Measurement of the Drell-Yan differential cross section $d\sigma/dM$ at $\sqrt{s} = 7$ TeV

Stoyan Stoynev

Northwestern University

On behalf of the CMS Collaboration
Outline

- Introduction
  - motivation and theoretical expectations
  - analysis procedure

- Event selection
  - samples
  - baseline selections

- Backgrounds and background estimation
  - data vs. MC comparison
  - background estimation procedures and results

- Unfolding

- Acceptance and efficiency
  - definitions
  - data-based efficiency corrections

- Systematic uncertainties

- Results
Motivation and theoretical expectations

Drell-Yan (DY) process

- Important SM benchmark process
- Can be used to constraint PDFs and test pQCD
- Important background for BSM searches

Our goal is to measure the differential \( \frac{1}{\sigma} \frac{d\sigma}{dM(ll)} \) cross section. We normalize to the cross section around the Z peak, 60-120 GeV \( (\sigma) \), which cancels part of the systematic uncertainties.

We calculate the theoretical expectations up to NNLO using (updated) FEWZ (2.0) http://gate.hep.anl.gov/fpetriello/FEWZ.html, arXiv:1011.3540v1

Three PDF sets employed:
- CT10, Phys. Rev. D82, 074024 (2010)

EWK corrections considered by HORACE, no QED final state radiation (FSR) in the model.
To measure \( R = \frac{1}{\sigma_i} \frac{d\sigma}{dM(ll)} \) we use the formula:

\[
\sigma_i = \frac{N^U_i}{A_i \epsilon_i \rho_i L_{\text{int}}}
\]

- \( N_i^U \) – unfolded (and background corrected) yield
- \( \sigma_i \) – mass bin, \( M(ll) \) – invariant mass
- \( N_{\text{NORM}}^U \) – (and background corrected) yield
- \( A_i \) – acceptance
- \( \epsilon_i \) – efficiency
- \( \rho_i \) – efficiency (and FSR) correction
- \( L_{\text{int}} \) – integrated luminosity

**Procedure:**

- Event selection
- Background subtraction
  - data-driven methods for dominant sources
- Unfolding
  - correcting for resolution effects
  - based on migration matrices from MC
- Acceptance and efficiency calculation using (NNLO matched) MC
- Efficiency correction using data-driven methods
- FSR correction (for the pre-FSR measurements)
  - based on the (NNLO matched) MC
- Cross section shape calculations (NORM refers to 60-120 GeV):

\[
R_i = \frac{N^U_i}{A_i \epsilon_i \rho_i} / \frac{N_{\text{NORM}}^U}{A_{\text{NORM}} \epsilon_{\text{NORM}} \rho_{\text{NORM}}}
\]

\[
r_i = \frac{R_i}{\Delta M_i}
\]

**Inv. mass binning (GeV):**

- 15 - 20
- 20 - 30
- 30 - 40
- 40 - 50
- 50 - 60
- 60 - 76
- 76 - 86
- 86 - 96
- 96 - 106
- 106 - 120
- 120 - 150
- 150 - 200
- 200 - 600

(resolution and statistics driven)
Analysis – types of results

**GEN level mass distribution**

![Graph showing mass distribution](image)

- **For illustration:** this Pythia sample starts at 20 GeV propagator mass

FSR changes the observed spectra.

We make both FSR corrected and FSR not corrected measurements.

Thus we report four different types of results.

- **Direct measurement:** post-FSR cross-section within the detector acceptance
- **Measurement with acceptance corrections:** post-FSR measurement in the full phase space
- **Measurements with FSR corrections:** pre-FSR measurements within the detector acceptance and in the full phase space
  - FSR corrections from Pythia
  - makes possible comparisons to models which do not have FSR
The measurements are based on 2010 data recorded by CMS: total integrated luminosity: $35.9 \pm 1.4 \text{ pb}^{-1}$

We use data triggered by single muons or electrons with transverse momentum thresholds of $< 15 \text{ GeV}$ and $< 17 \text{ GeV}$, respectively.

There is a special check made with double-muon triggers where thresholds are at 3 GeV.

Centrally produced MC samples:

- **DY signal**: PYTHIA (v. 6.422) + POWHEG with CT10 PDF and Z2 tune
- **Backgrounds**: PYTHIA / and MadGraph (v. 4.4.12)/
  - QCD (genuine or mis-identified leptons)
  - EWK (DY$\rightarrow\tau\tau$, W$\rightarrow l\nu$, diboson production)
  - Top quark pairs

The signal sample corresponds to $\sim 30$ times the statistics in data. Backgrounds correspond to at least few times the statistics in data.
## Baseline selections

### Muons

**Kinematic**
- two muons with opposite charges
  - $p_T(\mu_1) > 16$ GeV, $|\eta(\mu_1)| < 2.1$,
  - $p_T(\mu_2) > 7$ GeV, $|\eta(\mu_1)| < 2.4$
- at least one of the muons triggers ($p_T(\mu) > 16$ GeV, $|\eta(\mu)| < 2.1$)

**ID selection**
- minimal hits in the tracker to insure good $p_T$ measurement
- minimal muon hits, maximal $\chi^2$/ndf to avoid bad reconstructed muons
- impact parameter (beam spot) $|dxy| < 0.2$ cm

**Isolation**
- relative isolation (no ECAL) in $\Delta R < 0.3$
  $$I_{rel} = \left( \sum p_T(\text{tracks}) + \sum E_T(\text{had}) \right) / P_T(\mu) < 0.15$$

**Di-muon**
- di-muon vertex probability $> 0.02$
- $(\pi-)$ 3D angle between muons $> 5$ mrad to further suppress cosmic contamination

### Electrons

**Kinematic**
- two ECAL-driven electrons
  - $E_T(e_1) > 20$ GeV, $E_T(e_2) > 10$ GeV, $|\eta(e)| < 1.44$ OR $1.57 < |\eta(e)| < 2.5$
- the leading electron matches the trigger

**ID selection**
- minimal hits in the tracker
- track-ECAL cluster matching quality
- HCAL energy fraction restriction
- conversion removal
- impact parameter (primary vertex) $|dxy| < 0.02$ cm, $|dz| < 1$ cm

**Isolation**
- relative isolation in $\Delta R < 0.3$
  $$\left( \sum p_T(\text{tracks}) + \sum E_T(\text{em}) + \sum E_T(\text{had}) \right) / P_T(\mu) < 0.1$$
Backgrounds and background estimation

**Muon channel**
- QCD up to the Z peak - dominant
  - opposite sign (OS)/same sign (SS) method
  - template fits (muon isolation based)
- MC estimations for the rest
  - EWK backgrounds (in particular DY(\(\tau\tau\)))
  - top quark pairs
  - QCD at higher masses only

**-electron channel**
- true di-electron (DY(\(\tau\tau\)), ttbar, WW, tW) - dominant: use data-driven e-mu method
- fake electron backgrounds (QCD, W+jets) - relatively small: use data-driven “fake rate” method
- true Z backgrounds (WZ, ZZ) - non-dominant: use MC prediction

Details - in the backup slides.
Unfolding correction

- The unfolding procedure “removes” the effects of the resolution on the mass spectrum
  - FSR correction is done at a different stage
- We use the technique of matrix inversion to unfold the spectrum
  - in the limit of no background, the “ideal” (infinitely good resolution) mass spectrum \( N_{true} \) is related to the observed one, \( N_{obs} \), by response matrix \( T \):
    \[
    N_{i}^{obs} = \sum_{k} T_{ik} N_{k}^{true}
    \]
  - with \( T_{ik} \) –the probability that an event with true mass bin \( k \) is reconstructed in mass bin \( i \), which is extracted from the signal MC sample

- Then by inverting the response matrix we can recover the initial spectrum:
    \[
    N_{k}^{true} = \sum_{k} (T^{-1})_{ki} N_{i}^{obs}
    \]
<table>
<thead>
<tr>
<th>Inv. mass bin (GeV)</th>
<th>(N_{obs}^{\muons})</th>
<th>(N_{obs}^{\muons} - N_{bg}^{\muons})</th>
<th>(N_{unfolded}^{\muons})</th>
<th>(N_{obs}^{el})</th>
<th>(N_{obs}^{el} - N_{bg}^{el})</th>
<th>(N_{unfolded}^{el})</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-20</td>
<td>253 ± 16</td>
<td>241 ± 18</td>
<td>243 ± 18</td>
<td>16 ± 4</td>
<td>16 ± 4</td>
<td>16 ± 6</td>
</tr>
<tr>
<td>20-30</td>
<td>809 ± 28</td>
<td>735 ± 36</td>
<td>736 ± 36</td>
<td>91 ± 10</td>
<td>88 ± 10</td>
<td>94 ± 12</td>
</tr>
<tr>
<td>30-40</td>
<td>986 ± 31</td>
<td>910 ± 36</td>
<td>907 ± 37</td>
<td>179 ± 13</td>
<td>163 ± 14</td>
<td>164 ± 17</td>
</tr>
<tr>
<td>40-50</td>
<td>684 ± 26</td>
<td>632 ± 29</td>
<td>631 ± 30</td>
<td>243 ± 16</td>
<td>208 ± 18</td>
<td>219 ± 22</td>
</tr>
<tr>
<td>50-60</td>
<td>471 ± 22</td>
<td>435 ± 24</td>
<td>436 ± 26</td>
<td>211 ± 15</td>
<td>187 ± 16</td>
<td>234 ± 25</td>
</tr>
<tr>
<td>60-76</td>
<td>797 ± 28</td>
<td>768 ± 29</td>
<td>752 ± 31</td>
<td>455 ± 21</td>
<td>428 ± 22</td>
<td>620 ± 45</td>
</tr>
<tr>
<td>76-86</td>
<td>1761 ± 42</td>
<td>1755 ± 42</td>
<td>1471 ± 49</td>
<td>1599 ± 40</td>
<td>1588 ± 40</td>
<td>1277 ± 89</td>
</tr>
<tr>
<td>86-96</td>
<td>11786 ± 109</td>
<td>11761 ± 109</td>
<td>12389 ± 119</td>
<td>6998 ± 84</td>
<td>6981 ± 84</td>
<td>7182 ± 117</td>
</tr>
<tr>
<td>96-106</td>
<td>909 ± 30</td>
<td>904 ± 30</td>
<td>591 ± 38</td>
<td>587 ± 24</td>
<td>581 ± 24</td>
<td>441 ± 36</td>
</tr>
<tr>
<td>106-120</td>
<td>194 ± 14</td>
<td>191 ± 30</td>
<td>178 ± 17</td>
<td>132 ± 11</td>
<td>127 ± 12</td>
<td>127 ± 15</td>
</tr>
<tr>
<td>120-150</td>
<td>145 ± 12</td>
<td>141 ± 12</td>
<td>142 ± 13</td>
<td>67 ± 8</td>
<td>57 ± 9</td>
<td>53 ± 10</td>
</tr>
<tr>
<td>150-200</td>
<td>53 ± 7</td>
<td>49 ± 8</td>
<td>47 ± 9</td>
<td>34 ± 6</td>
<td>27 ± 7</td>
<td>25 ± 7</td>
</tr>
<tr>
<td>200-600</td>
<td>30 ± 6</td>
<td>27 ± 6</td>
<td>28 ± 6</td>
<td>26 ± 5</td>
<td>22 ± 6</td>
<td>21 ± 5</td>
</tr>
</tbody>
</table>
Modification of the original DY MC samples

- The DY samples we use are effectively ~NLO (Powheg + Pythia parton showers).

- At low invariant masses, the two high $p_T$ leptons in the analysis, indirectly impose the existence of a hard gluon in the process (in other words, the cross section in LO is vanishing for such a selection).

- The lepton kinematic distributions in this region are very sensitive to the exact description and the acceptance differs by ~50% (NLO vs NNLO) for the lowest invariant mass bin (and less than few % elsewhere).

- Thus for proper description of the low invariant mass region NNLO is mandatory!

- We have applied weights to the original MC samples determined from the ratio between the differential cross sections calculated at NNLO with FEWZ and at NLO with the Powheg MC.

- We use this "corrected MC" for calculating acceptance, efficiency and FSR corrections.

- Binning effects (limitations(validity of perturbative QCD +.statistical restrictions) are considered as an additional source of systematic uncertainty (~10% at lowest masses).
Acceptance and efficiency

- Acceptance * efficiency is derived from simulation according to:
  
  \[
  A \times \varepsilon = \frac{N_{\text{ACC}}}{N_{\text{GEN}}} \times \frac{N_{\text{SEL}}}{N_{\text{ACC}}} = \frac{N_{\text{SEL}}}{N_{\text{GEN}}} \quad (\leq 1)
  \]

- Post-FSR lepton quantities are used to calculate the di-lepton invariant mass and apply kinematic cuts.
- The acceptance accounts for the \( p_T \) and \( \eta \) cuts, the efficiency reflects the full selection.

\[ N_x \] – number of generated events, with \( X \):
\[ \text{GEN} \] – initially generated
\[ \text{ACC} \] – in the acceptance
\[ \text{SEL} \] – (RECO) selected

### Muons

\[ \gamma^*/Z \rightarrow \mu\mu \]

### Electrons

\[ \gamma^*/Z \rightarrow ee \]
Efficiency factorization and efficiency correction

We need to correct the MC efficiency to “match” the data. We factorize the event efficiency:

\[ \varepsilon(\text{event}) = \varepsilon(l_1) \varepsilon(l_2) \varepsilon(\text{dilepton}|l_1, l_2) \varepsilon(\text{event}, \text{trigger}|\text{dilepton}) \]

\[ \varepsilon(\text{event},\text{trigger}|\text{dimuon}) = \varepsilon(\mu_1, \text{trigger}|\mu_1) + \varepsilon(\mu_2, \text{trigger}|\mu_2) - \varepsilon(\mu_1, \text{trigger}|\mu_1) \varepsilon(\mu_2, \text{trigger}|\mu_2) \]

\[ \varepsilon(\mu) = \varepsilon(\text{track}) \varepsilon(\text{reco}+\text{id}|\text{track}) \varepsilon(\text{iso}|\text{reco}+\text{id}) \]

\[ \varepsilon(\text{event},\text{trigger}|\text{dielectron}) = \varepsilon(\text{leading } e, \text{trigger}|e) \]

\[ \varepsilon(e) = \varepsilon(\text{reco}|\text{ECAL E deposit}) \varepsilon(\text{ID}|\text{reco}) \]

From the deviations observed between data and MC we extract correction factors:

\[ \rho_{\text{eff}}(p_T, \eta) = \frac{\varepsilon_{\text{data}}(p_T, \eta)}{\varepsilon_{\text{mc}}(p_T, \eta)} \]

The correction factors are applied per lepton as weights in MC, following the efficiency factorization.
The lepton efficiencies are estimated by a tag and probe method with the exception of the muon isolation efficiency where the LKTC ("random cones") algorithm is applied.
Systematic uncertainties

- Energy scale: 2% uncertainty per electron, very significant effect on the mass shape

- Efficiency correction: statistically dominated, significant for lower energy electrons

- Backgrounds: more significant for lower masses (~4%), dominant for higher masses
  - based on data in the regions of significance

- Unfolding: significant around the Z region (~4%)
  - apart from error propagation, additional resolution effects contribute to the uncertainty; for muons, the small momentum scale uncertainty (~0.1%) is incorporated here

- FSR: significant below the Z peak (2%); irrelevant for electrons as other factors dominate
  - based on detailed comparisons between photon spectra properties in data and simulation

- Others: remaining non-dominant sources (pile-up, di-lepton selection)

- Acceptance: theory/PDF uncertainties on the ratio of acceptances (for a shape measurement with acceptance correction) – between 1 and 3% per invariant mass bin
  - based on a single PDF set, correlation between bins taken into account
  - theory uncertainties on the acceptance are small (typically below 1%)

Find the details in the backup slides.
The DY measurements are normalized to the Z region (60 GeV < M(ll) < 120 GeV) for each of the measurements (so as $R_{\text{NORM}}^* \equiv 1$):

$$R_{\text{det},i} = \frac{N_i^U}{\epsilon_i \rho_i} / \frac{N_{\text{NORM}}^U}{\epsilon_{\text{NORM}} \rho_{\text{NORM}}}$$

$$R_i = \frac{N_i^U}{A_i \epsilon_i \rho_i} / \frac{N_{\text{NORM}}^U}{A_{\text{NORM}} \epsilon_{\text{NORM}} \rho_{\text{NORM}}}$$

Separately, with and without FSR corrections (four measurements per channel)

(Pre-FSR) $r_i = \frac{R_i}{\Delta M_i}$

We combine the two lepton channels for it and make direct comparisons with predictions (NNLO)

Correlations between bins and between channels do not lead to a significant difference.
## Normalized DY differential cross sections (muons)

<table>
<thead>
<tr>
<th>Inv. Mass bin (GeV)</th>
<th>(R_{\text{det,post-FSR}}) ((10^{-3}))</th>
<th>(R_{\text{det}}) ((10^{-3}))</th>
<th>(R_{\text{post-FSR}}) ((10^{-3}))</th>
<th>(R) ((10^{-3}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-20</td>
<td>18 ± 2</td>
<td>19 ± 2</td>
<td>772 ± 67</td>
<td>780 ± 69</td>
</tr>
<tr>
<td>20-30</td>
<td>58 ± 3</td>
<td>58 ± 3</td>
<td>528 ± 33</td>
<td>533 ± 34</td>
</tr>
<tr>
<td>30-40</td>
<td>67 ± 3</td>
<td>67 ± 3</td>
<td>147 ± 8</td>
<td>147 ± 8</td>
</tr>
<tr>
<td>40-50</td>
<td>44 ± 2</td>
<td>41 ± 2</td>
<td>66 ± 4</td>
<td>62 ± 4</td>
</tr>
<tr>
<td>50-60</td>
<td>30 ± 2</td>
<td>23 ± 2</td>
<td>37 ± 3</td>
<td>30 ± 2</td>
</tr>
<tr>
<td>60-76</td>
<td>51 ± 2</td>
<td>28 ± 1</td>
<td>55 ± 3</td>
<td>32 ± 2</td>
</tr>
<tr>
<td>76-86</td>
<td>97 ± 4</td>
<td>56 ± 3</td>
<td>98 ± 5</td>
<td>58 ± 3</td>
</tr>
<tr>
<td>86-96</td>
<td>803 ± 14</td>
<td>861 ± 15</td>
<td>799 ± 23</td>
<td>857 ± 26</td>
</tr>
<tr>
<td>96-106</td>
<td>38 ± 3</td>
<td>43 ± 3</td>
<td>37 ± 3</td>
<td>41 ± 3</td>
</tr>
<tr>
<td>106-120</td>
<td>12 ± 1</td>
<td>12 ± 1</td>
<td>11 ± 1</td>
<td>12 ± 1</td>
</tr>
<tr>
<td>120-150</td>
<td>9.2 ± 0.9</td>
<td>9.7 ± 1.0</td>
<td>8.4 ± 0.8</td>
<td>8.8 ± 0.9</td>
</tr>
<tr>
<td>150-200</td>
<td>3.1 ± 0.6</td>
<td>3.2 ± 0.7</td>
<td>2.6 ± 0.5</td>
<td>2.7 ± 0.6</td>
</tr>
<tr>
<td>200-600</td>
<td>1.8 ± 0.4</td>
<td>1.9 ± 0.5</td>
<td>1.4 ± 0.3</td>
<td>1.5 ± 0.4</td>
</tr>
</tbody>
</table>

Only \(R\) (last column) is directly comparable between channels!
## Normalized DY differential cross sections (electrons)

<table>
<thead>
<tr>
<th>Inv. mass bin (GeV)</th>
<th>$R_{\text{det,post-FSR}}^{10^{-3}}$</th>
<th>$R_{\text{det}}^{10^{-3}}$</th>
<th>$R_{\text{post-FSR}}^{10^{-3}}$</th>
<th>$R^{10^{-3}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-20</td>
<td>6 ± 3</td>
<td>6 ± 3</td>
<td>487 ± 230</td>
<td>508 ± 238</td>
</tr>
<tr>
<td>20-30</td>
<td>13 ± 2</td>
<td>13 ± 2</td>
<td>536 ± 96</td>
<td>559 ± 97</td>
</tr>
<tr>
<td>30-40</td>
<td>24 ± 4</td>
<td>22 ± 4</td>
<td>129 ± 22</td>
<td>131 ± 21</td>
</tr>
<tr>
<td>40-50</td>
<td>28 ± 4</td>
<td>24 ± 4</td>
<td>52 ± 8</td>
<td>47 ± 7</td>
</tr>
<tr>
<td>50-60</td>
<td>30 ± 5</td>
<td>19 ± 3</td>
<td>39 ± 6</td>
<td>27 ± 4</td>
</tr>
<tr>
<td>60-76</td>
<td>78 ± 12</td>
<td>30 ± 4</td>
<td>84 ± 13</td>
<td>36 ± 5</td>
</tr>
<tr>
<td>76-86</td>
<td>144 ± 60</td>
<td>61 ± 25</td>
<td>147 ± 60</td>
<td>64 ± 26</td>
</tr>
<tr>
<td>86-96</td>
<td>722 ± 62</td>
<td>839 ± 60</td>
<td>715 ± 62</td>
<td>834 ± 60</td>
</tr>
<tr>
<td>96-106</td>
<td>44 ± 21</td>
<td>55 ± 26</td>
<td>43 ± 20</td>
<td>53 ± 25</td>
</tr>
<tr>
<td>106-120</td>
<td>13 ± 3</td>
<td>15 ± 3</td>
<td>12 ± 2</td>
<td>14 ± 3</td>
</tr>
<tr>
<td>120-150</td>
<td>5.4 ± 1.2</td>
<td>6.0 ± 1.3</td>
<td>4.8 ± 1.1</td>
<td>5.4 ± 1.2</td>
</tr>
<tr>
<td>150-200</td>
<td>2.5 ± 0.8</td>
<td>2.8 ± 0.8</td>
<td>2.1 ± 0.6</td>
<td>2.3 ± 0.7</td>
</tr>
<tr>
<td>200-600</td>
<td>2.1 ± 0.6</td>
<td>2.4 ± 0.7</td>
<td>1.5 ± 0.5</td>
<td>1.7 ± 0.5</td>
</tr>
</tbody>
</table>

Only $R$ (last column) is directly comparable between channels!
The measurements are in good agreement with the NNLO theoretical predictions, as computed with FEWZ.

The vertical error bar indicates the experimental (statistical and systematic) uncertainties summed in quadrature with the theory uncertainty resulting from the model-dependent kinematic distributions inside each bin.

Each data point is located on the horizontal axis at the position where the theoretical function has a value equal to its mean value over the bin.
The CMS Collaboration have measured the Drell-Yan differential cross section normalized to the Z region in the dilepton invariant mass range $15 \text{ GeV} < M(\ell\ell) < 600 \text{ GeV}$.

- It is based on $36 \text{ pb}^{-1}$.

We present results both inside the detector acceptance and in the full phase space.

The effect of final state QED radiation on the results is reported as well.

A correct description of the measurements requires modeling to NNLO for dilepton invariant masses below about 30 GeV.

The measurements are in good agreement with the NNLO theoretical predictions, as computed with FEWZ.

Find the paper in arXiv:1108.0566v1 [hep-ex], submitted to JHEP.
FEWZ@NNLO, MSTW2008:

<table>
<thead>
<tr>
<th>Inv. mass bin (GeV)</th>
<th>R ((10^{-3}))</th>
<th>PDF uncertainties (%)</th>
<th>Theory uncertainties (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-20</td>
<td>812</td>
<td>+4.3 / -3.3</td>
<td>+2.5 / -2.7</td>
</tr>
<tr>
<td>20-30</td>
<td>494</td>
<td>+3.6 / -2.8</td>
<td>+1.9 / -3.6</td>
</tr>
<tr>
<td>30-40</td>
<td>141</td>
<td>+2.7 / -2.3</td>
<td>+3.1 / -2.1</td>
</tr>
<tr>
<td>40-50</td>
<td>55</td>
<td>+2.1 / -1.9</td>
<td>+2.4 / -2.5</td>
</tr>
<tr>
<td>50-60</td>
<td>28</td>
<td>+1.6 / -1.5</td>
<td>+2.6 / -2.0</td>
</tr>
<tr>
<td>60-76</td>
<td>33</td>
<td>+0.9 / -0.9</td>
<td>+2.0 / -2.4</td>
</tr>
<tr>
<td>76-86</td>
<td>58</td>
<td>+0.2 / -0.2</td>
<td>+2.1 / -2.5</td>
</tr>
<tr>
<td>86-96</td>
<td>844</td>
<td>+0.1 / -0.1</td>
<td>+1.8 / -2.2</td>
</tr>
<tr>
<td>96-106</td>
<td>52</td>
<td>+0.2 / -0.2</td>
<td>+2.8 / -2.0</td>
</tr>
<tr>
<td>106-120</td>
<td>13</td>
<td>+0.5 / -0.5</td>
<td>+2.6 / -2.2</td>
</tr>
<tr>
<td>120-150</td>
<td>6.9</td>
<td>+0.9 / -0.9</td>
<td>+2.5 / -1.7</td>
</tr>
<tr>
<td>150-200</td>
<td>2.7</td>
<td>+1.5 / -1.6</td>
<td>+2.0 / -1.8</td>
</tr>
<tr>
<td>200-600</td>
<td>1.3</td>
<td>+2.8 / -2.9</td>
<td>+1.8 / -2.1</td>
</tr>
</tbody>
</table>
Backup

Data-driven background estimation methods:

- **OS/SS method**
  - define 6 categories of dilepton events n|X, where n = 0, 1, 2 is the number of isolated muons and X is OS or SS pair (per mass bin)
  - estimate the ratio OS/SS for the different categories; note that in MC $1|\text{OS} / 1|\text{SS} = 2|\text{OS} / 2|\text{SS}$ and assume it holds for data $N(2|\text{OS}) = N(2|\text{SS}) \frac{N(1|\text{OS})}{N(1|\text{SS})}$

- **Template fit method**
  - use the isolation variable ($I_{\text{rel}}$) shape to discriminate signal from background
  - the background shape is extracted from SS di-muon events
  - the signal shape is extracted from muons in the Z peak given a very tight selection
  - the isolation distribution in the signal region is fitted in bins of muon $p_T$ and di-muon mass

- **e-mu method**
  - predict true ee (or $\mu\mu$) backgrounds using data $e\mu$ candidates
  - $e\mu$ sample is virtually signal free
  - $e\mu \rightarrow ee$ extrapolation: use acceptance ratio from MC and 2x from branching fraction

All of them agree with the MC predictions within uncertainties.
### Background estimations, per channel, per source

<table>
<thead>
<tr>
<th>Inv. mass bin (GeV)</th>
<th>Electrons</th>
<th>Muons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>genuine $e^+e^-$</td>
<td>mis-id electrons</td>
</tr>
<tr>
<td>15-20</td>
<td>0.0 ± 0.2</td>
<td>0.4 ± 0.7</td>
</tr>
<tr>
<td>20-30</td>
<td>2.5 ± 1.7</td>
<td>0.9 ± 1.1</td>
</tr>
<tr>
<td>30-40</td>
<td>14.3 ± 4.6</td>
<td>1.5 ± 1.4</td>
</tr>
<tr>
<td>40-50</td>
<td>31.4 ± 6.9</td>
<td>3.7 ± 2.7</td>
</tr>
<tr>
<td>50-60</td>
<td>19.9 ± 5.2</td>
<td>3.9 ± 2.8</td>
</tr>
<tr>
<td>60-76</td>
<td>22.4 ± 5.3</td>
<td>4.9 ± 3.3</td>
</tr>
<tr>
<td>76-86</td>
<td>8.5 ± 2.8</td>
<td>2.5 ± 2.1</td>
</tr>
<tr>
<td>86-96</td>
<td>12.5 ± 1.8</td>
<td>4.4 ± 3.1</td>
</tr>
<tr>
<td>96-106</td>
<td>3.5 ± 1.8</td>
<td>2.1 ± 1.8</td>
</tr>
<tr>
<td>106-120</td>
<td>3.2 ± 1.9</td>
<td>1.5 ± 1.4</td>
</tr>
<tr>
<td>120-150</td>
<td>7.8 ± 3.1</td>
<td>2.0 ± 1.7</td>
</tr>
<tr>
<td>150-200</td>
<td>5.5 ± 2.5</td>
<td>1.6 ± 1.4</td>
</tr>
<tr>
<td>200-600</td>
<td>3.0 ± 1.9</td>
<td>1.4 ± 1.4</td>
</tr>
</tbody>
</table>

Different challenges suggest different approaches for the two channels.
Backup

Systematic uncertainties *(Muons / Electrons)* In % :

<table>
<thead>
<tr>
<th>Inv. mass bin (GeV)</th>
<th>Energy scale</th>
<th>Efficiency correction</th>
<th>Backgrounds</th>
<th>Unfolding</th>
<th>FSR</th>
<th>Others</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-20</td>
<td>- / 23.4</td>
<td>1.1 / 9.2</td>
<td>3.6 / 6.2</td>
<td>0.4 / 8.7</td>
<td>1.5 / -</td>
<td>1.0 / -</td>
<td>4.2 / 27.3</td>
</tr>
<tr>
<td>20-30</td>
<td>- / 3.6</td>
<td>1.1 / 8.5</td>
<td>3.1 / 2.8</td>
<td>0.2 / 2.1</td>
<td>1.1 / -</td>
<td>1.0 / -</td>
<td>3.6 / 9.9</td>
</tr>
<tr>
<td>30-40</td>
<td>- / 2.7</td>
<td>1.2 / 9.4</td>
<td>1.9 / 4.0</td>
<td>0.1 / 1.5</td>
<td>0.7 / -</td>
<td>1.0 / -</td>
<td>2.6 / 10.6</td>
</tr>
<tr>
<td>40-50</td>
<td>- / 3.3</td>
<td>1.2 / 7.5</td>
<td>1.7 / 5.2</td>
<td>0.2 / 1.4</td>
<td>0.7 / -</td>
<td>1.0 / -</td>
<td>2.4 / 9.9</td>
</tr>
<tr>
<td>50-60</td>
<td>- / 3.3</td>
<td>0.8 / 5.2</td>
<td>2.1 / 4.6</td>
<td>0.2 / 1.9</td>
<td>0.5 / -</td>
<td>0.5 / -</td>
<td>2.4 / 7.9</td>
</tr>
<tr>
<td>60-76</td>
<td>- / 10.3</td>
<td>0.6 / 3.3</td>
<td>1.0 / 2.2</td>
<td>0.2 / 2.0</td>
<td>1.4 / -</td>
<td>0.5 / -</td>
<td>1.9 / 11.2</td>
</tr>
<tr>
<td>76-86</td>
<td>- / 39.5</td>
<td>0.4 / 2.5</td>
<td>0.2 / 0.8</td>
<td>1.7 / 3.1</td>
<td>2.0 / -</td>
<td>0.5 / -</td>
<td>2.7 / 39.7</td>
</tr>
<tr>
<td>86-96</td>
<td>- / 3.9</td>
<td>0.3 / 1.9</td>
<td>0.05 / 0.2</td>
<td>0.2 / 0.6</td>
<td>0.5 / -</td>
<td>0.5 / -</td>
<td>0.8 / 4.4</td>
</tr>
<tr>
<td>96-106</td>
<td>- / 45.6</td>
<td>0.3 / 2.0</td>
<td>0.4 / 0.9</td>
<td>3.8 / 3.6</td>
<td>0.5 / -</td>
<td>0.5 / -</td>
<td>3.9 / 45.8</td>
</tr>
<tr>
<td>106-120</td>
<td>- / 13.2</td>
<td>0.3 / 2.1</td>
<td>1.4 / 2.6</td>
<td>0.7 / 2.4</td>
<td>0.5 / -</td>
<td>3.0 / -</td>
<td>3.4 / 13.9</td>
</tr>
<tr>
<td>120-150</td>
<td>- / 6.0</td>
<td>1.1 / 2.4</td>
<td>2 / 8.2</td>
<td>0.4 / 2.6</td>
<td>0.5 / -</td>
<td>1.0 / -</td>
<td>2.6 / 10.8</td>
</tr>
<tr>
<td>150-200</td>
<td>- / 5.7</td>
<td>2.1 / 2.8</td>
<td>6 / 12.9</td>
<td>0.9 / 2.4</td>
<td>0.5 / -</td>
<td>1.0 / -</td>
<td>6.5 / 14.5</td>
</tr>
<tr>
<td>200-600</td>
<td>- / 4.6</td>
<td>2.1 / 3.2</td>
<td>10 / 11.8</td>
<td>0.1 / 1.6</td>
<td>0.5 / -</td>
<td>1.0 / -</td>
<td>10.3 / 13.1</td>
</tr>
</tbody>
</table>
In the full phase space, in %:

<table>
<thead>
<tr>
<th>Inv. mass bin (GeV)</th>
<th>(μ) FSR correction</th>
<th>(e) FSR correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-20</td>
<td>97.28±0.02</td>
<td>93.8±0.1</td>
</tr>
<tr>
<td>20-30</td>
<td>97.28±0.02</td>
<td>93.9±0.2</td>
</tr>
<tr>
<td>30-40</td>
<td>98.43±0.02</td>
<td>96.8±0.3</td>
</tr>
<tr>
<td>40-50</td>
<td>104.0±0.1</td>
<td>107.7±0.6</td>
</tr>
<tr>
<td>50-60</td>
<td>120.2±0.3</td>
<td>139.3±1.0</td>
</tr>
<tr>
<td>60-76</td>
<td>166.4±0.5</td>
<td>230.7±1.4</td>
</tr>
<tr>
<td>76-86</td>
<td>167.1±0.4</td>
<td>224.1±1.0</td>
</tr>
<tr>
<td>86-96</td>
<td>91.63±0.03</td>
<td>83.9±0.1</td>
</tr>
<tr>
<td>96-106</td>
<td>88.0±0.1</td>
<td>78.5±0.5</td>
</tr>
<tr>
<td>106-120</td>
<td>91.3±0.2</td>
<td>83.9±1.0</td>
</tr>
<tr>
<td>120-150</td>
<td>93.2±0.3</td>
<td>87.9±1.4</td>
</tr>
<tr>
<td>150-200</td>
<td>94.3±0.4</td>
<td>89.1±2.2</td>
</tr>
<tr>
<td>200-600</td>
<td>92.8±0.7</td>
<td>87.5±3.2</td>
</tr>
</tbody>
</table>
Examine data and MC simulation with respect to the FSR
- sum energy in a cone of $\Delta R<0.3$ around the muon
- difference in R between the muon and photon with various cuts on the photon energy

No significant systematic bias
- the FSR modeling in MC shows a remarkable agreement
- possible systematic effects estimated by proper statistical variations of the fraction of FSR events as well as the energy and angular distributions of FSR photons