Measurement of the Z boson transverse momentum spectrum on ATLAS with early data

L. Kashif
Department of Physics, Harvard University, Cambridge, MA 02138, USA

One of the benchmark analyses to be performed with the first data at the CERN Large Hadron Collider will be the measurement of the Z boson transverse momentum spectrum. In this talk, I present a prospective analysis for this measurement in the dimuon channel on the ATLAS experiment. The analysis uses simulated datasets at a center-of-mass energy of 10 TeV. After summarizing the motivations for the measurement, I discuss the Z boson selection criteria, possible physics backgrounds, and background removal techniques with a focus on data-driven background determination. I then present my results, and conclude with an outlook toward the collision data expected in 2009-10.

1. Introduction

The first run of the Large Hadron Collider (LHC) at CERN is scheduled to start very soon. A number of benchmark analyses are being prepared by ATLAS (A Toroidal LHC Apparatus) and CMS (Compact Muon Solenoid), the two general-purpose experiments at the LHC. These analyses will help commission the detectors as well as establish known Standard Model physics in the new energy regime.

One of the benchmark analyses is the measurement of the transverse momentum spectrum of the Z boson. The Harvard group is developing an analysis to perform this measurement, which is presented in this article. In Section 2, we summarize the motivations for the measurement. In Section 3, we give the salient details of our analysis, including data-driven estimations of the major backgrounds and background removal. Section 4 presents the results of the analysis. We conclude in Section 5.

2. Motivations for measuring the Z \( p_T \) spectrum

The transverse momentum spectrum of the Z boson is important for a number of reasons, both theoretical and practical. The principal among these are briefly discussed below.

- Test of QCD predictions: The high-\( p_T \) region of the Z \( p_T \) spectrum is affected by perturbative QCD corrections, while non-perturbative corrections modify the low-\( p_T \) region. As an example, Figure 1 shows the Z differential cross-section as a function of the Z \( p_T \) from two Monte Carlo programs: Pythia and MC@NLO. Pythia uses leading-order matrix elements, while MC@NLO uses next-to-leading-order matrix elements. The difference in the two distributions is clearly visible.

Recently, a number of Monte Carlo tools have become available which combine electroweak and QCD corrections. Predictions from the state-of-the-art in theory can be tested by comparison with the measured Z \( p_T \) distribution.

- Discovery physics: Various physics scenarios beyond the Standard Model predict heavy particles that couple to the electroweak sector, and therefore can decay to high-\( p_T \) Z bosons. These include technicolor condensates, squarks and gluinos from supersymmetric models, right-handed quarks etc. If any of these scenarios is realized in nature, the decay signature of the heavy fields may well show up in the Z \( p_T \) spectrum.

- Inferring properties of the Z \( \rightarrow \nu\nu \) decay: The invisible Z \( \rightarrow \nu\nu \) decay is a background to many processes involving missing transverse energy, such as Supersymmetry and extra-dimension models where a graviton escapes into the higher-dimensional 'bulk'. The Z \( \rightarrow \mu\mu \) decay can be used to indirectly measure the rate and properties of the invisible mode. In order to do this, we need to measure the \( p_T \) spectrum of Z bosons in the muon channel.

- Both the ATLAS and CMS experiments will eventually use the total Z production cross-section to monitor luminosity on a run-by-run basis. Before we can do that, however, we must understand the Z \( p_T \) spectrum very well, especially the low-\( p_T \) region.

3. Details of the analysis

Our analysis is based on official ATLAS Monte Carlo datasets corresponding to a center-of-mass energy of 10 TeV. The datasets were generated with
Figure 1: $Z$ differential cross-section vs $p_T$ from two different MC generators: Pythia (black), which uses leading-order matrix elements, and MC@NLO (red), which uses next-to-leading order matrix elements.

The Pythia and MC@NLO generators\(^1\), and were fully simulated using the GEANT4 ATLAS detector simulation. The sample we use for the signal channel corresponds to an integrated luminosity of 40 $pb^{-1}$, which is a realistic estimate of the amount of data expected in the first year of LHC running.

Table I summarizes our $Z$ boson selection criteria.

| Selection criterion | Events passing cut (%) |
|---------------------|------------------------|
| 20 GeV single muon trigger | 69.7 $\pm$ 0.4 |
| At least one $\mu^+$ and one $\mu^-$ reconstructed | 49.8 $\pm$ 0.3 |
| $|\eta| < 2.5$ for both muons | 47.1 $\pm$ 0.3 |
| $p_T > 20$ GeV for both muons | 41.6 $\pm$ 0.3 |
| 76 GeV $< M_{\mu\mu} < 106$ GeV | 38.5 $\pm$ 0.3 |

3.1. Background reduction

Possible physics backgrounds to the signal channel include:

- $W \rightarrow \mu\nu + jet$
- $b\bar{b} \rightarrow \mu\mu + X$
- $t\bar{t}$ dimuon decays

- $Z \rightarrow \tau\tau \rightarrow \mu\nu_{\tau} \mu\nu_{\tau}$
- $WW, ZZ, WZ$

The first three channels have the largest cross-sections, which are the ones we are studying at this time. We estimate our backgrounds using Monte Carlo samples. However, we are investigating methods to extract the $W$ and $b\bar{b}$ backgrounds from data, as we will discuss shortly. To minimize backgrounds, we use two isolation criteria defined in a cone of size $\Delta R = 0.4$ around the muon track, where

$$\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$$ (1)

and $\eta$ and $\phi$ are the pseudorapidity and the azimuthal coordinate respectively. The isolation criteria are:

- number of tracks in cone, and
- the total transverse momentum of tracks, $\Sigma p_T$, in cone.

Figure 2 shows the distribution of these variables for the signal channel and the three main background channels. We optimize the cuts on these quantities using distributions of $S/\sqrt{S+B}$, where $S$ and $B$ refer respectively to the signal and the total contribution from all three background channels. Figure 3 shows the $S/\sqrt{S+B}$ distributions for the two isolation criteria.

The optimized values of the cuts are given by the position of the peaks:

- number of tracks in cone $< 4$,
- $\Sigma p_T$ of tracks in cone $< 8$ GeV.

Figure 4 shows the distribution of the dimuon invariant mass and the dimuon $p_T$ spectra from the signal channel and the three background channels. Distributions for the quantities are shown before any cuts, after the $Z$ boson selection cuts and before the isolation cuts, and after the isolation cuts. $t\bar{t}$ dimuon decays are the largest background contribution after all cuts. This is expected, since muons from this source are very similar to those from $Z$ decays. Following the isolation cuts, the remaining total background contamination is $\approx 0.2\%$ of the signal, which is small enough that we ignore it in the remainder of the analysis.

3.2. Estimation of $W \rightarrow \mu\nu$ background from data

In a fraction of the $W \rightarrow \mu\nu$ events, the muon from the $W$ decay combines with a muon in a jet to mimic the $Z \rightarrow \mu\mu$ signal. To estimate this background in data, we make the following assumption:

---

\(^1\)The PHOTOS tool was interfaced with both Pythia and MC@NLO to generate final-state photon radiation.
Figure 2: Number of tracks (top) and $\Sigma p_T$ of tracks (bottom) in cone of size $R = 0.4$. Distributions for the muon with the larger value of the quantity is shown on the left, and those for the muon with the smaller value on the right.

Figure 3: Optimization of the muon isolation cuts: (left) number of tracks in cone of size $R = 0.4$ around the muon track, and (right) $\Sigma p_T$ of tracks in cone. On each plot, the arrow indicates the value of the quantity where the peak in the distribution occurs.

The fraction of $Z \to \mu\mu$ events in which three muons pass our selection and isolation cuts is equal to the fraction of $W \to \mu\nu$ events in which two muons pass the same cuts.

To implement this assumption in a data sample, we will first apply the selection and isolation cuts on the sample, and then count the number of events in which three muons pass all the cuts. Multiplying this number by the ratio of the total $W$ cross-section to the total $Z$ cross-section will give an estimate of the $W \to \mu\nu$ events remaining after the cuts.

We tested the validity of our assumption on the
Monte Carlo samples. In the signal sample, $0.029 \pm 0.008\%$ events (13 events) have a third muon passing all the cuts, while $0.027 \pm 0.003\%$ events in the $W \rightarrow \mu \nu$ sample have a second muon passing the cuts. These numbers are consistent with each other, suggesting that our assumption is valid.

### 3.3. Estimation of $b \bar{b}$ background from data

In this case, the idea is to use a sample of $b \bar{b}$ events with non-isolated muons to estimate the number of $b \bar{b}$ events remaining after isolation cuts. In other words, we want to use a sample of $b \bar{b}$ events with non-isolated muons as a template for $b \bar{b}$ events with isolated muons. Figure 5 shows the dimuon invariant mass distribution in $b \bar{b}$ events with our isolation cuts and with reversed isolation cuts. Reversing the isolation cuts on a data sample selects a very pure sample of $b \bar{b}$ events with non-isolated muons. Within statistics, the two invariant mass distributions have similar shape, indicating that the distribution with non-isolated muons can be used as a template for $b \bar{b}$ background after isolation cuts. The signal template is a dimuon invariant mass distribution from $Z \rightarrow \mu \mu$ events after isolation cuts. A template fit now gives the fractions in which the signal and background templates must be combined to obtain the observed distribution.

We test this method on our Monte Carlo samples. The template fit gives $96 \pm 1\%$ for signal fraction and $3.9 \pm 0.4\%$ for $b \bar{b}$ background fraction. The known inputs to our 'data' distribution were $97\%$ signal and $3\%$ $b \bar{b}$ events, so the template fit seems to yield the correct numbers. The left-hand plot in Figure 6 shows the dimuon invariant mass distributions from 'data' and from the template fit result for invariant mass $> 20$ GeV; good agreement is observed. The right-hand plot in Figure 6 shows the distribution of signal fraction from the template fit in 1000 pseudo-experiments.
In each experiment, the numbers of signal and $b\bar{b}$ events in each bin of the 'data' distribution were varied within their Poisson fluctuations, and the template fit was done using the same signal and $b\bar{b}$ templates. The width of the Gaussian fit gives an estimate of the resolution that can be expected from the template fit.

4. Results

The measured $Z p_T$ spectrum must be corrected for various detector and reconstruction effects, including:

- trigger and reconstruction efficiencies,
- geometric and kinematic acceptance,
- resolution smearing.

We derive the correction factors in $Z p_T$ bins using Monte Carlo truth information (Figure 7). Figure 8 shows the measured $Z p_T$ spectrum from our sample before and after the above corrections, as well as the true spectrum. Since the correction factors in this case were obtained using the true spectrum, the good agreement between the corrected measurement and the prediction is, of course, expected.

The differential cross-section as a function of $Z p_T$ is given by:

$$\frac{d\sigma}{dp_T} = c_i \frac{s_i - b_i}{\Delta p_T, L_{int}, A_i}$$

where

- $s_i =$ signal events in $i$-th $p_T$ bin,
- $b_i =$ background events in $i$-th $p_T$ bin,
- $\Delta p_T =$ width of $i$-th $p_T$ bin,
- $L_{int} =$ integrated luminosity,
- $A_i =$ acceptance for $i$-th bin,
- $c_i =$ efficiency for $i$-th bin,
- $c_i =$ smearing correction for $i$-th bin.

Figure 9 shows the $Z$ differential cross-section before and after corrections, together with the expected distribution from Monte Carlo truth. As before, the good agreement between the corrected measurement and the prediction is expected.

5. Conclusion and outlook

We are developing an analysis to measure the $Z p_T$ spectrum in the dimuon channel on the ATLAS experiment. During the first LHC run, we expect $\approx 100pb^{-1}$ of delivered integrated luminosity, so that we should have a few tens of $pb^{-1}$ of luminosity on tape, assuming realistic data-taking efficiency. Our analysis is geared toward that amount of data. We are exploring methods to extract the major backgrounds to our signal channel with minimal dependence on Monte Carlo.

In the first run, the instantaneous luminosity will be in the range $10^{29-31} cm^{-2}s^{-1}$, so that cavern background and event pileup are not likely to be major...
Correction factors as a function of $Z_{pt}$, extracted using truth information:

*Top left:* efficiency (reconstruction and trigger)

*Top right:* acceptance (geometric and kinematic)

*Bottom left:* correction for resolution smearing

Figure 7: Corrections to the $Z_{pt}$ spectrum.

Figure 8: $Z_{pt}$ distribution from Monte Carlo truth (black), measured (uncorrected) (blue) and corrected (red). The plot corresponds to 40 pb$^{-1}$ of integrated luminosity. Errors are statistical.

Figure 9: $Z$ differential cross-section vs $p_T$ from Monte Carlo truth (black), measured (uncorrected) (blue) and corrected (red). The plot corresponds to 40 pb$^{-1}$ of integrated luminosity. Errors are statistical.

issues. The main challenge for this analysis will be understanding the detector, including muon trigger acceptance and efficiency, offline reconstruction efficiency, muon momentum scale and resolution, and the systematics associated with all of them.