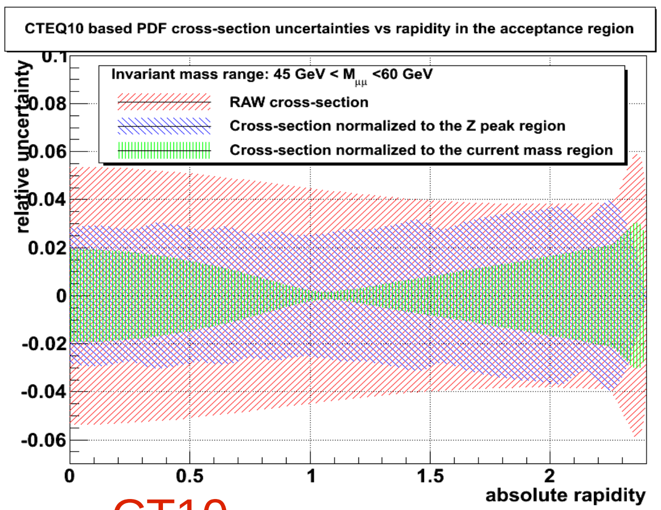
(Off-Resonance DY) Physics program at CMS - 2

- Double differential $d^2\sigma/dMdY$ cross-section measurements
  - start the measurement at invariant masses of 20 GeV and upto the highest reachable
  - for binning - take into account statistics, resolution, background, theory considerations

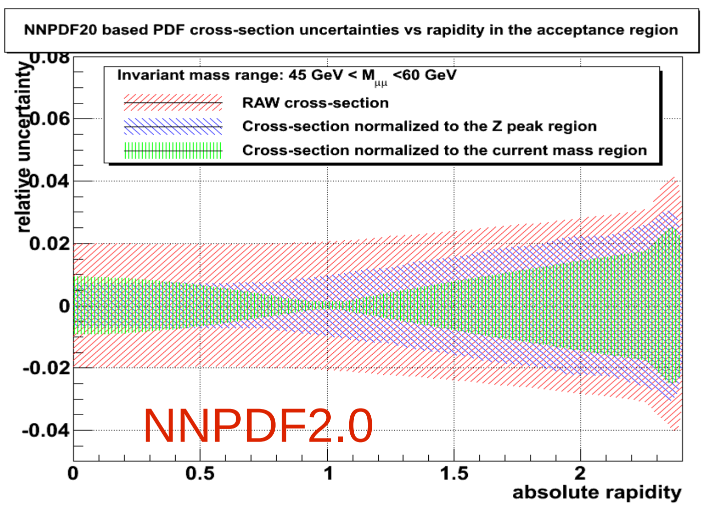
PDF uncertainties for different PDF sets vs rapidity in a given invariant mass range (45-60 GeV)

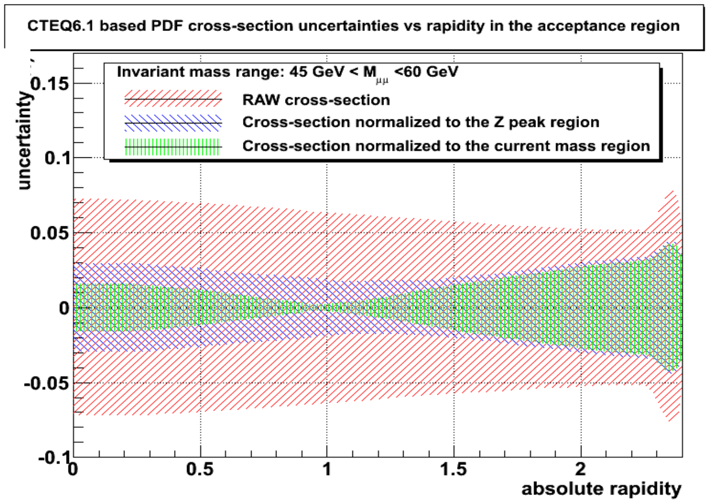


MSTW2008NLO



CT10





CTEQ6.1

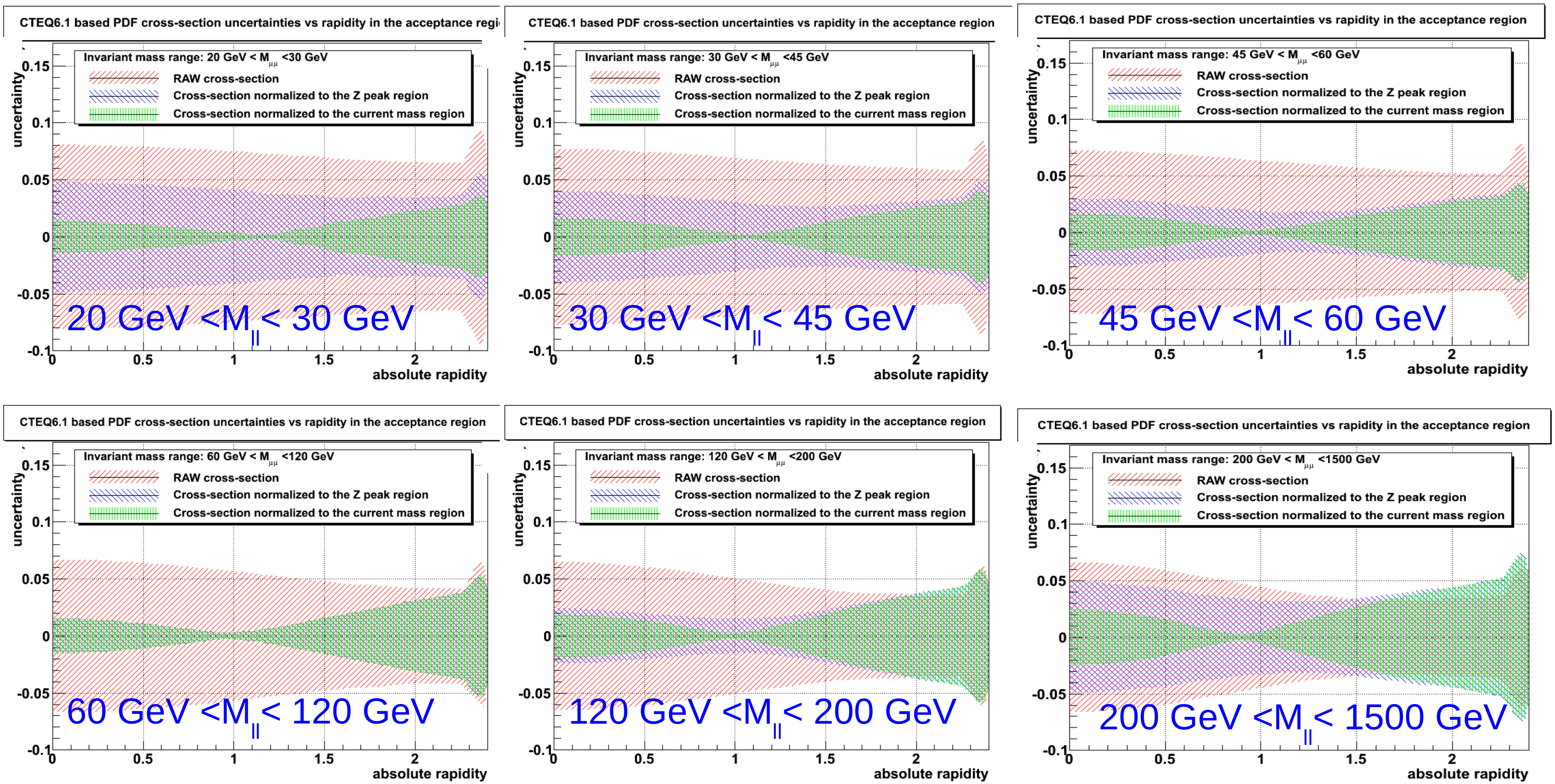
These plots show that normalization to the Z peak is better than normalization to the "current" mass region.

The plots are for predictions in the acceptance region.

(Off-Resonance DY) Physics program at CMS - 3

These plots show the trends vs mass and rapidity (also compare with the previous slide). Ultimately NNLO based PDFs are the primary goal.

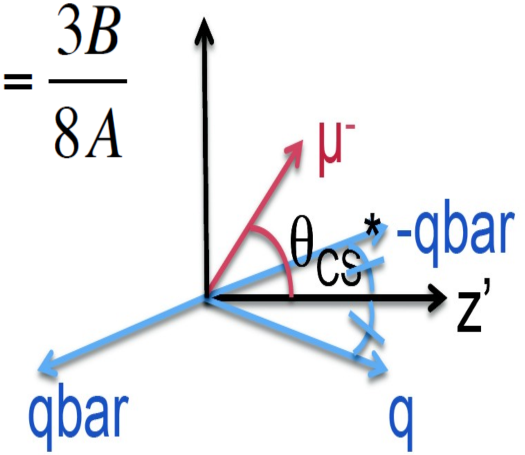
PDF uncertainties for CTEQ6.1 set vs rapidity in the considered mass ranges



(Off-Resonance DY) Physics program at CMS - 4

- Forward-Backward Asymmetry
  - start the measurement at invariant masses of 40 GeV and upto the highest reachable
  - for binning - take into account statistics, resolution, background, theory considerations

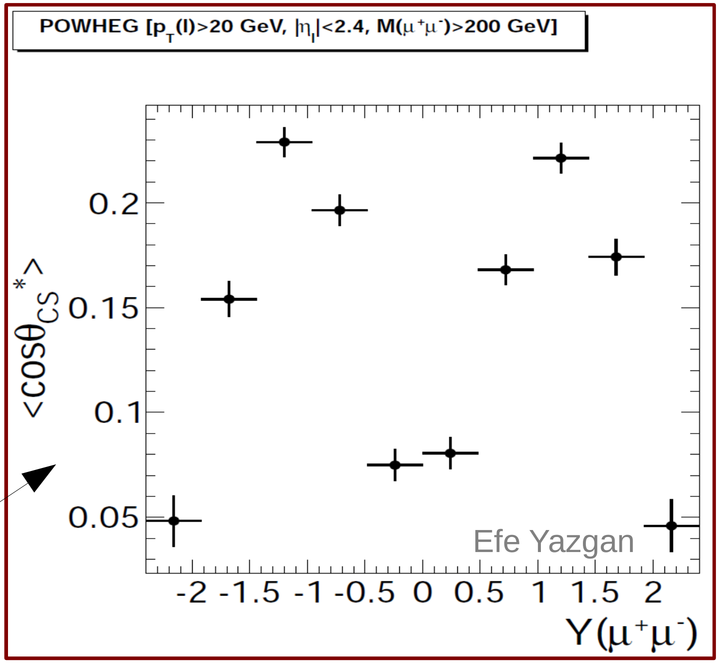
$$\frac{d\sigma}{d(\cos\theta^*)} = A(1+\cos^2\theta^*) + B\cos\theta^* \qquad A_{FB} = \frac{\sigma_F - \sigma_B}{\sigma_F + \sigma_B} = \frac{N_F - N_B}{N_F + N_B} = \frac{3B}{8A}$$



$$A_{FB} = A_{FB}(M, \sin^2\theta_W, q)$$

The measurement determines the forward-backward asymmetry of DY pairs vs invariant mass and rapidity.

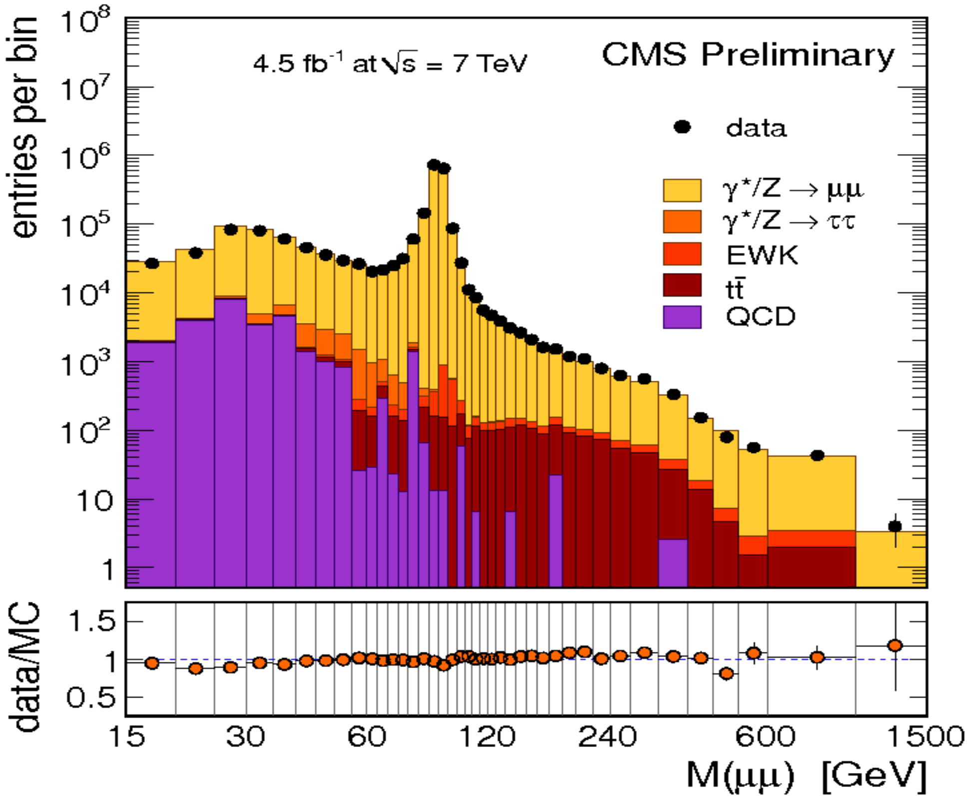
With systematic uncertainties under control it could be used in PDF fits together with the rest of the DY measurements. Fine binning in rapidity will need more statistics.

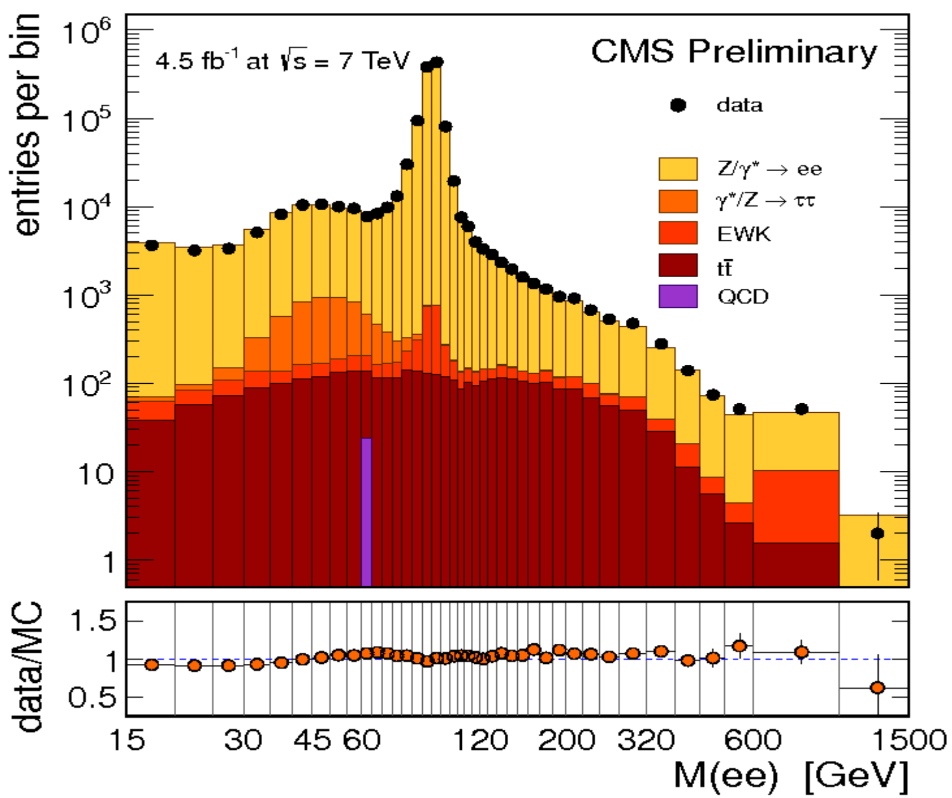


At high rapidity asymmetry decreases due to Y-acceptance

Signal and background yields

- The most significant sources of backgrounds are QCD processes, ttbar and DY$(\tau\tau)$
- The main backgrounds are estimated with data-driven techniques and compared to MC predictions –we always find them consistent within uncertainties
- The rapidity dependences are also well described by MC
- The background contamination is most significant at high invariant masses (upto 20%)



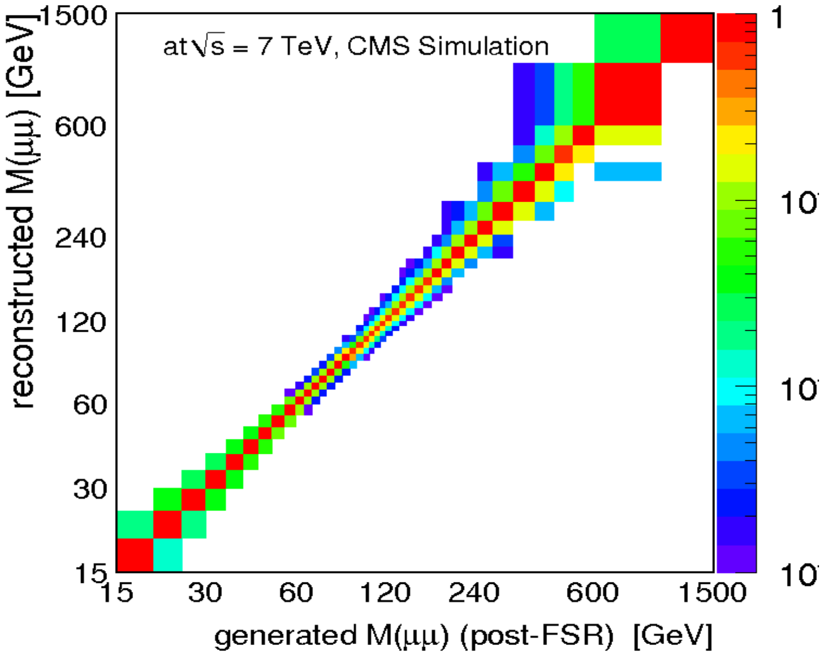


The electron yields are "cleaner" for the price of lower efficiency.

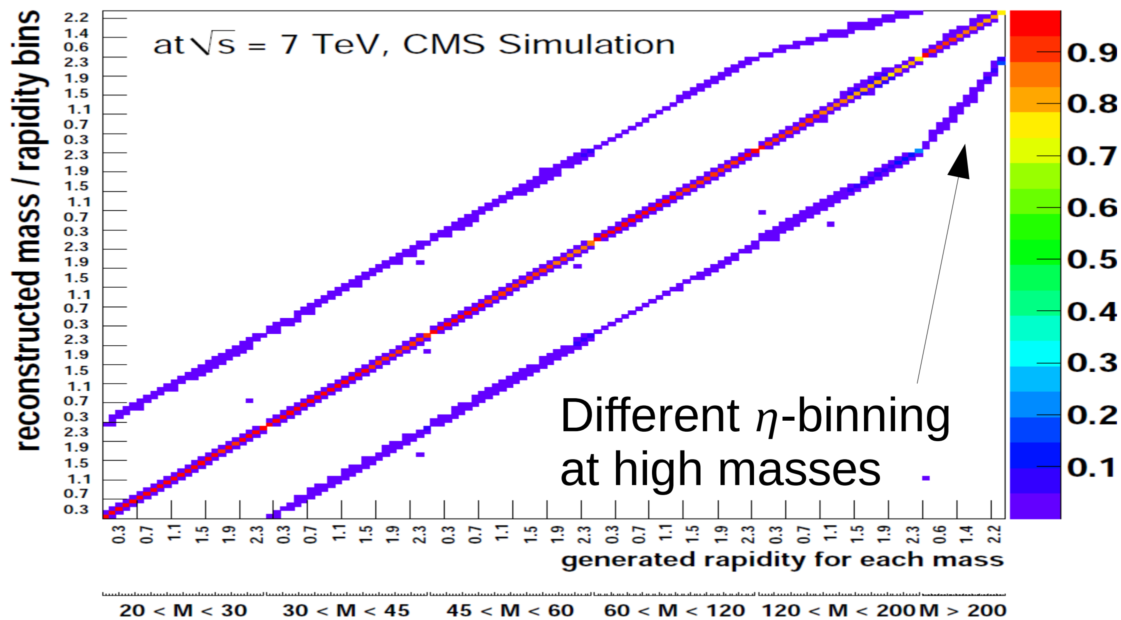
Resolution and unfolding

- We chose the binning so as to avoid significant resolution effects
- To account for these we use a response matrix from MC (its inverse is the "unfolding" matrix)

Response matrix – 1D muon analysis



Response matrix – 2D muon analysis



The response (unfolding) matrix is used to "unsmear" the resolution and obtain the "true" values.

It is near diagonal for the invariant mass muon analysis which is to say that most of the events do not migrate (due to resolution) outside their bins. For electrons the situations is a bit worse with few regions showing upto 30% contributions from off-diagonal elements. The same method is used in the forward-backward analysis though it is technically more complicated there.

In the mass-rapidity analysis we group the bins to construct the matrix representation above.

Note that the unfolding matrix is the "core" of the covariant matrix needed for the final results.

Main systematic sources

- When acceptance correction is applied this introduces theoretical and PDF uncertainties on the result –upto 3% in most of the mass spectrum covered, smaller around the Z peak
- Uncertainties on measurement efficiencies contribute by less than 1% to the results except for high masses (upto 3.5%)
- Resolution and scale effects and their modeling and correction (unfolding) become important at high masses (upto tens of percent uncertainty; it scales with statistics)
- Background is generally of no concern (upto 0.5% or dominated by statistics)
- Other investigated sources are with much smaller contribution

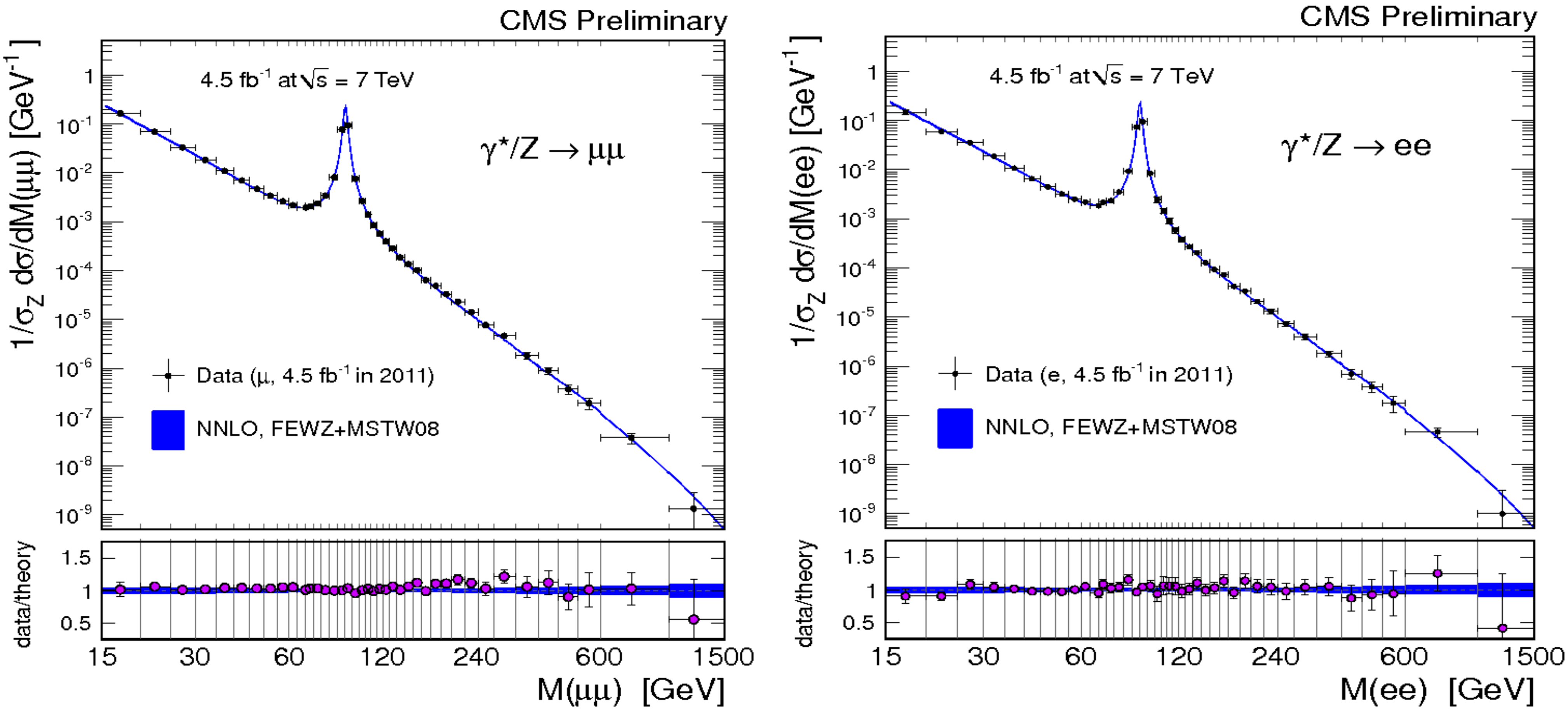
There are correlations between the different regions of phase space we are covering.

The core of the covariance matrix is the unfolding matrix in the analysis.

However the full covariance matrix also includes other contributions, noticeably from the reconstruction inefficiencies. It is not yet available.

DY differential cross-section measurement*

- Performed in 40 invariant mass bins starting from 15 GeV and **normalized to the Z peak**
- It is given in the full phase space but will be available in the acceptance region as well

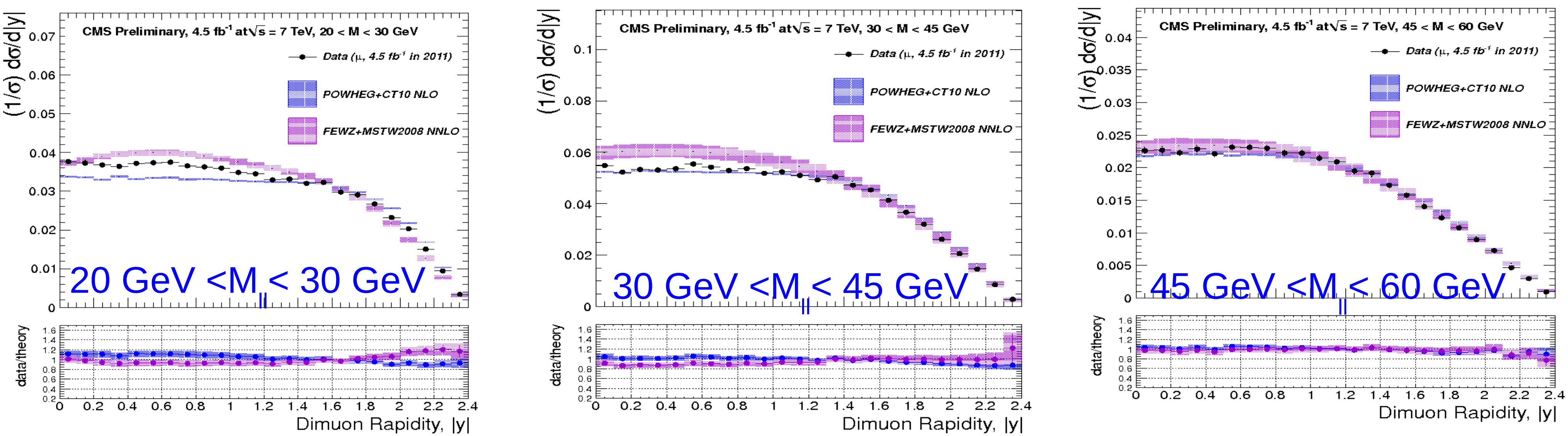


The data distribution is well described by NNLO (fluctuations are under scrutiny). At low invariant masses there is a difficult to control "modeling" uncertainty, at high invariant masses - the statistics plays a major role.

* Soon to appear CMS-PAS-EWK-11-007; CMS-PAS-EWK-11-020, J. High Energy Phys.10 (2011) 007

DY double differential cross-section measurement

- Performed in 24 rapidity bins between 0 and 2.4 (12 Y-bins for the highest mass bin) and 6 mass ranges: (20-30), (30,45), (45,60), **(60,120)**, (120, 200), (200,1500) GeV
- It is in the CMS acceptance region and for muons only, **normalized to the Z peak**



There is expected sensitivity to the order of calculations and modeling at lowest masses but we also see significant differences upto 40 GeV.

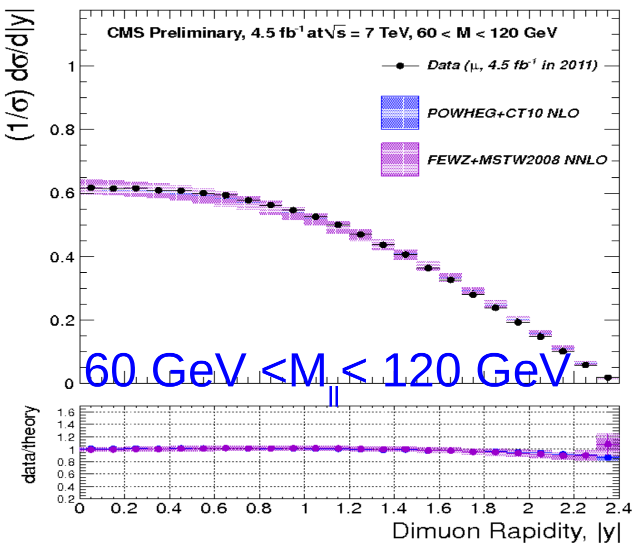
When the analysis is finalized, data is clearly to disfavor models**.

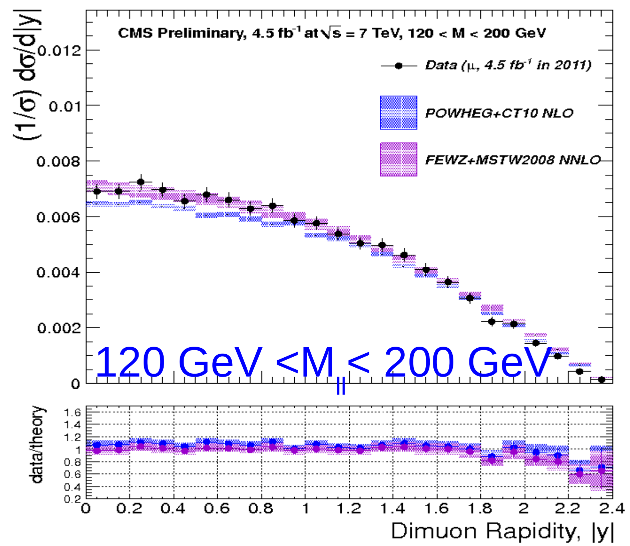
* Soon to appear CMS-PAS-EWK-11-007;

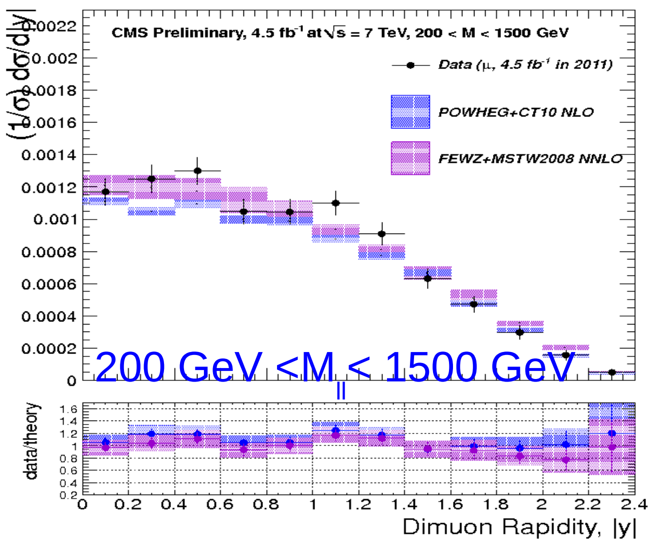
** Assuming no New Physics, this is a separate important topic.

DY double differential cross-section measurement

- Currently we only compare to MSTW2008NNLO and POWHEG(CT10)+Pythia MC.
- No extensive theoretical uncertainties are yet available in such details (we expect them to be smaller than the PDF uncertainties though)







Moving away from the Z peak we gain discrimination power again.

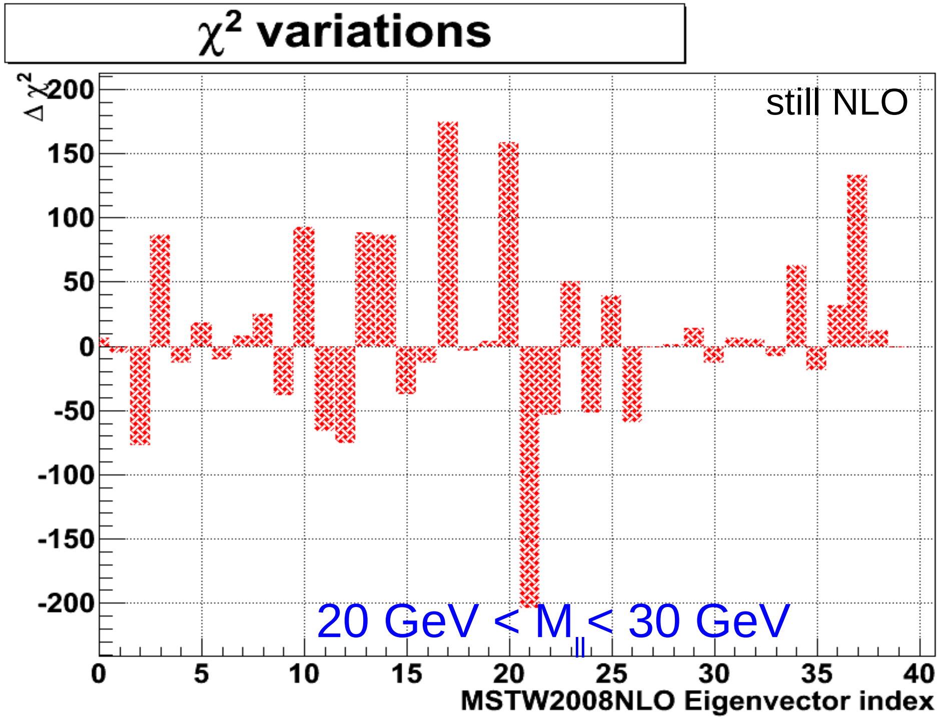
When the full covariance matrix becomes available we can test the influence of different parts of the data on global PDF fits – for now no correlations are discussed.

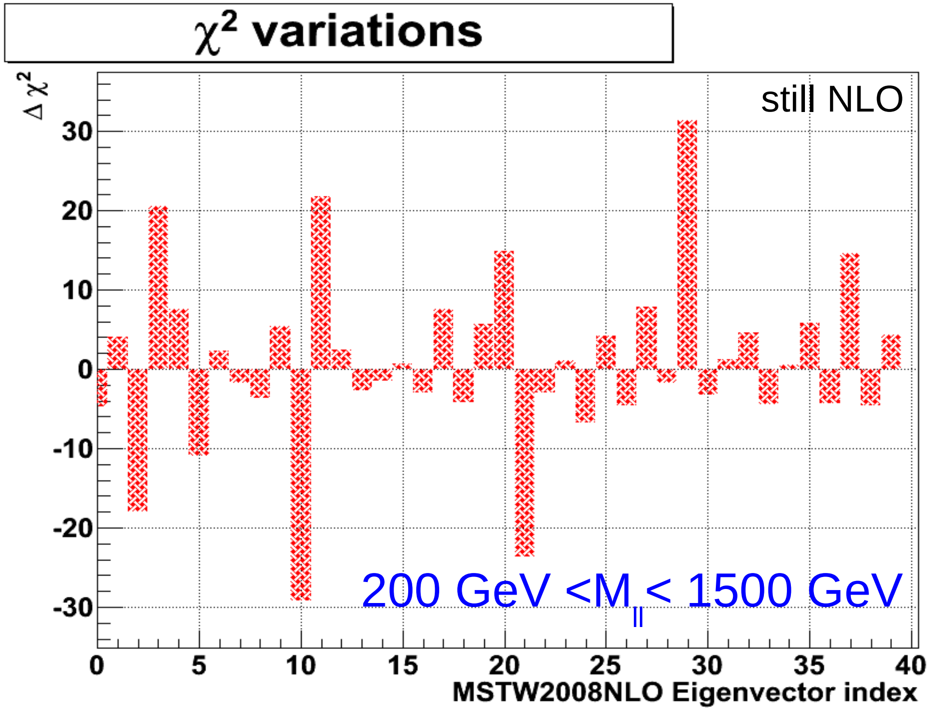
* Soon to appear CMS-PAS-EWK-11-007;

Sensitivity to PDFs

To test the sensitivity of our results to parameters in PDF sets we calculate the χ^2 between data and the predictions from a given PDF set and obtain the difference in χ^2 when we vary each of the eigenvectors by one/two standard deviation (or use individual eigenvectors).

MSTW2008NLO





Our tests do not include proper handling of correlations in the measurements and our error estimations are not yet finalized. For now these test could be only qualitative (not absolute).

The effect of unaccounted correlations would be smallest in the most distant mass bins.

On the other hand lowest invariant masses can be reasonably explored by NNLO sets only.

Sensitivity to PDFs

- Are these kind of studies useful in a DY measurement paper
  - only PDF specialists know the meaning of the Eigenvectors (or replicas)
  - the Eigenvectors are only a linear combination of the actual parameters
  - what is the most useful info we could provide (except the DY measurements)?

- We are testing substitution of DY fixed target data with CMS DY data (NNPDF2.1 first)
  - confronting different mass ranges or sets of mass ranges
  - full scale fit vs subsets
  - comparisons to DY fixed target data
  - all is at initial stage

Compute the predictions for **2D DY cross sections in the data bins** with **FEWZ at NNLO** for all the 1000 replicas of the following NNPDF2.1 NNLO PDF sets:

FEWZ allows to compute the predictions for all 1000 NNPDF replicas **within a single run**

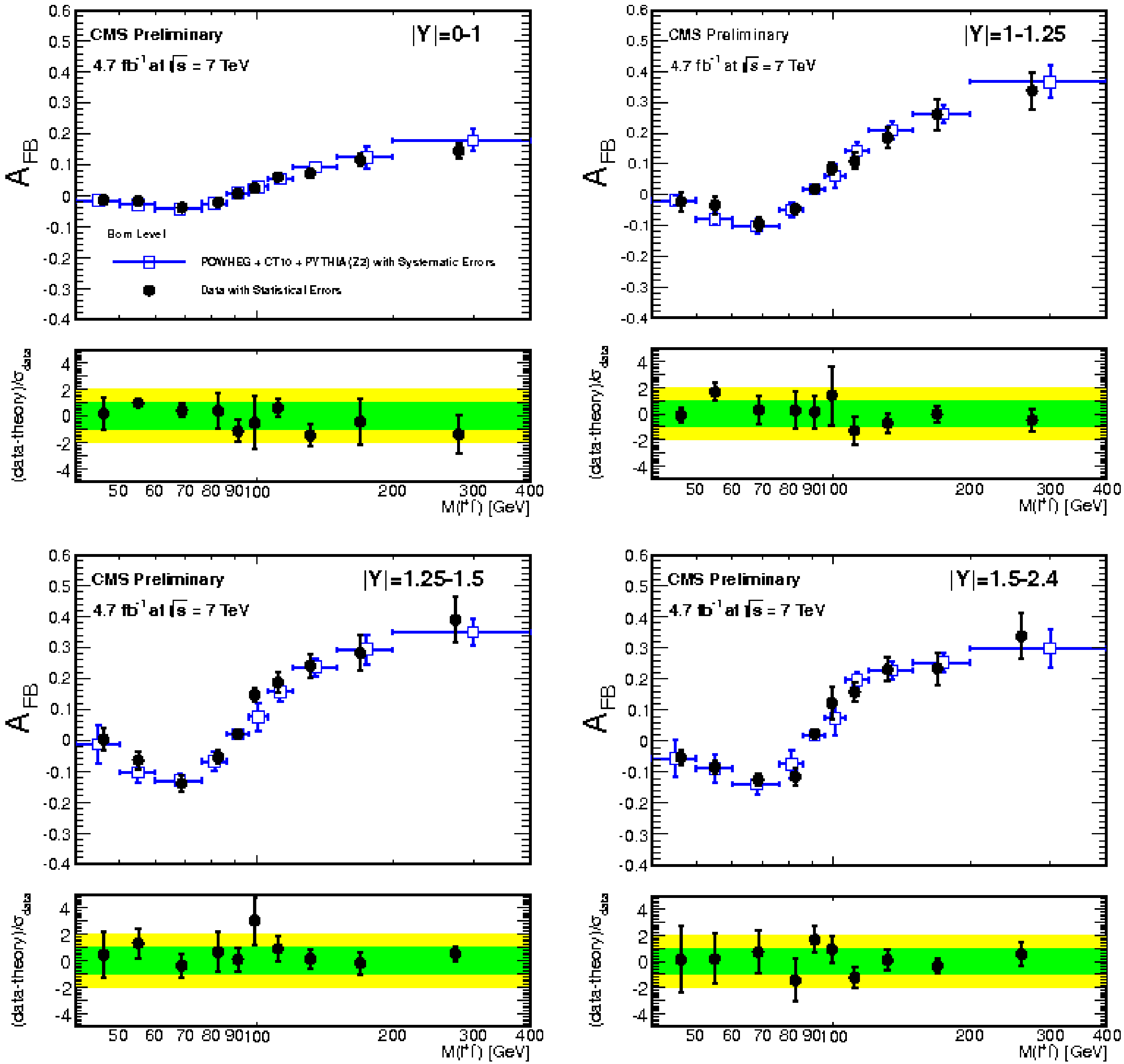
Compute the **χ^2 for data vs theory** for each of the 1000 replicas, and compute the **new weights***of each PDF

This information is enough to **quantify the impact of the 2D DY data into PDFs**

Juan Rojo

* arxiv:1012.0836,1108.1758

DY forward-backward asymmetry*



These are the combined results from both lepton channels and in the CMS acceptance region.

There are experimental reasons the rapidity ranges to be large.

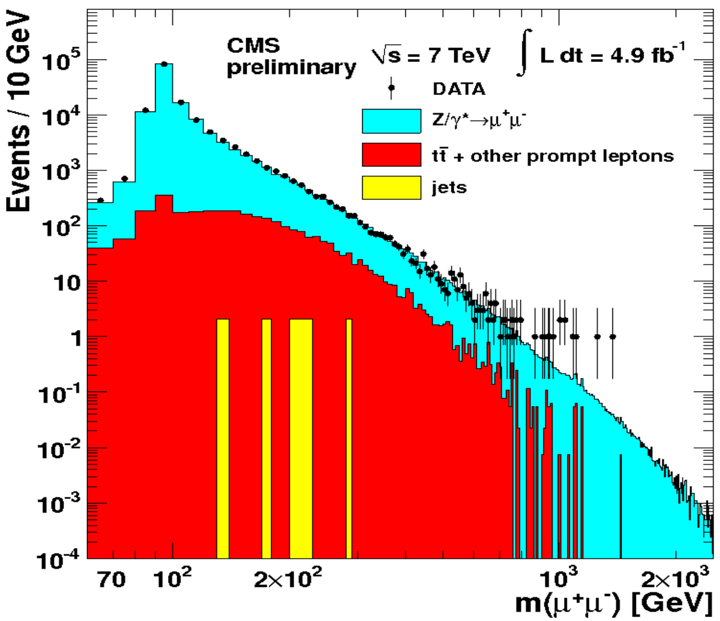
The measurements are in agreement with expectations though more data will be required for constraints on PDFs (unrealistic?).

Could FBA be useful for PDF constraints in any way?

* CMS-PAS-EWK-11-004

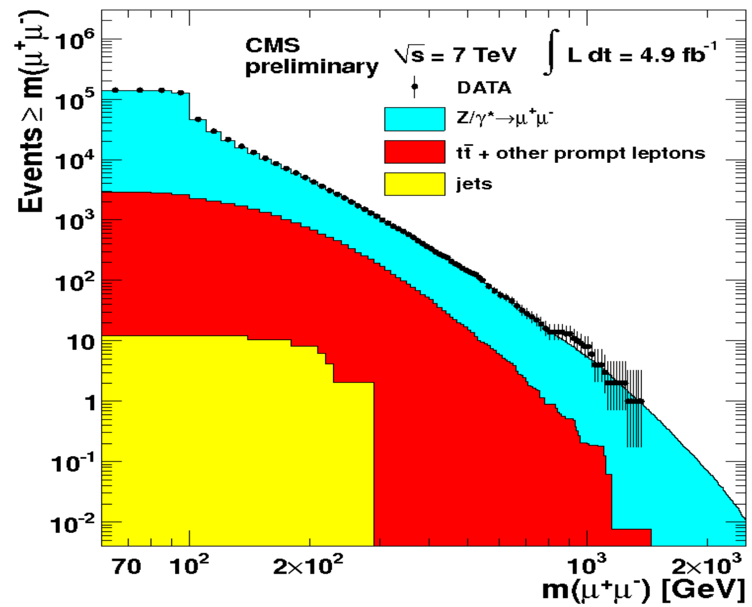
DY related studies (not a complete list, CMS)

- DY is the overwhelming background for various physics searches
- Inevitably such analyses contribute indirectly to precision measurements



These are the di-muon invariant mass spectrum and the integrated spectrum.

Di-electron spectra are also measured.



High mass di-lepton resonance searches are an independent and different view on the high end of the DY spectrum (selection is optimized for high momenta). No new resonances are discovered, limits set*.

At low invariant masses Higgs searches (through non-SM decays) also test the precision of the DY spectrum**. No deviations are observed, limits set.

* https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsEXO11019Winter2012 ; Documentation to become public soon
** CMS-PAS-HIG-12-004

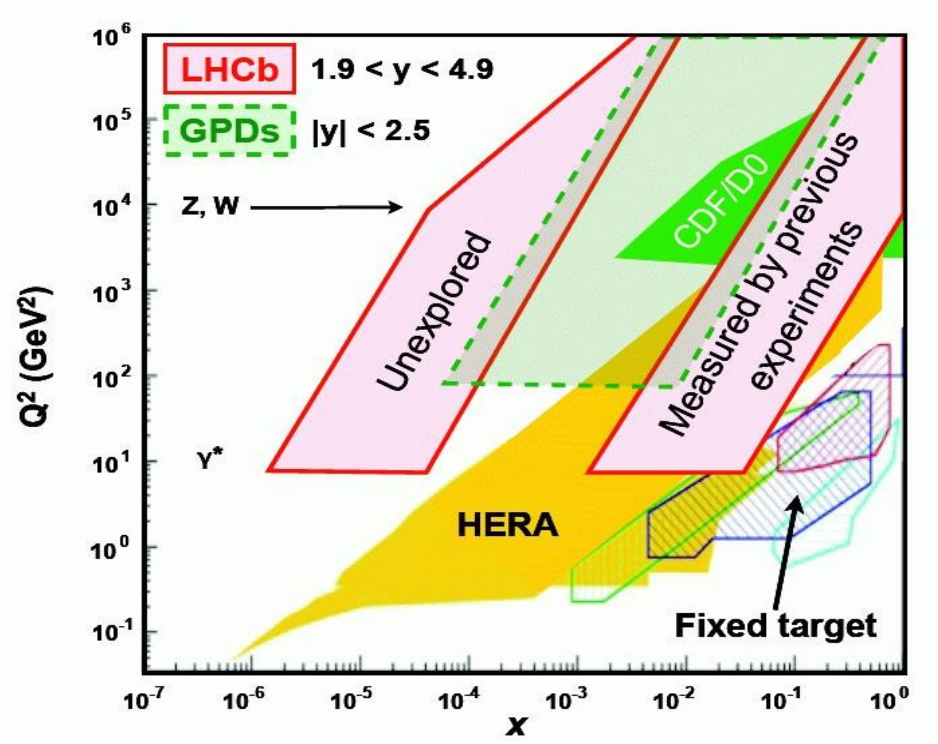
Perspectives

- Optimistically CMS will publish the differential and double-differential DY cross-section measurements this Summer together with the forward-backward asymmetry paper
- ATLAS are probably to deliver their DY results soon and these are closely comparable to CMS results
- LHCb are also working on DY and they have unique extended and high precision coverage particularly important for PDFs (CMS could partially compete though at lower resolution quality)

The datataking in 2012 will continue at sqrt(s) = 8 TeV.

To the end of the year we expect to have first preliminary results on DY at the new energy frontier.

On the other hand, it is realistically to think that to the end of this year collision DY data (@7 TeV) can be successfully applied in global PDF fits.



Jonathan Anderson and LHCb

Other analyses like $d^2\sigma/dMdq_T$ and jet/MET association to DY are of potential interest.

Summary

- CMS has presented its first public result on the double-differential DY cross-section $d^2\sigma/dMdY$ and updated results on the differential cross-section $d\sigma/dM$ - both at ~4.5 fb^{-1}

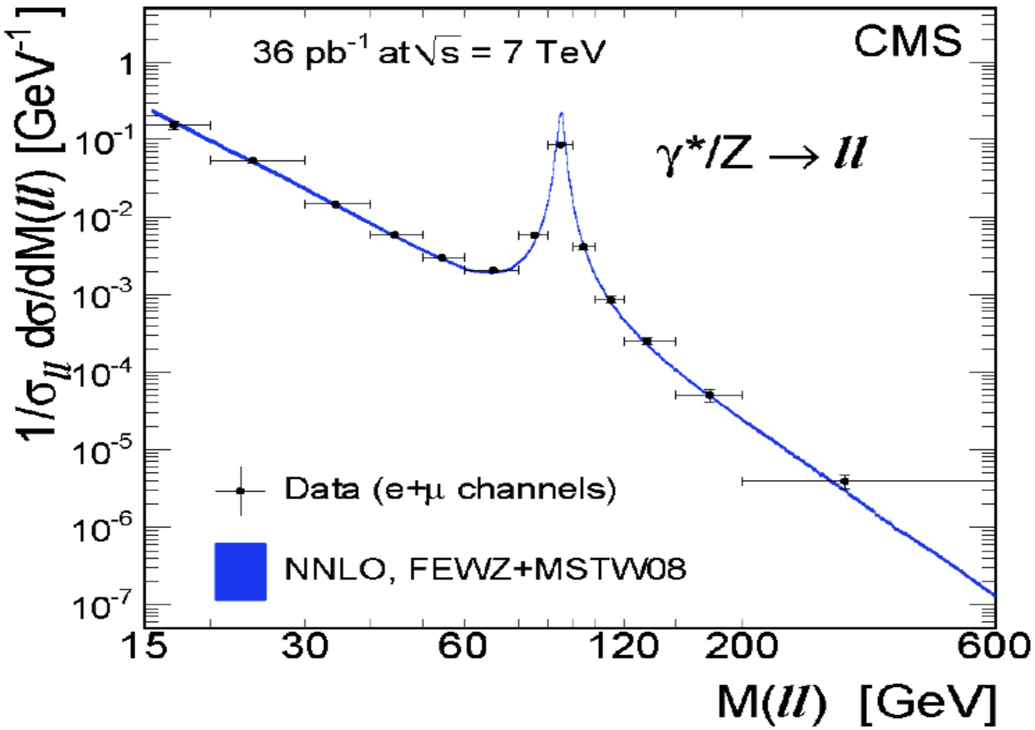
- CMS has also presented off-resonance measurement on the forward-backward asymmetry

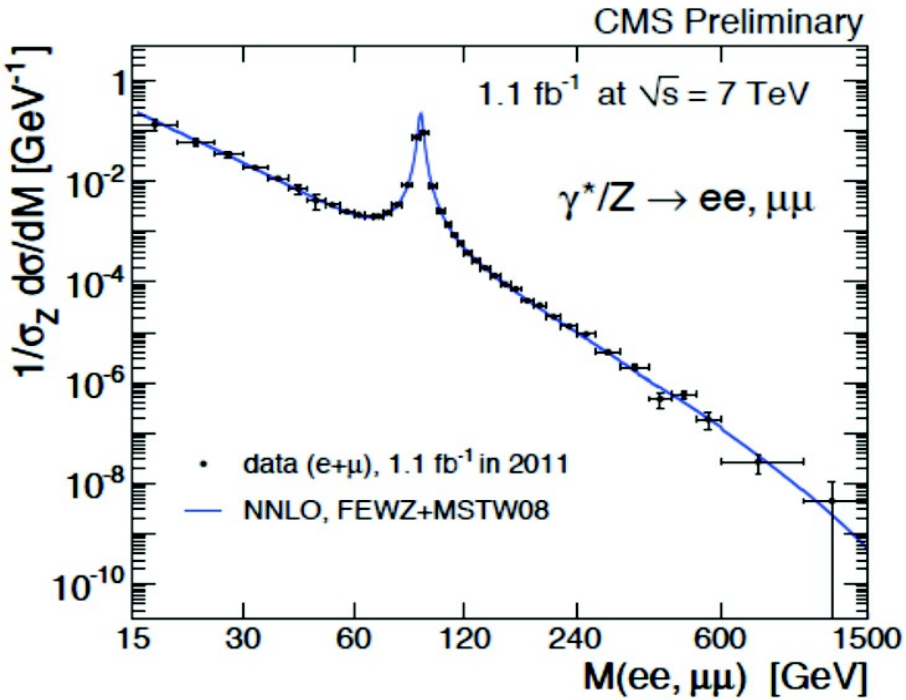
- ATLAS and LCHb will have their results available soon

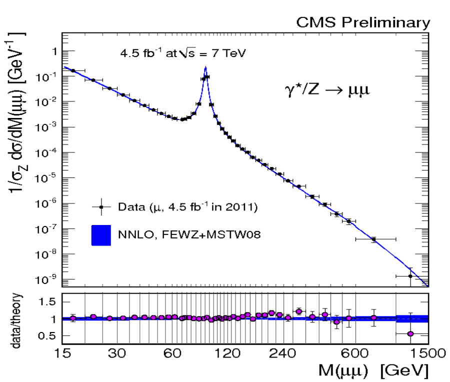
- The preliminary CMS results should shed light on the limitations or problems we may have on PDFs extraction

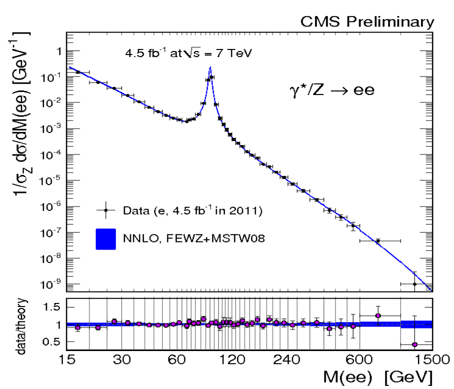
- The interaction between experimentalists and theoreticians <u>at this stage</u> should benefit us all and help us to improve the picture of PDFs we have later this year

Back up









Each measurement shown is in the full phase but the detector acceptance is different for each (transverse momentum requirements changed).

$$r_i = \frac{R_i}{\Delta M_i}$$

where

$$R_i = \frac{\dfrac{N_i^u}{A_i \cdot \varepsilon_i \cdot C_i}}{\dfrac{N_{norm}^u}{A_{norm} \cdot \varepsilon_{norm} \cdot C_{norm}}}$$

Back up

1-D analysis

- **Goal:**
  - Measure σ_i for each mass bin
  - Calculate shape $R=(1/\sigma_z)d\sigma/dM$
- **Procedure**
  1. Event selection
  2. Background subtraction
  3. Unfolding corrections
  4. Acceptance and efficiency correction
  5. Correction for Final State radiations
  6. Systematic Errors estimation
  7. Cross section (normalized to Z peak) measurement (R)
  8. Comparison with theoretical prediction

Analysis procedure is same for electron and muon channel

$$\sigma_i = \frac{N_i^u}{A_i \cdot \varepsilon_i \cdot C_i \cdot L}$$

$$R_i = \frac{\sigma_i \cdot L}{\sigma_{norm} \cdot L}$$

i: mass index

N^u: background subtracted, unfolded yield

A: acceptance

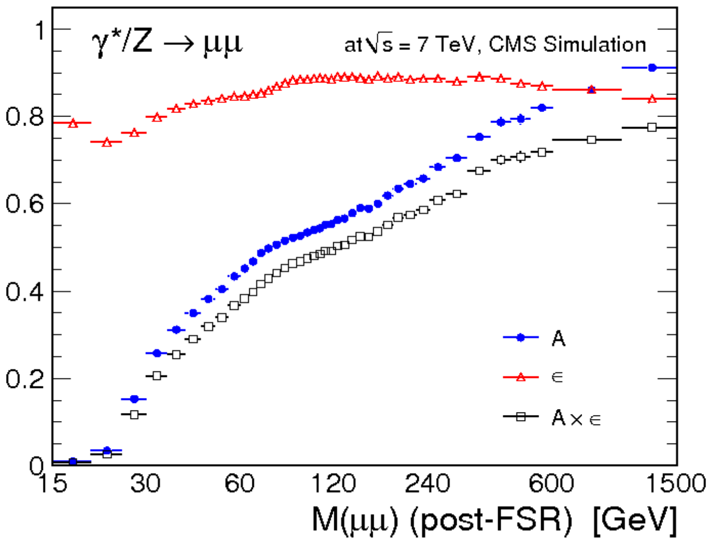
ε: efficiency

C: Correction for efficiency, FSR

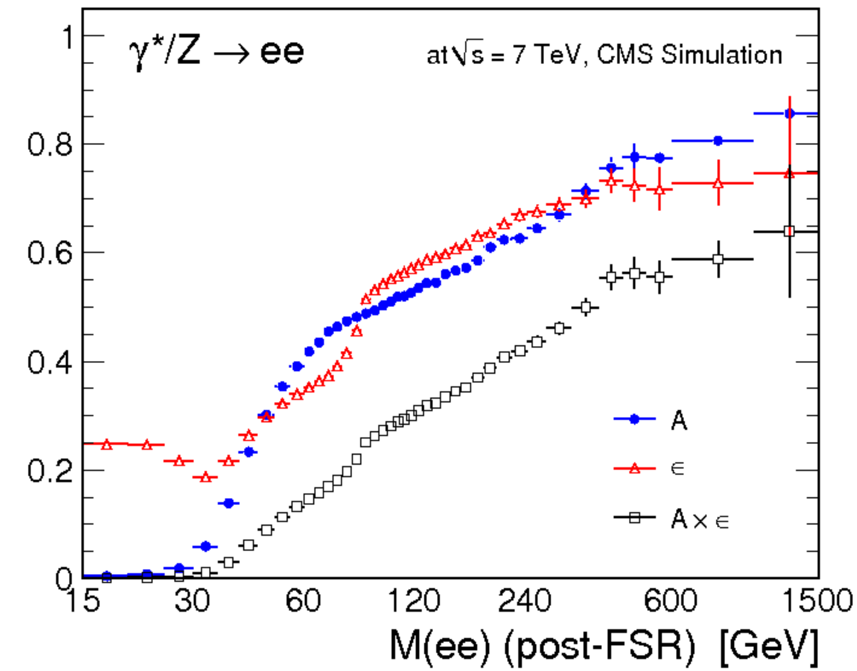
L: integrated luminosity

Mass Bins={15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 64, 68, 72, 76, 81, 86, 91, 96, 101, 106, 110, 115, 120, 126, 133, 141,150, 160, 171, 185, 200, 220, 243, 273, 320, 380, 440, 510, 600, 1000, 1500}; // 40 bins

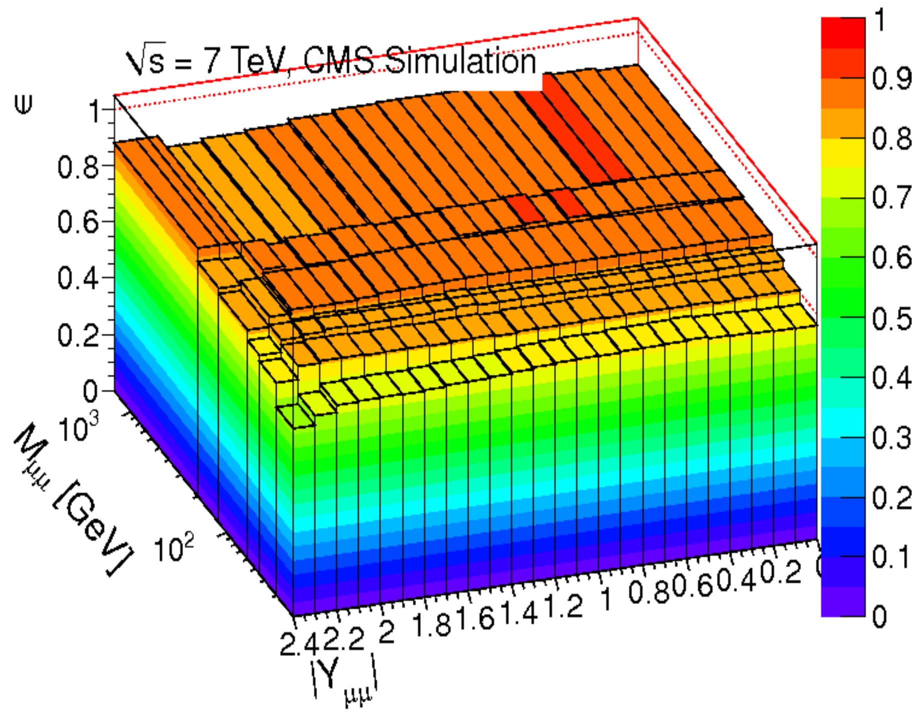
Back up



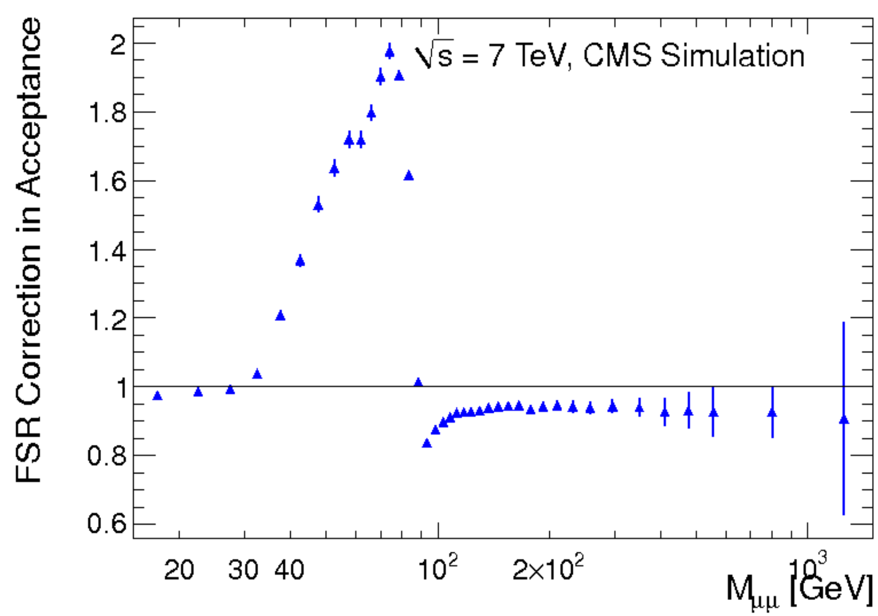
γ*/Z → μμ
at √s = 7 TeV, CMS Simulation
A
∊
A × ∊
M(μμ) (post-FSR) [GeV]



γ*/Z → ee
at √s = 7 TeV, CMS Simulation
A
∊
A × ∊
M(ee) (post-FSR) [GeV]

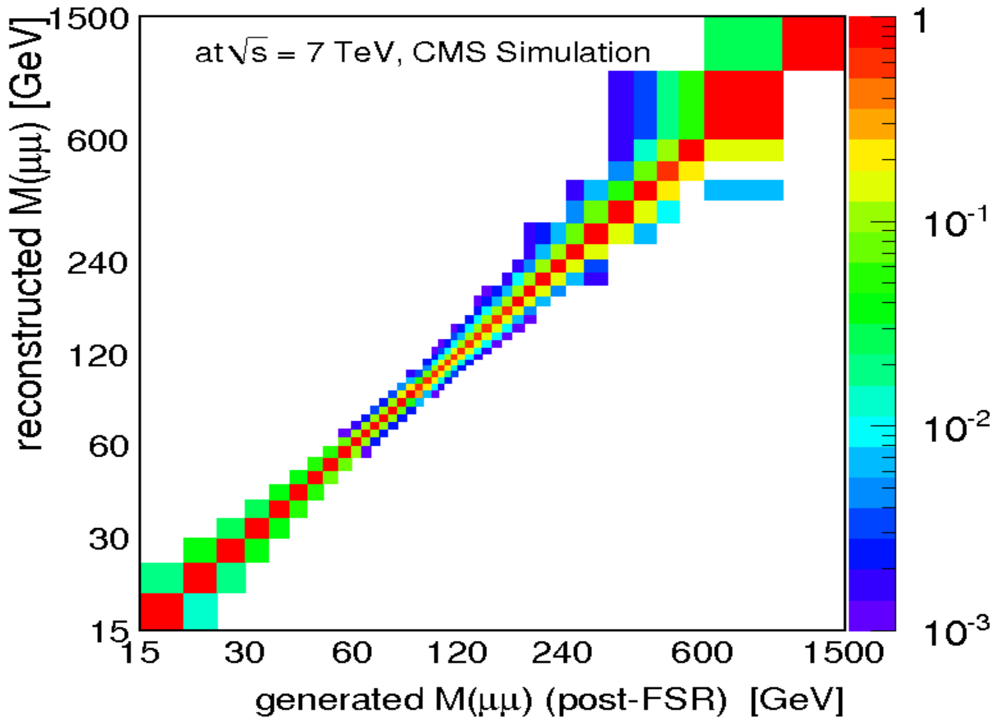


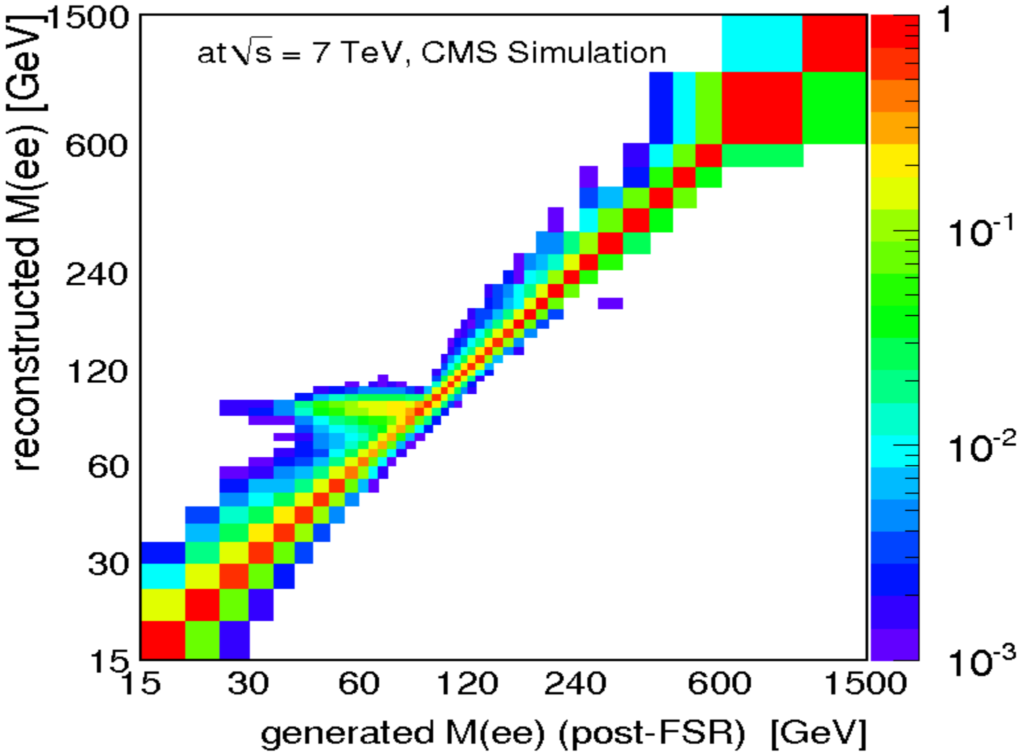
√s = 7 TeV, CMS Simulation
∊
M_μμ [GeV]
|Y_μμ|



√s = 7 TeV, CMS Simulation
FSR Correction in Acceptance
M_μμ [GeV]

Back up


at $\sqrt{s}$ = 7 TeV, CMS Simulation
reconstructed M(μμ) [GeV]
generated M(μμ) (post-FSR) [GeV]
15
30
60
120
240
600
1500
1
10^{-1}
10^{-2}
10^{-3}


at $\sqrt{s}$ = 7 TeV, CMS Simulation
reconstructed M(ee) [GeV]
generated M(ee) (post-FSR) [GeV]
15
30
60
120
240
600
1500
1
10^{-1}
10^{-2}
10^{-3}

Back up

$$\sigma_{ij} = \frac{N_{ij}^{u}}{\varepsilon_{ij} \cdot C_{ij} \cdot L}$$

- Goal:
  - Measure σ_{ij} for each mass-rapidity bin
  - Calculate shape $R=(1/\sigma_z)d^2\sigma/dMdY$ within $|Y| < 2.4$
- Procedure
  1. Event selection
  2. Background subtraction
  3. Unfolding corrections
  4. Efficiency correction
  5. Correction for Final State radiations
  6. Systematic Errors estimation
  7. Cross section (normalized to Z peak) measurement (R) within detector acceptance
  8. Comparison with theoretical prediction

Same as 1D

i: mass index

j: rapidity index

N^u: background subtracted, unfolded yield

ε: efficiency

C: Correction for efficiency, FSR

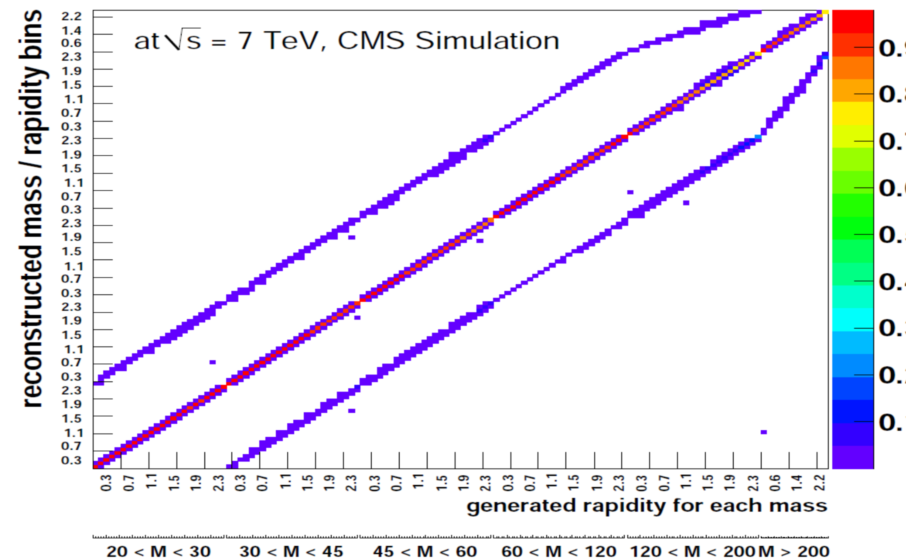
L: integrated luminosity

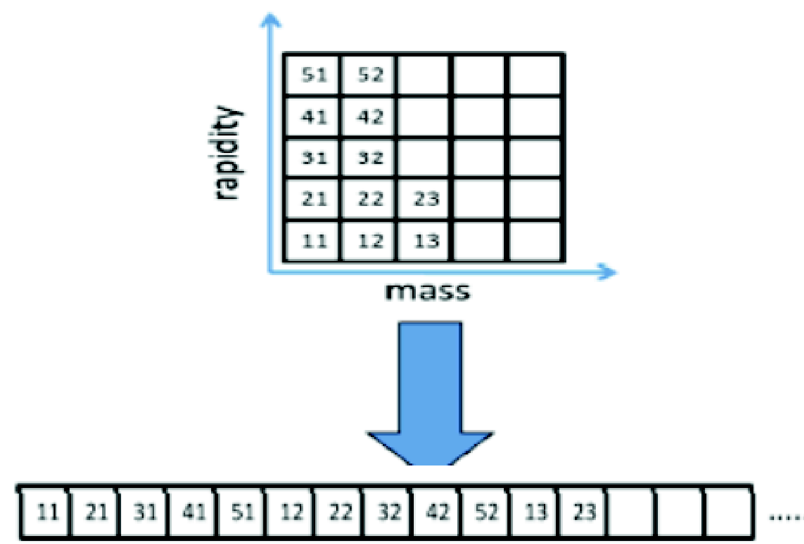
Mass Bins={20, 30, 45, 60, 120 , 200, 1500}; // 6 bins

Rapidity Bins={0, 0.1, 0.2, ..., 2.3, 2.4 }; // 24 bins for low masses
12 bins for mass>200

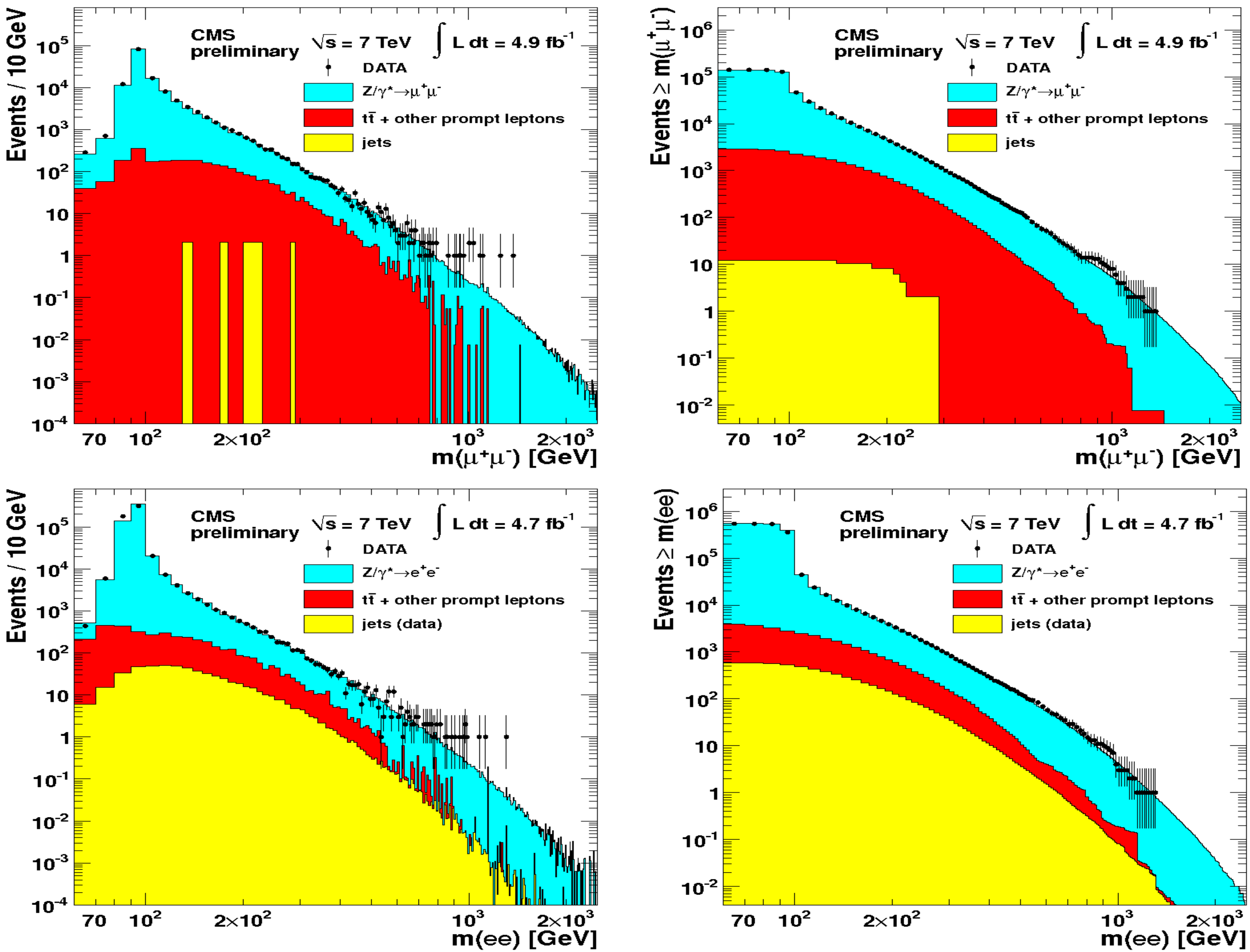
Back up

- The response matrix for 2D has following features
  - Block-diagonal structure
  - Each block corresponding to a given mass-rapidity bin.
  - The entries off the main diagonal are due to migration effects.
- To use conventional unfolding procedure for 2D measurement
  - Convert 2D array of mass-rapidity unfolding matrix into a 1D array
  - apply usual 1D unfolding procedure to the flattened spectrum





Back up


CMS preliminary
√s = 7 TeV
∫ L dt = 4.9 fb⁻¹
DATA
Z/γ*→μ⁺μ⁻
tt̄ + other prompt leptons
jets
Events / 10 GeV
m(μ⁺μ⁻) [GeV]
Events ≥ m(μ⁺μ⁻)
∫ L dt = 4.7 fb⁻¹
Z/γ*→e⁺e⁻
jets (data)
m(ee) [GeV]
Events ≥ m(ee)