CDF/PHYS/ELECTROWEAK/PUB/9910
September 30, 2009
Version 0

Observation of $Z^0Z^0 \rightarrow lll'l'$ at CDF

The CDF Collaboration
Dated: August 20, 2009

Abstract

The Z^0Z^0 cross section measurement is important to test Standard Model predictions of Electro-Weak couplings. In addition the Z^0Z^0 reconstruction is an useful training for Higgs searches in the high mass region. In 1.9fb^{-1} of data CDF has observed a Z^0Z^0 signal with a significance of 4.4σ when combining $Z^0Z^0 \rightarrow lll'l'$ and the $Z^0Z^0 \rightarrow ll\nu\nu$ decay channels. Here, the search for $Z^0Z^0 \rightarrow lll'l'$ with 4.8 fb^{-1} of data exploiting an extended lepton coverage is presented. The number of expected signal events is 4.68 ± 0.76, we observe 5 events with $4.15 \pm 1.62(stat.) \pm 2.87(syst.) \times 10^{-2}$ background events corresponding to a significance of 5.70σ. This is the first observation of a Z^0Z^0 signal using CDF data. With these events we measure a cross section of $1.56^{+0.80}_{-0.63}(stat.) \pm 0.25(syst.)$ pb, in agreement with Standard Model value 1.4 ± 0.1 pb at Next to Leading Order.

1 Introduction

The ZZ events are produced at Tevatron with a cross section the order of few pb, similar to what is expected in the Standard Model (SM) for certain Higgs masses. Moreover at high mass the Higgs boson is expected to decay to ZZ with a significant branching ratio and the reconstruction of the ZZ is a pre-requisite for this search. Finally the ZZ production cross section is sensitive to new physics through anomalous trilinear gauge coupling [1] and large extra-dimensions [2]. The first measurement of ZZ production cross section at the Tevatron was performed at CDF exploiting $ZZ \rightarrow lll'l'$ and $ZZ \rightarrow ll\nu\nu$ decay channels using 1.9 fb^{-1} of data [3]. Here we describe an update of the previous measurement using 4.8 fb^{-1} of data. We search only for $ZZ \rightarrow lll'l'$ decay channel with extended lepton categories. This decay channel has a small branching ratio but it has a very low background contamination and it can be measured quite precisely. These characteristics will allows a ZZ discovery, ie. a signal with at least 5σ significance, at CDF.

2 Data and Monte Carlo description

The data used in the $ZZ \rightarrow lll'l'$ search has been collected using the high momentum muon and high energy electron triggers corresponding to a integrated luminosity of 4.8 fb^{-1}. The evaluation of the background contamination to the searched signal has been performed by using the jet20, jet50, jet70 and jet100 datasets as it will be discussed later. The Monte Carlo (MC) simulation based on Pythia as generator and CDFSim as emulator of the detector response has been used to produce ZZ signal and some physics processes used in the background study

The Monte Carlo ZZ events, generated using Pythia, have both Z's that decay inclusively. A filter at generator level requires two leptons with p_T greater than 1 GeV/c. The ZZ Monte Carlo includes γ^* component with M_{ll} >15 GeV/c^2.

The WZ, $t\bar{t}$ and $Z\gamma$ samples are used in the background studies. The top sample has been generated with a mass of 175 GeV/c^2 with no minimum bias added to the event. The Bauer generator was used for $Z\gamma$. The events are filtered requiring the presence of a γ with $p_T >$ 4GeV/c and two leptons with M_{ll} >15 GeV/c^2.

The Monte Carlo samples are scaled on an event-basis in order to account for the luminosity of the data sample, the generated process cross-section as well as, in case, the generator-level filter efficiency. We also account for the measured trigger efficiency, different lepton reconstruction and identification efficiencies in data and simulation and the vertex position requirement. The used formula is:

$$\frac{\sigma \times \mathcal{B} \times \epsilon_{\text{filter}} \times \epsilon_i^{\text{trig}} \times s_i^{\text{lep}} \times \epsilon_{\text{vtx}} \times \mathcal{L}}{N_i^{\text{gen}}(|Z_0| < 60 \text{ cm})} \tag{1}$$

where

σ	is the cross-section for the Monte Carlo process
$\mathcal{B}$	is any branching fraction for the Monte Carlo process
ϵ_{filter}	is the filter efficiency applied for any filter used in the generation process
ϵ_i^{tr}	is the effective trigger efficiency for the event i (see below)
s_i^{lep}	is the effective lepton id scale factor for the event i (see below)
ϵ_{vtx}	is the run dependent efficiency of the z-vertex position requirement ($\|z_0\| < 60$ cm)
$\mathcal{L}$	is the luminosity of the dataset used

The trigger efficiencies are evaluated using a sample of $Z \to ll$ data events as well as the lepton scale factor using the *tag & probe* method that will be described later.

3 Lepton Selection

The lepton categories used in this analysis are:

- Electron: central electron and plug electron.
- Muon: central muon categories and the new categories; muon identified only using the outermost detector (Central Muon uPgrade), muons with stub in the extension detector and in the forward detector.
- Track: track that end ups in a not instrumented detector region.

The selection cuts, the procedures for the scale factor determination and the fake rate calculations are the same described in [4]

3.1 Lepton Identification

The lepton identification criteria are the same used in [4] with the exception of the track isolation that is not required in order to increase the signal acceptance. The new muon categories added to this analysis are described in the following section.

The central muon detector and the outermost detector detect the major part of the central muons but leave some areas in η-ϕ not instrumented by one of the detectors. We recover these muons by selecting tracks with good quality that are fiducial to these detectors. The central muon extensions detector extends the muon coverage up to $|\eta| < 1.0$. It mainly consists of two arches with two holes on top and bottom that at the beginning of RunII were filled with chambers 90° wide, *Miniskirts* at the bottom of the detector and one 30° wide *Keystone* at the top of one side. The extension-based trigger paths now include information from these chambers to trigger on high-p_T muon events. We define a separate category for these kind of events, which basically differs from the previous ones only by ϕ regions requirements..

Muon coverage has been further extended in the region $1.0 < |\eta| < 1.5$ by the Intermediate Muon detector, consisting of the barrel muon detector, the barrel scintillator upgrade and the toroid scintillator upgrade. We define only one category based on finding a stubs in any of the above detectors. There is no trigger path connected with this category that can be exploited in this analysis, but its usage can increase the purity of the current stubless forward muon category. Since the stubless forward muons have the highest fake rate among the muon categories, requiring a stub we select a subset of events with higher purity.

A visual representation of the η-ϕ coverage of the final muon categories is given in Figure 1 for a small subset of data. Finally, for the new categories we require a p_T (for muons) or E_T (for electrons) greater than 10 GeV/c, or 20 GeV/c^2 for triggerable objects as for the old ones.

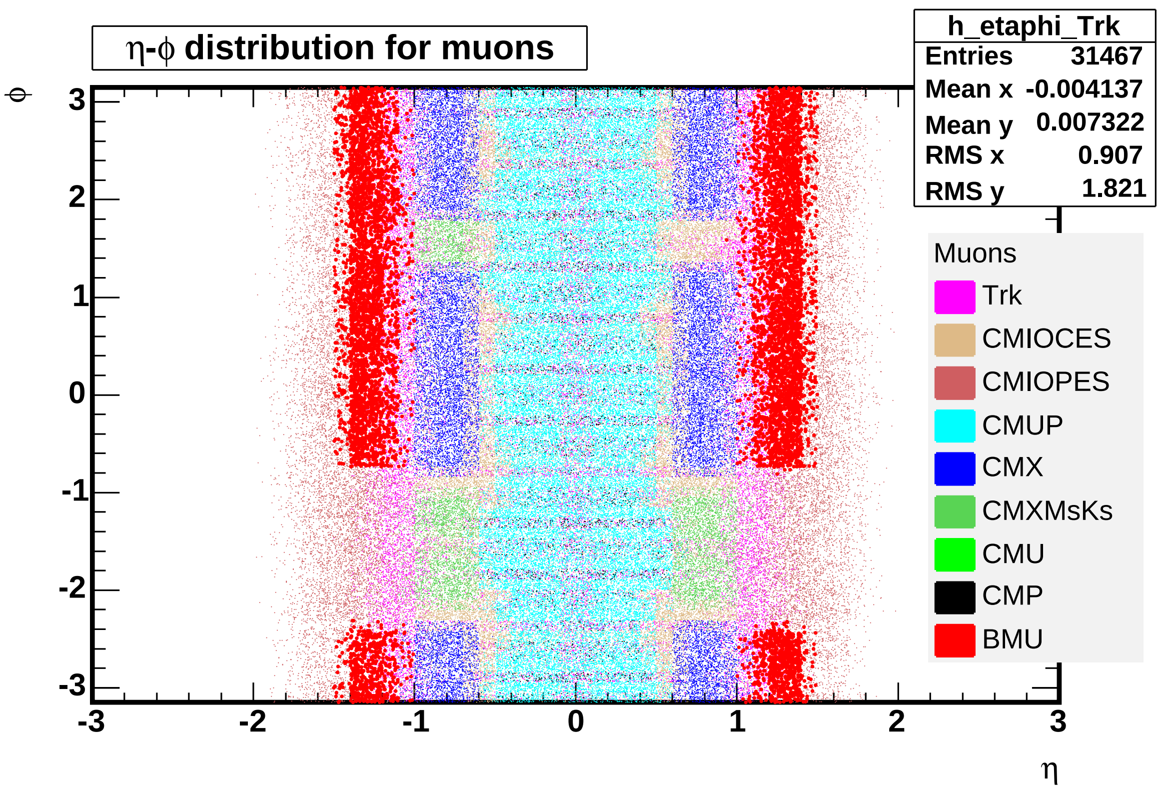

Figure 1: The η-ϕ coverage of all the muon categories used in this analysis.

3.2 Lepton Identification efficiency Scale Factors

The lepton identification efficiency scale factors (shortened in lepton ID or SF) are defined as the ratio of the lepton ID efficiencies in data and MC data. We use the *tag & probe* method. Z candidates are selected with the $67 < m(ll) < 106$ GeV/c^2 dilepton invariant mass window, and leptons are required to have a $p_T(E_T)$ greater than 20 GeV/c. The efficiency of the probe objects are assumed to be well reproduced by the Monte Carlo and then collider events are corrected to match the MC values using the SF. Generic probes are defined for central and forward muons and then used with the fiduciality requirement to define the actual probe used in the efficiency calculation. The background is subtracted using data from the Z-mass sidebands. The scale factors are taken from the $H \rightarrow WW$ analysis [5] since we use the same lepton categories.

4 ZZ Event Selection

The event selection consists in the reconstruction of two Zs each one that decays in two leptons. We summarize here the cuts:

- Exactly 4 leptons.
- At least one trigger Lepton ($p_T \geq 20 GeV/c$).
- Each other lepton must have $p_T \geq\ 10\ GeV/c$.
- min $dR_{Leptons} > 0.1$.
- M_{ll1} in (76-106) GeV/c^2 (the closest to the nominal Z mass), M_{ll2} in (40-140) GeV/c^2.
- Each lepton pair from Z must have Same Flavor and Opposite Charge.

The cuts described above have been applied to the ZZ Monte Carlo data sample to extract the number of expected events: **4.68 ± 0.02(stat.) ± 0.76(syst.)**

5 Background Determination

The dominant backgrounds are Z+jets and $Z\gamma$+jets where two jets or a jet and a photon are misreconstructed as leptons. Contributions from WZ and $t\bar{t}$ are negligible. Since the Monte Carlo simulations can not reproduce these process with the necessary accuracy the contamination of these events to the searched signal has been evaluated using the data. The procedure involves the selection of identified leptons (*numerator objects*) and fakeable objects (*denominator objects*) in a data sample enriched in jets after the electroweak contribution removal. The fake rate as function of the p_T is determined using the following formula:

$$f_i(p_T) = \frac{N_i(Identified\ Leptons)\ -\ \sum_{j \in EWK} N_{ij}(Identified\ Leptons)}{N_i(Denominator\ Objects)\ -\ \sum_{j \in EWK} N_{ij}(Denominator\ Objects)} \quad (2)$$

for each lepton category i. Fake rates have been measured using different jet samples, `JET20`, `JET50`, `JET70` and `JET100` to avoid any possible bias due to the sample transverse momentum. The final value is obtained averaging the results obtained in each jet data set and taking as systematic uncertainty the maximum spread between the different measurements.

Then, in our data sample we look for events containing at least one denominator object, we weight each denominator object with the appropriate fake rate $f_i(p_T)$.

In order to measure the Z+jets and $Z\gamma$+jets contribution to the signal we evaluate the number of events with three real leptons and one fake lepton and the number of events with two real leptons and two fake leptons. This method does not take into account the possible contamination of real ZZ events to the lepton+denominator(s) sample. The amount of real ZZ events with one (or more) lepton(s) failing the lepton selections could be a consistent fraction of our background and we cannot neglect that contribution. We can calculate the ZZ contamination to the background estimation from a MC sample of ZZ events but we decide to use a background selection that reduce them as much as possible, in order to avoid a complex iterative correction that would be needed to proper subtract it. The total number of expected background events is $(4.1\pm1.6\pm2.9)\cdot10^{-2}$ Since the data driven background determination has shown some critical points we decided to cross check the results by using the Monte Carlo data. We evaluated the contribution to our data sample from different physics process applying the measured fake rates to the denominator objects to WZ,$t\bar{t}$ and $Z\gamma$ Monte Carlo data scaled according to Eq. 1. The result obtained is in agreement within the errors with the data driven method output.

6 Systematic Errors Determination

The systematics uncertainties considered in this analysis are summarized in Table 1. These

Fractional Uncertainty (%)	
NLO Acceptance	10.0 %
Cross - section	10.0 %
PDF uncertainty	2.7 %
Luminosity	6.0 %
LeptonID $\pm1\ \sigma$	3.6 %
Trigger Efficiency	2.1 %
Background Modeling	67 %

Table 1: Summary of the systematic errors.

systematic errors are used to determine the signal sensitivity and then they are propagated to the determination of the systematics uncertainty on the measured cross section. The dominant contribution comes from the fake rates calculation uncertainties described in the previous section. This error contributes mainly to the sensitivity determination while it has a limited impact to the cross section measurement given the small number of background events.

The systematic uncertainties due to the lepton ID efficiencies are calculated varying coherently the lepton ID scale factors by 1σ for each lepton and counting the number of expected events. From a signal MC sample we found a variation of $\pm 3.6\%$ and we take this as systematic error.

The uncertainties due to trigger efficiency are calculated by varying trigger scale factor for the triggerable(s) lepton(s). We found a variation of $\pm 2.1\%$ We evaluated also the contribution of the E_T threshold of the electron trigger by lowering it to 16 GeV/c^2 instead of the nominal 20 GeV/c^2 in the triggerable definition. We found anyway this correction to be negligible in the total trigger efficiency systematics (<0.1 %) and so we didn't included.

A systematic of $\pm 6\%$ is used on the total luminosity, as done in all the CDF analysis.

We assigned an uncertainty of $\pm 10\%$ to the theoretical ZZ production cross section following the theoretician calculations when we evaluated the expected number of events while an error of $\pm 10\%$ is caused by the variation of the acceptance when ZZ signal is generated with MC at the Next to Leading Order . A systematic of 2.7% is added due to the Parton Distribution Functions (PDF) limited knowledge.

7 Results

In $\mathcal{L} = 4.8\ fb^{-1}$ we expect 4.68 ± 0.76 events; we observed **5 events** with $4.15 \pm 1.62(\text{stat}) \pm 2.87(\text{syst}) \cdot 10^{-2}$ expected background events. The scatter plot in figure 2 shows the invariant mass distribution of the two lepton for sub-leading P_T Z^0 versus the invariant mass distribution of the two lepton for leading P_T Z^0 for data (blue stars) and expected signal from MC (black box) with the small background contribution (red box). Table 2 summarizes the number of events expected as signal and background and the observed ones.

Events in $\mathcal{L} = 4.8$ fb^{-1}	
Signal	4.68 ± 0.02(stat.) ± 0.76(syst.)
Z(γ)+jets	0.041 ± 0.016(stat.) ± 0.029(syst.)
Total expected	4.72 ± 0.03(stat.) ± 0.76(syst.)
Observed	**5**

Table 2: Summary of the number of expected and observed signal and background events.

The significance of the signal is 5.70 σ and it is calculated by using the probability to have a number of events equal or greater than those observed as statistical and systematic fluctuations of the estimated background. Figure 3 shows the probability distribution that the background fluctuates to give n number of events while figure 4 displays the probability distribution to observe 1,2..N events with the actual data. Table 3 reports the probability of a 2,3,5 sigmas significance observation and table 4 shows the statistical significance and the p-value obtained with the actual data. The characteristics of each

Probability of Observing a Signal	
Significance	Probability
2σ	0.99
3σ	0.95
5σ	0.70

Table 3: Discovery probability.

Observed Results	
Events	5
P-Value	1.2×10^{-8}
Significance	$5.7\ \sigma$

Table 4: Summary of statistical significance

event that pass the selections are summarized in table 5. The first two events are the same found by the analysis performed on 1.9 fb^{-1}.

Candidate	**leptons**	$\mathbf{M_{ll-1}}$	$\mathbf{M_{ll-2}}$	**4 lepton invariant mass**
1	$trk\mu/\mu\mu$	90.5 GeV/c^2	88.5 GeV/c^2	324.8 GeV/c^2
2	$trk\mu/\mu\mu$	91.6 GeV/c^2	94.2 GeV/c^2	169.4 GeV/c^2
3	$ee/\mu\mu$	93.0 GeV/c^2	86.4 GeV/c^2	191.9 GeV/c^2
4	$ee/\mu\mu$	93.3 GeV/c^2	79.7 GeV/c^2	229.2 GeV/c^2
5	$\mu\mu/\mu\mu$	91.7 GeV/c^2	55.1 GeV/c^2	325.0 GeV/c^2

Table 5: List of events that pass the selection cuts

Even with very low statistic we made a first attempt to study the kinematics properties of this signal. Figure 5 compares the p_T of Zs in the ZZ MC (left) and on data (right). In figure 6 the four leptons invariant mass is shown for data (blue stars) and for Monte Carlo.

Using these 5 ZZ events we can calculate the production cross section:

$$\sigma = \frac{N_{Obs} - N_{Bck}}{\mathcal{L} \cdot \epsilon \cdot \mathcal{B}} \tag{3}$$

where ϵ is the total efficiency, $\mathcal{B}$ the $Z \to ll$ branching fraction, N_{obs} the observed events. N_{bck} the expected number of background events and $\mathcal{L}$ the total integrated luminosity:

$$\sigma_{ZZ} = 1.56^{+0.80}_{-0.63}(stat.) \pm 0.25(syst.) \quad (4)$$

in agreement with the previous measurements done by CDF [3] and D0 [6] and with the theoretical expectation: $\sigma_{ZZ} = 1.4 \pm 0.1$ pb.

References

[1] K. Hagiwara, R. D. Peccei, D. Zeppenfeld, and K. Hikasa, Nuclear Physics B 282, 253 1987

[2] M. Kober, B. Koch, and M. Bleicher, Phys. Rev. D76, 125001 (2007)

[3] T. Aaltonen et al. CDF Coll. *First Measurement of ZZ Production in $p\bar{p}$ Collision at* $\sqrt{s} = 1.96\ TeV$ Phys.Rev.Lett.100:201801,2008

[4] A. Abulencia et al. CDF Coll. *Observation of WZ Production* Phys.Rev.Lett.98:161801,2007

[5] Doug Benjamin, Dean Andrew Hidas, Mark Kruse, Eric James, Sergo Jindariani, Britney Rutherford, Roman Lysak, Aidan Robson, Rick St. Denis, Peter Bussey, Matthew Herndon, Jennifer Pursley, Simone Pagan Griso, Donatella Lucchesi *Search for H-¿WW Production at CDF using 3.6 fb-1 of Data* $/cdf/www/pub/cdf9764_pubnoteHWW36fb.ps$

[6] V. M. Abazov et al. D0 Coll. *Observation of ZZ Production in $p\bar{p}$ Collision at* $\sqrt{s} = 1.96\ TeV$ Phys. Rev. Lett. 101, 171803 (2008).

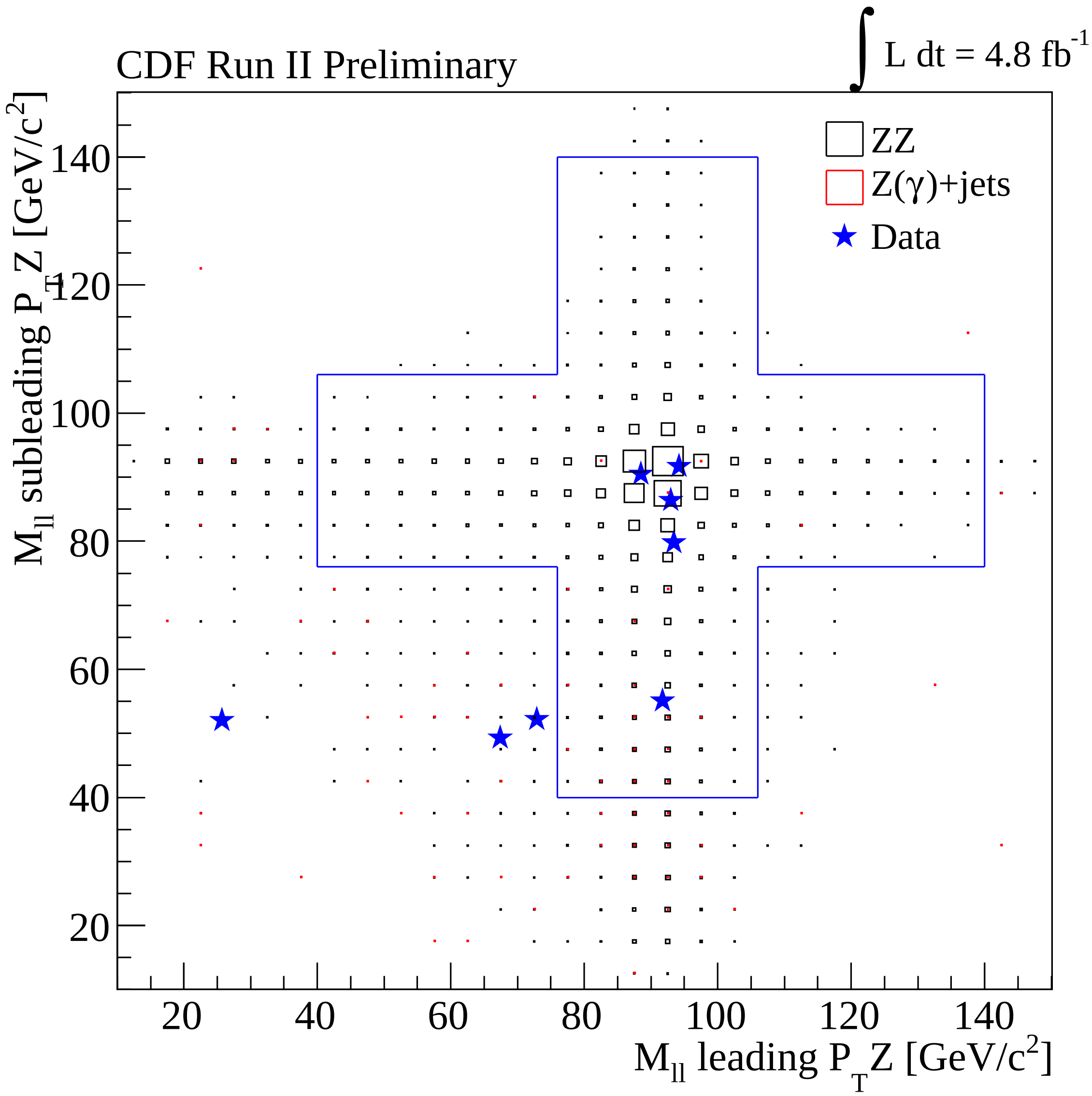

Figure 2: Invariant mass distribution for the two Z's

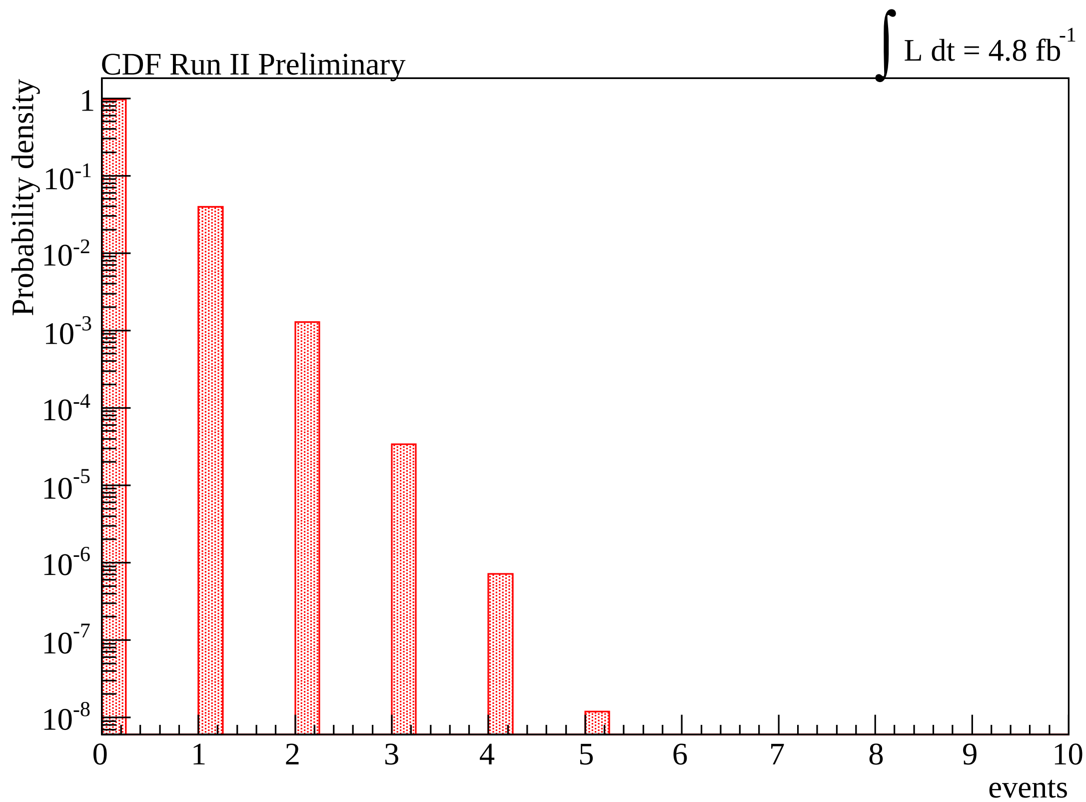

Figure 3: Probability distribution that the background fluctuates to give n events of signal.

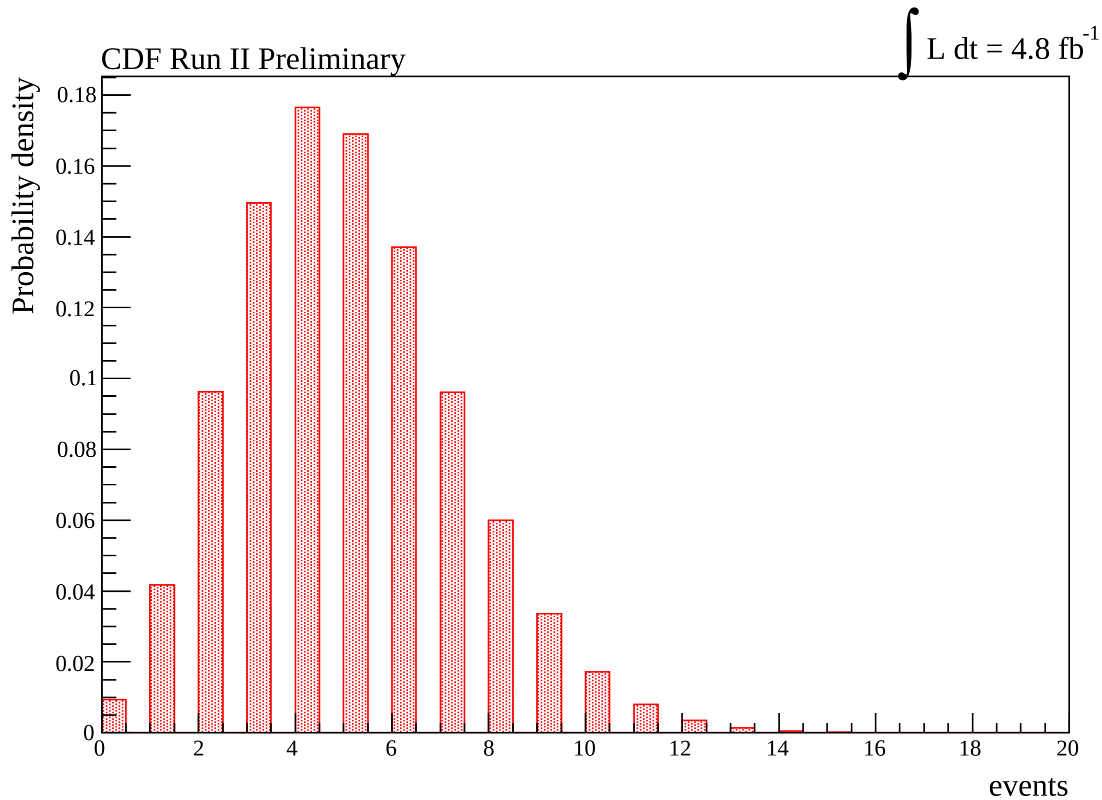

Figure 4: Probability distribution to observe 1,2,... N events in the dataset used for this search

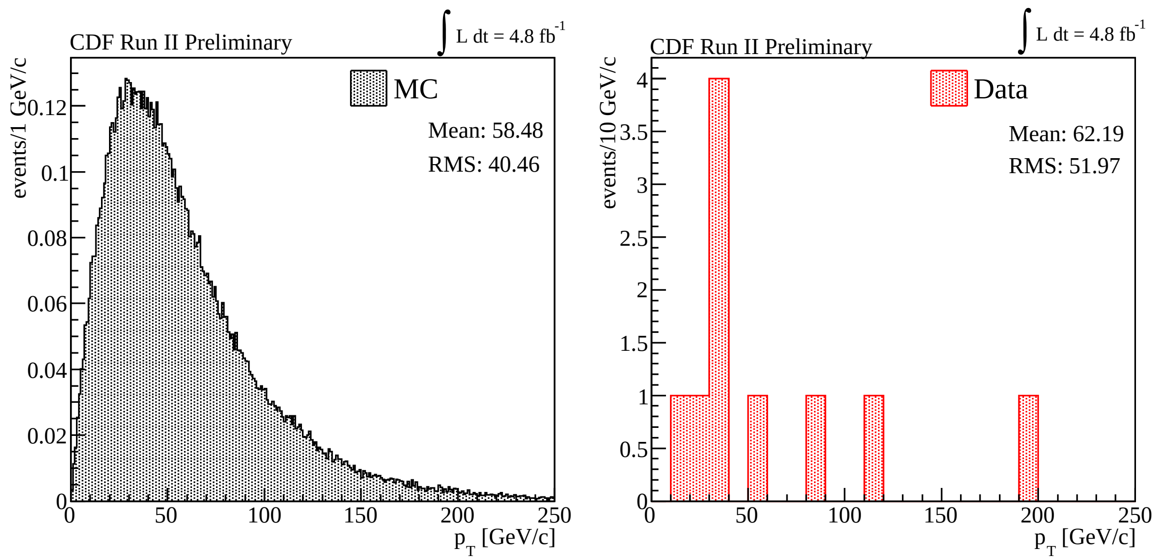

Figure 5: p_T of the two Z for each events, in black MC sample of ZZ, in red the events found in the Data sample

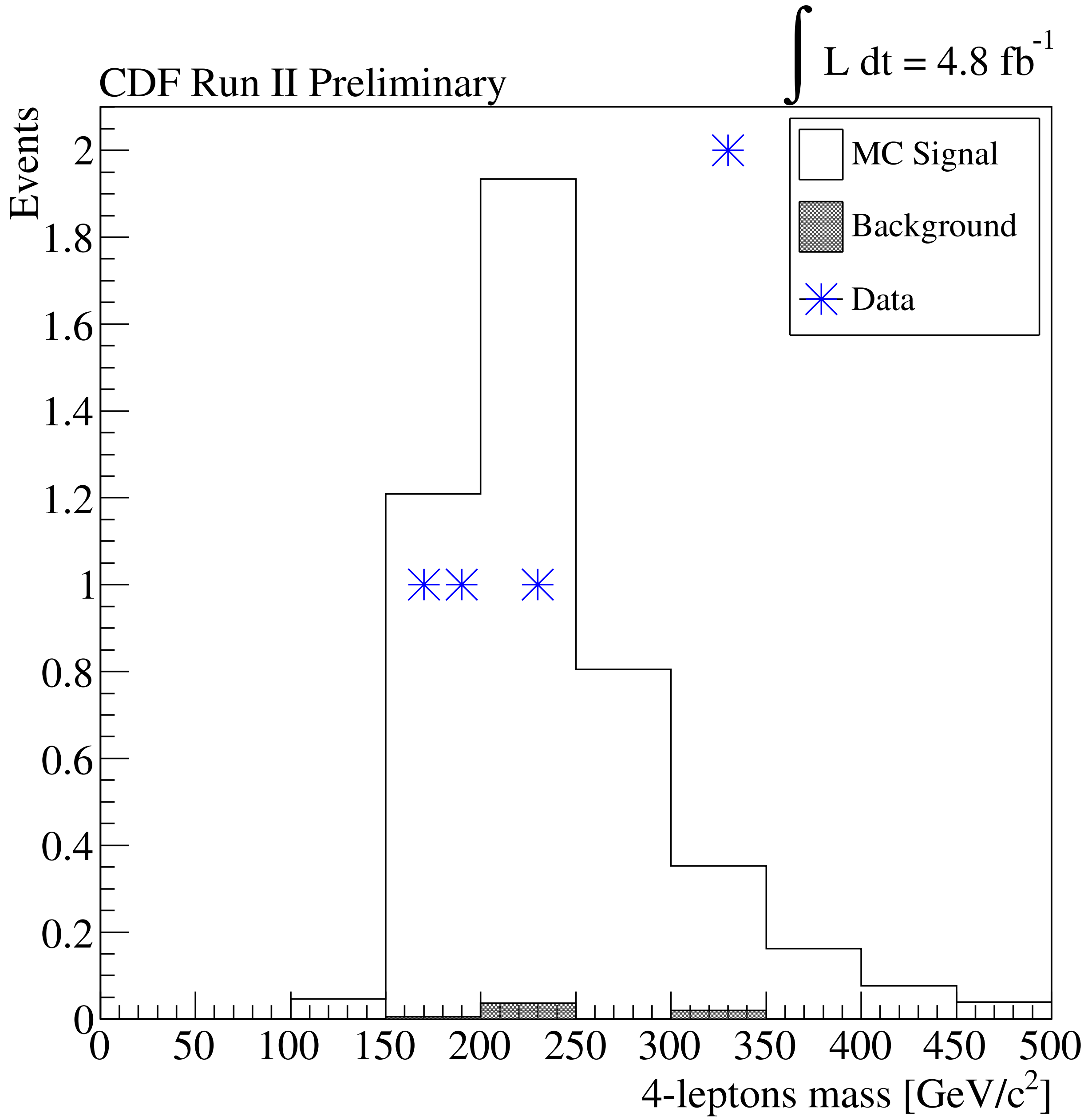

Figure 6: Four lepton invariant mass distribution for data (blue starts with a bin size of 25 GeV/c^2) and for Monte Carlo (solid line with bin size of 50 GeV/c^2).