CDF/TOP/CDFR/7583
Draft 2.0
May 9, 2005

Measurement of the Top Quark Mass in the Dilepton Channel using the Leading-Order Differential Cross-Section

A. Kovalev, D. Whiteson, H.H. Williams (Penn)
B. Jayatilaka, M. Tecchio, D. Gerdes (Michigan)

Abstract

We present a measurement of the top quark mass in the dilepton channel using a technique in which we form a posterior probability for the mass as a product of the per-event differential cross-section for leading order top quark pair production. The calculation of the differential cross section for processes which produce background events are used to reduce the impact of background events in the sample.

In $\int L = 335pb^{-1}$, we expect a statistcal uncertainty of 9.2 GeV ($M_t = 178.0$) and a systematic uncertainly of 3.7 GeV.

Contents

List of Figures

List of Tables

1 Introduction

The mass of the top quark is a free parameter of the Standard Model; it is of great interest due to its exceedingly large value and the constraints it places on the mass of the Higgs boson. Though the number of top quark pairs observed at the Fermilab Tevatron is small, they comprise the entire sample available for study in advance of the completion of the Large Hadron Collider, making this small sample extremely valuable. We seek to extract the maximal information from these events; to do so, we employ a natural and intuitive, though calculationally intensive technique that uses our detailed understanding of the physical production process [1].

In Section 2, we describe the general structure of the method. Section 3 describes the parametrization of detector response and jet resolution. Section 4 details the calculation of the per-event differential cross-section for $t\bar{t}$ production. Section 5 describes the calculation for background processes, and Section 6 describes the prescription for extracting the top mass from an ensemble of events which contain both signal and background contributions. Section 7 describes the event selection, and Section 8 measures the response of the technqiue and provides a calibration. Section 9 contains the results from the accumulated data sample.

2 Overview of the Technique

The complete information contained in an event $(\mathbf{x})$ regarding the top mass can be expressed compactly as the conditional probability:

$$P(\mathbf{x}|M_t)$$

where M_t is the top pole mass and $\mathbf{x}$ is a vector of measured event quantities. We note that the conditional probability $P(\mathbf{x}|M_t)$ is well known from our theoretical understanding of the top quark production process:

$$P(\mathbf{x}|M_t) = \frac{1}{\sigma(M_t)} \frac{d\sigma(M_t)}{d\mathbf{x}}$$

where $\frac{d\sigma}{d\mathbf{x}}$ is the per-event differential cross section. Making the explicit assumption of a uniform prior in top mass, we can set

$$P(M_t|\mathbf{x}) \propto P(\mathbf{x}|M_t)$$

If the momentum of each parton could be exactly deduced from final-state particles, then the calculation of $\frac{d\sigma}{d\mathbf{x}}$ would be simple. Instead, we must integrate over quantities which are unknown because they are unmeasured by the detector, such as neutrino energies. In addition, quark energies are not directly measured, but are estimated from the observed energies of the corresponding jets. We parametrize this uncertainty using a transfer function between quark and jet energies. These are typically expressed as functions $f(p, j)$, the probability of measuring jet energy j given parton energy p. The total expression for the probability of a given pole mass for a specific event can be written as

$$P(\mathbf{x}|M_t) = \frac{1}{N}\int d\Phi_6 |\mathcal{M}_{t\bar{t}}(p;M_t)|^2 \prod_{jets} f(p_i, j_i) f_{PDF}(q_1) f_{PDF}(q_2)$$

where the integral is over the entire six-particle phase space, $\mathbf{x}$ is the vector of measured quantities, charged lepton and jet momenta, $\mathbf{x} = (l, j_0, ..., j_i)$, q is the vector of incoming parton-level quantities and p is the vector of resulting parton-level quantities, lepton and quark momenta. The constant term in front of the integral satisfies the normalization condition for the probability:

$$\int d\mathbf{x}\ P(\mathbf{x}|M_t) = 1$$

Constraints such as conservation of momentum which appear as delta functions and modify the integration are here implicity included in the phase-space integration and are discussed in detail in the following sections.

2.1 Signal and Background processes

The probability $P(\mathbf{x}|M_t)$ is sufficient to extract the top quark mass in an unpolluted sample. However, the top quark candidate events collected by CDF have a small fraction of background events which mimic the top quark signature. To reduce the effect of these events on the measurement, we calculate the probability $P_{bg}(\mathbf{x})$ that they were produced by a background process; we can form the generalized per-event probability as

$$P(\mathbf{x}|M_t) = P_s(\mathbf{x}|M_t)P_s + P_{bg1}(\mathbf{x})P_{bg1} + P_{bg2}(\mathbf{x})P_{bg2}...$$

simply a sum of the probabilities for each process, weighted by their respective prior. Section 5 describes the likelihood formalism in detail.

3 Transfer Functions

The transfer functions $f(x, y)$ describe the probability of reconstructing an object of energy y given its energy x immediately after the hard-scattering process. Of particular importance to this measurement is the transfer function, $f(E_{parton}, E_{jet})$, which describes the probability of a parton with energy E_{parton} being measured as a jet with energy E_{jet} in the CDF detector.

3.1 Parameterization of $f(E_{parton}, E_{jet})$

We extract the jet energy transfer functions from simulated events by parametrizing the measured jet energies, E_{jet}, as a function of the quark energies, E_{parton}. We parameterize this transfer function, $f(E_{parton}, E_{jet})$ as $f(\delta)$ (in a manner similar to D0 in Run I [**?**]), where $\delta \equiv E_{parton} - E_{jet}$. $f(\delta)$ is parameterized as two gaussians, one to account for the peak of the δ distribution, and one for the tails:

$$f(\delta) = \frac{1}{\sqrt{2\pi}(p_2 + p_3 p_5)} \Big[\exp \frac{-(\delta - p_1)^2}{2p_2^2} + p_3 \exp \frac{-(\delta - p_4)^2}{2p_5^2} \Big] \tag{1}$$

Where each p_i depends linearly on E_{parton}:

$$p_i = a_i + b_i E_{parton} \tag{2}$$

The parameters p_i are extracted by expressing $f(E_{parton}, E_{jet})$ as:

$$n(E_{jet}, E_{parton}) dE_{jet} dE_{parton} = n(E_{parton}) dE_{parton} f(E_{parton}, E_{jet}) \tag{3}$$

Where $n(E_{jet}, E_{parton}) dE_{jet} dE_{parton}$ is the number of events with jet energy between E_{jet} and $E_{jet} + dE_{jet}$ and parton energy between E_{parton} and $E_{parton} + dE_{parton}$ and $n(E_{parton})$ is the number of partons with energy between E_{parton} and $E_{parton} + dE_{parton}$. To extract the parameters of $f(E_{parton}, E_{jet})$, we perform an un-binned maximum likelihood fit using $t\bar{t}$ MC events. The likelihood is defined as:

$$\begin{aligned} \ln L &= -\sum_{i=1}^{N} \ln n(E_{jet}^{(i)}, E_{parton}^{(i)}) \\ &= -\sum_{i=1}^{N} \ln n(E_{parton}^{(i)}) - \sum_{i=1}^{N} \ln f(E_{parton}^{(i)}, E_{jet}^{(i)}) \end{aligned} \tag{4}$$

Since the first term does not depend on $f(E_{parton}, E_{jet})$, it can be dropped from the minimization. Using matched b-quarks and jets from `HERWIG` $t\bar{t}$ events for $M_t =$

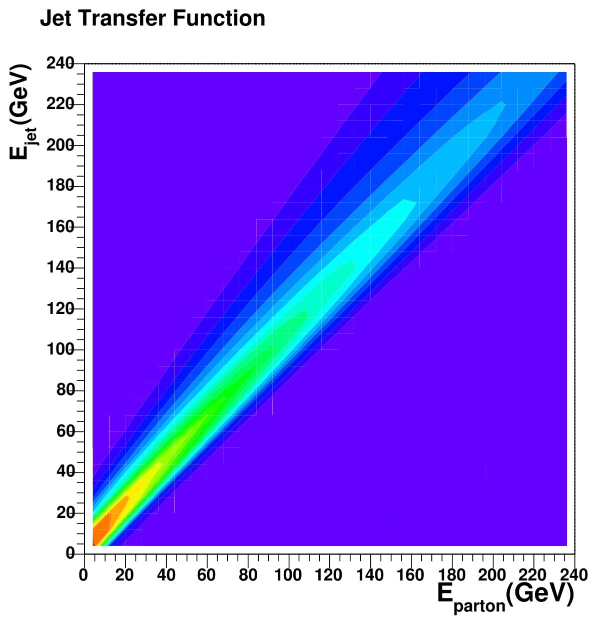

Figure 1: Response of the Transfer functions.

178 GeV/c^2, we extract the parameters shown in Table 1. The response of the transfer function from these extracted parameters is shown in Figure 1

p_i	a_i	b_i
p_1	-1.896 ± 0.623	-0.023 ± 0.008
p_2	2.832 ± 0.540	0.075 ± 0.005
p_3	0.699 ± 0.075	0.0 ± 0.001
p_4	-1.789 ± 0.789	-0.187 ± 0.012
p_5	8.036 ± 0.665	0.095 ± 0.008

Table 1: Parameters for $f(E_{parton}, E_{jet})$ extracted using matched b-jet and quarks from HERWIG $t\bar{t}$ MC.

3.2 Tests of $f(E_{parton}, E_{jet})$

A simple test of the jet transfer function can be done by attempting to predict the jet energy distribution resulting from a known distribution of parton energies from MC run through a full CDF detector simulation. This is done by integrating over the expression for $n(E_{jet}, E_{parton})dE_{jet}dE_{parton}$ in equation 3 over a slice of parton energy:

$$H(E_{jet}) = \int_{E_{parton_2}}^{E_{parton_1}} n(E_{parton})dE_{parton}f(E_{parton}, E_{jet}) \tag{5}$$

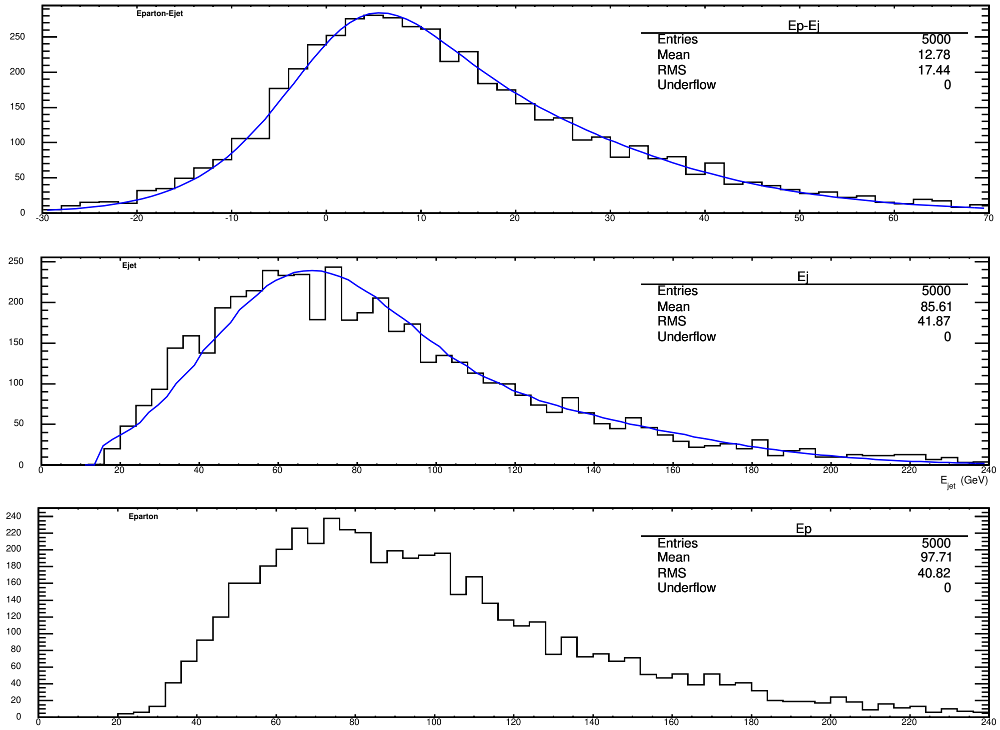

Figure 2: A simple test of $f(E_{parton}, E_{jet})$: simulated E_{jet} and $E_{parton} - E_{jet}$ distributions from a known distribution of E_{parton}. The lines shows the predicted distributions for E_{jet} and $E_{parton} - E_{jet}$ as derived from E_{parton} using $f(E_{parton}, E_{jet})$.

Such a test using $H(E_{jet})$ and a similar test using $H(E_\delta)$ is shown in Figure 2 for the full parton energy range and in Figure 3 for smaller slices in parton energy.

Since the jet transfer function models the detector response to partons, it should be independent of the physics process producing these partons. We test this by using $f(E_{parton}, E_{jet})$ parameterized from $M_t = 178$ GeV/c^2 events to predict the jet energy distribution resulting from b-quarks in MC top decays of varied top masses. Figure 4 shows that the jet transfer function derived using partons from $M_t = 178$ GeV/c^2 top decays can accurately predict jet energies from top decays of M_t ranging from 150 GeV/c^2 through 200 GeV/c^2.

A more sophisticated test of the jet energy transfer functions can be done by using the transfer function to calculate a physical variable, such as the invariant top quark mass from MC $t\bar{t}$ decay:

$$m_x^2 = E_{jet}E_l(1 - cos\theta_{jl}) + E_{jet}E_\nu(1 - cos\theta_{j\nu}) + E_lE_\nu(1 - cos\theta_{l\nu}) \quad (6)$$

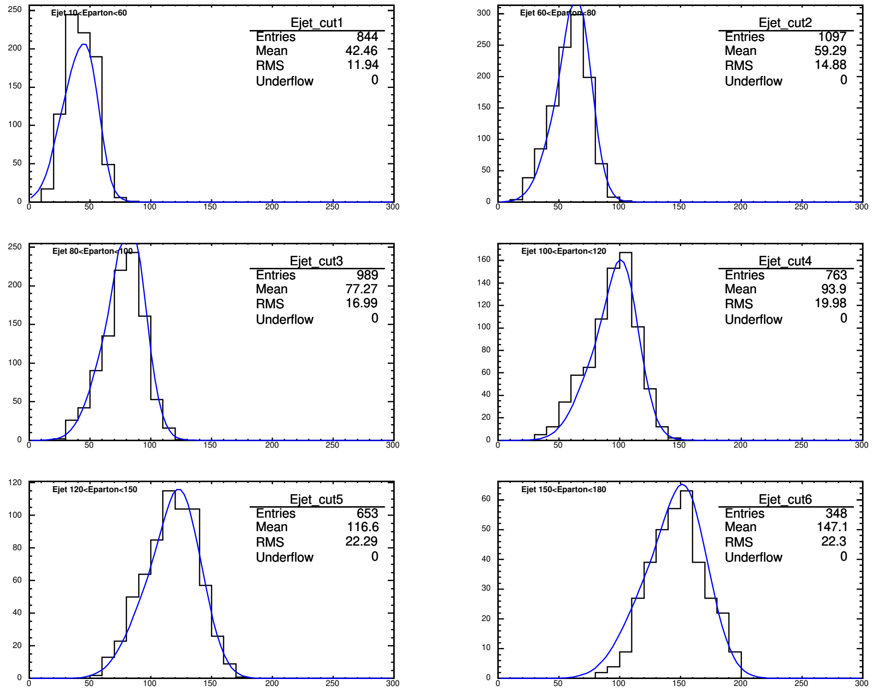

Figure 3: A simple test of $f(E_{parton}, E_{jet})$ over smaller ranges of E_{parton}: simulated E_{jet} from a known distribution of E_{parton} over small sub-ranges of that E_{parton} distribution. The lines shows the predicted distributions for E_{jet} as derived from E_{parton} using $f(E_{parton}, E_{jet})$.

To calculate m_x^2 using $f(E_{parton}, E_{jet})$, we calculate the integral:

$$\frac{dN}{dm_x} \propto \int dE_{jet}^{(2)} J\left(\frac{E_{jet}^{(1)} E_{jet}^{(2)}}{m_x E_{jet}^{(2)}}\right) \prod_{i=1}^{2} f(E_{parton}^{(i)}, E_{jet}^{(i)}) \tag{7}$$

Where J is the Jacobian factor necessary in the variable change. The result of this test using `HERWIG` $t\bar{t}$ MC is shown in Figure 5.

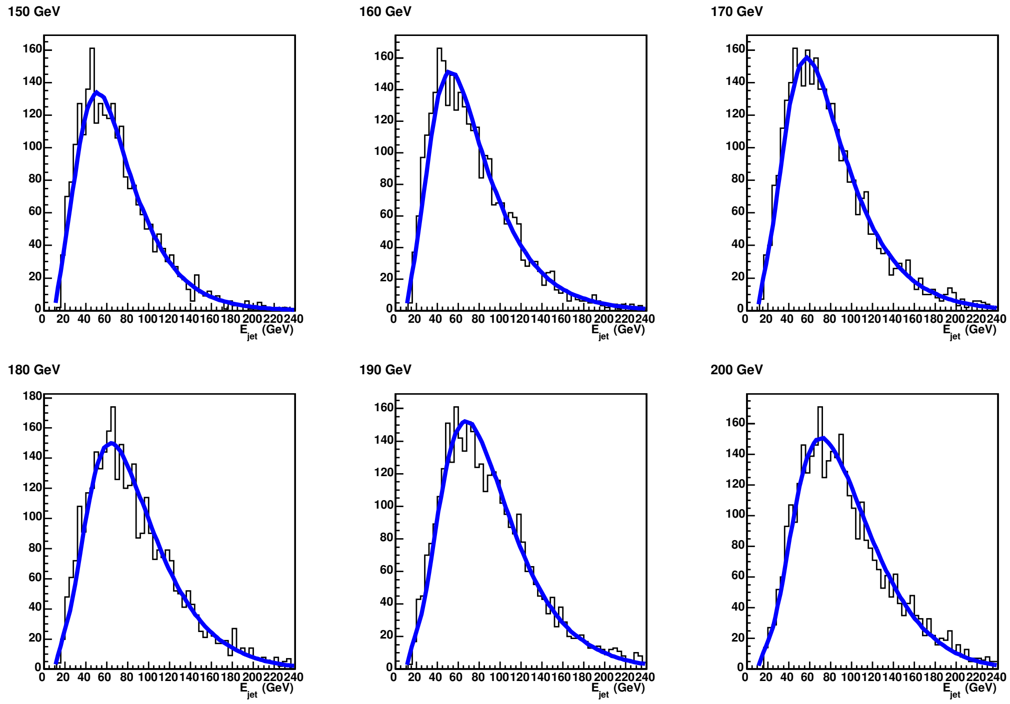

Figure 4: Tests of $f(E_{parton}, E_{jet})$ stability. The jet energy distributions shown are from detector simulation of b-quarks from MC $t\bar{t}$ decay of $M_t = 150$, 160, 170, 180, 190, and 200 GeV/c^2. The line shows the predicted jet energy distribution for each case using the parton energy distribution (not shown) and $f(E_{parton}, E_{jet})$ derived from b-quarks from MC $t\bar{t}$ decay of $M_t = 178$ GeV/c^2

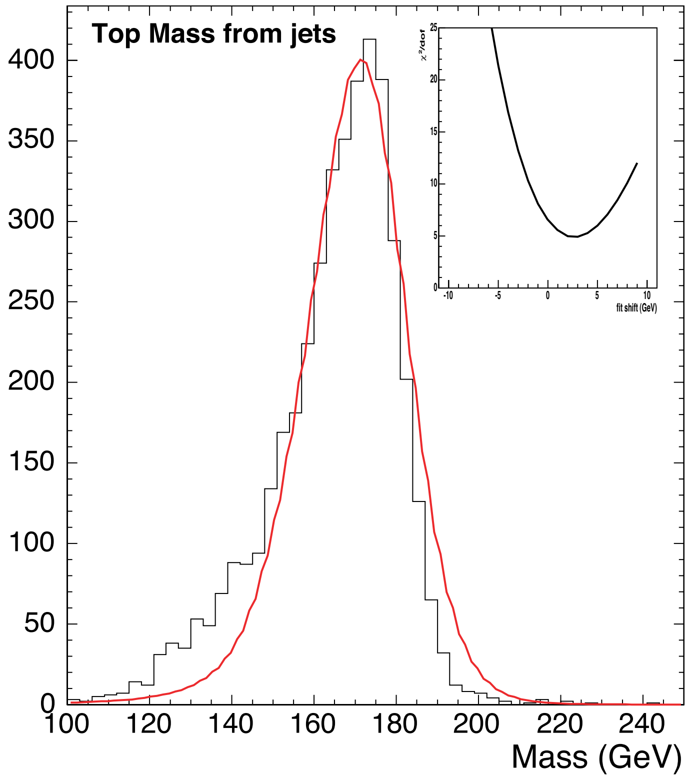

Figure 5: Invariant mass distribution of the jet, lepton and neutrino from the decay of the t in `HERWIG` $t\bar{t}$ MC. The histogram corresponds to MC events with jet and lepton energies after CDF detector simulation and parton-level neutrino energies. The line corresponds to the mass distribution dN/dm_t (equation 7) using the parameters of $f(E_{parton}, E_{jet})$ given in table 1. The inner plot shows the χ^2 per degree of freedom as the histogram is shifted relative to the curve.

4 Differential cross-section for $t\bar{t}$ production

The expression for the cross section for a $2 \to 2$ process can be written as

$$\sigma = \sum_{i,j} \int d\Phi_t dx_i dx_j \frac{(2\pi)^4 |\mathcal{M}|^2}{4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}} d\Phi_6 f_{\mathtt{PDF}}(x^i) f_{\mathtt{PDF}}(x^j) \tag{8}$$

At the heart of the calculation is the expression for the matrix element. Following [3], we will assume that the squared matrix element can be factorized:

$$|\mathcal{M}(q_1, q_2 \to t_1, t_2)|^2 \; |\mathcal{M}(t_1 \to \ell_1, \nu_1, p_1)|^2 \; |\mathcal{M}(t_2 \to \ell_2, \nu_2, p_2)|^2$$

and we use their expression for the matrix element:

$$|\mathcal{M}(q_1, q_2 \to t_1, t_2)|^2 = \frac{g_s^4}{9}(2 - \beta \sin(\theta^*))$$

$$|\mathcal{M}(t_1 \to \ell_1, \nu_1, p_1)|^2 = \frac{g_W^4}{4} \frac{[S_t^2 - M_S^2][S_t^2(1 - \cos(\theta_{l\nu})^2) + S_W^2 * (1 + \cos(\theta_{l\nu})^2)]}{[(S_W^2 - M_W^2)^2 + M_W^2 \Gamma_W^2][(S_t^2 - M_t^2) + M_t^2 \Gamma_t^2)]}$$

where

$$S_W^2 = (l + \nu)^2, S_t^2 = (b + l + \nu)^2$$

are the invariant masses of the t and W particles in the event.

Expression (8) describes the total cross section for this process; we must calculate the differential cross-section for a particular event. The expression for $|\mathcal{M}|^2$ is a function of parton-level four-momenta of incoming particles and outgoing particles, and the integration runs over all of parton-level phase space. The information reagarding the measurements made in an individual event is introduced as conditional probability densities for parton-level quantities.

The matrix element depends on two initial state momenta — incoming q_1 and q_2. Outgoing parton-level variables are: b-quarks p_1 and p_2, leptons ℓ_1 and ℓ_2, neutrinos ν_1 and ν_2. Observed quantities consist of jets j_1 and j_2, measured leptons L_1 and L_2 and two components of missing transverse energy.

The set of assumptions regarding the transfer between parton-level quantities and the observables consists of the following:

- Leptons are measured perfectly. This assumption effectively translates into "transfer function" for leptons being a three-dimensional δ-function

$$\delta^3(l_1 - L_1)\, \delta^3(l_2 - L_2)$$

- Jet angles are measured perfectly. Equivalently,

$$\delta(\theta_{j_1} - \theta_{p_1})\delta(\phi_{j_1} - \phi_{p_1})f(E_{p_1}, E_{j_1})\,\delta(\theta_{j_2} - \theta_{p_2})\delta(\phi_{j_2} - \phi_{p_2})f(E_{p_2}, E_{j_2})$$

  is the b-parton to jet transfer function. We will normally call the energy part "jet transfer function" for brevity[1].

- We assume incoming partons to be massless and to have transverse momentum components $q_x \ll q_z$, $q_y \ll q_z$, where the coordinate system is chosen so that z-axis is set along the colliding beam, and x and y are in the transverse plane.

- Masses of the final state leptons are zero, masses of the b-quarks are known.

4.0.1 Differential cross section and notation

We begin with an expression for the differential cross-section that is given by (9).

$$\begin{aligned} d\sigma = &\left(\frac{1}{(2\pi)^3}\right)^6 \iint \sum_{i,j} |M(q_1^i, q_2^j \to \ell_1, \nu_1, p_1, \ell_2, \nu_2, p_2)|^2 \\ &(2\pi)^4 \delta^4(q_1 + q_2 - \ell_1 - \nu_1 - p_1 - \ell_2 - \nu_2 - p_2) \\ &\delta(q^x)\,\delta(q^y)\,dq^x\,dq^y\; f^i_{\mathtt{PDF}}(\xi_1^i, Q^2)\; f^j_{\mathtt{PDF}}(\xi_2^j, Q^2)\; f_{\mathtt{flux}}(q_1^z, q_2^z)\,d\xi_1^i\,d\xi_2^j \\ &\delta^3(\ell_1 - L_1)\,\delta^3(\ell_2 - L_2)\,\frac{d^3\ell_1}{2E_{\ell_1}}\frac{d^3\ell_2}{2E_{\ell_2}} \\ &f(E_{p_1}, E_{j_1})\,f(E_{p_2}, E_{j_2})\delta(\theta_{p_1} - \theta_{j_1})\,\delta(\theta_{p_2} - \theta_{j_2})\,\delta(\phi_{p_1} - \phi_{j_1})\,\delta(\phi_{p_2} - \phi_{j_2})\frac{d^3p_1}{2E_{p_1}}\frac{d^3p_2}{2E_{p_2}} \\ &f_{\mathtt{utf}}\frac{d^3\nu_1}{2E_{\nu_1}}\frac{d^3\nu_2}{2E_{\nu_2}} \end{aligned} \tag{9}$$

The Bjorken variables ξ_1^i and ξ_2^j can be considered as fractions of the beam energy carried by the interacting incoming partons. We express them in terms of q_1^z and q_2^z as $\xi_1^i = q_1^z/E_{\mathtt{beam}}$ and $\xi_2^i = q_2^z/E_{\mathtt{beam}}$. The sum is over possible incoming parton flavors.

[1] there is some potential ambiguity here. When dealing with δ-functions in spherical coordinates one would normally need to include extra factors, for instance:

$$\delta^3(\vec{r} - \vec{r'}) = \frac{1}{r^2}\,\delta(r - r')\,\delta(\cos\theta - \cos\theta')\,\delta(\phi - \phi')$$

We have some freedom in definition of the transfer functions and we choose to *define* it so that by integration in spherical coordinates Jacobian factors are absorbed in $f(E_j, E_p)$, angular δ-functions are $\delta(\theta_j - \theta_p)\delta(\phi_j - \phi_p)$, while $p_R^2\sin(\theta_p)dp_R d\theta_p d\phi_p$ is the volume element. The validity of the implicit assumption that jet energy transfer functions are independent of θ or, equivalently, pseudo-rapidity η is supported by the observation that no strong angular dependence is present. This, of course, can be refined to allow for angular dependence but is not perceived necessary at for our initial analysis.

We integrate over all the final state parton-level three-momenta (final state particles are stable).

The δ-functions in q_x and q_y constrain transverse momenta of the incoming partons to be zero. This is not strictly true or necessary. For simplicity we use this in the derivation of the differential cross-section and show later that by a trivial change this constraint can be removed and this extra integration can be reintroduced later together with appropriate prior distributions for q_x, q_y. [2]

The $f_{\texttt{utf}}$ factor is the unclustered energy transfer function. The motivation for it as well as a definition is given below. Loosely speaking, this function quantifies the dependence between parton-level momenta, measured jet energies and measured missing transverse energy.

4.0.2 Phase space transformation and Integration

In order to efficiently perform integrations over the parton-level variables, we perform a phase-space transformation which splits the original phase space into one or more subspaces, but introduces the equivalent number of extra variables and integrations. These variables are the relevant invariant masses. Some details on the formula used are provided in Appendix A and in [2]. We introduce invariant masses that correspond to intermediate t and $\bar{t}$ quarks and W-bosons.

The entire integrand after the phase space transformation results in (10). Each additional integration over invariant mass of the intermediate particle has a corresponding δ-function in squared invariant mass and each intermediate particle four-momenta has corresponding δ^4-function for the momentum conservation at the intermediate vertex. It would be easy to integrate over zero-components of intermediate four-momenta that leaves only three-momenta integrals for the intermediate particles, but the algebra

[2] Consider two incoming partons and their transverse momentum components: q_1^x, q_2^x, q_1^y and q_2^y. Let us change the variables to $q_x = \frac{1}{\sqrt{2}}(q_1^x + q_2^x)$, $q_y = \frac{1}{\sqrt{2}}(q_1^y + q_2^y)$, $q_u = \frac{1}{\sqrt{2}}(q_1^x - q_2^x)$ and $q_v = \frac{1}{\sqrt{2}}(q_1^y - q_2^y)$. Inspecting all the expressions that contain the transverse components we may find out that they only depend on the q_x and q_y. We may further assume that the initial momenta components are independent and Gaussian distributed. The q_u and q_v in this case are trivially integrated away. The resulting distribution of the q_x and q_y are also Gaussian, moreover in the absence of correlation between q_1^x and q_2^x, q_x will even have the same σ.

becomes quite involved and we chose a different approach.

$$\begin{aligned}
d\sigma =&(2\pi)^4\,|M|^2\left(\frac{1}{(2\pi)^3}\right)^6 \delta^4(q_1+q_2-t_1-t_2)\,\delta(t_1^2-M_{t_1}^2)\,d^4t_1\,\delta(t_2^2-M_{t_2}^2)\,d^4t_2\\
&\delta^4(t_1-w_1-p_1)\,\delta^4(t_2-w_2-p_2)\,\delta^4(w_1-\nu_1-\ell_1)\,\delta^4(w_2-\nu_2-\ell_2)\\
&\delta(w_1^2-M_{w_1}^2)\,d^4w_1\,\delta(w_2^2-M_{w_2}^2)\,d^4w_2\,\frac{d^3p_1}{2E_{p_1}}\frac{d^3p_2}{2E_{p_2}}\frac{d^3\nu_1}{2E_{\nu_1}}\frac{d^3\nu_2}{2E_{\nu_2}}\frac{d^3\ell_1}{2E_{\ell_1}}\frac{d^3\ell_2}{2E_{\ell_2}}\\
&\delta(q^x)\,\delta(q^y)\,dq^x\,dq^y\,f_{\texttt{flux}}(q_1^z,q_2^z)\,f^i_{\texttt{PDF}}(\frac{q_1^z}{E_{\texttt{beam}}},Q^2)\,f^j_{\texttt{PDF}}(\frac{q_2^z}{E_{\texttt{beam}}},Q^2)\,\frac{dq_1^z\,dq_2^z}{E^2_{\texttt{beam}}}\,f_{\texttt{utf}}\\
&\delta^3(\ell_1-L_1)\,\delta^3(\ell_2-L_2)\,f(E_{p_1},E_{j_1})\,f(E_{p_2},E_{j_2})\\
&\delta(\theta_{p_1}-\theta_{j_1})\,\delta(\theta_{p_2}-\theta_{j_2})\,\delta(\phi_{p_1}-\phi_{j_1})\,\delta(\phi_{p_2}-\phi_{j_2})\,dM_{t_1}^2\,dM_{t_2}^2\,dM_{W_1}^2\,dM_{W_2}^2
\end{aligned} \tag{10}$$

First we integrate over $d^3\ell_1\,d^3\ell_2$ and [optionally][3] over $dq^x\,dq^y$, making use of the corresponding δ-functions. This step results in the expression (11)

$$\begin{aligned}
d\sigma =&(2\pi)^4\,|M|^2\left(\frac{1}{(2\pi)^3}\right)^6 \delta\left((q_1+q_2)^0-t_1^0-t_2^0\right)\delta\left((q_1+q_2)^z-t_1^z-t_2^z\right)\delta(t_1^x+t_2^x)\,\delta(t_1^y+t_2^y)\\
&\delta(t_1^2-M_{t_1}^2)\,\delta(t_2^2-M_{t_2}^2)\,\delta(w_1^2-M_{W_1}^2)\,\delta(w_2^2-M_{W_2}^2)\,d^4t_1\,d^4t_2\,d^4w_1\,d^4w_2\\
&\delta^4(t_1-w_1-p_1)\,\delta^4(t_2-w_2-p_2)\,\delta^4(w_1-\nu_1-L_1)\,\delta^4(w_2-\nu_2-L_2)\\
&\frac{d^3p_1}{2E_{p_1}}\frac{d^3p_2}{2E_{p_2}}\frac{d^3\nu_1}{2E_{\nu_1}}\frac{d^3\nu_2}{2E_{\nu_2}}\frac{1}{2E_{L_1}}\frac{1}{2E_{L_2}}\\
&f_{\texttt{flux}}(q_1^z,q_2^z)\,f^i_{\texttt{PDF}}(\frac{q_1^z}{E_{\texttt{beam}}},Q^2)\,f^j_{\texttt{PDF}}(\frac{q_2^z}{E_{\texttt{beam}}},Q^2)\,\frac{dq_1^z\,dq_2^z}{E^2_{\texttt{beam}}}\,f_{\texttt{utf}}\,f(E_{p_1},E_{j_1})\,f(E_{p_2},E_{j_2})\\
&\delta(\theta_{p_1}-\theta_{j_1})\,\delta(\theta_{p_2}-\theta_{j_2})\,\delta(\phi_{p_1}-\phi_{j_1})\,\delta(\phi_{p_2}-\phi_{j_2})\,dM_{t_1}^2\,dM_{t_2}^2\,dM_{W_1}^2\,dM_{W_2}^2
\end{aligned} \tag{11}$$

Next step consists of integration over

$$\delta\left((q_1+q_2)^0-t_1^0-t_2^0\right)\;\delta\left((q_1+q_2)^z-t_1^z-t_2^z\right)\;dq_1^z\,dq_2^z \tag{12}$$

The energy components in this expression are then expressed in terms of q_1^z and q_2^z in the collider approximation mentioned earlier [4]. The expressions inside these two δ-functions are used to solve for q_1^z and q_2^z in terms of the remaining variables. The entire expression gets an additional factor of 1/2 [5].

We carry on the integration with

$$\delta^4(w_1-\nu_1-L_1)\,\delta^4(w_2-\nu_2-L_2)\,d^4w_1\,d^4w_2 \tag{13}$$

[3] integrations over q_x, q_y can be postponed until the end and integrated with non-δ-function priors.

[4] massless incoming partons with transverse momenta of the incoming partons $q_t \ll q_z$. Here it means that without loss of generality we may substitute $q_1^0 = q_1^z$ and $q_2^0 = -q_2^z$ and make sure energies are positive.

[5] this is easily shown by either doing integrations over $dq_1^z\,dq_2^z$ one by one or by noticing that the system of coupled equations inside the two δ-functions has a Jacobian of 2.

that imposes the constraint

$$\begin{aligned} w_1 &= (L_1 + \nu_1) \\ w_2 &= (L_2 + \nu_2) \end{aligned} \tag{14}$$

Next integration is over t variables:

$$\delta^4(t_1 - w_1 - p_1)\ \delta^4(t_2 - w_2 - p_2)\ d^4t_1\ d^4t_2 \tag{15}$$

that results in:

$$\begin{aligned} t_1 &= (L_1 + \nu_1 + p_1) \\ t_2 &= (L_2 + \nu_2 + p_2) \end{aligned} \tag{16}$$

Further integration over angular variables of b-partons is done by first transforming the parton variables into spherical coordinates:

$$d^3p \to p_R^2 \sin\theta_p d\theta_p d\phi_p dp_R \tag{17}$$

And the expression takes the form:

$$\begin{aligned} d\sigma =& (2\pi)^4 \frac{|M|^2}{2} \left(\frac{1}{(2\pi)^3}\right)^6 f^i_{\texttt{PDF}}(\frac{q_1^z}{E_{\texttt{beam}}}, Q^2)\ f^j_{\texttt{PDF}}(\frac{q_2^z}{E_{\texttt{beam}}}, Q^2)\ \frac{1}{E^2_{\texttt{beam}}} \\ & \delta\left((L_1+\nu_1+p_1)_x + (L_2+\nu_2+p_2)_x\right)\ \delta\left((L_1+\nu_1+p_1)_y + (L_2+\nu_2+p_2)_y\right) \\ & \delta\left((L_1+\nu_1+p_1)^2 - M^2_{t_1}\right)\ \delta\left((L_2+\nu_2+p_2)^2 - M^2_{t_2}\right) \\ & \delta\left((L_1+\nu_1)^2 - M^2_{W_1}\right)\ \delta\left((L_2+\nu_2)^2 - M^2_{W_2}\right) \\ & \frac{p^2_{1R}\sin\theta_{p_1} dp_{1R}}{2E_{p_1}}\ \frac{p^2_{2R}\sin\theta_{p_2} dp_{2R}}{2E_{p_2}}\ \frac{d^3\nu_1}{2E_{\nu_1}}\ \frac{d^3\nu_2}{2E_{\nu_2}}\ \frac{1}{2E_{L_1}}\ \frac{1}{2E_{L_2}} \\ & f_{\texttt{flux}}(q_1^z, q_2^z)\ f_{\texttt{utf}}\ f(E_{p_1}, E_{j_1})\ f(E_{p_2}, E_{j_2})\ dM^2_{t_1}\ dM^2_{t_2}\ dM^2_{W_1}\ dM^2_{W_2} \end{aligned} \tag{18}$$

Dirac δ-functions have a number of properties that are used extensively in simplification of the cross-section integrals. A well known property of a δ-function in a single variable is given by (19)

$$\delta(g(x)) = \sum_i \frac{\delta(x - x_i)}{|g'(x_i)|} \tag{19}$$

where x_i are the zeros of $g(x)$. For an integral of two-dimensional function of two variables, we think it should be reasonably modified to:

$$\delta(\mathbf{f}(x, y)) = \sum_i \frac{\delta(x - x_i)\delta(y - y_i)}{|J(x_i, y_i)|} \tag{20}$$

where J is the determinant of the Jacobian matrix of $\mathbf{f}$ evaluated at the zero-points. Skipping the mathematically rigorous proof we will use (20) keeping in mind that degenerate zeroes should in theory receive special treatment.

Further integrations over 6 components of neutrino momenta and 6 "free" δ-functions has a twofold effect. First, the integration in question is equivalent to solving the system of equations that is obtained by requiring the expressions inside δ-function to equal zero. The neutrino solutions obtained in the process should then be substituted into all the expressions that depend on them and expressing the rest of the kinematic variables in term of measured variables and components of neutrino momenta. Second, the expressions inside the δ-functions are not the neutrino momenta themselves but some functions of those. Thus the entire expression is combined with the Jacobian terms that originate from this fact in a way similar to (20). There remain 6 extra integrations to be done at this point: 2 over the magnitudes of b-parton momenta, and 4 over the assumed invariant masses of t-quarks and W-bosons. Consequently for the purpose of finding neutrino solutions these quantities are assumed known, as the final stage integration will vary them and solve for neutrino momenta at each point of evaluation.

It should be emphasized that the sum over neutrino solutions comes naturally from the integration of δ-function (19); it makes little sense to undertake separation of neutrino solutions or to "rate" them by which one is perceived as "more likely".

4.0.3 Neutrino Solutions

The system of equations for neutrino momenta is the following:

$$\begin{aligned}
&(L_1+\nu_1+p_1)^2 = M_{t_1}^2\\
&(L_2+\nu_2+p_2)^2 = M_{t_2}^2\\
&(L_1+\nu_1)^2 = M_{W_1}^2\\
&(L_2+\nu_2)^2 = M_{W_2}^2\\
&(L_1+\nu_1+p_1)_x + (L_2+\nu_2+p_2)_x = 0\\
&(L_1+\nu_1+p_1)_y + (L_2+\nu_2+p_2)_y = 0
\end{aligned} \tag{21}$$

The first two equations of (21) can be rewritten in the following way:

$$\begin{aligned}
(\nu_1\cdot p_1) &= \frac{M_{t_1}^2 - M_{W_1}^2 - m_b^2}{2} - (L_1\cdot p_1) = a_{01}\\
(\nu_2\cdot p_2) &= \frac{M_{t_2}^2 - M_{W_2}^2 - m_b^2}{2} - (L_2\cdot p_2) = a_{02}
\end{aligned} \tag{22}$$

where we introduced a_{01} and a_{02} as convenience notation. The second pair of equations from (21) take form:

$$\begin{aligned}
(L_1\cdot\nu_1) &= \frac{M_{W_1}^2}{2} = b_{01}^2\\
(L_2\cdot\nu_2) &= \frac{M_{W_2}^2}{2} = b_{02}^2
\end{aligned} \tag{23}$$

with additional notation introduced. Expanding the four-vector products in first four equation of (21) we can express ν_{1z} and ν_{2z} in terms of x and y components and thus exclude them from the system:

$$\begin{aligned} E_{\nu 1} &= \frac{a_{01} + \nu_{1x}p_{1x} + \nu_{1y}p_{1y} + \nu_{1z}p_{1z}}{E_{p1}} \\ E_{\nu 1} &= \frac{b_{01} + \nu_{1x}L_{1x} + \nu_{1y}L_{1y} + \nu_{1z}L_{1z}}{E_{L1}} \end{aligned} \tag{24}$$

These two result in:

$$\nu_{1z} = \alpha + \beta\nu_{1x} + \gamma\nu_{1y} \tag{25}$$

where

$$\begin{aligned} \alpha &= \left(\frac{p_{1z}}{E_{p1}} - \frac{L_{1z}}{E_{L1}}\right)^{-1}\left(\frac{b_{01}^2}{E_{L1}} - \frac{a_{01}}{E_{p1}}\right) \\ \beta &= \left(\frac{p_{1z}}{E_{p1}} - \frac{L_{1z}}{E_{L1}}\right)^{-1}\left(\frac{L_{1x}}{E_{L_1}} - \frac{p_{1x}}{E_{p1}}\right) \\ \gamma &= \left(\frac{p_{1z}}{E_{p1}} - \frac{L_{1z}}{E_{L1}}\right)^{-1}\left(\frac{L_{1y}}{E_{L1}} - \frac{p_{1y}}{E_{p1}}\right) \end{aligned} \tag{26}$$

By making the $1 \to 2$ index substitution one gets corresponding expression for ν_{2z}. By introducing the notation:

$$\begin{aligned} A &= \frac{b_{01}^2 + \alpha L_{1z}}{E_{L1}} \\ B &= \frac{L_{1x} + \beta L_{1z}}{E_{L1}} \\ C &= \frac{L_{1y} + \gamma L_{1z}}{E_{L1}} \end{aligned} \tag{27}$$

the original first four equations in (21) transform into two equations of the form:

$$\begin{aligned} (1+\beta^2 - B^2)\nu_{1x}^2 + (1+\gamma^2 - C^2)\nu_{1y}^2 + 2(\beta\gamma - BC)\nu_{1x}\nu_{1y} + \\ 2(\alpha\beta - AB)\nu_{1x} + 2(\alpha\gamma - AC)\nu_{1y} + (\alpha^2 - A^2) = 0 \end{aligned} \tag{28}$$

or in yet another notation:

$$C_{1x}\nu_{1x}^2 + C_{1y}\nu_{1y}^2 + 2C_{1xy}\nu_{1x}\nu_{1y} + 2D_{1x}\nu_{1x} + 2D_{1y}\nu_{1y} + G_1 = 0 \tag{29}$$

Now using two remaining equations from (21), we substitute ν_{1x} through ν_{2x} and ν_{2y} through ν_{1y} and arrive at the following system:

$$\begin{aligned} a_1\nu_{2x}^2 + a_2\nu_{2x}\nu_{1y} + a_3\nu_{1y}^2 + a_4\nu_{2x} + a_5\nu_{1y} + a_6 = 0 \\ b_1\nu_{2x}^2 + b_2\nu_{2x}\nu_{1y} + b_3\nu_{1y}^2 + b_4\nu_{2x} + b_5\nu_{1y} + b_6 = 0 \end{aligned} \tag{30}$$

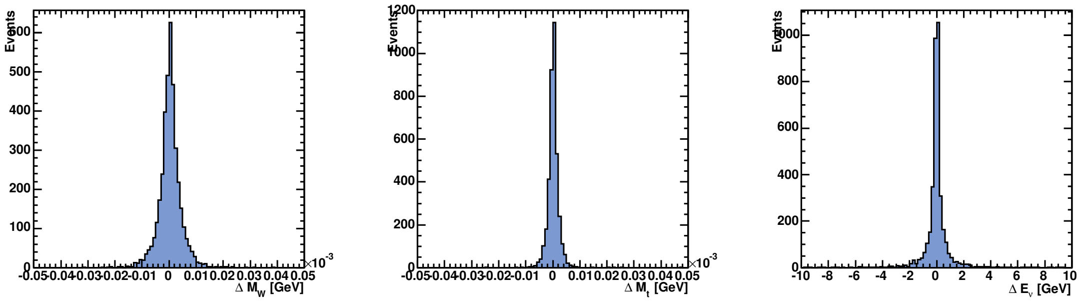

Figure 6: Residual in M_W, M_t and E_ν for numerical solutions to the equations required for a change of variables.

with the substitutions given by: $\nu_{1x} = -K_x - \nu_{2x}$ and $\nu_{2y} = -K_y - \nu_{1y}$ and

$$
\begin{aligned}
K_x &= (L_{2x} + p_{2x} + L_{1x} + p_{1x}) \qquad (-q_x) \\
K_y &= (L_{2y} + p_{2y} + L_{1y} + p_{1y}) \qquad (-q_y)
\end{aligned}
\tag{31}
$$

The q_x and q_y terms in (31) are optional and can be included if the initial system transverse momentum is non-zero. This refinement will not affect the neutrino solutions themselves, the only adjustment that needs to be made in this case is in the aforementioned definitions of K_x and K_y. Coefficients in the bi-quadratic system of equations (30) are given by:

	first equation	second equation
ν_{2x}^2	C_{1x}	C_{2x}
$\nu_{2x}\nu_{1y}$	$-2C_{1xy}$	$-2C_{2xy}$
ν_{1y}^2	C_{1y}	C_{2y}
ν_{2x}	$2C_{1x}K_x - 2D_{1x}$	$2D_{2x} - 2C_{2xy}K_y$
ν_{1y}	$2D_{1y} - 2C_{1xy}K_x$	$2C_{2y}K_y - 2D_{2y}$
const	$C_{1x}K_x^2 - 2D_{1x}K_x + G_1$	$C_{2y}K_y^2 - 2D_{2y}K_y + G_2$

(32)

The two coupled equations are converted to a single fourth order polynomial which can be rapidly solved using numerical techniques. We employ the Sturm Sequence technique; Figure 6 demonstrates the accuracy of the solutions for parton-level quantities. Alternative analytical solutions can be found in Appendix C.

4.0.4 Jacobian Factor

Jacobian factor that appears from integrations over neutrino momenta components has the form:

$$
J = \left| \frac{\partial f_i}{\partial \nu_{(1,2)(x,y,z)}} \right| \tag{33}
$$

The f_i are the functions of neutrino momenta that occur inside δ-functions:

$$\begin{aligned}
f_1 &= ((L_1+\nu_1+p_1)_x + (L_2+\nu_2+p_2)_x) \\
f_2 &= ((L_1+\nu_1+p_1)_y + (L_2+\nu_2+p_2)_y) \\
f_3 &= \left((L_1+\nu_1+p_1)^2 - M_{t_1}^2\right) \\
f_4 &= \left((L_2+\nu_2+p_2)^2 - M_{t_2}^2\right) \\
f_5 &= \left((L_1+\nu_1)^2 - M_{W_1}^2\right) \\
f_6 &= \left((L_2+\nu_2)^2 - M_{W_2}^2\right)
\end{aligned} \tag{34}$$

The determinant has the form:

$$\begin{vmatrix}
1 & 0 & J_{3x} & 0 & J_{5x} & 0 \\
0 & 1 & J_{3y} & 0 & J_{5y} & 0 \\
0 & 0 & J_{3z} & 0 & J_{5z} & 0 \\
1 & 0 & 0 & J_{4x} & 0 & J_{6x} \\
0 & 1 & 0 & J_{4y} & 0 & J_{6y} \\
0 & 0 & 0 & J_{4z} & 0 & J_{6z}
\end{vmatrix}$$

that evaluates to:

$$\begin{aligned}
J = |&J_{3z}J_{4z}J_{5y}J_{6x} - J_{3y}J_{4z}J_{5z}J_{6x} - J_{3z}J_{4z}J_{5x}J_{6y} + \\
&+ J_{3x}J_{4z}J_{5z}J_{6y} + J_{3z}J_{4y}J_{5x}J_{6z} - J_{3z}J_{4x}J_{5y}J_{6z} + \\
&+ J_{3y}J_{4x}J_{5z}J_{6z} - J_{3x}J_{4y}J_{5z}J_{6z}|
\end{aligned} \tag{35}$$

with:

$$
\begin{aligned}
J_{3x} &= \frac{\partial f_3}{\partial \nu_{1x}} = \frac{2(L_1^0 + p_1^0)\nu_{1x}}{\sqrt{\nu_{1x}^2 + \nu_{1y}^2 + \nu_{1z}^2}} - 2(L_{1x} + p_{1x}) \\
J_{3y} &= \frac{\partial f_3}{\partial \nu_{1y}} = \frac{2(L_1^0 + p_1^0)\nu_{1y}}{\sqrt{\nu_{1x}^2 + \nu_{1y}^2 + \nu_{1z}^2}} - 2(L_{1y} + p_{1y}) \\
J_{3z} &= \frac{\partial f_3}{\partial \nu_{1z}} = \frac{2(L_1^0 + p_1^0)\nu_{1z}}{\sqrt{\nu_{1x}^2 + \nu_{1y}^2 + \nu_{1z}^2}} - 2(L_{1z} + p_{1z}) \\
J_{4x} &= \frac{\partial f_4}{\partial \nu_{2x}} = \frac{2(L_2^0 + p_2^0)\nu_{2x}}{\sqrt{\nu_{2x}^2 + \nu_{2y}^2 + \nu_{2z}^2}} - 2(L_{2x} + p_{2x}) \\
J_{4y} &= \frac{\partial f_4}{\partial \nu_{2y}} = \frac{2(L_2^0 + p_2^0)\nu_{2y}}{\sqrt{\nu_{2x}^2 + \nu_{2y}^2 + \nu_{2z}^2}} - 2(L_{2y} + p_{2y}) \\
J_{4z} &= \frac{\partial f_4}{\partial \nu_{2z}} = \frac{2(L_2^0 + p_2^0)\nu_{2z}}{\sqrt{\nu_{2x}^2 + \nu_{2y}^2 + \nu_{2z}^2}} - 2(L_{2z} + p_{2z}) \\
J_{5x} &= \frac{\partial f_5}{\partial \nu_{1x}} = \frac{2(L_1^0 \nu_{1x})}{\sqrt{\nu_{1x}^2 + \nu_{1y}^2 + \nu_{1z}^2}} - 2L_{1x} \\
J_{5y} &= \frac{\partial f_5}{\partial \nu_{1y}} = \frac{2(L_1^0 \nu_{1y})}{\sqrt{\nu_{1x}^2 + \nu_{1y}^2 + \nu_{1z}^2}} - 2L_{1y} \\
J_{5z} &= \frac{\partial f_5}{\partial \nu_{1z}} = \frac{2(L_1^0 \nu_{1z})}{\sqrt{\nu_{1x}^2 + \nu_{1y}^2 + \nu_{1z}^2}} - 2L_{1z} \\
J_{6x} &= \frac{\partial f_6}{\partial \nu_{2x}} = \frac{2(L_2^0 \nu_{2x})}{\sqrt{\nu_{2x}^2 + \nu_{2y}^2 + \nu_{2z}^2}} - 2L_{2x} \\
J_{6y} &= \frac{\partial f_6}{\partial \nu_{2y}} = \frac{2(L_2^0 \nu_{2y})}{\sqrt{\nu_{2x}^2 + \nu_{2y}^2 + \nu_{2z}^2}} - 2L_{2y} \\
J_{6z} &= \frac{\partial f_6}{\partial \nu_{2z}} = \frac{2(L_2^0 \nu_{2z})}{\sqrt{\nu_{2x}^2 + \nu_{2y}^2 + \nu_{2z}^2}} - 2L_{2z}
\end{aligned}
\tag{36}
$$

4.0.5 Integral: final

After neutrino solutions are found and the δ-function induced Jacobian is evaluated the expression for the unnormalized probability density takes the following form:

$$\begin{aligned}\frac{d\sigma}{d\mathbf{x}} = &\left(\frac{1}{(2\pi)^3}\right)^6 (2\pi)^2 \iint \sum_{i,j} \frac{|M|^2}{2E^2_{\mathtt{beam}}} \frac{p^2_{1R}\sin\theta_{p_1}dp_{1R}}{2E_{p_1}} \frac{p^2_{2R}\sin\theta_{p_2}dp_{2R}}{2E_{p_2}} \frac{J^{-1}}{E_{\nu_1}E_{\nu_2}E_{L_1}E_{L_2}} \\ &f^i_{\mathtt{PDF}}(\frac{q^z_1}{E_{\mathtt{beam}}}, Q^2)\ f^j_{\mathtt{PDF}}(\frac{q^z_2}{E_{\mathtt{beam}}}, Q^2)\ f_{\mathtt{flux}}(q^z_1, q^z_2)\ f_{\mathtt{utf}} \\ &f(E_{p_1}, E_{j_1})\ f(E_{p_2}, E_{j_2})\ M_{t_1}\ M_{t_2}\ M_{W_1}\ M_{W_2}\ dM_{t_1}\ dM_{t_2}\ dM_{W_1}\ dM_{W_2}\end{aligned} \tag{37}$$

In order to obtain the differential cross-section the 6-dimensional integral (37) over b-parton momenta and four invariant masses should be evaluated. The sum is performed over the incoming parton flavors. The additional sum that needs to be done is over two possible jet assignments[6].

4.0.6 Unclustered energy

Some information about the event is available in the form of missing transverse energy. In order to include it in the analysis we need a probability of certain values for x and y components of the missing transverse energy given the parton quantities in the event *and* the measured jet energies. We construct a "pseudo"-transfer function that depends on the neutrino momenta, b-quark momenta and measured jet energies, with these factors adding to the $\not{E}_T$ value deterministically:

$$\begin{aligned} q^x &= E^x_T + \sum \nu^x - \sum p^x_j + \sum p^x \\ q^y &= E^y_T + \sum \nu^y - \sum p^y_j + \sum p^y \end{aligned} \tag{38}$$

where $E^{x,y}_T$ denotes components of total "detectable" transverse momenta, $\nu^{x,y}$ — components of "undetectable" transverse momenta, $p^{x,y}_j$ are the jet momenta, $p^{x,y}$ are the momenta of the partons that produce the jets. The components of $\not{E}_T$ in an ideal detector would be given then by:

$$\begin{aligned} \not{E}_T{}^x &= \sum \nu^x + \sum (p^x - p^x_j) - q^x \\ \not{E}_T{}^y &= \sum \nu^y + \sum (p^y - p^y_j) - q^y \end{aligned} \tag{39}$$

Mis-measurement of the jet energies is factored out in the jet energy transfer function. Neutrino momenta will be integrated over at some point. Thus our desired transfer

[6] For each event it is not known which jet corresponds to which lepton. In our case we have only two possibilities. If the jet assignment is considered a discrete variables that is not measured, the total probability is the sum over two possible values.

function will encode the resolution in the unclustered energy, or, which is equivalent, the relation between "ideal" $\not{E}_T$ of Equation (39) and the "real" $\not{E}_T$, measured in the detector.

The q_x and q_y components are unknown and thus should be integrated over. They are not entirely arbitrary. There is additional knowledge available that describes some of their values as more likely than others. We already used it implicitly when we assumed that $q_{x,y} \ll q_z$. When the integration over invariant masses of intermediate W-s and t-s is performed, the solutions for corresponding $q_{x,y}$ vary from likely to improbable. To quantify this effect and restrict $q_{x,y}$ to have "reasonable" values we introduce "prior" $p(q_x, q_y|I)$ that summarizes our knowledge about the momentum of the initial parton system. Naively it can be modeled as Gaussian-distributed deviation from $(0, 0)$ with appropriately chosen $\sigma_{q_{x,y}}$. Our expression for matrix element multiplied by the appropriate transfer functions and PDF-s gets additional factor — initial transverse momentum prior.

Unclustered Energy transfer function characterizes the probability to measure a particular amount of unclustered energy for a given value of "true" unclustered energy, $p(\not{E}_{Tm}{}^x, \not{E}_{Tm}{}^y \,|\, \not{E}_T{}^x, \not{E}_T{}^x)$, where the "true" unclustered energy is given by equation (39). Naively modeling this conditional probability density as a Gaussian, we effectively de-weight the angular and invariant mass solutions tried by the integration routine that correspond to the unclustered energy which differs from the measured by a large amount.

These two functions[7], as explained above are the prior for transverse momentum of the system distribution and the likelihood function for the measured unclustered energy, conditional on the "true" unclustered energy. In order to obtain the posterior probability (or differential cross-section) it is necessary to integrate over the variables of ignorance q_x and q_y. In the extreme case of a very narrow prior, these conditions effectively degenerate into constraints, removing two of the integrations.

[7]It should be noted that although the effects of the introduction of these functions may appear accounting for ISR and FSR, it is not, at least in a strict sense. Although one may speculate that ISR may be considered roughly equivalent to giving the system of initial partons a recoil and FSR as giving a contribution to mis-measurement of $\not{E}_T$, the reasoning behind such "intuitive" picture is not straightforward. First of all, the matrix element used in the calculation does not account for either. Second, the FSR kicks in *before* the measurement of final state particle momenta are done and thus the solutions for $q_{x,y}$ would give a momentum of $t\bar{t}$ system *after* recoil from FSR. This in turn means that the wrong momenta are used in calculation of matrix element for both final and initial state momenta. This makes our interpretation of the model as a leading order calculation with probabilistic marginalization over unmeasured quantities the preferred interpretation.

4.1 Tests of $d\sigma(qq \to tt)/dx$

The calculation of the differential cross section describes our understanding of the physical process of the production of top quark pairs and subsequent top-quark decays. However, to limit computational complexity, we introduce some assumptions that simplify the model. It is possible to measure the effect of these assumptions by testing our ability to extract the top mass in different sub-classes of events in which these simplifying assumptions are broken to increasing degrees.

4.1.1 Parton-level tests

We begin with simulated $p\overline{p} \to t\bar{t}$ events generated with Pythia in which all of the parton-level information from the hard scattering and subsequent decay is known. Evaluating the matrix element for these events as a function of top-quark pole mass, see Figure 7, we see the sharp peaks which correspond to the invariant mass of each of the top quarks in the event. Note that with perfect parton-level information, we do not need to perform any integration.

Individual events are not very instructive; we wish to probe whether the calculation contains a bias, and whether the error can be accurately estimated from the shape of the probability distribution. We form pseudo-experiments by constructing the joint probability of ten events, see Figure 8. Choosing the most probable value as the measured mass and extracting the error from the shape of the probability distribution gives us an unbiased measure of the top mass and an accurate error estimate.

To make the test more realistic, we smear the energy of the b-quarks by 10% to produce simple jets and throw away the b-quark momentum and all of the information from the ν and $\overline{\nu}$. This tests our integration procedure as well as the numerical kinematic solutions. Figure 9 shows four example event probabilities. Figure 10 shows that the extracted measurement from ten event pseudo-experiments is unbiased and that the error is accurate.

4.1.2 Tests with simulated events

Events which are measured by the CDF detector are far less ideal than smeared parton-level events. While the lepton energy resolution is very good, especially for electrons, it is not perfect; the angles of jets may occasionally be incorrect; the sample also contains events with secondary leptons from $W \to \tau$ decays.

Figure 11 shows the calculation of the probability for four representative events generated with Herwig and passed through the full chain of CDF simulation. The calculation uses the transfer functions derived above. The shapes are quite smooth, which indicates that the calculation is well behaved. The probabilities peak in roughly

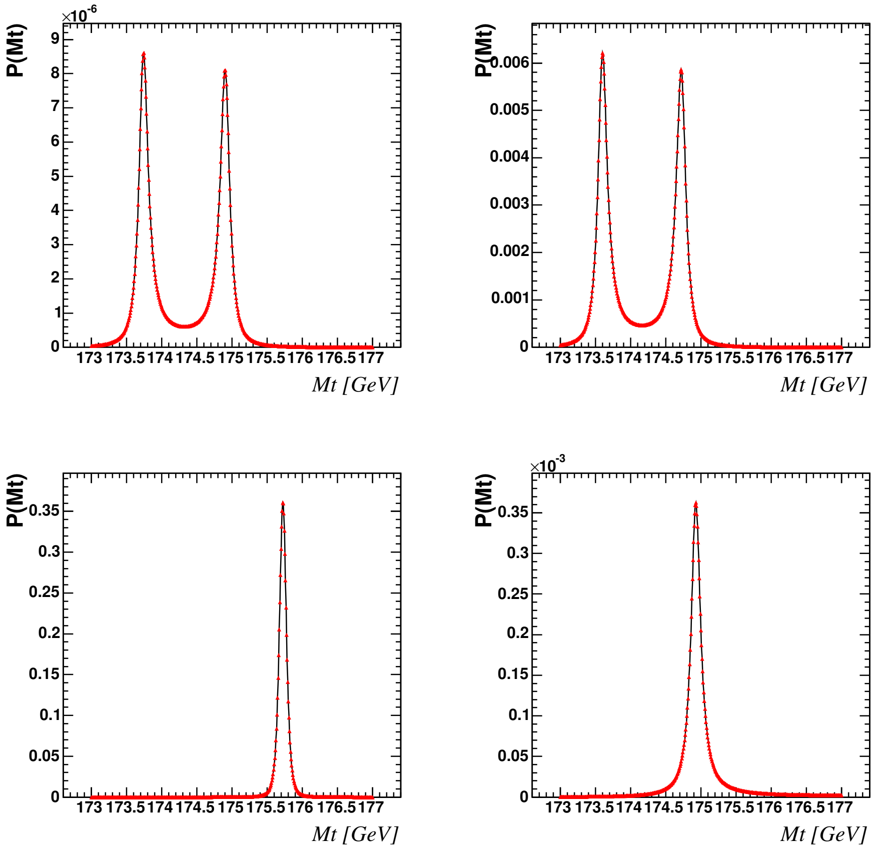

Figure 7: $|M|^2$ as a function of pole mass for four parton-level top quark pair events with $M_t = 175$ GeV.

the region one would expect and are fairly broad. The variation of peak positions and the widths of the probability distributions can be qualitatively attributed to the amount of relevant information on the top mass contained in the observed quantities.

To measure the effect of aforementioned simplifying assumptions on the top-mass measurement, we begin with relatively well-measured events and relax our requirements in stages.

We focus first on events where our assumptions are satisfied in order to test the machinery in an understood environment. Events in which jets are well matched to quarks and both W-bosons decay to an electron constitute a sample for which our assumption that quark angles, lepton angles and lepton energies are well measured are most likely true. These events give a well-behaved pull distribution, see Table 2. Events where W-bosons decay muonically may stretch our assumption that the lepton momenta are perfectly measured. Muon momentum is measured by the central tracking

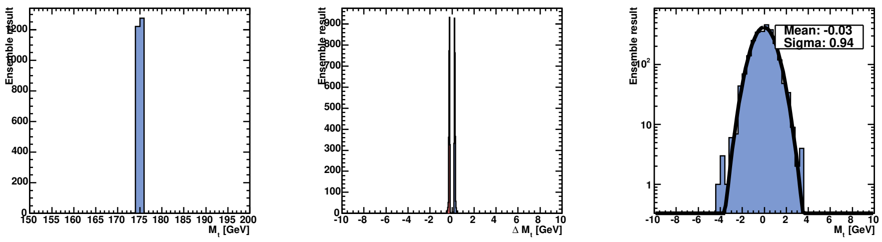

Figure 8: An ensemble of pseudo-experiments using perfect information for parton-level top quark pair events with $M_t = 175$ GeV. The extracted mass is the peak of the joint probability. As these shapes are not necessarily Gaussian, the positive and negative errors $\Delta M_t^{\pm}$ are defined such that $\int_{M_t - \Delta M_t^-}^{M_t + \Delta M_t^+} P(x|M_t) = 0.68$.

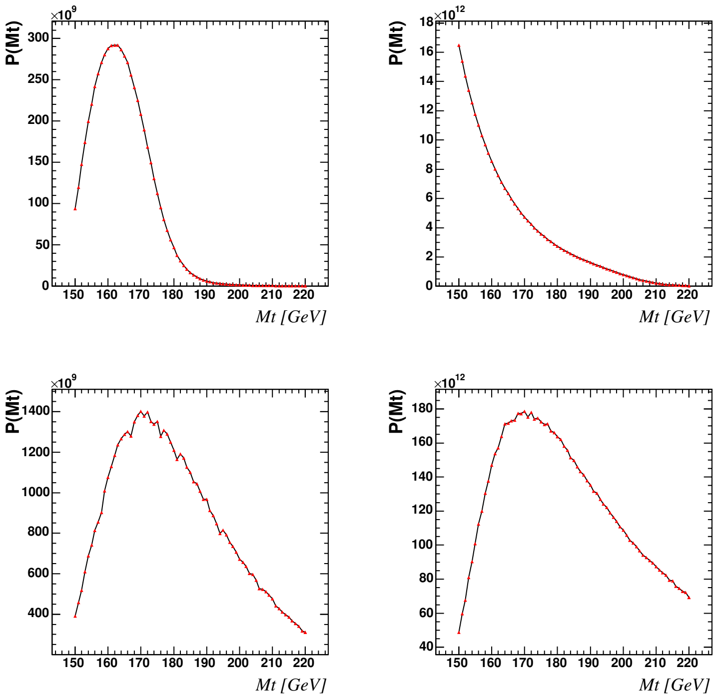

Figure 9: Posterior probability as a function of pole mass for four parton-level top quark pair events where the quark energies are smeared by 10% with $M_t = 175$ GeV.

chamber, whose resolution degrades at higher momentum.

We also examine events where at least one W-boson decays to a τ and subsequently

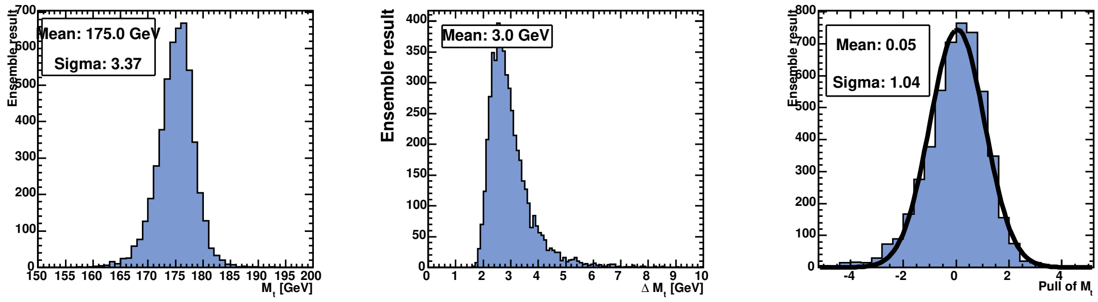

Figure 10: An ensemble of pseudo-experiments of 10 parton-level top quark pair events where the quark energies are smeared by 10% with $M_t = 175$ GeV

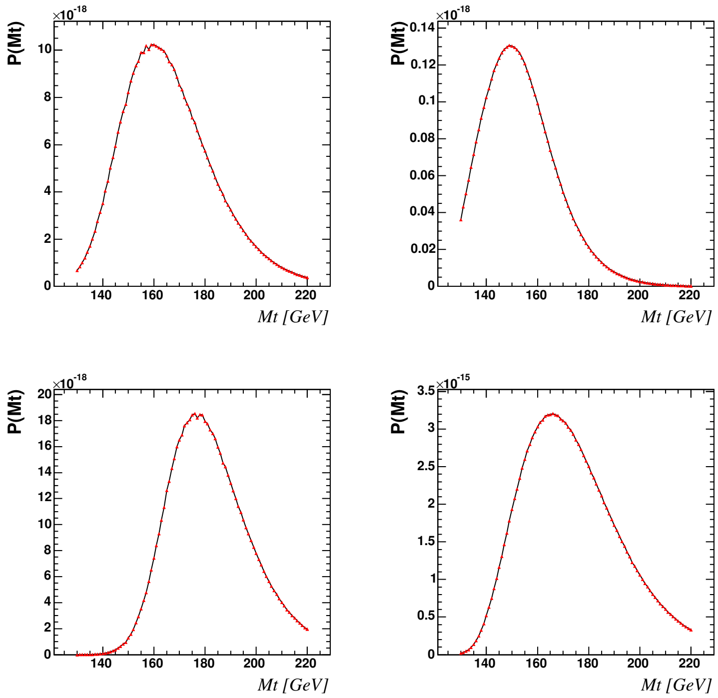

Figure 11: Posterior probability as a function of pole mass for four simulated top quark pair events with $M_t = 180$ GeV.

to an electron or muon. Here, the lower energy of the resulting lepton and the loss of information due to the two additional neutrinos lead to a significant widening of the pull distribution. If we relax the requirement that the quarks be well-matched to jets, and mix the dileptons in the expected proportions, we notice that the effect of our approximations is to introduce a widening of the pull distribution. These can be

Leptons	Jets	Fitted Pull
Ideal	TF-smeared	$\mathbf{0.98 \pm 0.03}$
P_T-matched	Tightly Matched	$\mathbf{1.00 \pm 0.03}$
Ideal	Matched	1.12 ± 0.05
P_T-matched	Matched	1.08 ± 0.03
ee	Matched	1.10 ± 0.03
$e\mu$	Matched	1.18 ± 0.03
$\mu\mu$	Matched	1.21 ± 0.03
$\geq 1\tau$	Matched	1.28 ± 0.03
All	Matched	1.24 ± 0.03
Ideal	All	1.36 ± 0.05
ee	All	1.31 ± 0.03
$\mu\mu$	All	1.43 ± 0.03
$\geq 1\tau$	All	1.66 ± 0.03
All	2 jets	1.32 ± 0.03
All	All	$\mathbf{1.44 \pm 0.03}$

Table 2: Tests of the extraction of the fitted pull width as restrictions on the sample are relaxed.

measured with simulated events and corrected for; see Section 5.

5 Backgrounds

While the di-lepton sample is fairly clean, it is not a pure sample. The major backgrounds are:

- Drell-Yan production with two jets,
- pair production of W bosons with two jets
- associated production of W and Z bosons with two jets
- production of W with three jets, where one is misreconstructed as a lepton

In order to reduce the effect of these events on the measurement of the mass, we calculate the differential cross-section for the Z+2jet processes and WW+2jet processes. Section 5.1 discusses the calculation of the matrix element using the ALPGEN routine, and the subsequent sections describe the procedure to integrate the matrix element over the phase space.

5.1 Matrix Element calculation with Alpgen

Background processes are different from the signal process in an important way. While the signal process can be well described using a small number of diagrams (*e.g.* top quark pair production), the background to these events with hard jets typically require more than an order of magnitude more diagrams to accurately describe the process. Further, there is no closed-form expression for the QCD processes which dominate this production.

We choose to extract the matrix-element calculation machinery from a well-known Monte Carlo event generator which employs effective approximations to estimate the matrix element. This has several advantages:

- It is trivial to generate events from the matrix element in order to test its evaluation
- The matrix element code is well known and has been extensively tested
- Authors of the code are likely to provide assistence and technical knowledge

There exists a wide variety of Monte Carlo event generators to choose from. Recently, generators such as CompHEP and Madgraph have provided generation based directly on the explicit construction and evaluation of the Feynman diagrams. These generators offer a transparent and clean interface to the matrix element. However, in the case of a final state with more than a single jet, the number of diagrams very quickly becomes unwieldy and the calculation very slow. For this reason, we have chosen to use the matrix element evaluation of the ALPGEN[?] generator as it offers the following advantages:

- The ALPHA algorithm avoids explicit generation and evaluation of every diagram
- Large, well-studied samples of generated events are available and form the basis of understanding of the experimentally collected data.

In the next sections, we discuss the prescription that we have developed for the evaluation of the matrix element by adapting a portion of the ALPGEN program.

5.2 Interface to Alpgen

The matrix element calculated by ALPGEN is a function of the final and initial particles spin, color, flavor as well as momentum. In the generation of events, a single spin, color and flavor assignment is randomly generated. As this information is not available to us from detector measurements, we must perform a sum over all possible configurations.

As the number of final state particles increases, the number of spin and color configurations grows quickly; to rigorously sum over each of the possibilities would be impractical and unnecessary. However, it appears to be possible to approximate the total sum with a statistical sampling. Figure 5.2 shows the distribution of $|M|^2$ in generated events using sums of a specific number of terms. it is clear that the matrix element calculation is sensitive to the configuration and that it converges fairly rapidly.

We sample spin and color configurations until the sum has converged, satisfying the criteria

$$\frac{RMS(|M|^2)}{Mean(|M|^2)\sqrt{N}} < \epsilon$$

where ϵ is the convergence tolerance, set to $\epsilon = 0.25$ for these calculations.

The calculation of the matrix element is well behaved; as an example, Figure 13 shows the variation of $|M_{Zjj}|^2$ with the di-lepton invariant mass. It shows the expected strong peak at $M_{ll} = M_Z$.

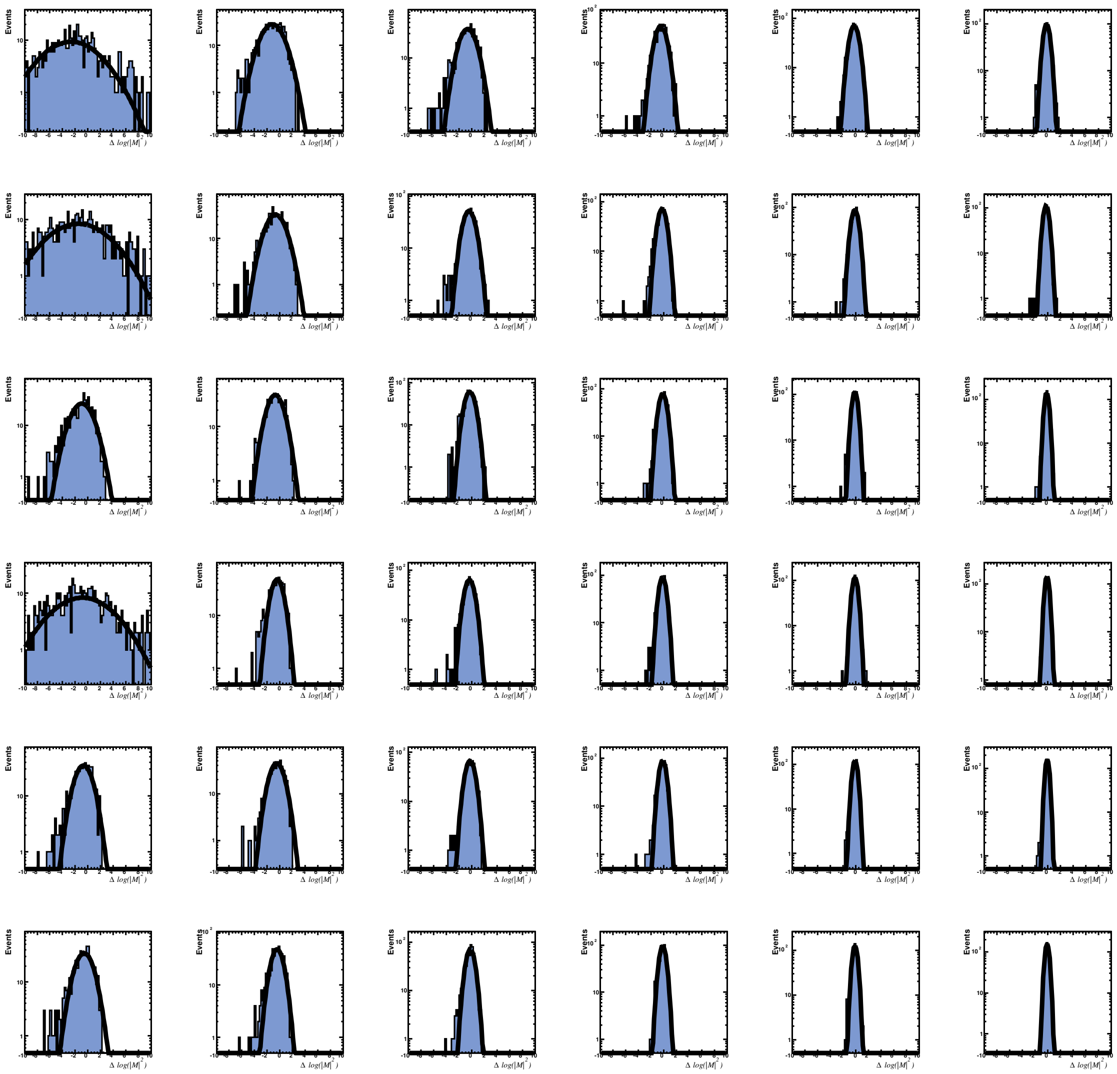

Figure 12: Variation in $\log(|M|^2)$ with increasing terms in the spin and color sum. From top, number of spin terms sampled increases by powers of 2 from 1 to 32. From left, number of color terms sampled increases by powers of 2 from 1 to 32.

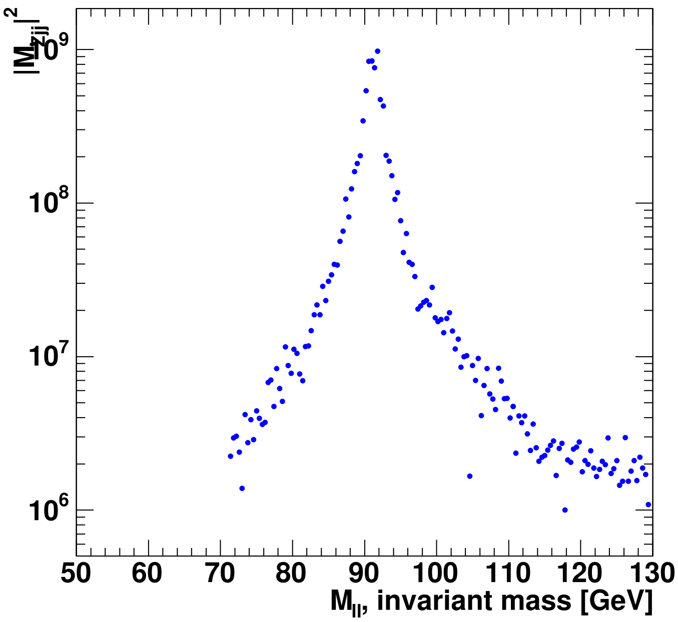

Figure 13: Variation of the squared matrix element for $pp \to Z(\to ll) + 2jets$ as a function of the leptonic pair invariant mass.

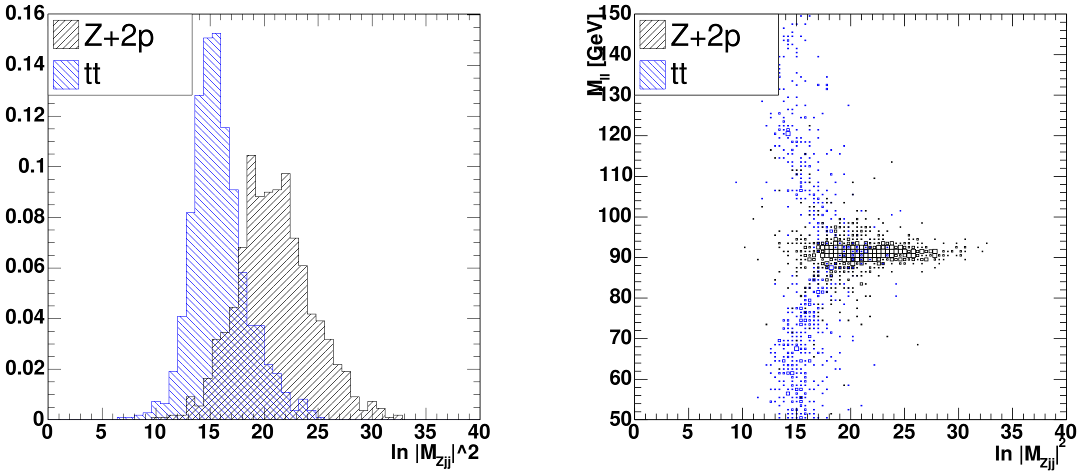

Figure 14: Evaluation of the squared matrix element for $pp \to Z(\to ll) + 2jets$ for parton-level $t\bar{t}$ and $Z + 2p$ events.

5.3 Drell-Yan : differential cross-section

Calculation of the differential cross-section for Z+2jet production requires integration of the matrix element over the phase space. In this case, the number of constraints from the measurement of the final state objects reduces the number of integrals to exactly zero.

We use the following notation: q_1, q_2 are the four-momenta of the final state partons which lead to creation of extra jets, L_1, L_2 are the four-momenta of the final state leptons, p_1, p_2 — the four-momenta of incoming partons. The expression for the differential cross-section of this process is of the form:

$$\begin{aligned} d\sigma = & \sum_{a,b} \iint \frac{1}{(2\pi)^{12}} \frac{1}{16} \frac{1}{E_{L_1} E_{L_2} E_{q_1} E_{q_2}} \\ & f(E_{q_1}, E_{j_1})\, f(E_{q_2}, E_{j_2})\, q_1^{R^2} \sin\theta_{q_1}\; q_2^{R^2} \sin\theta_{q_2} \\ & \delta(\phi_{q_1} - \phi_{j_1})\, \delta(\theta_{q_1} - \theta_{j_1})\, \delta(\phi_{q_2} - \phi_{j_2})\, \delta(\theta_{q_2} - \theta_{j_2}) \\ & dq_1^R\, d\theta_{q_1}\, d\phi_{q_1}\, dq_2^R\, d\theta_{q_2}\, d\phi_{q_2} \\ & f^a_{\mathtt{PDF}}(\frac{p_1^z}{E_{\mathtt{beam}}}, Q^2)\; f^b_{\mathtt{PDF}}(\frac{p_2^z}{E_{\mathtt{beam}}}, Q^2)\; \frac{dp_1^z\, dp_2^z}{E^2_{\mathtt{beam}}} \\ & \delta^4(p_1 + p_2 - L_1 - L_2 - q_1 - q_2) \frac{(2\pi)^4}{4(p_1^{\mathtt{cm}} \sqrt{s})} |\mathcal{M}|^2 \end{aligned} \tag{40}$$

In the above expression we set the transverse components of incoming parton momenta to zero. This simplification can be removed if necessary by the method similar to one used in derivation of the signal differential cross-section.

Using the same collider approximation as before, without the loss of generality $p_1^0 = p_1^z$ and $p_2^0 = p_2^z$. Integration over $dp_1^z\, dp_2^z$ and t and z parts of the four-dimensional δ-function result in the extra factor of $1/2$ and the following substitution:

$$\begin{aligned} p_1^z &= \frac{1}{2}\left((L_1^z + L_2^z + q_1^z + q_2^z) + (L_1^0 + L_2^0 + q_1^0 + q_2^0)\right) \\ p_2^z &= \frac{1}{2}\left((L_1^z + L_2^z + q_1^z + q_2^z) - (L_1^0 + L_2^0 + q_1^0 + q_2^0)\right) \end{aligned} \tag{41}$$

Evaluation of the integrals over $d^3q_1\, d^3q_2$ is performed in spherical coordinates, in order to take advantage of the angular δ-functions. After the angular integrations, all $\phi_{q_{1,2}}$ and $\theta_{q_{1,2}}$ are replaced by corresponding jet angles. Only two integrals and two δ-functions remain — $\delta(L_1^x + L_2^x + q_1^x + q_2^x)\, \delta(L_1^y + L_2^y + q_1^y + q_2^y)\, dq_1^R\, dq_2^R$. This leads to the following system of equations for radial components of the quarks:

$$\begin{aligned} \begin{pmatrix} (\sin\theta_{j_1} \sin\phi_{j_1}) & (\sin\theta_{j_2} \sin\phi_{j_2}) \\ (\sin\theta_{j_1} \cos\phi_{j_1}) & (\sin\theta_{j_2} \cos\phi_{j_2}) \end{pmatrix} \begin{pmatrix} q_1^R \\ q_2^R \end{pmatrix} \equiv \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} q_1^R \\ q_2^R \end{pmatrix} = \\ = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \equiv \begin{pmatrix} -(L_1^x + L_2^x)[+p^x] \\ -(L_1^y + L_2^y)[+p^y] \end{pmatrix} \end{aligned} \tag{42}$$

where the terms in square brackets have to be added if we allow non-zero transverse momenta of the incoming partons. As in the signal case, the expressions inside the δ-functions are some functions of the integration variables. The Jacobian factor is required. Using (20), we write down this factor as:

$$\begin{aligned}|J| &= \left|\frac{\partial f_i}{\partial x_j}\right| = \\ &= \begin{vmatrix}(\sin\theta_{j_1}\ \sin\phi_{j_1}) & (\sin\theta_{j_2}\ \sin\phi_{j_2}) \\ (\sin\theta_{j_1}\ \cos\phi_{j_1}) & (\sin\theta_{j_2}\ \cos\phi_{j_2})\end{vmatrix} = \\ &= 2\sin\theta_{j_1}\ \sin\theta_{j_2}\ \sin(\phi_{j_1}-\phi_{j_2})\end{aligned} \tag{43}$$

The entire expression for the differential cross-section:

$$\begin{aligned}d\sigma =& \frac{1}{(2\pi)^{12}}\frac{1}{16}\frac{1}{E_{L_1}E_{L_2}E_{q_1}E_{q_2}}f(E_{q_1},E_{j_1})f(E_{q_2},E_{j_2})\,q_1^{R^2}q_2^{R^2}\sin\theta_{j_1}\sin\theta_{j_2} \\ & f^a_{\texttt{PDF}}(\frac{p_1^z}{E_{\texttt{beam}}},Q^2)f^b_{\texttt{PDF}}(\frac{p_2^z}{E_{\texttt{beam}}},Q^2)|J|^{-1}\frac{(2\pi)^4}{4(p_1^{\texttt{cm}}\sqrt{s})}\frac{|\mathcal{M}|^2}{2}\end{aligned} \tag{44}$$

The flux factor $4(p_1^{\texttt{cm}}\sqrt{s})$ can be rewritten as $2|p_1^z p_2^z|$. With some trivial trigonometrical transformations:

$$d\sigma = \frac{1}{(2\pi)^8}\frac{1}{16^2}\frac{f_{\texttt{PDF}}(\frac{p_1^z}{E_{\texttt{beam}}},Q^2)\,f_{\texttt{PDF}}(\frac{p_2^z}{E_{\texttt{beam}}},Q^2)}{E_{L_1}E_{L_2}E_{q_1}E_{q_2}}\frac{|\mathcal{M}|^2}{2}\frac{f(E_{q_1},E_{j_1})\,f(E_{q_2},E_{j_2})\,q_1^{R^2}\,q_2^{R^2}}{|p_1^z p_2^z||\sin(\phi_{j_1}-\phi_{j_2})|} \tag{45}$$

and the q_1^R, q_2^R are the solutions of (42):

$$q_1^R = \frac{c_1 a_{22} - c_2 a_{12}}{a_{11}a_{22} - a_{12}a_{21}} \qquad q_2^R = \frac{a_{11}c_2 - a_{21}c_1}{a_{11}a_{22} - a_{12}a_{21}} \tag{46}$$

with

$$\begin{aligned}& c_1 = -(L_1^x + L_2^x) \qquad [+p^x] \\ & c_2 = -(L_1^y + L_2^y) \qquad [+p^y] \\ & (a_{11}a_{22} - a_{12}a_{21}) = 2\sin\theta_{q_1}\sin\theta_{q_2}\sin(\phi_{q_1}-\phi_{q_2}) \\ & (c_1 a_{22} - c_2 a_{12}) = \sin\theta_{q_2}\left((L_1^y + L_2^y)\sin\phi_{q_2} - (L_1^x + L_2^x)\cos\phi_{q_2}\right) \\ & (a_{11}c_2 - a_{21}c_1) = \sin\theta_{q_1}\left((L_1^x + L_2^x)\cos\phi_{q_2} - (L_1^y + L_2^y)\sin\phi_{q_2}\right)\end{aligned} \tag{47}$$

where, as before, the terms in square brackets should be included if non-zero incoming transverse momentum is considered.

5.3.1 Performance

We evaluate the Zjj probability on events which have passed the final dilepton selections. [8] The detector simulation degrades our ability to separate the signal from the

[8] refer to the Section 7 for the description of the event selection

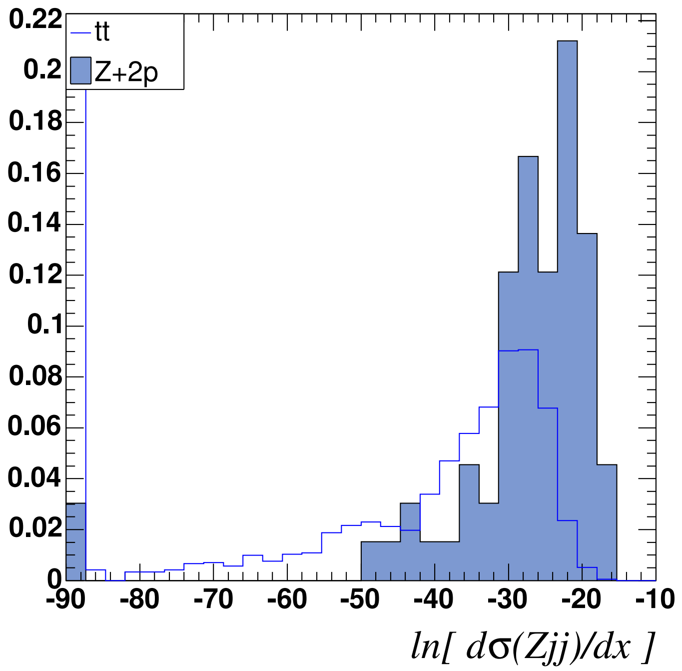

Figure 15: Evaluation of the Zjj probability for Z and $t\bar{t}$ events.

background events, and the event selection sculpts both sub-samples so that they are nearly identical kinematically. However, there is still visible discrimination between the Z events and the other constituents. Figures 15 and **??** show the distribution of the natural logarithm of the Zjj probability densities for both selections. Note that $t\bar{t}$ and WW events peak significantly lower than the Zjj events, while the $Z \to \tau\tau$ events are more similar to the Z events.

5.4 Cross section for W pair production

As with the case of $t\bar{t}$ and Drell-Yan, the calculation of the differential cross-section for WW+2 jets production requires integration of the matrix element over the phase space. As with Drell-Yan, an `ALPGEN` matrix element is used for the calculation. Unlike Drell-Yan, the presence of two neutrinos in the events results in there being six integrals to evaluate.

5.4.1 Differential cross section and notation

We use the following notation: L_1, L_2 are the measured four-momenta and ℓ_1, ℓ_2 are the parton-level four-momenta of the final state leptons, ν_1, ν_2 are the four-momenta of the final state neutrinos, p_1, p_2 are the four-momenta of the final state partons that lead to creating extra jets and q_1, q_2 are the four-momenta of the incoming partons. The same set of assumptions made in the signal case about the measured lepton energies, jet angles, and transverse components of the initial parton energies are made here. The differential cross-section for the process then takes the form:

$$
\begin{aligned}
d\sigma = \left(\frac{1}{(2\pi)^3}\right)^6 \iint \sum_{i,j} |M|^2 (2\pi)^4 \delta^4(q_1 + q_2 - \ell_1 - \nu_1 - p_1 - \ell_2 - \nu_2 - p_2) \\
\delta(q^x)\,\delta(q^y)\,dq^x\,dq^y\, f^i_{\mathtt{PDF}}\left(\frac{q_1^i}{E_{\mathtt{beam}}}, Q^2\right) f^j_{\mathtt{PDF}}\left(\frac{q_2^j}{E_{\mathtt{beam}}}, Q^2\right) f_{\mathtt{flux}}(q_1^z, q_2^z)\,\frac{dq_1^i\,dq_2^j}{E_{\mathtt{beam}}} \\
\delta^3(\ell_1 - L_1)\,\delta^3(\ell_2 - L_2)\,\frac{d^3\ell_1}{2E_{\ell_1}}\frac{d^3\ell_2}{2E_{\ell_2}} f(E_{p_1}, E_{j_1})\, f(E_{p_2}, E_{j_2}) \\
\delta(\theta_{p_1} - \theta_{j_1})\,\delta(\theta_{p_2} - \theta_{j_2})\,\delta(\phi_{p_1} - \phi_{j_1})\,\delta(\phi_{p_2} - \phi_{j_2})\frac{d^3p_1}{2E_{p_1}}\frac{d^3p_2}{2E_{p_2}}\frac{d^3\nu_1}{2E_{\nu_1}}\frac{d^3\nu_2}{2E_{\nu_2}}
\end{aligned}
\tag{48}
$$

Here, $\delta^3(\ell_i - L_i)$, $\delta(\theta_{p_1} - \theta_{j_1})$, and $\delta(\phi_{p_i} - \phi_{j_i})$ enforce the assumptions about lepton momenta and jet angles, as in Equation 9. As with the signal case, the δ-functions over q_x and q_y impose the collider approximation $q_i^0 = q_i^z$ which we can make without loss of generality.

5.4.2 Phase space transformation and integration

As with the signal case, we perform a phase-space transformation in order to make the integration over parton-level variables more efficient. We introduce invariant masses of intermediate particles, in this case, the two W bosons, using the procedure described in Appendix A. The integration procedure is very similar to that used in the signal case (described in Section 5). After integrating over $d^3\ell_i$ and the δ-functions introduced by the phase-space transformation, we transform the parton variables into spherical coordinates, integrate over the parton angles and are left with:

$$
\begin{aligned}
d\sigma = & \int (2\pi)^4 \frac{|M|^2}{2} \left(\frac{1}{(2\pi)^3}\right)^6 f^i_{\texttt{PDF}}(\frac{q_1^z}{E_{\texttt{beam}}}, Q^2)\ f^j_{\texttt{PDF}}(\frac{q_2^z}{E_{\texttt{beam}}}, Q^2)\ \frac{1}{E^2_{\texttt{beam}}} \\
& \delta\left((L_1+\nu_1+p_1)_x + (L_2+\nu_2+p_2)_x\right)\ \delta\left((L_1+\nu_1+p_1)_y + (L_2+\nu_2+p_2)_y\right) \\
& \delta\left((L_1+\nu_1)^2 - M^2_{W_1}\right)\ \delta\left((L_2+\nu_2)^2 - M^2_{W_2}\right) \\
& \frac{p^2_{1R}\sin\theta_{p_1} dp_{1R}}{2E_{p_1}}\ \frac{p^2_{2R}\sin\theta_{p_2} dp_{2R}}{2E_{p_2}}\ \frac{d^3\nu_1}{2E_{\nu_1}}\ \frac{d^3\nu_2}{2E_{\nu_2}}\ \frac{1}{2E_{L_1}}\ \frac{1}{2E_{L_2}} \\
& f_{\texttt{flux}}(q_1^z, q_2^z)\ f(E_{p_1}, E_{j_1})\ f(E_{p_2}, E_{j_2})\ dM^2_{W_1}\ dM^2_{W_2}
\end{aligned}
\tag{49}
$$

As with the signal case, we are left with integration over the six neutrino momentum components. Unlike the signal case, we only have two intermediate state invariant masses to integrate over. Thus, we must retain two neutrino momentum components as integration variables. We arbitrarily chose the z-components. We are then left with 6 integrations: 2 over the neutrino components ν_{1z} and ν_{2z}, 2 over the magnitudes of the partons resulting in jets p_{1R} and p_{2R}, and 2 over the assumed invariant masses of the W-bosons m_{W_1} and m_{W_2}. We now require a Jacobian term (from Equation 20) and solutions for the remaining four neutrino momentum components in terms of either known quantities or the integration variables.

5.4.3 Neutrino Solutions

The system of equations for neutrino momenta for W pairs is the following:

$$
\begin{aligned}
(L_1+\nu_1)^2 &= M^2_{W_1} \\
(L_2+\nu_2)^2 &= M^2_{W_2} \\
(L_1+\nu_1+p_1)_x + (L_2+\nu_2+p_2)_x &= 0 \\
(L_1+\nu_1+p_1)_y + (L_2+\nu_2+p_2)_y &= 0
\end{aligned}
\tag{50}
$$

The need to solve for only four of the six neutrino components makes the derivation of neutrino solutions simpler than in the signal case. We define:

$$
\begin{aligned}
Q_x &\equiv -L_{1x} - L_{2x} - p_{1x} - p_{2x} \\
Q_y &\equiv -L_{1y} - L_{2y} - p_{1y} - p_{2y}
\end{aligned}
\tag{51}
$$

Then, using the last two equations in (50), we make the substitutions $\nu_{2x} = Q_x - \nu_{1x}$ and $\nu_{2y} = Q_x - \nu_{1y}$ in the first two equations. We expand the first two equations in (50) and derive two coupled quadratic equations in terms of ν_{2x} and ν_{2y} that are of the form:

$$\begin{aligned} a_1\nu_{2x}^2 + a_2\nu_{2x}\nu_{1y} + a_3\nu_{1y}^2 + a_4\nu_{2x} + a_5\nu_{1y} + a_6 = 0 \\ b_1\nu_{2y}^2 + b_2\nu_{2x}\nu_{1y} + b_3\nu_{1x}^2 + b_4\nu_{2y} + b_5\nu_{1x} + b_6 = 0 \end{aligned} \tag{52}$$

The two coupled equations are then converted to a single fourth order polynomial which can be solved numerically using the Sturm Sequence technique. Alternatively, an analytical solution similar to the one shown in Appendix C can be employed.

5.4.4 Jacobian Factor

Jacobian factor that appears from integrations over neutrino momenta components has the form:

$$J = \left| \frac{\partial f_i}{\partial \nu_{(1,2)(x,y,z)}} \right| \tag{53}$$

The f_i are the functions of neutrino momenta that occur inside δ-functions:

$$\begin{aligned} f_1 &= ((L_1 + \nu_1 + p_1)_x + (L_2 + \nu_2 + p_2)_x) \\ f_2 &= ((L_1 + \nu_1 + p_1)_y + (L_2 + \nu_2 + p_2)_y) \\ f_3 &= \left((L_1 + \nu_1)^2 - M_{W_1}^2\right) \\ f_4 &= \left((L_2 + \nu_2)^2 - M_{W_2}^2\right) \\ f_5 &= \nu_{1z} \\ f_6 &= \nu_{2z} \end{aligned} \tag{54}$$

The determinant of the Jacobian has the form:

$$\begin{vmatrix} 1 & 0 & J_{3x} & 0 & 0 & 0 \\ 0 & 1 & J_{3y} & 0 & 0 & 0 \\ 1 & 0 & 0 & J_{4x} & 0 & 0 \\ 0 & 1 & 0 & J_{4y} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

This evaluates to:

$$J = |J_{3y}J_{4x} - J_{3x}J_{4y}| \tag{55}$$

Where:

$$
\begin{aligned}
J_{3x} &= \frac{\partial f_3}{\partial \nu_{1x}} = \frac{2L_1^0 \nu_{1x}}{\sqrt{\nu_{1x}^2 + \nu_{1y}^2 + \nu_{1z}^2}} - 2L_{1x} \\
J_{3y} &= \frac{\partial f_3}{\partial \nu_{1y}} = \frac{2L_1^0 \nu_{1y}}{\sqrt{\nu_{1x}^2 + \nu_{1y}^2 + \nu_{1z}^2}} - 2L_{1y} \\
J_{4x} &= \frac{\partial f_3}{\partial \nu_{2x}} = \frac{2L_2^0 \nu_{2x}}{\sqrt{\nu_{1x}^2 + \nu_{1y}^2 + \nu_{1z}^2}} - 2L_{2x} \\
J_{4y} &= \frac{\partial f_3}{\partial \nu_{2y}} = \frac{2L_2^0 \nu_{2y}}{\sqrt{\nu_{1x}^2 + \nu_{1y}^2 + \nu_{1z}^2}} - 2L_{2y}
\end{aligned}
\tag{56}
$$

5.4.5 Final integral

As with the signal case, we evaluate the Jacobian factor after finding neutrino solutions and obtain a final expression for the unnormalized probability for WW+2 jets events:

$$
\begin{aligned}
\frac{d\sigma}{d\mathbf{x}} = \left(\frac{1}{(2\pi)^3}\right)^6 (2\pi)^2 \iint \sum_{i,j} \frac{|M|^2}{2E_{\mathtt{beam}}^2} \frac{p_{1R}^2 \sin\theta_{p_1} dp_{1R}}{2E_{p_1}} \, \frac{p_{2R}^2 \sin\theta_{p_2} dp_{2R}}{2E_{p_2}} \\
\frac{J^{-1}}{E_{\nu_1} E_{\nu_2} E_{L_1} E_{L_2}} d\nu_{1z} d\nu_{2z} f_{\mathtt{PDF}}^i\left(\frac{q_1^z}{E_{\mathtt{beam}}}, Q^2\right) \, f_{\mathtt{PDF}}^j\left(\frac{q_2^z}{E_{\mathtt{beam}}}, Q^2\right) \\
f_{\mathtt{flux}}(q_1^z, q_2^z) f(E_{p_1}, E_{j_1}) \, f(E_{p_2}, E_{j_2}) \, M_{W_1} \, M_{W_2} \, dM_{W_1} \, dM_{W_2}
\end{aligned}
\tag{57}
$$

This resulting 6-dimensional integral is very similar to the one obtained in the signal case. This integral is evaluated over the parton momenta, the W boson invariant masses, and the neutrino momenta z components to obtain the differential cross-section. The sum over i, j is over incoming parton flavors.

5.4.6 Performance

A test of the ALPGEN matrix element for $WWjj$ is shown in Figure 16 using generated, parton-level events for both $WWjj$ and $t\bar{t}$. We also evaluate the $WWjj$ probability on simulated events that have passed final dilepton selections for the DIL sample. As with the Zjj case, the event selection results in the passing background events to be very similar to $t\bar{t}$ kinematically. There is still, however, visible discrimination between $WWjj$ events and $t\bar{t}$ events using the $WWjj$ probability calculation, as seen in Figure 17.

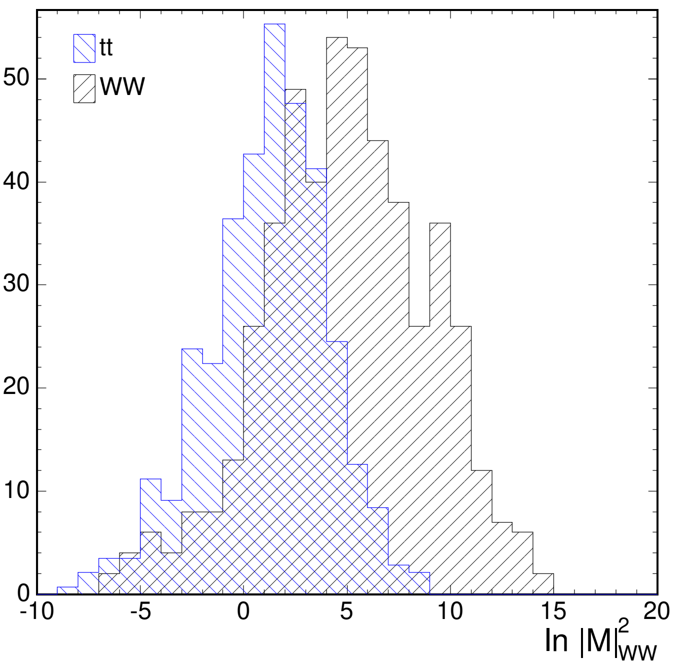

Figure 16: Evaluation of the squared matrix element for $pp \to WW(\to l\nu l\nu) + 2jets$ for parton-level $t\bar{t}$ and $WW + 2p$ events.

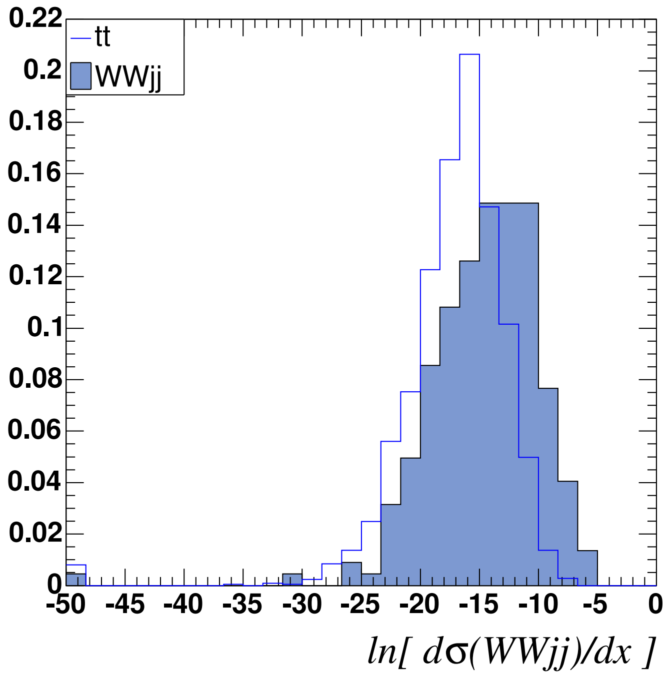

Figure 17: Evaluation of the $WWjj$ probability for WW and $t\bar{t}$ events

5.5 Fake background modelling

Modelling of the fakes presents a particular challenge, as they do not correspond simply to a particular physical process. Many of the fake events creep into the sample when a jet fakes a lepton signature, most typically an electron. As the rate at which jets fake leptons is << 1, we expect the vast majority of the fake events to have a single real lepton and three jets. Events of this category are typically produced via the process

$$p\overline{p} \to Wjjj \to l\nu jjj$$

where the neutrino escapes without detection.

Using this process as the basis for our model, we write

$$P_{fake} = P_{Wjjj}(j_1, j_2, l_1, [l_2 \to j_3]) + P_{Wjjj}(j_1, j_2, [l_1 \to j_3], l_2)$$

where we sum over the possibility that either lepton is a fake originating from a jet. Calculation of P_{Wjjj} follows the style of the other probability calculations above.

5.5.1 Wjjj: differential cross section

Calculation of the differential cross-section for W+3jet production requires integration of the matrix element over the phase space. In this case, the number of constraints from the measurement of the final state objects reduces the number of integrals to four.

We use the following notation: q_1, q_2, q_3 are the four-momenta of the final state partons which lead to creation of extra jets, L is the four-momenta of the final state lepton,ν the four-momentum of the final state neutrino and p_1, p_2 — the four-momenta of incoming partons. The expression for the differential cross-section of this process is of the form:

$$\begin{aligned} d\sigma = & \sum_{a,b} \iint \frac{1}{(2\pi)^{12}} \frac{1}{32} \frac{1}{E_L E_\nu E_{q_1} E_{q_2} E_{q_3}} \\ & f(E_{q_1}, E_{j_1})\, f(E_{q_2}, E_{j_2})\, f(E_{q_e}, E_{j_e})\; q_1^{R2} \sin\theta_{q_1}\; q_2^{R2} \sin\theta_{q_2}\; q_3^{R2} \sin\theta_{q_3} \\ & \delta(\phi_{q_1} - \phi_{j_1})\, \delta(\theta_{q_1} - \theta_{j_1})\, \delta(\phi_{q_2} - \phi_{j_2})\, \delta(\theta_{q_2} - \theta_{j_2})\, \delta(\phi_{q_3} - \phi_{j_3})\, \delta(\theta_{q_3} - \theta_{j_3}) \\ & dq_1^R\, d\theta_{q_1}\, d\phi_{q_1}\, dq_2^R\, d\theta_{q_2}\, d\phi_{q_2}\, dq_3^R\, d\theta_{q_3}\, d\phi_{q_3} d^3\nu \\ & f^a_{\mathtt{PDF}}(\frac{p_1^z}{E_{\mathtt{beam}}}, Q^2)\; f^b_{\mathtt{PDF}}(\frac{p_2^z}{E_{\mathtt{beam}}}, Q^2)\; \frac{dp_1^z\, dp_2^z}{E^2_{\mathtt{beam}}} \\ & \delta^4(p_1 + p_2 - L_1 - q_1 - q_2 - q_3) \frac{(2\pi)^4}{4(p_1^{\mathtt{cm}}\sqrt{s})} |\mathcal{M}|^2 \end{aligned} \tag{58}$$

In the above expression we set the transverse components of incoming parton momenta to zero. This simplification can be removed if necessary by the method similar to one used in derivation of the signal differential cross-section.

Using the same collider approximation as before, without the loss of generality $p_1^0 = p_1^z$ and $p_2^0 = p_2^z$. Integration over $dp_1^z\, dp_2^z$ and t and z parts of the four-dimensional δ-function result in the extra factor of 1/2 and the following substitution:

$$
\begin{aligned}
p_1^z &= \frac{1}{2}\left((L^z + q_1^z + q_2^z + q_3^z) + (L^0 + q_1^0 + q_2^0 + q_3^0)\right) \\
p_2^z &= \frac{1}{2}\left((L^z + q_1^z + q_2^z + q_3^z) - (L^0 + q_1^0 + q_2^0 + q_3^0)\right)
\end{aligned}
\tag{59}
$$

Evaluation of the integrals over $d^3q_1\, d^3q_2, d^3q_3$ is performed in spherical coordinates, in order to take advantage of the angular δ-functions. After the angular integrations, all $\phi_{q_{1,2,3}}$ and $\theta_{q_{1,2,3}}$ are replaced by corresponding jet angles.

The entire expression for the differential cross-section:

$$
\begin{aligned}
d\sigma = \sum_{a,b} \iint & \frac{dq_{R_1} dq_{R_2} dq_{R_3} d\nu_z}{(2\pi)^{12} 32 E_L E_\nu E_{q_1} E_{q_2} E_{q_3}} \\
& f(E_{q_1}, E_{j_1}) f(E_{q_2}, E_{j_2}) f(E_{q_3}, E_{j_3})\, q_1^{R^2} q_2^{R^2} q_3^{R^2} \sin\theta_{j_1} \sin\theta_{j_2} \sin\theta_{j_3} \\
& f^a_{\mathtt{PDF}}\left(\frac{p_1^z}{E_{\mathtt{beam}}}, Q^2\right) f^b_{\mathtt{PDF}}\left(\frac{p_2^z}{E_{\mathtt{beam}}}, Q^2\right) \frac{(2\pi)^4}{4(p_1^{\mathtt{cm}}\sqrt{s})} \frac{|\mathcal{M}|^2}{2}
\end{aligned}
\tag{60}
$$

The flux factor $4(p_1^{\mathtt{cm}}\sqrt{s})$ can be rewritten as $2|p_1^z p_2^z|$.

5.5.2 Performance

A test of the matrix element for $Wjjj$ using parton-level events is shown in Figure 18; the matrix element is able to separate $t\bar{t}$ events from $Wjjj$ events quite well. In Figure 19, we have discarded the neutrino information, Gaussian smeared the parton-level quantities and performed the four-dimensional integration using Gauassian transfer functions.

Finally, we test the full fake probability using the sample of fakeable events gathered from the data. Figure 20 shows the resulting distributions.

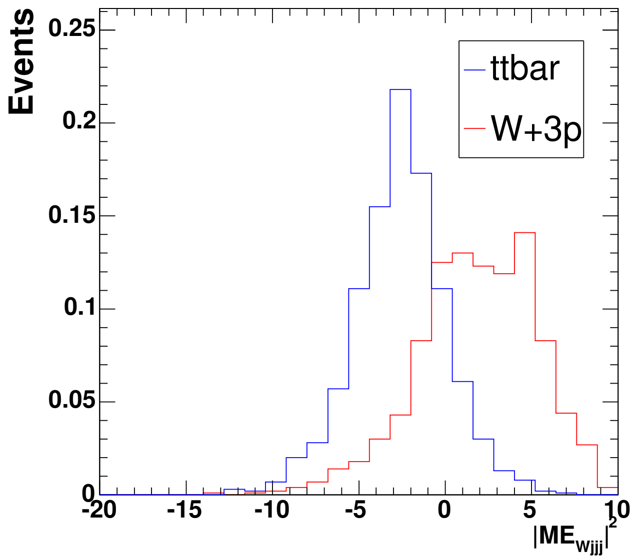

Figure 18: Evaluation of the squared matrix element for $pp \to W(\to l\nu) + 3jets$ for parton-level $t\bar{t}$ and $W + 3p$ events.

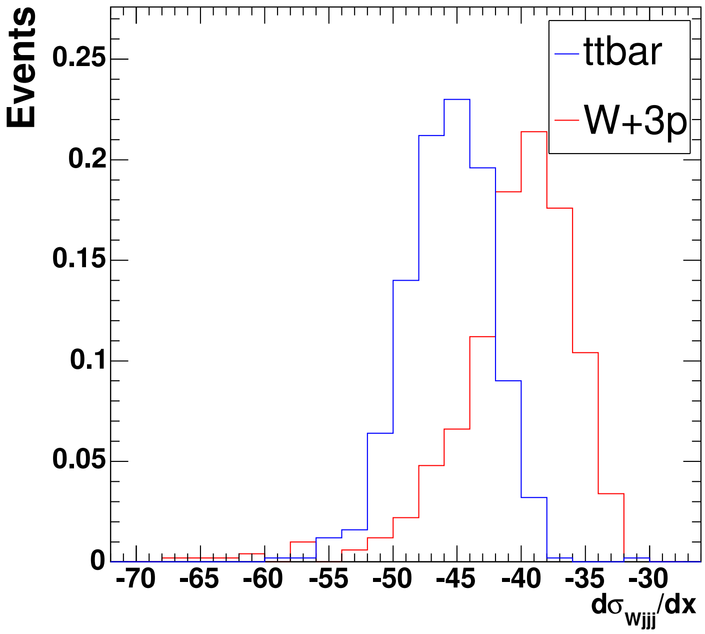

Figure 19: Evaluation of the $Wjjj$ probability for Gaussian smeared $Wjjj$ and $t\bar{t}$ events

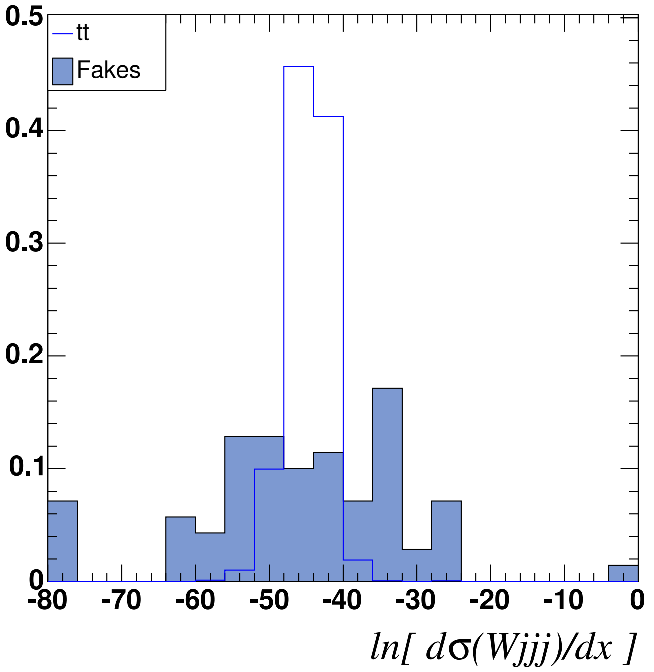

Figure 20: Evaluation of the $Wjjj$ probability for fully simulated $t\bar{t}$ and fakeable events from the data.

6 Combination of Signal and Background

There are at least two different methods of extracting the mass parameter in the presence of background. Available information includes kinematic properties of the selected events as well as their total number. The first method utilized all the available information including total number of events in the sample and corresponding luminosity. Loosely speaking it combines mass measurement with cross-section measurement. The second method is less ambitious and the total number of events in the sample is considered to be a part of experimental setup, while information about luminosity is disregarded. In this case only the kinematic properties of the selected events are used.

Let us introduce several facts and notations that lead to expressions for the likelihoods. Single-event likelihood is given by (61)

$$p(\mathbf{x}_i|M_t) = p(\mathbf{x}_i|s, M_t)p(s) + p(\mathbf{x}_i|b)p(b) \tag{61}$$

Properly normalized conditional probabilities from (61) can be expressed in terms of differential cross-sections for signal and background (62, 63) and acceptance-adjusted total cross-sections (64, 65).

$$p(\mathbf{x}_i|s, M_t) = \frac{1}{\sigma_s \epsilon_s}\left(\frac{d\sigma_s}{d\mathbf{x}_i}\right) \tag{62}$$

$$p(\mathbf{x}_i|b) = \frac{1}{\sigma_b \epsilon_b}\left(\frac{d\sigma_b}{d\mathbf{x}_i}\right) \tag{63}$$

$$\sigma_s \epsilon_s = \int\limits_{\texttt{selection}} \left(\frac{d\sigma_s}{d\mathbf{x}}\right) d\mathbf{x} \tag{64}$$

$$\sigma_b \epsilon_b = \int\limits_{\texttt{selection}} \left(\frac{d\sigma_b}{d\mathbf{x}}\right) d\mathbf{x} \tag{65}$$

The expected number of signal and background events are expressed as (66, 67)

$$\lambda_s = \mathcal{L}\sigma_s\epsilon_s \tag{66}$$

$$\lambda_b = \mathcal{L}\sigma_b\epsilon_b \tag{67}$$

where $\mathcal{L}$ is the total luminosity.

It is important to note that actual number of signal and background events are independently Poisson-distributed with parameters given by the above expected numbers (66, 67). The expected number of background events does not depend on the parameter of interest. The signal fraction of the sample, on the other hand, depends on the mass parameter through the mass-dependent expected number of signal events. A particular value of "estimated" signal-to-background ratio is thus bound to a particular assumed mass. More formally, $R(M_t^0) = S/B = \lambda_s(M_t^0)/\lambda_b$ and $R(M_t) = \lambda_s(M_t)/\lambda_b$ changes with M_t.

6.1 Method 1: combined

The likelihood term in this method is the probability of observing n events with certain sets of observable quantities $\{\mathbf{x}\}$, given assumed pole mass M_t and luminosity $\mathcal{L}$. The equation (68) shows this probability expressed in terms of differential cross-sections and expected number of signal and background events.

$$\begin{aligned} & p(\{\mathbf{x}\}, n \mid M_t, \mathcal{L}) = \\ = & \frac{\exp\left(-(\lambda_s + \lambda_b)\right)}{n!} (\lambda_s + \lambda_b)^n \prod_i \left(p(\mathbf{x}_i | s, M_t) \frac{\lambda_s}{\lambda_s + \lambda_b} + p(\mathbf{x}_i | b) \frac{\lambda_b}{\lambda_s + \lambda_b} \right) = \\ & = \frac{\exp\left(-(\lambda_s + \lambda_b)\right)}{n!} \mathcal{L}^n \prod_i \left(\frac{d\sigma_s}{d\mathbf{x}_i} + \frac{d\sigma_b}{d\mathbf{x}_i} \right) \end{aligned} \tag{68}$$

If one chooses to follow the log-likelihood prescription the equation (69) shows the relevant expression.

$$\begin{aligned} \ln L &= -(\lambda_s + \lambda_b) + n \ln \mathcal{L} - \ln n! + \sum_i \ln \left(\frac{d\sigma_s}{d\mathbf{x}_i} + \frac{d\sigma_b}{d\mathbf{x}_i} \right) = \\ &= -\mathcal{L}(\sigma_s \epsilon_s + \sigma_b \epsilon_b) + n \ln \mathcal{L} - \ln n! + \sum_i \ln \left(\frac{d\sigma_s}{d\mathbf{x}_i} + \frac{d\sigma_b}{d\mathbf{x}_i} \right) \end{aligned} \tag{69}$$

Note that signal-to-background ratio does not enter in the per-event sum on the right, which is simply a sum of natural logarithms of "total" differential cross-sections of individual events. This sum does not depend on event selection (acceptance) directly either.

Correct relative normalization of differential cross-sections is essential for reliable extraction of the mass parameter. The total acceptance corrected cross-sections in (69) can be obtained from:

- direct integration of differential cross-section over accepted region of $\mathbf{x}$
- direct integration of differential cross-section over entire $\mathbf{x}$ space with subsequent acceptance correction where acceptance is estimated from Monte Carlo generated events
- Monte Carlo generated events for signal and using measured or estimated signal-to-background ratio for background

Note that for the signal the entire $(\sigma_s\epsilon_s)(M_t)$ curve is needed.

The first approach to total cross-section evaluation is the most self-consistent but technically challenging. Only one free parameter is involved in log-likelihood optimization. The second relies on the separate and accurate acceptance estimate. The third requires parametrization of the total signal cross section dependence on the mass.

Let us focus on the third approach and consider a particular parametrization of the acceptance-corrected cross-section:

$$\sigma_s\epsilon_s = A\exp\left(-\frac{M_t}{m}\right) \tag{70}$$

The total background cross-section term will be expressed in terms of known sample composition for a certain value of M_t:

$$\sigma_b\epsilon_b = \beta A\exp\left(-\frac{M_t^0}{m}\right) \tag{71}$$

Substitution of (70, 71) into (69) and subsequent partial differentiation with respect to M_t as well as the introduced parameters shows that the only viable strategy is to estimate the parameters as good as possible before the log-likelihood optimization and do a one-parameter optimization on M_t. This observation follows from the fact that the kinematic sum term does not depend on either A, m or β and that some of the partial derivatives of the log-likelihood with respect to these parameters are always non-zero. It gives us a reason to believe that simultaneous optimization is not possible. Consequently the third approach requires accurate parametrization of the acceptance-corrected signal cross-section as well as a good estimate of a sample composition for a particular M_t value.

One more observation. If the kinematic sum was the only term in (69), the entire log-likelihood would maximize at a certain value of M_t. With extra dependence on M_t introduced by the total cross-section terms the mass estimate shifts with respect to this value. The exact amount of the shift is determined by the total cross-section terms. Thus, it is important to calculate or parametrize them correctly.

6.2 Method 2: minimal

As mentioned before this method does not strive to include the extra information provided by the luminosity measurement and number of events in the sample. Number of events n is considered a part of pre-conditions. The kinematic properties of the selected events constitute all the essential information in this case. The likelihood in question is given by (72):

$$\begin{aligned} p(\{\mathbf{x}\} \mid M_t, n) &= \\ &= \prod_i \left(p(\mathbf{x}_i|s, M_t) \frac{\sigma_s \epsilon_s}{\sigma_s \epsilon_s + \sigma_b \epsilon_b} + p(\mathbf{x}_i|b) \frac{\sigma_b \epsilon_b}{\sigma_s \epsilon_s + \sigma_b \epsilon_b} \right) = \\ &= (\sigma_s \epsilon_s + \sigma_b \epsilon_b)^{-n} \prod_i \left(\frac{d\sigma_s}{d\mathbf{x}_i} + \frac{d\sigma_b}{d\mathbf{x}_i} \right) \end{aligned} \tag{72}$$

The transformations were carried out with the help of expressions (62, 63, 64, 65, 66, 67). The expression for the log-likelihood is given by (73)

$$\ln L = -n \ln (\sigma_s \epsilon_s + \sigma_b \epsilon_b) + \sum_i \ln \left(\frac{d\sigma_s}{d\mathbf{x}_i} + \frac{d\sigma_b}{d\mathbf{x}_i} \right) \tag{73}$$

As in the previous subsection, the total acceptance-corrected cross-section terms can be obtained by integration of the differential cross-section over accepted region, by integration over all space with acceptance correction applied to the result and by parametrization of the signal cross-section and independent estimate of the sample composition for a particular value of M_t. Thus, most of the observations made in the previous subsection apply here as well. The same reasoning as before leads us to the conclusion that when the parametrization of the acceptance-corrected total cross-sections is done the simultaneous log-likelihood optimization with respect to M_t and parameters of (70, 71) will not be possible. As before, the parametrization needs to be accurate because it introduces the necessary shift into the mass estimate.

6.3 Comparison

The log-likelihoods given by (69) and (73) share many similarities. The kinematic sum term is exactly the same. It does not carry any explicit dependence on sample composition or any extra parameters. It is the source of essential dependence on M_t.

The common and quite unfortunate feature of both methods is that although the total cross-section terms are different, both have to be either calculated from the differential cross-sections or parametrized. The former is computationally intense, the latter has to be accurate as well as consistent with the differential cross-sections. The simultaneous log-likelihood optimization with respect to the mass and the parameters of the total cross-section does not seem to be possible.

Both methods make the mass estimate shift with respect to the value obtained from the kinematic sum term only. The shift amounts are different as in the first method there is an additional “force” in the form of expected number of events that is taken into account. This is perceived as the only essential difference between the two methods.

6.4 Approach taken

In our current analysis we decided to pursue a variation of Method 2, with a slight modification, where total acceptance corrected cross-sections from (72) are expressed instead in terms of expected number of signal and background events. The total acceptance corrected signal cross-section is parametrised and parameters are determined using Monte Carlo.

Let us denote acceptance corrected total signal cross-section as $(\sigma_s\epsilon_s)(M_t)$. The sample composition is given by the expected number of signal events $\lambda_s(M_t)$ and background λ_b. For a certain assumed top mass m_0 the sample composition is estimated by the $t\bar{t}$ production cross-section measurement group and are presented in [?]. The expected numbers of events of each type are dependent on the selection used. Introducing the normalisation factors $\mathcal{N}_s$ and $\mathcal{N}_b$ which multiplied by the corresponding differential cross-sections produce the proper probability densities:

$$\begin{aligned} p_s(\mathbf{x}_i|M_t) &= \frac{d\sigma_s}{d\mathbf{x}_i}\frac{(\sigma_s\epsilon_s)(m_0)}{(\sigma_s\epsilon_s)(M_t)\,\mathcal{N}_s(m_0)} \\ p_b(\mathbf{x}_i) &= \frac{d\sigma_b}{d\mathbf{x}_i}\frac{1}{\mathcal{N}_b} \\ \lambda_s(M_t) &= \frac{\lambda_s(m_0)}{(\sigma_s\epsilon_s)(m_0)}(\sigma_s\epsilon_s)(M_t) \end{aligned} \tag{74}$$

The per-event probability density is then given by:

$$\begin{aligned} p(\mathbf{x}_i|M_t) &= p_s(\mathbf{x}_i|M_t)\frac{\lambda_s(M_t)}{\lambda_s(M_t)+\lambda_b} + p_b(\mathbf{x}_i)\frac{\lambda_b}{\lambda_s(M_t)+\lambda_b} \\ &= \left[\frac{d\sigma_s}{d\mathbf{x}_i}\frac{\lambda_s(m_0)}{\mathcal{N}_s(m_0)} + \frac{d\sigma_b}{d\mathbf{x}_i}\frac{\lambda_b}{\mathcal{N}_b}\right]\left(\frac{1}{\frac{\lambda_s(m_0)}{(\sigma_s\epsilon_s)(m_0)}(\sigma_s\epsilon_s)(M_t)+\lambda_b}\right) \end{aligned} \tag{75}$$

The likelihood expression for the entire dataset is, of course, simply a product of per-event probability densities given by (75).

7 Selection

There are two complementary selection schemes for dilepton decays of top quarks. The traditional DIL selection requires two tight leptons, while the LTRK selection requires a single tight lepton and an isolated track. The DIL selection has less background, but the LTRK selection has a higher signal acceptance. Preliminary results suggested that the DIL selection would provide a more sensitive and robust result.

In this section, we briefly describe the event selection and extract the acceptance as a function of the top mass, which is needed for the normalization of the $t\bar{t}$ probability.

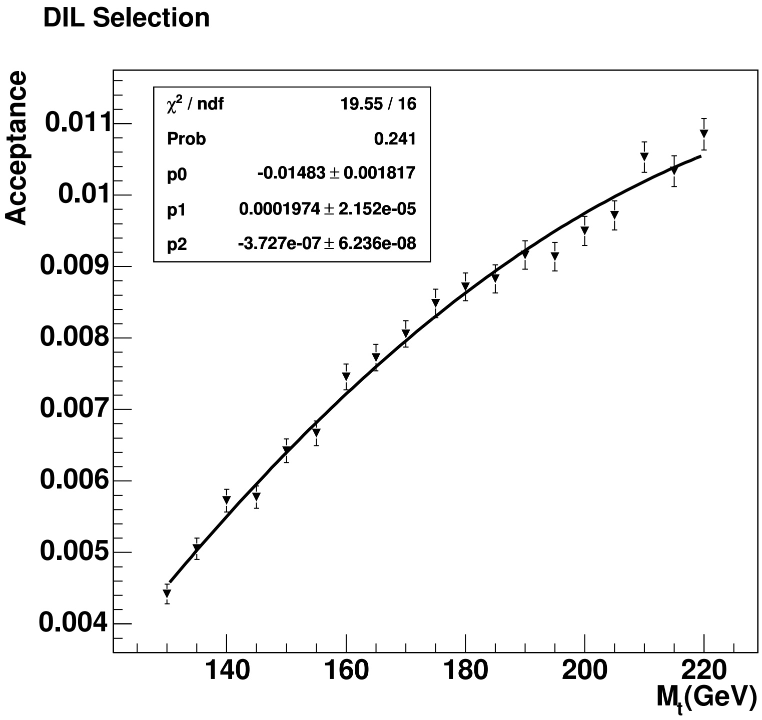

Figure 21: Acceptance of the DIL selection as a function of top pole mass.

Process	N_{DIL}
$t\bar{t}$	16.94 ± 0.19
WW	1.26 ± 0.21
WZ	0.41 ± 0.07
DY$\to ll$	4.44 ± 1.30
Fakes	3.66 ± 1.46
DY$\to \tau\tau$	0.77 ± 0.16
Total	27.5 ± 2.4

Table 3: Expected sample composition in $\int L = 335pb^{-1}$. The number of expected $t\bar{t}$ events assumes $\sigma_{t\bar{t}} = 6.7$ pb and $M_t = 175$ GeV.

7.1 DIL selection

The official DIL selection criteria [?] is applied to samples with varying top mass. Figure 21 shows the acceptance as a function of top mass, and a fit to the curve.

From preliminary estimates from the DIL group, we expect the sample composition shown in Table 3, which we use in all pseudo-experiments in the following sections.

8 Calibration of Response and Error

For the class of events which are well described by our model of top quark pair production and decay, the mass extraction technique is well behaved. It measures the mass without significant bias and with a well-estimated error [9].

The events in our sample are not restricted to this class of events. The response of the method is slightly degraded by the presense of poorly reconstructed jets, events with secondary leptons from $W \to \tau$ decays and background processes. Some of the factors that impair our ability to extract the true value of mass can be eliminated by proper modelling of corresponding probabilities. Unfortunately, not all of these factors yield resonable probability expressions. Thus, our expression of the posterior probability distribution of top-quark mass lacks terms for hard to model components. We explicitly correct for the incomplete nature of the method by shifting the extracted value of the top quark mass and increasing the statistical uncertainty of the measurement. In this section, we measure the response of the method and derive a calibration correction.

8.1 Signal only tests with P_s

We first measure the response of the method using purely signal events. The number of events in each pseudo-experiment is a random number, Poisson distributed with expected number of events shown in Table 3. Probability expression in this case includes only the term that describes leading order $t\bar{t}$ production. As shown in Section 4, the extracted mass demonstrates small bias in this case. Figure 22 shows that the response is linear, with a slope just below 1.0. The bias is attributed to the W decay into taus and to the events with one of the leading jets not coming directly from a $t\bar{t}$ decay. The statistical uncertainty is underestimated for the same reason. Figure 23 shows the residual and pull parameters of the bias-corrected top quark mass.

8.2 Signal and Background tests with P_s

We add to our pseudo-experiments events from the background processes following the expected sample composition from Table 3, while still using only the signal term in the probability expression. Figure 24 shows that the performance is degraded significantly as the background smears out the top mass information. The degradation is manifested by the decrease in the slope of the response curve. Reasons for it are quite obvious — none of the added backgrounds are properly described by signal-only probability expression.

[9] refer to Section 4.1

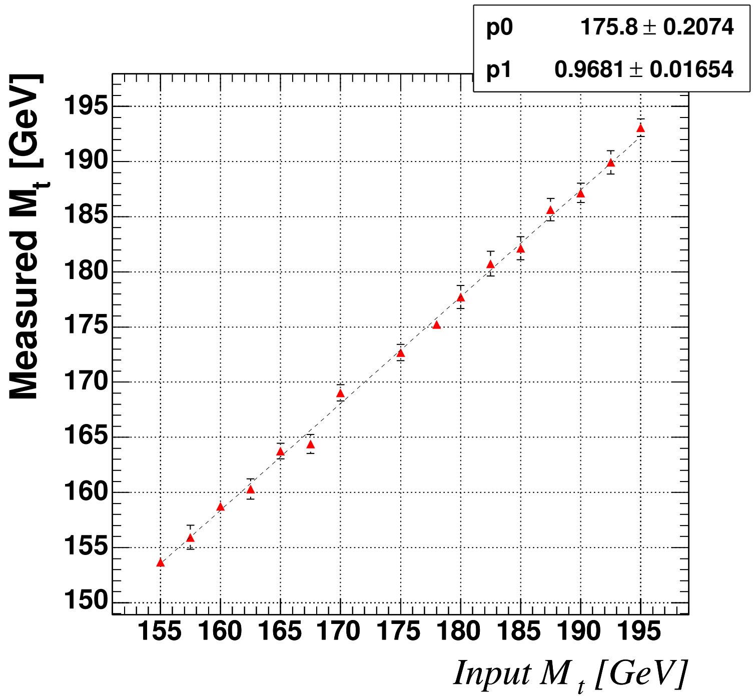

Figure 22: Response curve for pseudo-experiments of signal events, using only P_s to extract the mass.

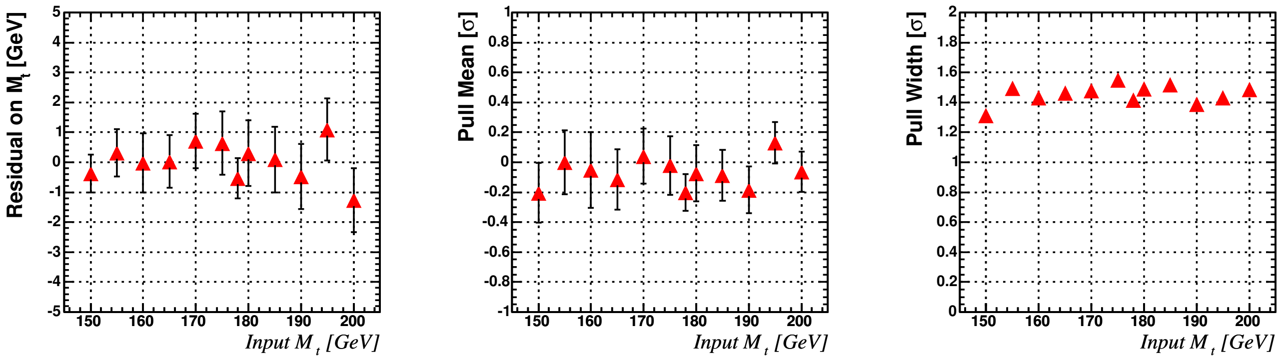

Figure 23: Residual in top mass, mean and width of pull distribution for varying simulated M_t, in signal-only pseudo-experiments.

8.3 Signal and Background tests with P_s and P_b

When we add to our likelihood machinery the description of the background processes in the form of P_b, we recover a significant amount of sensitivity to the top mass. It should

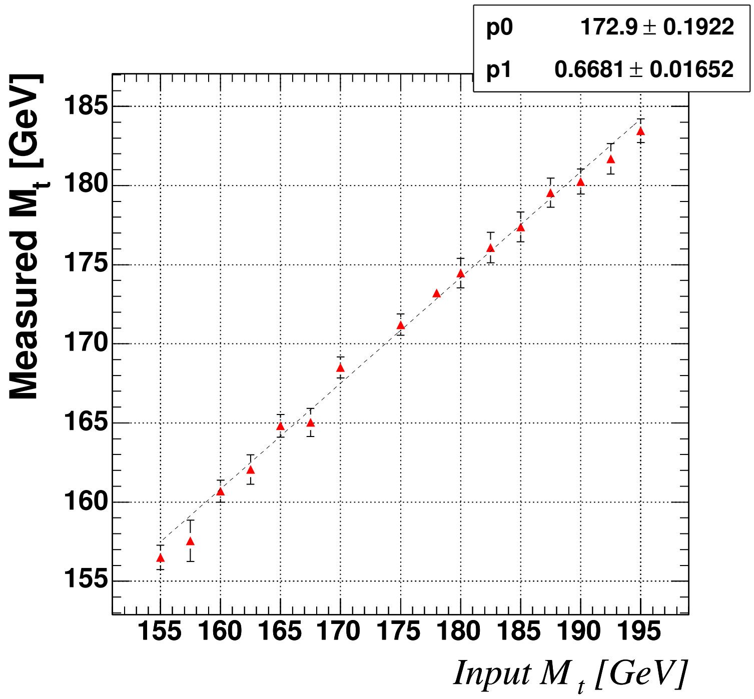

Figure 24: Response curve for pseudo-experiments of signal and background events, using only P_s to extract the mass.

be noted that the only backgrounds that are described are Z-boson production with two additional jets and W-boson pair production. This is the reason why recovered sensitivity is not on par with signal-only case. The generated pseudo-experiments contain backgrounds that are not described and thus the recovery is partial. Figure 25 shows the response with the full probability expression.

Table 8.3 gives the estimated statistical error as each background is added to the pseudo-experiments and likelihood. Again, the introduction of background events increases the estimated statistical uncertainty, while including the background terms into the probability expression we are able to recover part of the sensitivity. Figure 27 shows residuals and pull parameters of the slope-corrected and error-corrected top quark mass.

8.3.1 Derivation of Response correction

The necessary bias correction is based on the numbers presented in Figure 25. We transform the raw mass into the corrected mass using a linear correction of the form

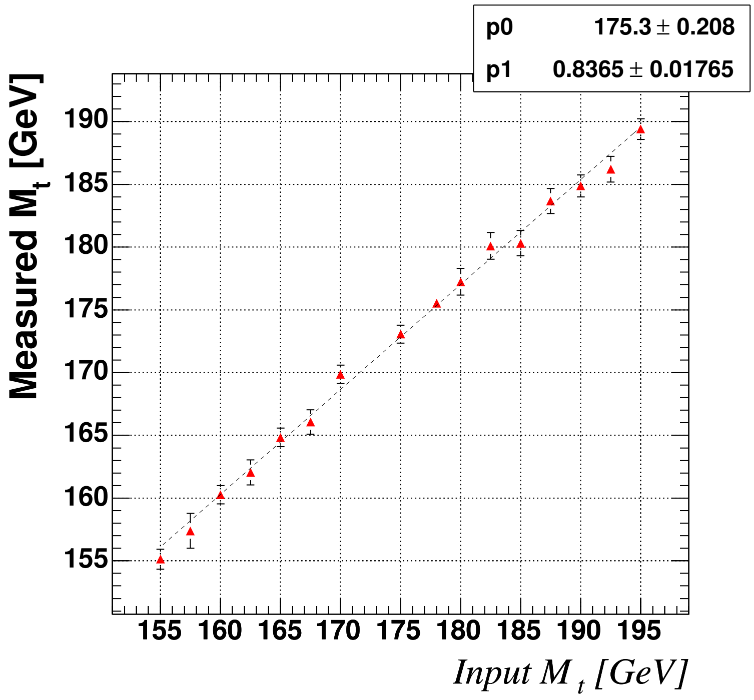

Figure 25: Response curve for pseudo-experiments of signal and background events, using P_s and P_b to extract the mass.

$$M_t = 178.0 + (M_0 - M_t^{raw}) * s$$

where $M_0 = 175.7 \pm 0.29$ and $s = 0.86 \pm 0.02$. The statisical error in this transformation is propogated to a systematic error in the final analysis.

8.3.2 Derivation of Error correction

A significant contribution to the underestimate of the error comes from the contribution of unmodelled background events. In Figure 26, we can see the effect on the sensitivity, measured error and fitted pull width as a function of the number of background events in pseudo-experiments. If the background probability is not included, the unmodelled background will increase the width of the extracted mass distribution but not the error, resulting in an underestimate of the actual error and a larger pull width. When the background probability is included, the width of the extracted mass is relatively flat, and the result pull widths are stable.

In Table 8.3, we can see the effect on the fitted pull width in pseudo-experiments with successively increase backgrounds. Backgrounds for which we have a model do not

Experiments	Likelihood	Response slope	Mass RMS
Signal	Signal	0.97	8.33 GeV
Signal+Z	Signal	0.83	9.43 GeV
Signal+Z	Signal+Z	0.91	8.57 GeV
Signal+Fakes	Signal	0.80	9.19 GeV
Signal+Fakes	Signal+Fake	0.95	8.41 GeV
Signal+WW	Signal	0.90	8.80 GeV
Signal+WW	Signal+WW	0.91	8.70 GeV
Signal+Z+WW+Fakes	Signal	0.68	10.58 GeV
Signal+Z+WW+Fakes	Signal+Z+WW+Fakes	0.85	9.12 GeV
Signal+Bg	Signal	0.67	10.79 GeV
Signal+Bg	Signal+Z+WW	**0.84**	**9.29** GeV

Table 4: Degradation of the error as background events are added to the simulated sample, and recovery of the sensitivity as the background probability is included. These estimates are obtained for the expected sample composition given by Table **??**.

Experiments	Likelihood	Fitted Pull Width
Sig	Sig	1.44 ± 0.03
Sig+Z	Sig+Z	1.40 ± 0.03
Sig+Z+WW	S+Z+WW	1.43 ± 0.03
Sig+Z+WW+Fakes	S+Z+WW+F	1.45 ± 0.03
Sig+Z+WW+Fakes+$Z \to \tau\tau$	S+Z+WW+F	1.46 ± 0.03
All	S+Z+WW+F	1.46 ± 0.03

Table 5: Effect on the fitted pull of modelled and unmodelled backgrounds.

increase the pull width, while unmodelled backgrounds, especially fakes, do widen the pull. In Table 8.3, we verify that the final pull width is not very sensitive to reasonable changes in the underlying model.

We derive a correction factor by measuring the average understimate, see Figure 27. We measure

$$S_{error} = 1.50$$

Sample	Fitted Pull Width
Herwig	1.49 ± 0.03
Pythia	1.54 ± 0.03
FSR less	1.53 ± 0.03
FSR more	1.53 ± 0.03
ISR less	1.52 ± 0.03
ISR more	1.53 ± 0.03

Table 6: Effect on the fitted pull variation of the signal sample

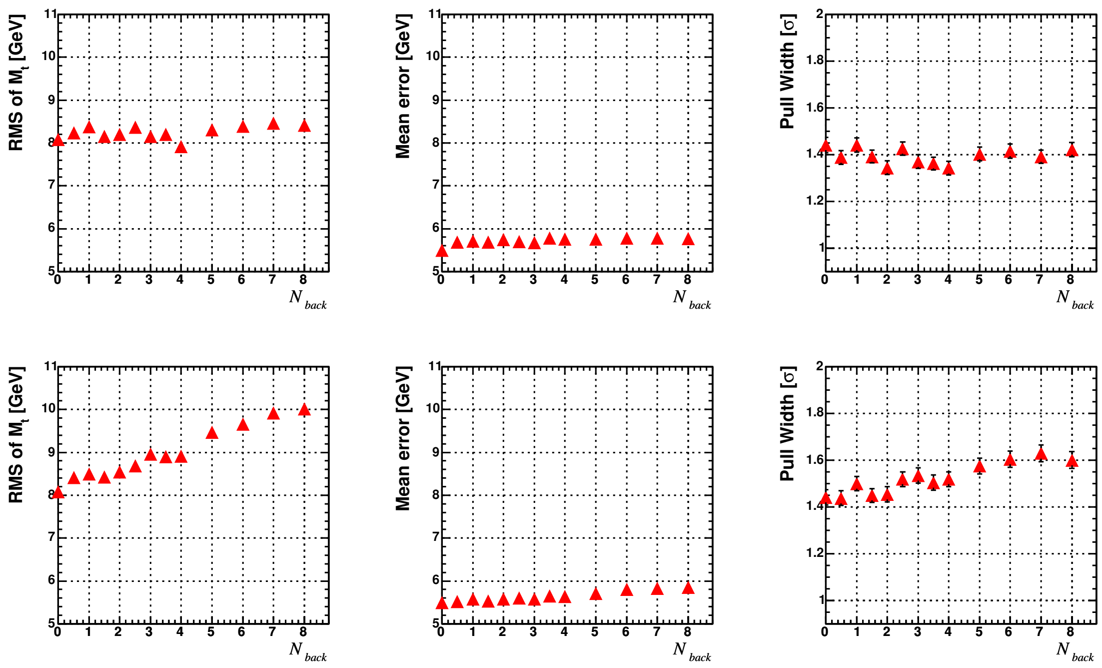

Figure 26: RMS of mass distribution, measured error and fitted pull width as a function of the number of expected Z background events with (*top*) and without (*bottom*) background probability.

and apply this to all samples independent of M_t. The final corrected and scaled performance as a function of M_t can be seen in Figure 28 and specifically for $M_t = 178$ in Figure 29.

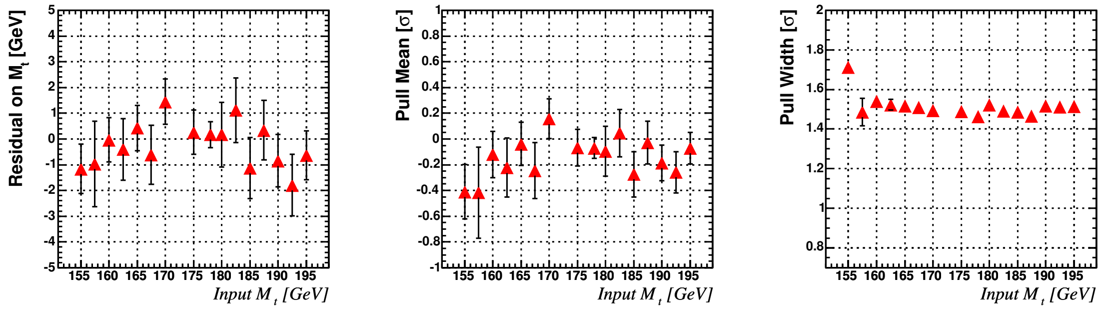

Figure 27: Residual, pull mean and pull width before error scaling, for varying top mass samples using the DIL selection

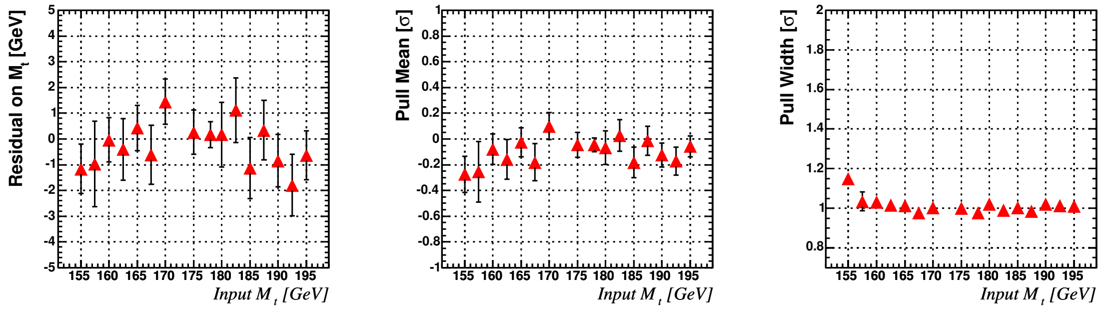

Figure 28: Residual, pull mean and pull width after all corrections for varying top mass samples using the DIL selection

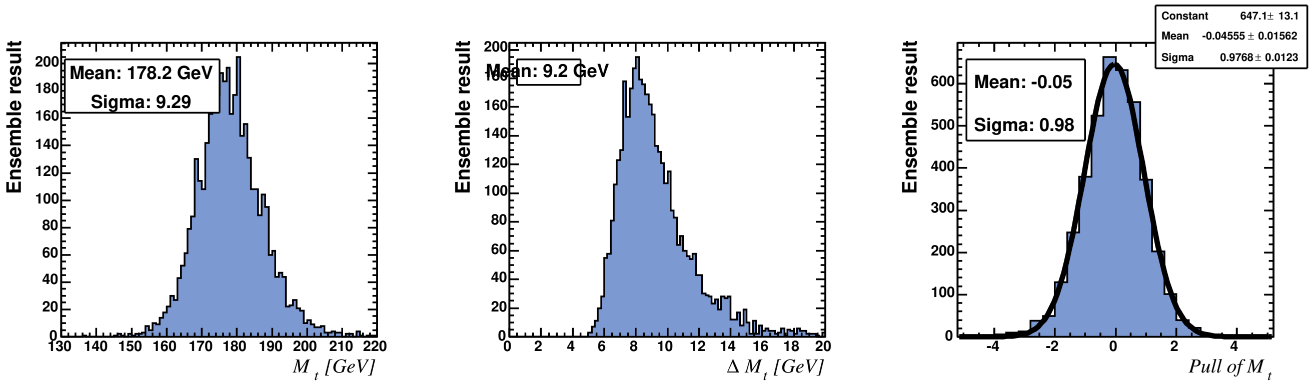

Figure 29: Extracted mass, error and pull for DIL selection and $M_t = 178$ GeV

9 Systematic Errors

Our measurement is calibrated using monte carlo events; the majority of systematic errors therefore come from our uncertainty in the accuracy of the simulation in modelling the data. In this Section, we describe the facets of the simulation which may not accurately describe the observed data and estimate the sensitivity of our measurement.

9.1 Jet Response

The measured energy of jets is calibrated using simulated and data events. Due to lack of knowlegde of the underlying processes in hadronization and limited data statistics, there is an uncertainty in the correction of the raw jet energy.

We vary the response of the jets in simulated events by one unit of the uncertainty in the response, and measure the difference in our extracted top mass.

Sample	Measured M_t for $M_t^{true} = 178.0$
JES $+1\sigma$	180.91 ± 1.26 GeV
JES -1σ	175.66 ± 1.29 GeV

We take half of the difference between the two measurements, $\Delta_{JES} = 2.62$ GeV.

9.2 Generator

We measure the difference in our extracted mass when using events simulated with Pythia and with Herwig.

Sample	Measured M_t for $M_t^{true} = 178.0$
Herwig	178.1 ± 0.65 GeV
Pythia	177.5 ± 0.70 GeV

We take the difference between the two measurements, $\Delta_{Gen} = 0.6 \pm 0.95$ GeV.

9.3 Response calibration

The final result depends on the correction of our response curve. Uncertainty in this response translated directly into uncertainty in our result. We vary the response by 1 sigma within the statistical errors of the fit, and measure the difference in our extracted mass.

Sample	Variation of (offset,slope) $(+\sigma,+\sigma)$	$(-\sigma,-\sigma)$	$(+\sigma,-\sigma)$	$(-\sigma,+\sigma)$
$M_t = 170$	171.3	171.6	171.1	172.5
$M_t = 178$	178.0	178.4	177.9	178.4
$M_t = 185$	183.5	184.2	183.7	184.0

In order to be conservative, we take half of the difference between the pair with the largest variation, $\Delta_{Resp} = 0.6$ GeV.

9.4 Sample composition

We require a knowledge of the expected signal and background composition of the sample in order to construct pseudo-experiments from which we extract our response curve. Given our response curve, we vary the background composition within the quoted uncertainties and measure the difference in the mean extracted mass of an ensemble of pseudo-experiments.

Varied Sample	$+\sigma$	$-\sigma$	Δ
$t\bar{t}$	177.7	177.8	0.1
$Fakes$	178.0	177.2	0.4
Z	177.3	177.9	0.3
WW	177.5	178.1	0.3
WZ	178.1	177.9	0.2
$Z \to \tau\tau$	178.1	177.6	0.3

We add these contributions in quadrature to measured $\Delta_{Sample} = 0.7$ GeV.

9.5 PDF uncertainties and α_s

To estimate the error due to the error on the PDFs, we reweight the Pythia sample according to 20 sets of positive and negative eigenvectors. We modify our prescription for forming pseudo-experiments to allow for event weights, in which events with higher weight are more likely to be chosen for a given pseudo-experiment. Figure **??** shows the difference between the positive and negative eigenvectors. The total error added in quadrature is $\Delta_{PDF} = 0.53$ GeV.

In addition, we compare samples with different PDFs, CTEQGL and MRST72 and confirm that their central values agree well.

Sample	M_t
Pythia, CTEQ5L	178.32 ± 0.70 GeV
Pythia, MRST72	178.6 ± 0.83 GeV
Pythia, MRST75	177.3 ± 1.08 GeV

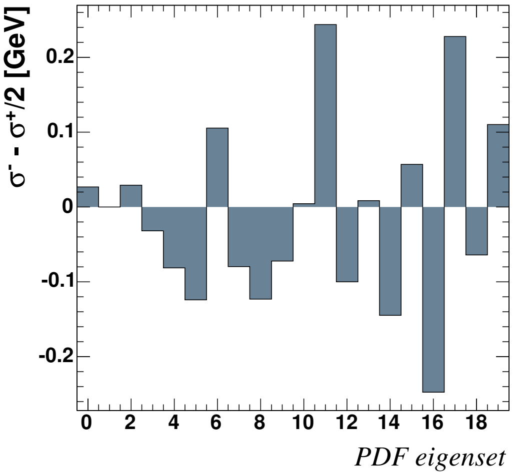

Figure 30: Difference in extracted mass between the positive and negative PDF eigenvectors for a Pythia sample.

We estimate our systematic errors due to variation of α_s to be $\Delta_{\alpha_s} = 1.3 \pm 1.36$ GeV.

9.6 Radiation

Our model does not contain any description of initial (ISR) or final state radiation (FSR), which may contribute significantly to the fraction of mismeasured events. ISR and FSR are poorly understood and there are significant uncertainties in their modelling. We vary the parameters controlling ISR and FSR production in Pythia simulation and measure the difference in the extracted top mass.

Sample	M_t
Pythia	177.5 ± 0.70 GeV
Pythia, ISR less	177.5 ± 0.68 GeV
Pythia, ISR more	178.4 ± 0.68 GeV
Pythia, FSR less	177.7 ± 0.82 GeV
Pythia, FSR more	178.1 ± 0.65 GeV

We therefore estimate the systematic error due to modelling of radiation to be $\Delta_{ISR} = 0.46 \pm 0.48$ GeV, and $\Delta_{FSR} = 0.15 \pm 0.52$ GeV.

9.7 Gluon contribution

The matrix element quoted earlier describes the production of $t\bar{t}$ pairs via qq annhilation. At the TeVatron, up to 15% of the $t\bar{t}$ pairs may be produced from gg anhilation. Our leading-order generators, Pythia and Herwig, contain only $\approx 5\% gg$ events. To estimate our sensitivity to this fraction, we vary the composition from 0 to 15% in the Herwig sample.

gg Fraction	M_t
0%	177.45 ± 0.85 GeV
5%	177.43 ± 0.85 GeV
15%	177.73 ± 1.27 GeV

we take as a systematic error, $\Delta_{gg} = 0.30$ GeV.

9.8 Background statistics

To measure our dependence on the relatively small background samples, we split each background sample into twenty pairs of disjoint sets. We measure the mass in the $M_t = 178$ GeV sample for each of the disjoint sets and take the average difference between them as an estimate of the error.

Background	$\overline{\Delta M_t}$
Z	0.90 GeV
Fakes	1.00 GeV
WW	0.39 GeV
WZ	0.16 GeV
$Z \to \tau\tau$	0.38 GeV

We take the sum in quadrature, $\Delta_{BGstat} = 1.5$ GeV.

9.9 Background Modelling

The modelling of the two largest backgrounds may be imperfect; we estimate our sensitivity to possible imperfections in these background models.

9.9.1 Drell-Yan

Drell-Yan production is difficult to model well, as it has a large cross-section and a small acceptance. Since there is no real missing energy due to escaping neutrinos, there is a small probability to pass the missing energy requirement. One might be concerned whether the modelling of large missing energy on the tails of these distribu-

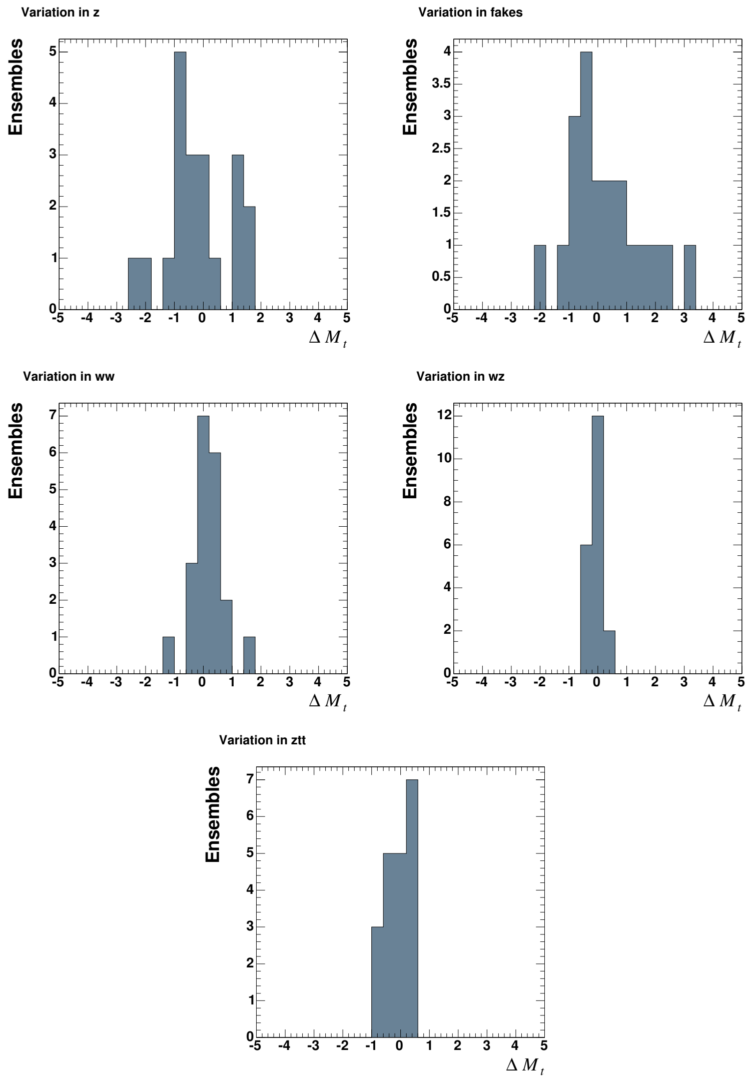

Figure 31: Variation in measured mass when background samples are split into disjoint sets. Top left, Z; Top right, Fakes. Middle left, WW; Middle right, WZ. Bottom, $Z \to \tau\tau$.

tions is accurate. Figures 32-34 show the distribution of missing energy in data and in simulation for events with two leptons in the Z window. In each case, the simulation describes the shape of the events seen in data over a range of missing energy. One would be particulary concerned if there were a considerable number of data events on the tails; no such effect is apparant.

Prescription I

To gauge our sensitivity to events on the tails of the distribution, we vary the composition of our pseudo-experiments by both enhancing events and suppressing on the tails and deweighting events. We assign a weight to each event as

$$w_+ = 1 + \frac{1}{\sqrt{n_+}}$$
$$w_0 = 1$$
$$w_- = 1 - \frac{1}{\sqrt{n_+}}$$

where n_+ represents the number of events further out on the tail of the missing energy distribution than the given event. The event with the largest missing energy will receive weights ($w_+ = 2, w_0 = 1, w_- = 0$). This enhances or suppresses the tail events.

Weights	M_t
w_+	177.8 ± 0.4 GeV
w_0	178.1 ± 0.4 GeV
w_-	177.7 ± 0.4 GeV

Prescription II

We use a second prescription to enhance and suppress events with large missing energy. We assign weights to the events proportional to the measured missing energy, and to the inverse of the measured missing energy:

$$w_+ \propto ME_T$$
$$w_0 = 1$$
$$w_- \propto ME_T^{-1}$$

The event with the largest missing energy will receive large weights for w_+ and small weights for w_-. This enhances or suppresses the tail events.

Weights	M_t
w_+	177.8 ± 0.4 GeV
w_0	178.1 ± 0.4 GeV
w_-	177.7 ± 0.4 GeV

The effect is nearly identical in both prescriptions, and within the statistical error. As demonstrated previously, the method successfully suppresses a good portion of

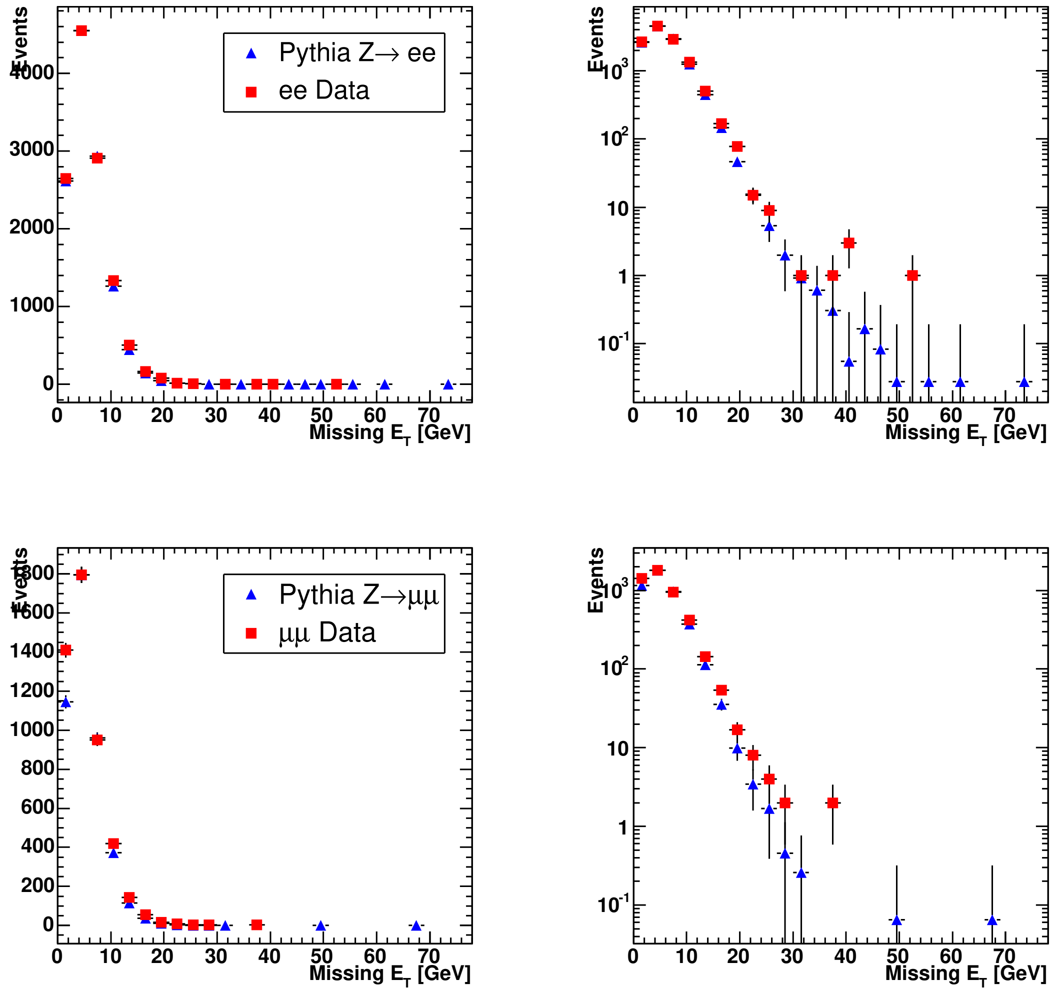

Figure 32: Missing transverse energy in Z+0jet events, where the Z decays to electrons (*top*) or muons (*bottom*).

contribution of Zjj events to the extracted mass; this makes the result less sensitive to variations in the background shape.

We take as the error due to the shape of this background $\Delta^{Z}_{Shape} = 0.4 \pm 0.6$ GeV.

9.9.2 Fake background

The fake background is very difficult to model well in the simulation, as it is sensitive to the smallest details of the detector performance. To avoid issues of modelling, the events which model this background are drawn from the data itself; the events are

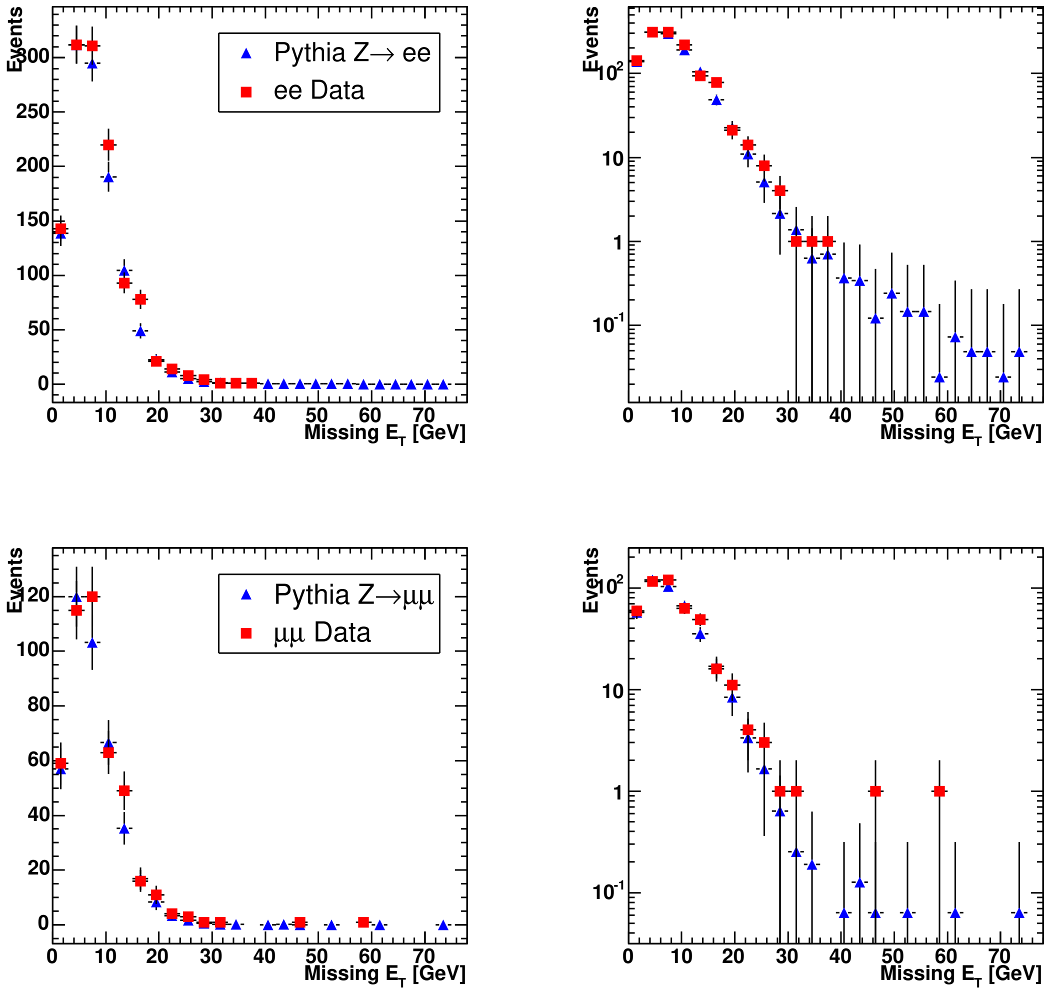

Figure 33: Missing transverse energy in Z+1jet events, where the Z decays to electrons (*top*) or muons (*bottom*).

selected with a looser requirement on one of the leptons in order to accumulate a sample of fake candidates, and then weighted by the probability that the loose lepton would pass lepton identification requirements. These weights are calculated as a function of the P_T and isolation of the fake candidate, and each candidate has its own weight and errors.

To gauge our sensitivity to the calculation of the fake rates, we vary the fake rates in two ways. First (a), we enhance those events with large fake rate, to exaggerate their effect; second (b), we enhance events with small fake rate, to exaggerate their effect:

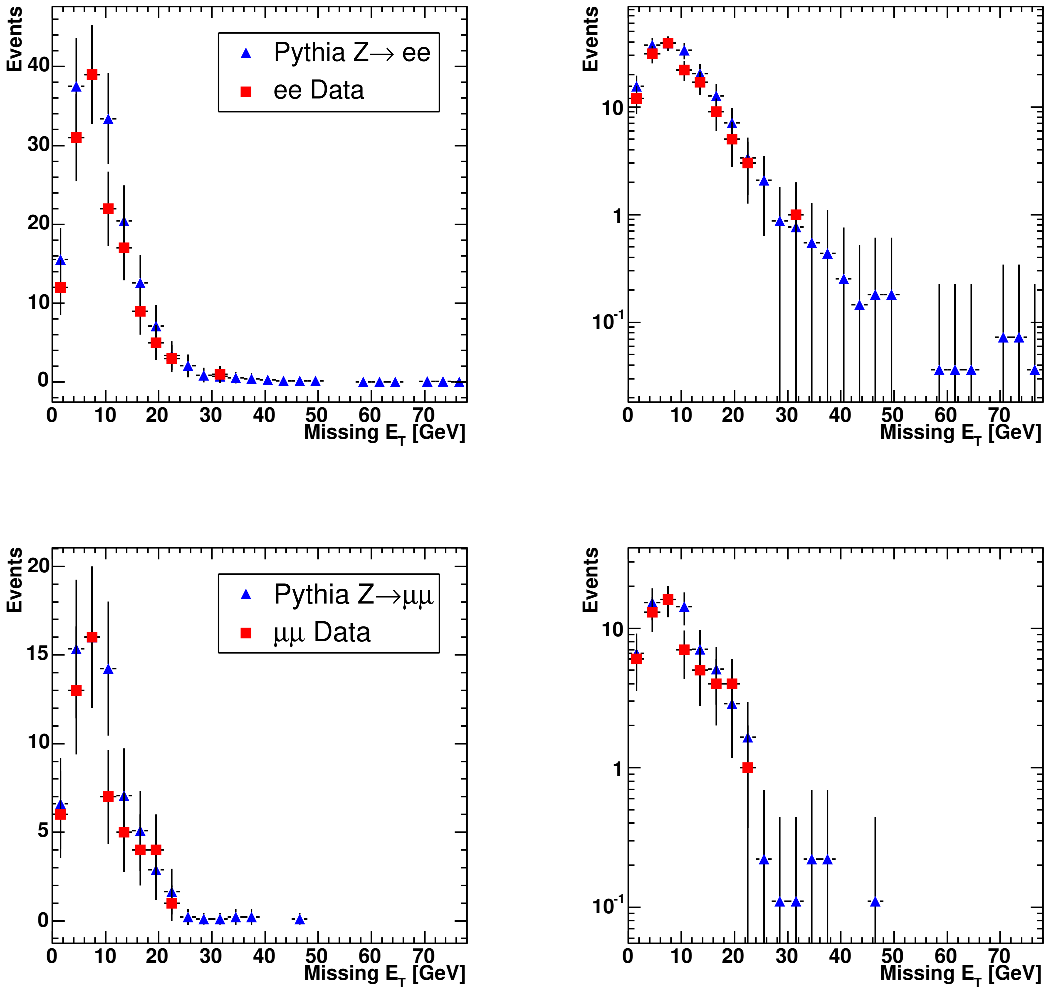

Figure 34: Missing transverse energy in Z+2jet events, where the Z decays to electrons (*top*) or muons (*bottom*).

Mode	$w > \overline{w}$	$w < \overline{w}$
a	$w \to w + \Delta w$	$w \to w - \Delta w$
b	$w \to w - \Delta w$	$w \to w + \Delta w$

We study the effect of this weighting on fully realistic pseudo-experiments:

Weights	M_t
w_a	178.2 ± 0.4 GeV
default	178.1 ± 0.4 GeV
w_b	177.6 ± 0.4 GeV

Source	Size
Jet Energy Scale	2.62 GeV
Generator	0.52 ± 0.96 GeV
Response uncertainty	0.6 GeV
Sample composition uncertainty	0.7 GeV
Background MC	1.5 GeV
Background modelling	0.7 ± 0.8 GeV
FSR modelling	0.15 ± 0.52 GeV
ISR modelling	0.46 ± 0.48 GeV
PDFs	0.53 GeV
α_s	1.3 ± 1.36 GeV
Gluon fraction	0.3 GeV
Total	3.7 GeV

Table 7: Summary of systematic errors

As argued above, the method is reasonably successful and reducing the effect of these background events, which makes the extractes mass less sensitive to their shapes.

We take the difference as our systematic error, $\Delta_{Shape}^{fakes} = 0.6 \pm 0.6$ GeV, and the total shape error to be the two in quadrature

$$\Delta_{Shape} = 0.7 \pm 0.8 GeV.$$

9.10 Summary

The systematic errors measured in the previous sections are summarized in Table 9.10. Our total systematic uncertainty is 3.70 GeV.

10 Data

The collected sample of events in $\int L = 335pb^{-1}$ contains 33 candidate dilepton events. Their run, event numbers and a summary of their kinematic properties can be found in Appendix D, Table 8.

10.1 Kinematic Properties of Observed Events

We examine the events observed in the data and compare their kinematic properties to that predicted from the simulation, with all expected contributions summed as expected. We note first that the number of observed events is higher than expected, 33 observed to 25 expected. Given the relatively small statistics, this is not an unlikely Poisson fluctuation. Examining the kinematic distributions in Figures 40-53 does not reveal any dramatic variation in shape for any variable.

In addition, we compare the distribution of $t\bar{t}$, Zjj, Fake and $WWjj$ probabilities for the observed and expected events, see Figures 35-38. These quantities are sensitive to all of the kinematic information in the event, and show no significant shape discrepancies. They all appear consistent with a positive fluctuation of events.

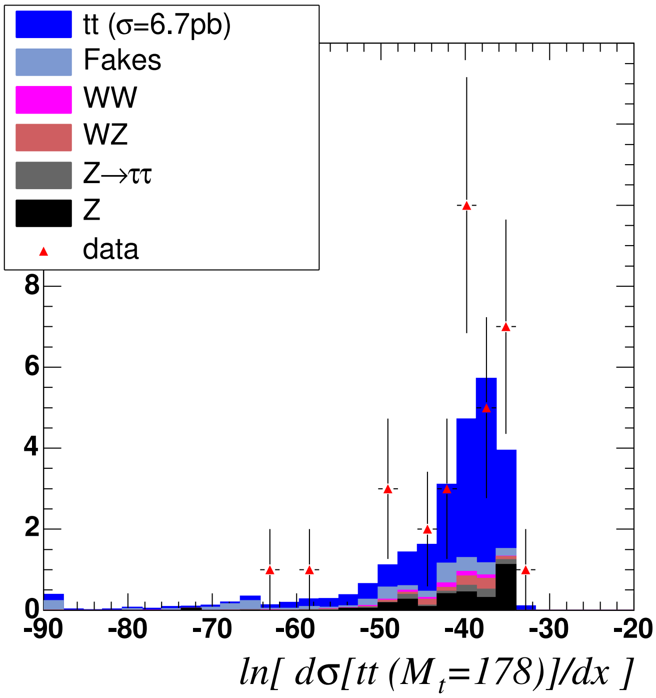

Figure 35: Evaluation of the $t\bar{t}$ probability for simulated events, scaled by expected contribution to the sample composition, and for observed events.

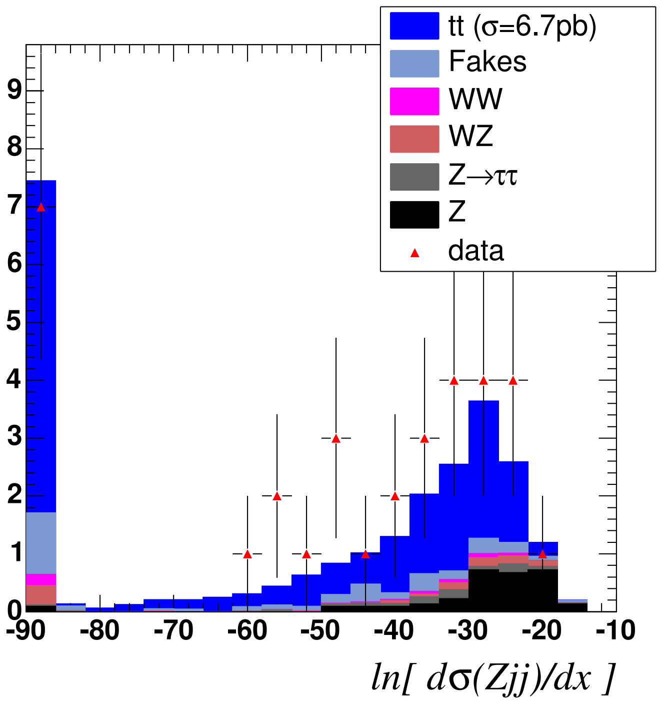

Figure 36: Evaluation of the Zjj probability for simulated events, scaled by expected contribution to the sample composition, and for observed events.

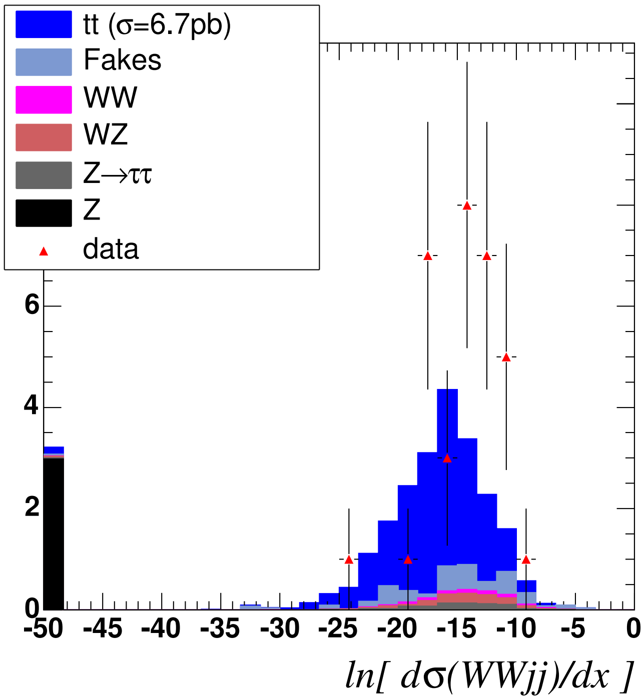

Figure 37: Evaluation of the $WWjj$ probability for simulated events, scaled by expected contribution to the sample composition, and for observed events.

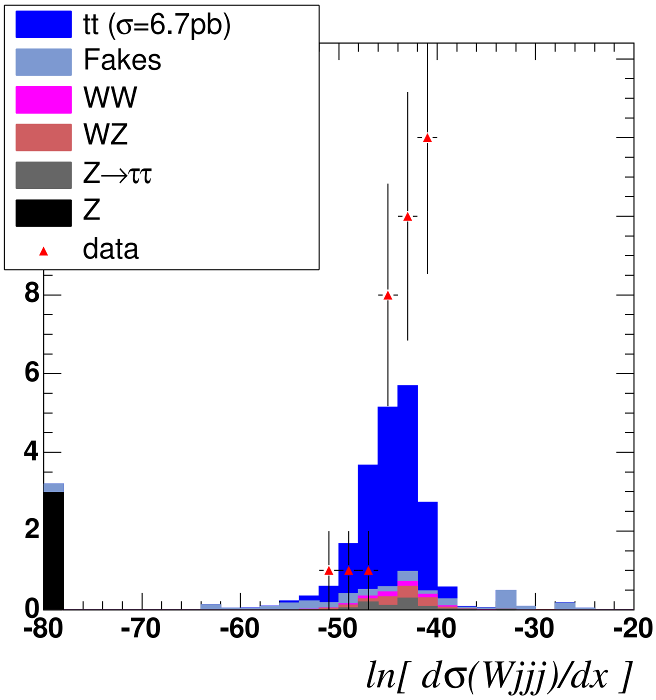

Figure 38: Evaluation of the Fake probability for simulated events, scaled by expected contribution to the sample composition, and for observed events.

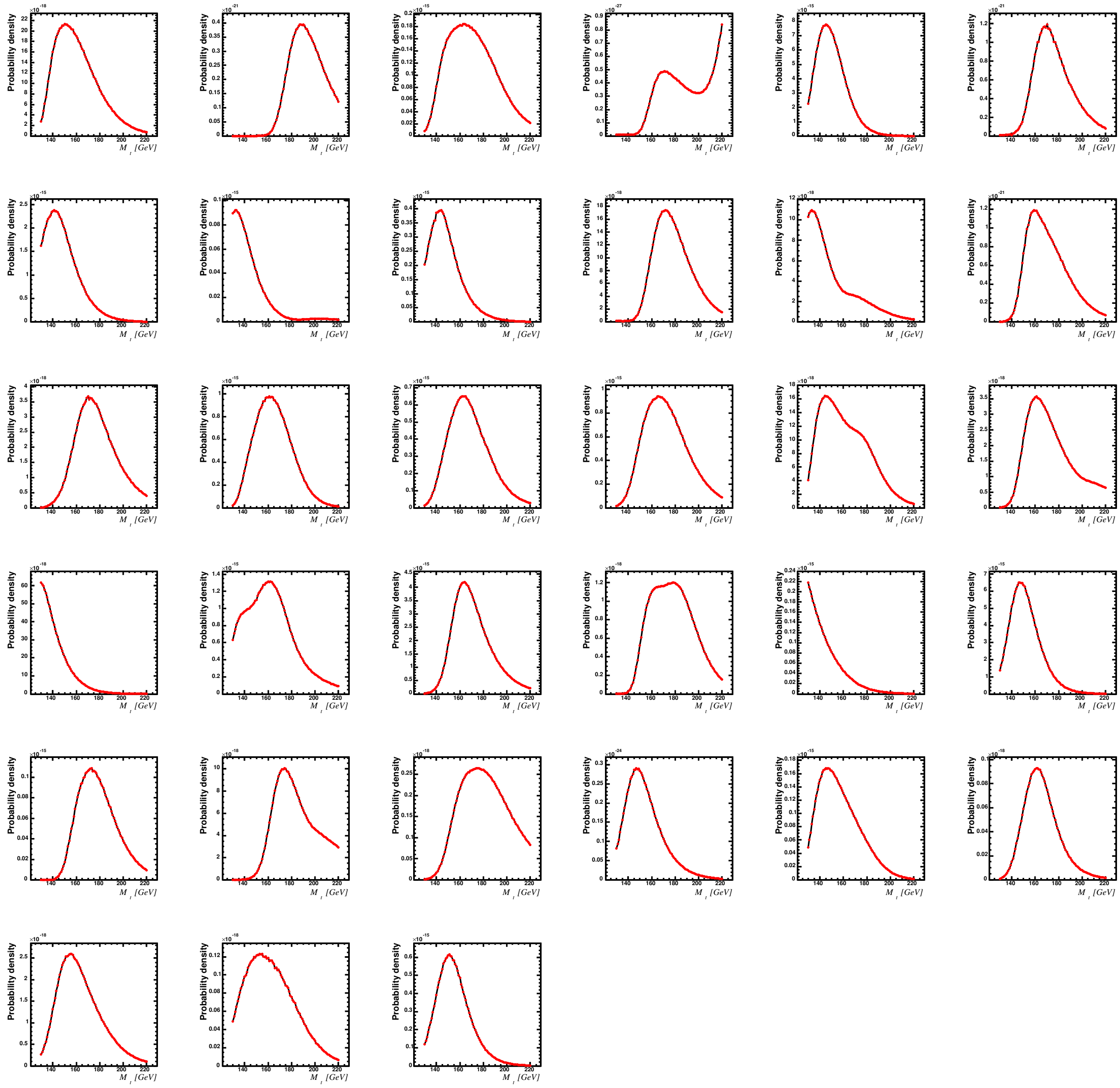

Figure 39: Unnormalized and uncorrected $d\sigma_{t\bar{t}}(M_t)/dx$ for the 33 candidate events.

Appendices

Appendix A: Invariant mass.

When a process has an intermediate particle, the integration over the phase space of the event can be simplified by splitting the phase space and introducing the integration over invariant mass of the decay products. We start out with:

$$\delta^4(P - \sum_{i=1}^{n} p_i) \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i} \tag{76}$$

— the phase space for the process. The following identity is very useful in reducing number of integrations. For final state particles, all m_i are equal to their real masses.

$$\delta(p_i^2 - m_i^2)\frac{d^4 p_i}{(2\pi)^3} = \frac{d^3 p_i}{(2\pi)^3 2E_i} \tag{77}$$

We will try to split phase space by introducing intermediate particle with momentum

$$q = \sum_{i=1}^{j} p_i \tag{78}$$

The phase space then takes the form:

$$(2\pi)^3 \delta^4(P - q - \sum_{i=j+1}^{n} p_i) \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i} \; \delta^4(q - \sum_{i=1}^{j} p_i) \; \frac{d^4 q}{(2\pi)^3} \tag{79}$$

Change from integration over q^0 to integration over M_q^2 is done by introducing

$$\delta(q^2 - M_q^2) \; dM_q^2 \tag{80}$$

and integrating over q^0 using equation (77). The following equation:

$$\delta^4(P - q - \sum_{i=j+1}^{n} p_i) \; \delta^4(q - \sum_{i=1}^{j} p_i) \; \delta(q^2 - M_q^2) \; \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i} \; \frac{d^4 q}{(2\pi)^3} \; (2\pi)^3 dM_q^2 \tag{81}$$

becomes:

$$\delta^4(P - q - \sum_{i=j+1}^{n} p_i) \; \delta^4(q - \sum_{i=1}^{j} p_i) \; \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i} \; \frac{d^3 q}{(2\pi)^3 2E_q} \; (2\pi)^3 dM_q^2 \tag{82}$$

or to make it look exactly as [2]:

$$\left[\delta^4(P - q - \sum_{i=j+1}^{n} p_i) \prod_{i=j+1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i} \; \frac{d^3 q}{(2\pi)^3 2E_q}\right] \left[\delta^4(q - \sum_{i=1}^{j} p_i) \prod_{i=j}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i}\right] (2\pi)^3 dM_q^2 \tag{83}$$

The point of the exercise in algebra of Dirac delta-functions and integration was partially to show that the cited formula from [2] makes sense and at the same time to emphasize the following fact — *additional integration over q three-momentum has to be performed.* Which is natural since introduction of a four-dimensional delta-function results in necessary integrations over three momentum components and invariant mass.

Appendix B: computation

The computer code that implements the analysis technique outlined in this note relies on numerical computation, namely Monte Carlo integration and finding roots of a polynomial. Monte Carlo integration was done using GNU Scientific Library [4]. Operations on Lorentz vectors were done using CLHEP library [5]. Polynomial root finder was implemented using code from [**?**].

The Monte Carlo integration algorithm VEGAS [6] is used to do the multi-dimensional integrals. Adaptive techniques employed by VEGAS adjust the integration grid using the histogram projections on the axes. When doing Monte Carlo integration with VEGAS one has to be aware of the number of potential pitfalls

- certain choices of integration variables lead to hard to integrate functions.
- grid adjustment rate has to be reduced for "feature-rich" functions.
- integrals of the Breit-Wigner distributions depend on the choice of integration limits.

Searching for solutions of the coupled system of quadratic equations or the equivalent fourth-order polynomial we found the following:

- although the analytic solution of the system can be obtained it was found to perform slower and be less forgiving with respect to roundoff errors.

Appendix C: analytic neutrino solutions

Analytic solutions to the system of bi-quadratic equations (30) were found using Mathematica system.

$$\begin{aligned} p_1x^2 + p_2xy + p_3y^2 + p_4x + p_5y + p_6 = 0 \\ q_1x^2 + q_2xy + q_3y^2 + q_4x + q_5y + q_6 = 0 \end{aligned} \tag{84}$$

As a first step we find the most generic solution for $y(x)$ with y expressed in terms of x:

$$y = \frac{-(p_6q_3) + p_3q_6 - p_4q_3x + p_3q_4x - p_1q_3x^2}{p_5q_3 - p_3q_5 - p_3q_2x + p_2q_3x} \tag{85}$$

For this equation on y to be obtainable the 4-th order equation on x must be satisfied:

$$x^4 + \frac{k_3}{k_4}x^3 + \frac{k_2}{k_4}x^2 + \frac{k_1}{k_4}x + \frac{k_0}{k_4} = 0 \tag{86}$$

where:

$$\begin{aligned}
&k_0 = (p_6^2q_3^2 + q_6(p_5^2q_3 - p_3p_5q_5 + p_3^2q_6) + p_6(-(p_5q_3q_5) + p_3(q_5^2 - 2q_3q_6)))/q_3^2 \\
&k_1 = (2p_4p_6q_3^2 + p_5^2q_3q_4 - 2p_3p_6q_3q_4 + 2p_3p_6q_2q_5 - p_2p_6q_3q_5 + p_3p_4q_5^2 + \\
&\quad p_3(-2p_4q_3 + 2p_3q_4 - p_2q_5)q_6 - p_5(p_6q_2q_3 + p_4q_3q_5 + p_3q_4q_5 + p_3q_2q_6 - 2p_2q_3q_6))/q_3^2 \\
&k_2 = (p_3^2q_4^2 + q_3(-(p_2p_6q_2) + p_4^2q_3 + 2p_1p_6q_3 + 2p_2p_5q_4 - \\
&\quad p_1p_5q_5 - p_4(p_5q_2 + p_2q_5) + p_2^2q_6) + p_3(p_6q_2^2 - p_5q_2q_4 - 2p_4q_3q_4 + 2p_4q_2q_5 - p_2q_4q_5 + \\
&\quad p_1q_5^2 - p_2q_2q_6 - 2p_1q_3q_6))/q_3^2 \\
&k_3 = (p_3(p_4q_2^2 - p_2q_2q_4 - 2p_1q_3q_4 + 2p_1q_2q_5) + q_3(-(p_1p_5q_2) + 2p_1p_4q_3 \\
&\quad + p_2^2q_4 - p_2(p_4q_2 + p_1q_5)))/q_3^2 \\
&k_4 = (p_1(p_3q_2^2 + q_3(-(p_2q_2) + p_1q_3)))/q_3^2
\end{aligned} \tag{87}$$

In terms of $a_0 = \frac{k_0}{k_4}$, $a_1 = \frac{k_1}{k_4}$, $a_2 = \frac{k_2}{k_4}$ and $a_3 = \frac{k_3}{k_4}$ the four solutions to the quartic equation in x are:

$$x_{1,2,3,4} = \left(-\frac{a_3}{4} \mp \frac{A}{4\sqrt{3}} \mp \frac{\sqrt{\left(-8a_2 + 3a_3^2 - \frac{(2\sqrt[3]{2}B_1)}{\Gamma} - 2^{\frac{2}{3}}\Gamma + \frac{3\sqrt{3}(8a_1 - 4a_2a_3 + a_3^3)}{A}\right)}}{2\sqrt{6}} \right) \tag{88}$$

with

$$
\begin{aligned}
\Gamma &= \sqrt[3]{B_2 + \sqrt{-4B_1^3 + B_2^2}} \\
A &= \sqrt{-8a_2 + 3a_3^2 + \frac{4\sqrt[3]{2}B_1}{\Gamma} + 2^{\frac{5}{3}}\Gamma} \\
B_1 &= (12a_0 + a_2^2 - 3a_1a_3) \\
B_2 &= (27a_1^2 - 72a_0a_2 + 2a_2^3 - 9a_1a_2a_3 + 27a_0a_3^2)
\end{aligned} \tag{89}
$$

Run	Event	P_T of reconstructed object [GeV] Jet 1	Jet 2	Lepton 1	Lepton 2	$\mathrm{M}E_T$
143257	760520	27.29	20.96	43.87	36.79	92.24
150418	960369	89.48	87.01	30.67	25.99	68.66
150431	368759	58.27	36.19	54.85	32.56	44.07
150435	2896171	45.37	42.88	88.21	44.22	64.71
151978	507773	35.89	34.72	47.04	45.7	91.35
153325	599511	37.11	25.23	50.78	34.32	46.47
153374	2276742	72.02	56.87	85.76	55.15	58.14
153447	2643751	87.45	26.3	106.4	20.98	35.52
154654	7344016	59.42	52.86	36.43	19.96	56.66
155114	478702	34.48	28.47	73.59	36.02	89.96
156484	3099305	35.17	34.31	66.91	34.22	75.91
160988	385505	26.05	21.02	87.15	56.91	32.19
161633	963604	40.24	26.49	62.27	36.36	60.52
162820	7050764	33.46	26.24	65.95	52.41	95.84
163012	1438203	119.1	98.54	103.8	73.67	80.55
165198	1827962	34.48	23.82	63.08	31.83	101.4
165364	592961	47.35	21.59	78.9	47.31	127.9
166063	2833132	41.36	38.04	49.73	49.62	64.97
167053	12011678	75.77	50.43	86.61	26.31	28.47
167629	180103	96.57	78.59	77.06	35.49	28.4
167631	2058969	33.97	31.34	38.33	36.65	61.32
168599	2964061	40.76	23.68	35.7	33.45	96.85
177491	3807306	25.94	21.03	67.44	42.8	79.22
178540	2208375	64.42	34.77	96.19	42.62	53.59
178738	10340757	50.05	21.39	75.71	36.06	88.4
178738	1660363	99.99	60.59	42.67	36.52	90.26
183963	1259645	84.72	27.04	142.1	28.9	63.14
184779	892809	51.1	44.26	46.05	20.01	48.86
185037	2287335	85.73	56.09	65.78	60.03	26.92
185377	103906	76	20.87	69.53	26.34	80.59
185594	3002817	80.5	26.08	139	22.28	199.2
186598	1618142	69.41	35.75	70.19	68.91	66.96
186598	4194951	54.5	36.05	90.35	58.62	42.65

Table 8: Run and event numbers and some kinematic quantities for candidate dilepton events.

Appendix D: Kinematic Details of Events

The run, event numbers and a summary of their kinematic properties can be found in Table 8. Distribution of kinematic variables can be found in Figures 40-53.

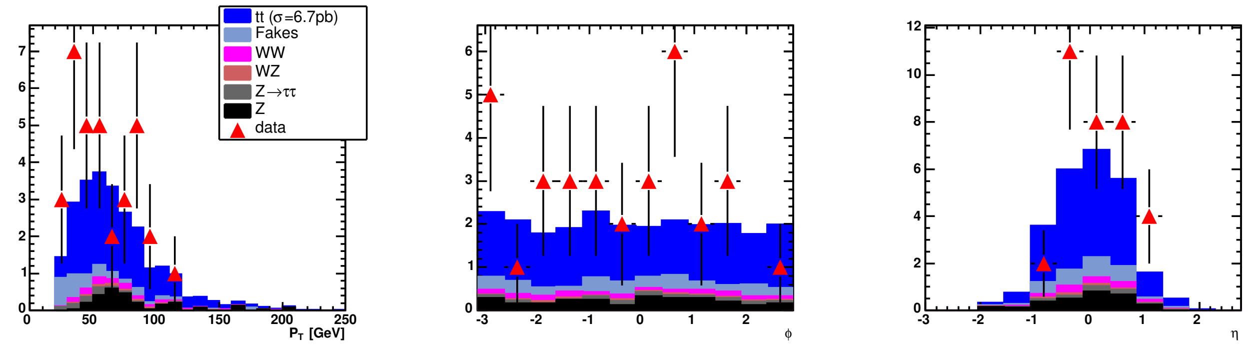

Figure 40: Transverse momentum (P_T), azimuthal angle (ϕ) and pseudo-rapidity (η) of the leading jet.

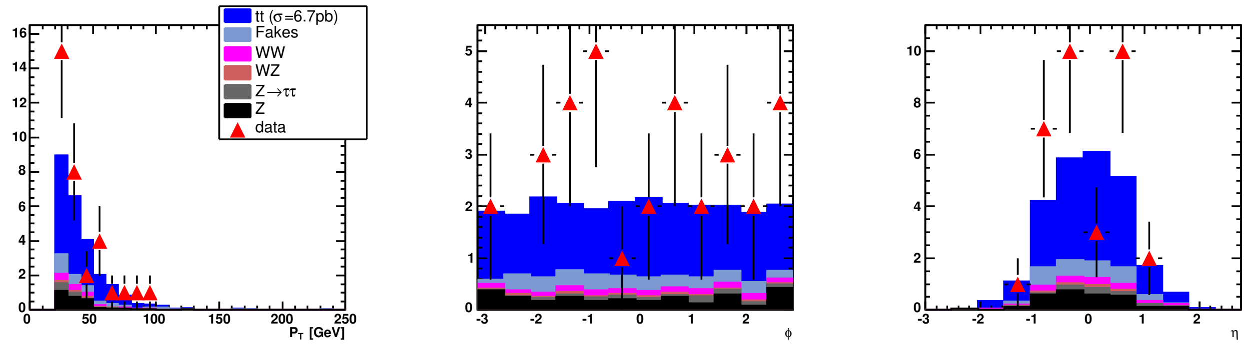

Figure 41: Transverse momentum (P_T), azimuthal angle (ϕ) and pseudo-rapidity (η) of the second jet.

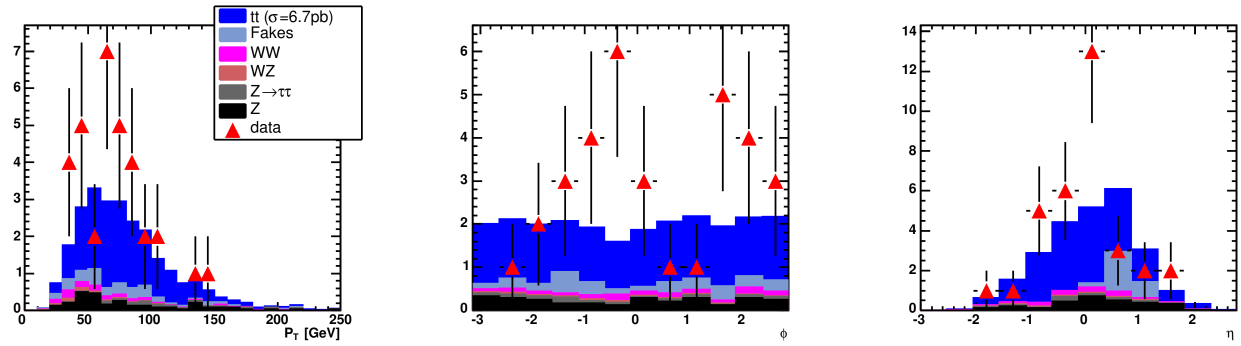

Figure 42: Transverse momentum (P_T), azimuthal angle (ϕ) and pseudo-rapidity (η) of the leading lepton.

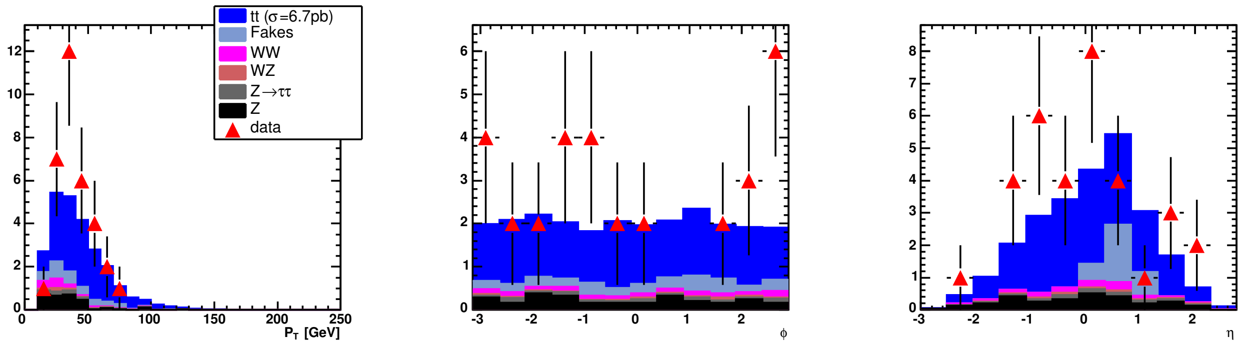

Figure 43: Transverse momentum (P_T), azimuthal angle (ϕ) and pseudo-rapidity (η) of the second lepton.

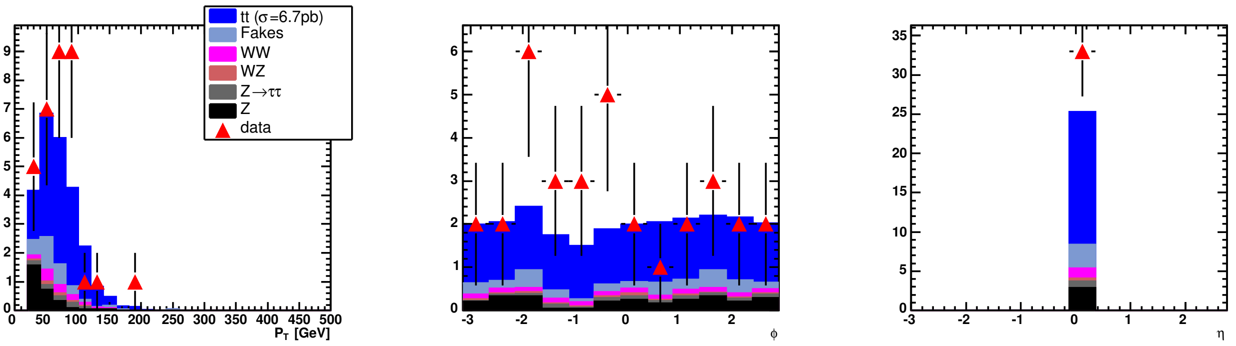

Figure 44: Transverse momentum (P_T), azimuthal angle (ϕ) and pseudo-rapidity (η) of the missing transverse energy.

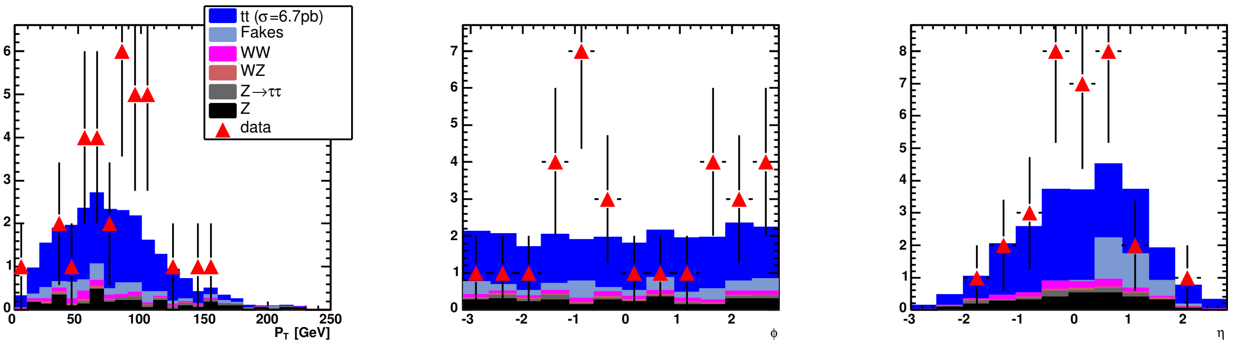

Figure 45: Transverse momentum (P_T), azimuthal angle (ϕ) and pseudo-rapidity (η) of the sum of the leptons.

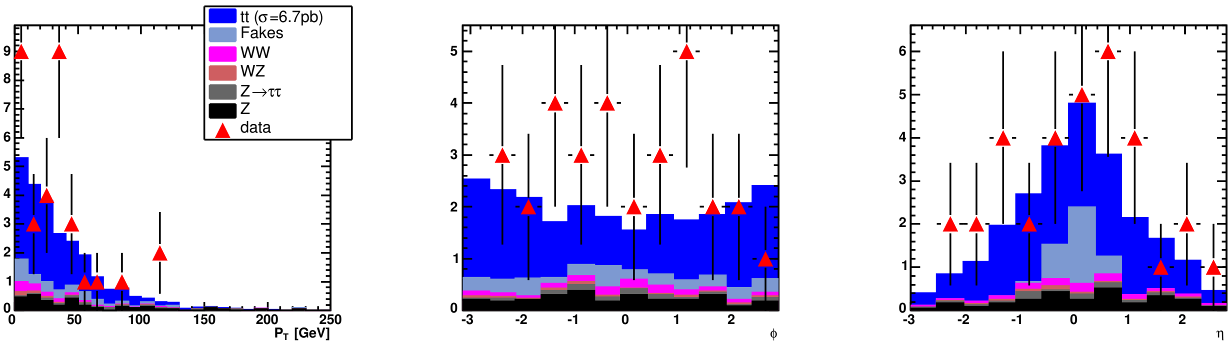

Figure 46: Difference in transverse momentum (P_T), azimuthal angle (ϕ) and pseudo-rapidity (η) of the two leptons.

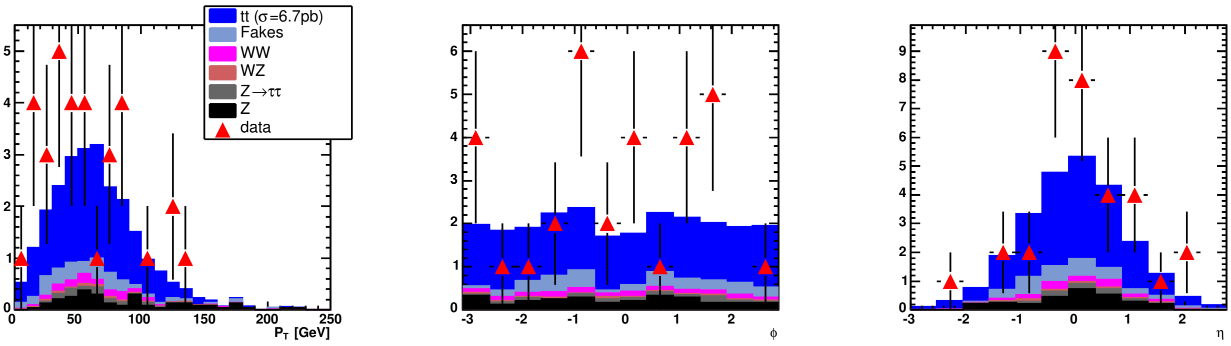

Figure 47: Transverse momentum (P_T), azimuthal angle (ϕ) and pseudo-rapidity (η) of the sum of the jets.

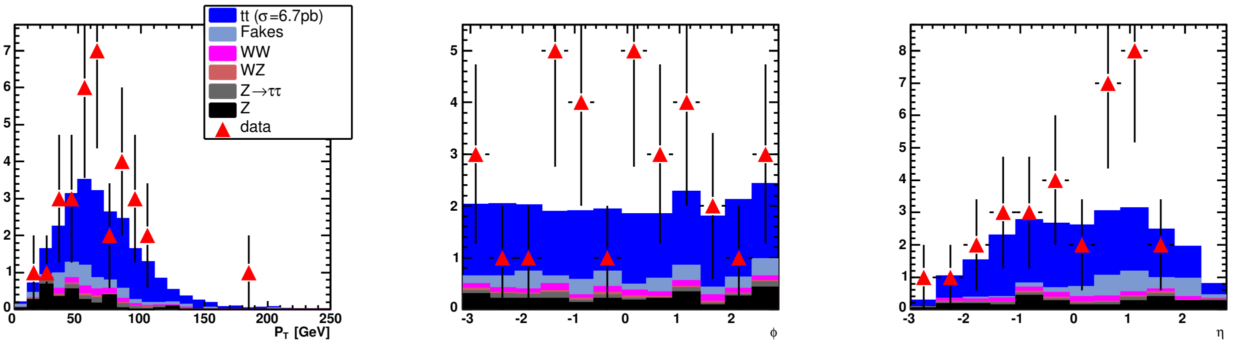

Figure 48: Transverse momentum (P_T), azimuthal angle (ϕ) and pseudo-rapidity (η) of the sum of both jets and leptons.

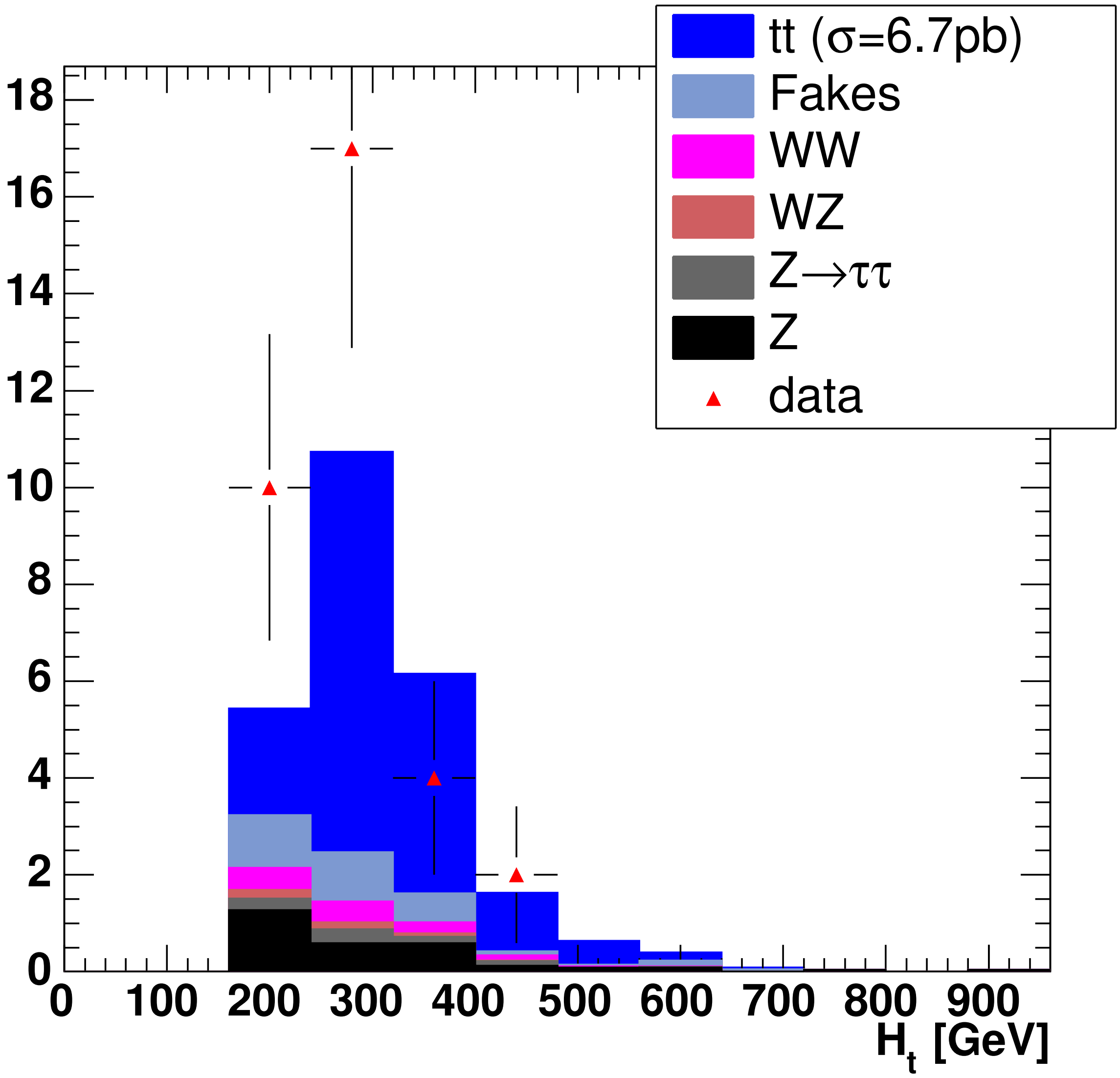

Figure 49: Distribution of H_T.

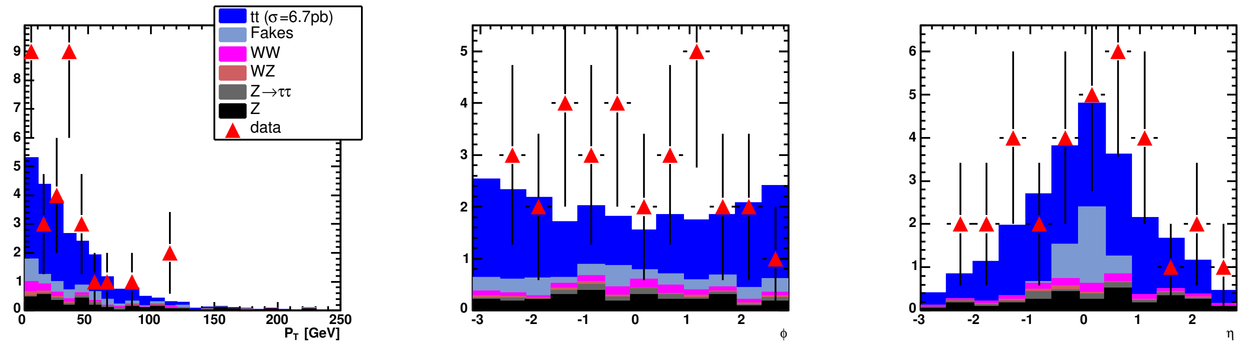

Figure 50: Transverse momentum (P_T), azimuthal angle (ϕ) and pseudo-rapidity (η) of difference between the leptons.

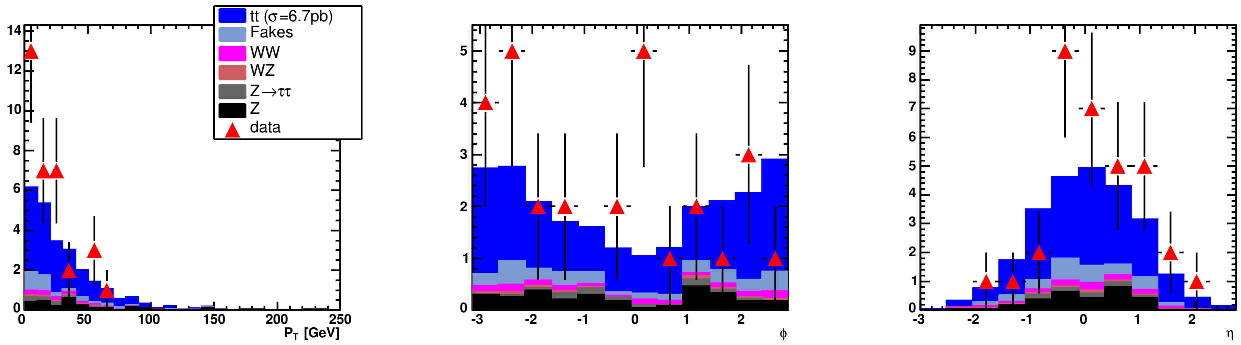

Figure 51: Transverse momentum (P_T), azimuthal angle (ϕ) and pseudo-rapidity (η) of the difference between the leading jets.

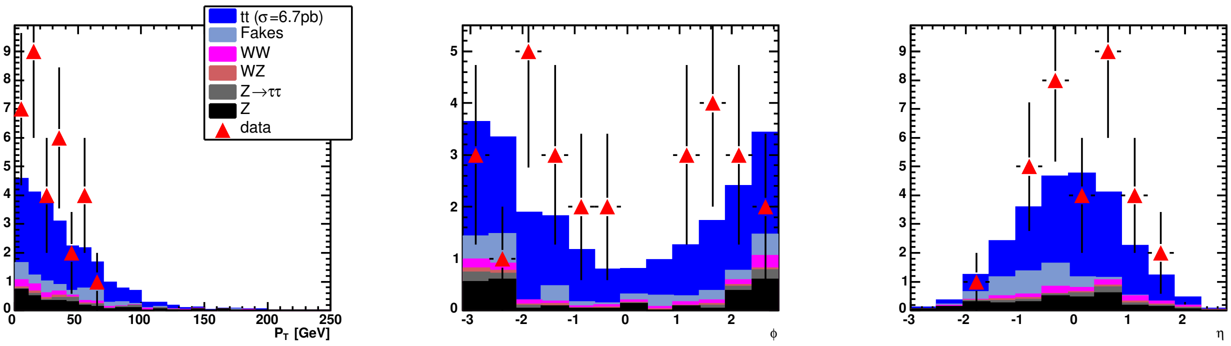

Figure 52: Transverse momentum (P_T), azimuthal angle (ϕ) and pseudo-rapidity (η) of the difference between the leading lepton and the leading jet.

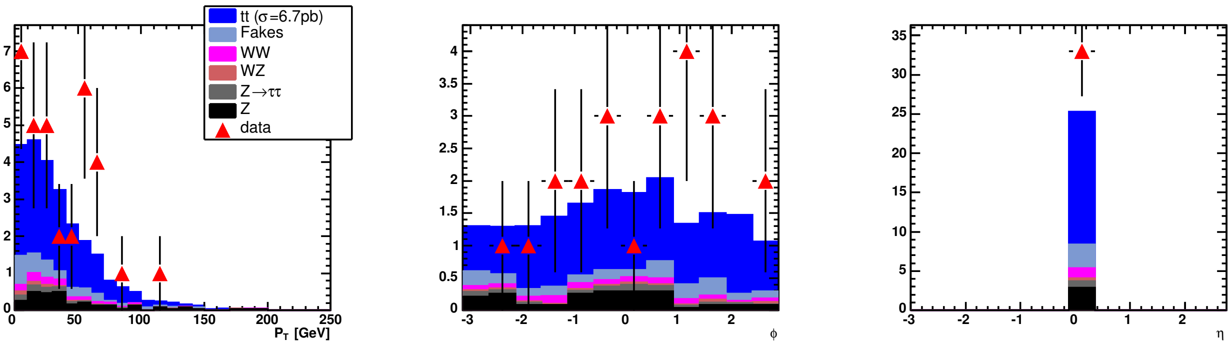

Figure 53: Transverse momentum (P_T), azimuthal angle (ϕ) and pseudo-rapidity (η) of the difference between the missing energy and the leading jet.

References

[1] Kondo, Estrada, Canelli

[2] Particle Data Book, Phys. Rev. D.

[3] Mahlon, Parke.

[4] `http://www.gnu.org/software/gsl/`.

[5] `http://wwwasd.web.cern.ch/wwwasd/lhc++/clhep/` and the ZOOM links therein.

[6] P. Lepage, VEGAS integration algorithm.