Higgs Searches in the $H \to WW$ Decay Mode Using the ATLAS Detector

by

William B. Quayle

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Physics)

at the

University of Wisconsin – Madison

2008

Abstract

The prospects for Higgs boson searches in the $H \rightarrow WW$ decay mode are investigated. Search strategies are presented for the channels $H + 0j$ ($H \rightarrow WW \rightarrow e\nu\mu\nu$), $H + 2j$ ($H \rightarrow WW \rightarrow e\nu\mu\nu$), and $H + 2j$ ($H \rightarrow WW \rightarrow \ell\nu qq$). Methods to extract the background normalization from data are discussed for all channels. With an integrated luminosity of 10 fb^{-1}, the $H + 0j$, $H \rightarrow WW \rightarrow e\nu\mu\nu$ channel can reach a 5σ discovery for a Standard Model Higgs boson in the range $145 < M_H < 180$ GeV; its mass can be measured with a precision of about 2 GeV if M_H is around 160 GeV, or with a precision of about 6 GeV if M_H is close to 140 GeV. The $H + 2j$, $H \rightarrow WW \rightarrow e\nu\mu\nu$ channel has a similar sensitivity, reaching the 5σ significance level for $150 < M_H < 180$ GeV and yielding a less precise mass determination. The statistical significance for $H + 2j$, $H \rightarrow \ell\nu qq$ peaks at about 350 GeV, where it is slightly below the 5σ level; the expected error on the mass determination at this mass is about 6 GeV. The combined sensitivity of the two $H \rightarrow WW \rightarrow e\nu\mu\nu$ channels is estimated, and a combined fit including the $H \rightarrow ZZ \rightarrow 4\ell$ channel is presented. With 1 fb^{-1} of integrated luminosity, the combined fit can exclude Standard Model Higgs bosons with masses above 145 GeV. At 10 fb^{-1} the expected significance is above the 5σ level for $M_H > 135$ GeV.

Acknowledgements

This work would not have been possible or even meaningful at all if not for the efforts of the *thousands* of people who have worked for more than a decade to make the Large Hadron Collider and the ATLAS detector a reality. When working in such a highly interconnected field, how can one properly thank all the people who have contributed?

I am deeply grateful to my advisor, Sau Lan Wu, for helping me to get started in Higgs physics and for the years of support and encouragement that followed. I would like to thank Bruce Mellado, Trevor Vickey, Luis Flores, Nathan Goldschmidt, Sanjay Padhi, and Michael Duehrssen for collaboration and feedback that played a crucial role in the development of this analysis. Many thanks also to Saleem Ali, Tapas Sarangi, Haimo Zobernig, Werner Wiedenmann, Pete McNamara, Kyle Cranmer, Yaquan Fang, Xin Chen, Tai Tsun Wu, and Yibin Pan, for numerous discussions that have helped shape my perspectives and understanding over the years since I started working at CERN. This thesis would not have been possible without Miron Livny, Sanjay Padhi, Neng Xu, and Annabelle Leung, whose immeasurable help in the area of computing was essential to the success of this analysis. I am also grateful to Antonella Lofranco, Catharine Noble, Teresa Sherron, Aimee Lefkow, and Jane Schimmel for all that they have done to make it possible for me to be based at CERN.

Above all, I would like to thank my parents and my sister for their unwavering support. They have helped me more than I can describe.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The most important achievement of particle physicists over the last 50 years was the demonstration that the interactions of all known leptons and hadrons are described by what has come to be known as the "Standard Model," a locally gauge invariant [1] theory based on an SU(3)⊗SU(2)⊗U(1) gauge symmetry [2, 3, 4, 5, 6]. Among the most dramatic successes of the theory were its predictions of the mass and couplings of the Z boson, which were verified experimentally with impressive precision at LEP. [7]

A crucial ingredient in the development of the Standard Model was the realization that a spontaneous symmetry breaking mechanism could give rise to massive particles that do not spoil the gauge invariance of the theory. [8, 9, 10, 11, 12, 13] In the first papers that used a spontaneous symmetry breaking mechanism to generate particle masses, the symmetry breaking was accomplished by introducing a single complex spin-0 field which transforms as a doublet under isospin transformations.

The Standard Model makes no strong prediction about the mass of the scalar field. Moreover, the number of scalar doublets is not uniquely predicted; it is straightforward to construct theories with, for example, two Higgs doublets. One of the highest priorities of modern high-energy physics is to place experimental constraints on the physics of the so-called Higgs sector. This is one of the major motivations for the Large Hadron Collider (LHC), presently under construction at CERN. [14]

This thesis investigates the prospects for Higgs boson searches in the $H \rightarrow WW$ channel using the ATLAS detector. [15] Chapter 2 reviews the the Standard Model, with emphasis on the motivation for $H \rightarrow WW$ searches. Chapter 3 gives a brief overview of the features of the LHC and the ATLAS detector most relevant to this analysis. Chapter 4 describes the event selection, background normalization, statistical treatment, and expected sensitivity for a Higgs boson search in the events where the Higgs boson is not accompanied by any jets $(H + 0j)$ and where it decays in the mode $H \rightarrow WW \rightarrow e\nu\mu\nu$. Chapter 5 describes the analysis of the $H + 2j$, $H \rightarrow WW \rightarrow e\nu\mu\nu$ channel, and Chapter 6 describes a search for heavy Higgs resonances in the channel $H + 2j$, $H \rightarrow WW \rightarrow l\nu qq$. Finally, Chapter 7 discusses the statistical treatment of a combined fit of the $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ channels with the $H \rightarrow ZZ \rightarrow 4l$ channel.

Chapter 2

Theoretical Background

The particle content of the Standard Model consists of three generations of fermions, plus several gauge and Higgs bosons that are required to ensure invariance under local SU(3)⊗SU(2)⊗U(1) gauge transformations. The SU(2)⊗U(1) part of the theory is most relevant to the present study; however, it is worthwhile to mention a few basic points about the SU(3) part of the theory as well. Fundamental fermions that carry no SU(3) color charge are referred to as leptons; those that do carry a color charge are termed quarks. Free quarks have not been observed; instead, the existence and properties of quarks must be inferred from the properties of color-neutral composite particles (hadrons) formed by combining pairs or triplets of quarks (mesons and baryons, respectively). "Up-type" quarks (u, c, and t) and antiquarks carry electrical charges of $\pm\frac{2}{3}e$, while "down-type" quarks and antiquarks carry electrical charges of $\mp\frac{1}{3}e$, where e is the electrical charge of the electron. Quarks are arranged in SU(2) doublets, much like the leptons that will be discussed in more detail in the next section; in fact, much of that discussion applies to quarks as well.

The salient aspects of the SU(2)⊗U(1) part of the Standard Model are discussed in Section 2.1, and phenomenology relevant to $H \to WW$ searches is reviewed in

Section 2.2.

2.1 Electroweak Physics and the Higgs Mechanism

The interactions that result from the SU(2)⊗U(1) part of the gauge symmetry are known as the Electroweak force. It is known experimentally that the Weak interaction violates Parity [16]. In particular, the strong polarization of the muons produced in pion decays [17] indicates that the Weak interaction only acts on left-handed leptons and neutrinos. Moreover, the neutrino mass is known to be quite small [18]. In the present discussion, neutrinos are therefore taken to be massless; in other words, right-handed neutrinos are ignored completely. Finally, neutrino-nucleon scattering experiments [19] have demonstrated that there are separate types of neutrinos corresponding to each of the charged leptons, and that the neutrino produced in association with, for example, a muon is a muon neutrino and not an electron neutrino. In other words, the gauge boson that mediates the Weak charged current interaction couples the left-handed charged lepton to the corresponding neutrino. In the Standard Model, therefore, the left-handed charged and neutral leptons in each generation transform as an SU(2) doublet, while the right-handed leptons transform as an SU(2) singlet. Thus, the leptons are denoted as

$$E_L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \qquad E_R = e^-_R \tag{2.1}$$

and the Lagrangian is required to be invariant under the transformations

$$E_L \to e^{i\left(\alpha^i(x)\frac{\sigma^i}{2} + \beta(x)/2\right)} E_L, \qquad E_R \to e^{i\beta/2} E_R \tag{2.2}$$

Here, α and β are continuous and differentiable but otherwise arbitrary functions, and the σ^i are the Pauli sigma matrices. The charged leptons are known to be massive, and so the Lagrangian will have to include mass terms for these particles. However, E_L and E_R have different quantum numbers, so a mass term of the form $m\overline{E_L}E_R + h.c.$ will violate charge conservation. Instead, in the Standard Model, one introduces a field ϕ and adds to the Lagrangian terms like $\overline{E_L}\phi E_R + h.c.$ In order to prevent such terms from being manifestly nonrenormalizable, it *must* be the case that:

- ϕ has spin 0; for couplings with three fields, this follows from angular momentum conservation. For couplings with larger numbers of fields, it is difficult to construct interaction terms that do not spoil the renormalizability of the theory.

- ϕ transforms as a doublet under SU(2) transformations. This is necessary because E_L transforms as an SU(2) doublet, while the Lagrangian must remain invariant under all transformations.

- one component of ϕ has an electrical charge opposite to that of the electron, and the other has 0 electrical charge. This is necessary so that the quantum numbers of the terms like $\overline{E_L}\phi E_R + h.c.$ add to zero.

The terms like $\overline{E_L}\phi E_R + h.c.$ will play the role of fermion mass terms below.

Given the above constraints, one can write down an expression for ϕ and its behavior under an SU(2)⊗U(1) gauge transformation

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \qquad \phi \to e^{i\alpha^a \sigma^a/2} e^{i\beta/2} \phi \tag{2.3}$$

as well as the corresponding covariant derivative

$$D_\mu \phi = \left(\partial_\mu - i\frac{g}{2} A^a_\mu \sigma^a - \frac{i}{2} g' B_\mu \right) \phi \tag{2.4}$$

The most general Lagrangian for the ϕ fields that preserves renormalizability is

$$\mathcal{L}_{Higgs} = |D_\mu \phi|^2 + \mu^2 \phi^\dagger \phi - \frac{\lambda}{2}(\phi^\dagger \phi)^2 \tag{2.5}$$

where the couplings to the fermions have been omitted. If $\mu^2 > 0$, then the $\phi = 0$ does not represent a stable vacuum; the minimum is instead the set of points with $|\phi| = v = \left(\frac{\mu^2}{\lambda}\right)^{1/2}$. The quantity v is known as the vacuum expectation value of ϕ. It is convenient to work with the parameterization

$$\phi = \frac{1}{\sqrt{2}} U(x) \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.6}$$

where $h(x)$ is real-valued and $U(x)$ is an operator representing an arbitrary SU(2) gauge transformation. Without loss of generality, it is possible to work in a gauge where $U(x)$ is 1.

The kinetic energy term of Equation 2.5 includes terms that couple the ϕ fields to the A and B fields. When the gauge symmetry is broken, these terms become mass terms for the A and B gauge boson fields and terms that describe the couplings of the physical states to the ϕ field. The mass terms are the ones that do not include $h(x)$:

$$\mathcal{L}_{Gauge\ Boson\ mass} = \frac{v^2}{8}\left[g^2\left(A_\mu^1\right)^2 + g^2\left(A_\mu^2\right)^2 + \left(-gA_\mu^3 + g' B_\mu\right)^2\right] \tag{2.7}$$

This contains terms that mix A_μ^3 and B_μ; in order to read off the physical particle masses directly from the Lagrangian, it is helpful to define fields that are mass eigenstates:

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}\left(A_\mu^1 \mp i A_\mu^2\right) \\ Z_\mu^0 &= \frac{1}{\sqrt{g^2 + g'^2}}\left(gA_\mu^3 - g' B_\mu\right) \\ A_\mu &= \frac{1}{\sqrt{g^2 + g'^2}}\left(g' A_\mu^3 + g B_\mu\right) \end{aligned} \tag{2.8}$$

With the additional definitions

$$
\begin{aligned}
T^{\pm} &= \frac{1}{2}\left(\sigma^1 \pm i\sigma^2\right) \\
T^3 &= \frac{\sigma^3}{2} \\
e &= \frac{gg'}{\sqrt{g^2+g'^2}} \\
Q &= T^3 + Y \\
\cos\theta_w &= \frac{g}{\sqrt{g^2+g'^2}}
\end{aligned}
\tag{2.9}
$$

the covariant derivative can be rewritten in terms of the mass eigenstate fields

$$
D_\mu = \partial_\mu - i\frac{g}{\sqrt{2}}\left(W_\mu^+ T^+ + W_\mu^- T^-\right) - i\frac{g}{\cos\theta_w} Z_\mu \left(T^3 - \sin\theta_w Q^2\right) - ieA_\mu Q \tag{2.10}
$$

The kinetic part of the Lagrangian in Equation 2.5 (including the terms with $h(x)$) becomes

$$
\mathcal{L}_K = \frac{1}{2}\left(\partial_\mu h\right)^2 + \left[M_W^2 W^{\mu+} W_\mu^- + \frac{1}{2} M_Z^2 Z_\mu Z^\mu\right]\left(1+\frac{h}{v}\right)^2 \tag{2.11}
$$

where $M_W = \frac{gv}{2}$ and $M_Z = \frac{v}{2}\sqrt{g^2+g'^2}$, and the A_μ field is massless. The Lagrangian describes a neutral scalar particle h, known as the Higgs boson, which couples to the W and the Z with a strength that is proportional to their masses.

From the $\phi^\dagger\phi$ and $\left(\phi^\dagger\phi\right)^2$ terms in Equation 2.5, one can obtain a mass term for the Higgs boson as well as three-point and four-point self-couplings:

$$
\mathcal{L}_V = -\frac{1}{2} M_H^2 h^2 - M_H \sqrt{\frac{\lambda}{2}} h^3 - \frac{\lambda}{4} h^4 \tag{2.12}
$$

where $M_H = v\sqrt{\frac{\lambda}{2}}$.

In terms of the vacuum parameterized by Equation 2.6, a Lagrangian term proportional to $\overline{E_L}\phi E_R + h.c.$ will lead to a term proportional to $v\overline{e_L}e_R + h.c.$, i.e. a mass term. The coupling strength of these terms is therefore proportional to the electron mass; similar terms exist for muons, taus, and quarks.

The couplings of the leptons to the gauge bosons arise from the kinetic terms for the leptons:

$$\mathcal{L}_K = \overline{E}_L(i\gamma^\mu D_\mu)E_L + \overline{e}_R(i\gamma^\mu D_\mu)e_R + h.c. \tag{2.13}$$

In terms of the mass eigenstate fields, this becomes

$$\begin{aligned} \mathcal{L}_K &= \overline{E}_L(i\gamma^\mu \partial_\mu)E_L + \overline{e}_R(i\gamma^\mu \partial_\mu)e_R + h.c. \\ &+g\left(W^+_\mu J^{\mu+}_W + W_\mu J^{\mu-}_W + Z^0_\mu J^\mu_Z\right) + eA_\mu J^\mu_{EM} \end{aligned} \tag{2.14}$$

where

$$\begin{aligned} J^{\mu+}_W &= \frac{1}{\sqrt{2}}\left(\overline{\nu}_L\gamma^\mu e_L\right) \\ J^{\mu-}_W &= \frac{1}{\sqrt{2}}\left(\overline{e}_L\gamma^\mu \nu_L\right) \\ J^\mu_Z &= \frac{1}{\cos\theta_w}\left[\frac{\overline{\nu}_L\gamma^\mu\nu_L}{2} + \overline{e}_L\gamma^\mu\left(-\frac{1}{2}+\sin^2\theta_w\right)e_L + \overline{e}_R\gamma^\mu\left(\sin^2\theta_w\right)e_R\right] \\ J^\mu_{EM} &= -\overline{e}\gamma^\mu e \end{aligned} \tag{2.15}$$

with similar expressions for the quarks.

The discussion up to this point has implicitly made the assumption that there is only one ϕ doublet. In practice, there is no reason (except for simplicity) to make this assumption; it is possible that there are two or more doublets, or that the Higgs sector is governed by more complicated dynamics. In such scenarios, most of the theoretical arguments presented above still hold: there must be at least one (possibly composite) field that transforms as a complex scalar SU(2) doublet, gauge bosons and fermions acquire their masses from their couplings to the Higgs sector when SU(2) symmetry is spontaneously broken, etc. However, there are typically important phenomenological consequences, such as multiple neutral and charged Higgs bosons, pseudoscalar Higgs bosons, and so on.

A concrete model of particular interest is the Two Higgs Doublet Model (2HDM), the minimal extension of the framework discussed above with interesting phenomenological consequences. In the 2HDM, there are two separate complex scalar SU(2) doublets, ϕ_1 and ϕ_2, which both acquire separate vacuum expectation values v_1 and v_2. The most general Lagrangian for the 2HDM ϕ fields that conserves CP symmetry and is invariant under the discrete symmetry $\phi_1 \to -\phi_1$, $\phi_2 \to +\phi_2$ is [20]

$$\begin{aligned} V(\phi_1,\phi_2) = {} & \lambda_1 \left(\phi_1^\dagger\phi_1 - v_1^2\right)^2 + \lambda_2 \left(\phi_2^\dagger\phi_2 - v_2^2\right)^2 \\ & + \lambda_3 \left[\left(\phi_1^\dagger\phi_1 - v_1^2\right) + \left(\phi_2^\dagger\phi_2 - v_2^2\right)\right]^2 \\ & + \lambda_4 \left[\left(\phi_1^\dagger\phi_1\right)\left(\phi_2^\dagger\phi_2\right) - \left(\phi_1^\dagger\phi_2\right)\left(\phi_2^\dagger\phi_1\right)\right] \\ & + \lambda_5 \left[\mathrm{Re}\left(\phi_1^\dagger\phi_2\right) - v_1 v_2\right]^2 \\ & + \lambda_6 \left[\mathrm{Im}\left(\phi_1^\dagger\phi_2\right)\right]^2 \end{aligned} \tag{2.16}$$

When SU(2) symmetry of the 2HDM is broken, one finds a total of five observable Higgs fields: a light neutral scalar denoted as h^0, a heavy scalar H^0, a pseudoscalar A^0, and two charged Higgs bosons, $H^\pm$. The h^0 and H^0 mix via the mass-squared matrix

$$\mathcal{M} = \begin{pmatrix} 4v_1^2\left(\lambda_1+\lambda_3\right) + v_2^2\lambda_5 & \left(4\lambda_3+\lambda_5\right) + v_1v_2 \\ \left(4\lambda_3+\lambda_5\right) + v_1v_2 & 4v_2^2\left(\lambda_2+\lambda_3\right) + v_1^2\lambda_5 \end{pmatrix} \tag{2.17}$$

Two parameters are of particular interest for the $H \to WW$ analysis: $\beta = v_2/v_1$ and α, which is defined by

$$\begin{aligned} \sin 2\alpha &= \frac{2\mathcal{M}_{12}}{\sqrt{\left(\mathcal{M}_{11}-\mathcal{M}_{22}\right)^2 + 4\mathcal{M}_{12}^2}} \\ \cos 2\alpha &= \frac{\mathcal{M}_{11}-\mathcal{M}_{22}}{\sqrt{\left(\mathcal{M}_{11}-\mathcal{M}_{22}\right)^2 + 4\mathcal{M}_{12}^2}} \end{aligned} \tag{2.18}$$

Of the various Higgs bosons in the 2HDM, only h^0 and H^0 have couplings to the WW final state. The structure of the $h^0W^+W^-$ and $H^0W^+W^-$ couplings is the same as the coupling of the W to the Higgs boson of the one Higgs doublet model, but the strength of these couplings is lowered by $\sin(\beta - \alpha)$ (in the case of h^0), or $\cos(\beta - \alpha)$ (in the case of H^0). On the other hand, the strengths of the couplings to the quarks and leptons are modified by factors involving $\sin\alpha$, $\cos\alpha$, $\sin\beta$, and $\cos\beta$, depending on the details of how the Higgs fields couple to the fermions. So while the couplings (and hence the production cross-sections and decay branching ratios) were predicted in the one Higgs Doublet model, they are not predicted in the 2HDM.

The current lower bound on the Higgs boson mass is 114.4 GeV [21]; this comes from direct searches at LEP. This bound applies to Standard Model Higgs bosons; the bounds on the Higgs boson masses in the most general 2HDM are considerably looser. Precision measurements of a variety of Electroweak parameters, such as the W and Z masses and forward-backward asymmetries, yield an upper bound of 285 GeV on the mass of a Standard Model Higgs boson. [7] However, this upper bound clearly does not apply in the case of the 2HDM; there, H^0 can be very heavy even if h^0 is light.

2.2 Phenomenology

The studies presented in this document deal with searches for Higgs bosons produced via two mechanisms. The so-called "gluon fusion" production mechanism is a gluon-initiated process where the Higgs boson is produced via its coupling to the top quark triangle. The other production mechanism is commonly referred to as "Vector Boson Fuson," although technically the term "Weak Boson Fusion" is more appropriate. This production is a quark scattering process mediated by a W or a Z;

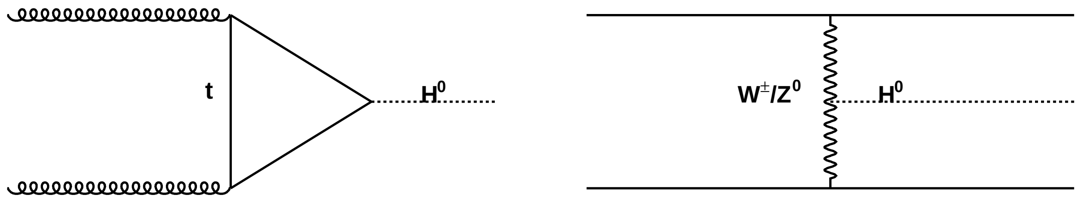

Figure 2.1: Feynman diagrams for the signal processes considered. Left: the gluon fusion process, $gg \rightarrow H$. Right: the Vector Boson Fusion process.

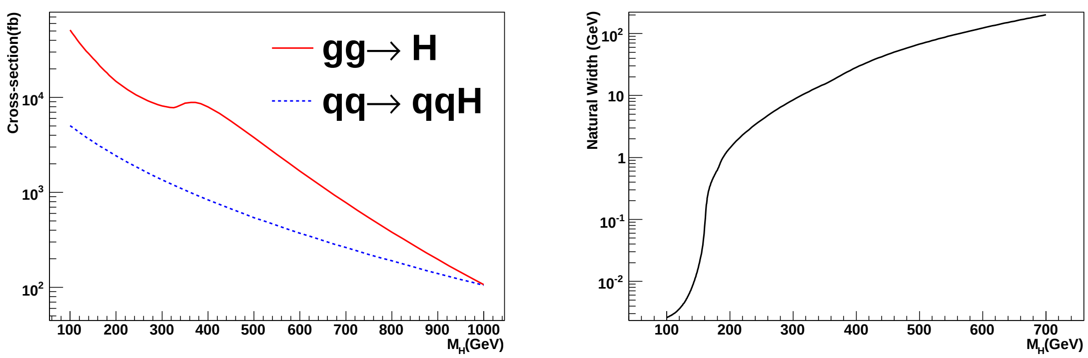

Figure 2.2: The cross-section for Higgs production in the Standard Model for the two production mechanisms considered (left), and the natural width of the Higgs boson in the Standard Model (right).

representative Feynman diagrams for both processes are shown in Figure 2.1.

The left plot in Figure 2.2 shows the cross-section for these two production mechanisms in the Standard Model as a function of M_H. [22] The gluon-fusion process is expected to be the dominant Higgs boson production mechanism at LHC, for the Standard Model at least. However, the Vector Boson Fusion process is expected to be a powerful discovery channel as well thanks to the struck quarks in the final state. It is possible to achieve a strong suppression against QCD backgrounds by taking advantage of this feature; this will have an important impact on the searches

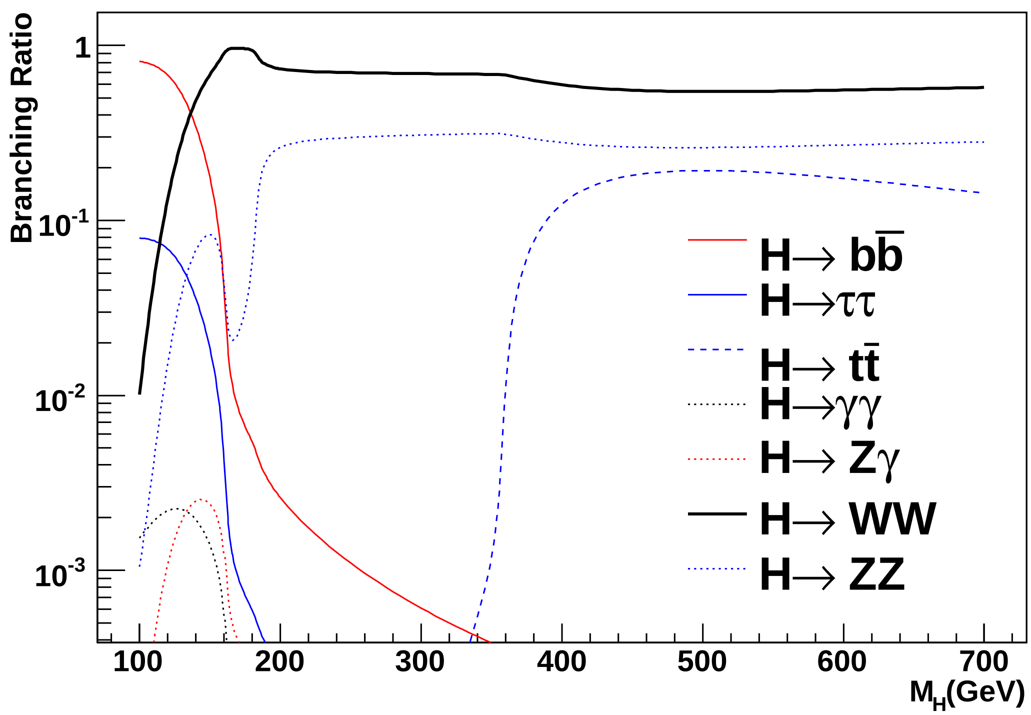

Figure 2.3: The branching ratio for $H \to WW$ (solid black line) and for several other decay modes.

presented in Chapters 5 and 6.

The plot on the right of Figure 2.2 shows the natural width of the Higgs boson in the Standard Model (computed with the HDECAY [23] package). For Higgs masses below 200 GeV, the natural width is very small compared to the experimental resolution in this mass range; the searches presented in Chapters 4 and 5 can therefore be regarded as model-independent searches for WW resonances. At around 300 GeV, the natural width becomes comparable to the experimental resolution of the $H \to WW \to l\nu qq$ channel; the fit presented in Chapter 6 uses a signal model whose signal shape depends on this prediction. The analysis of Chapter 6 should therefore be regarded as a search that is geared toward particles that are similar to the Higgs boson of the Standard Model; additional work would be necessary to remove this model dependence by introducing a free parameter to measure the natural width of the Higgs boson in this channel.

Figure 2.3 shows the branching ratio for $H \to WW$ as a function of M_H in the Standard Model, along with the corresponding curves for several other decay modes for comparison. (These branching ratios have also been computed using HDECAY.) The $H \to WW$ mode (the black line in the plot) is the dominant decay mode for Higgs boson masses larger than about 135 GeV.

Chapter 3

The LHC and the ATLAS Detector

The LHC will collide pairs of protons at a center-of-mass energy of 14 TeV at four points along its 27 km circumference. The proton beams are created by stripping the electrons from Hydrogen atoms in a plasma. They reach the final collision energy after passing through a series of accelerators, beginning with an RF Quadrupole which increases the beam energy to 750 keV. From there, the beams are accelerated to 50 MeV by the Linac-2 facility, to 1.4 GeV by the Proton Synchrotron Booster, and to 25 GeV by the Proton Synchrotron. The beams reach the LHC injection energy of 450 GeV after passing through the Super Proton Synchrotron, and finally, the LHC itself accelerates them the rest of the way to 7 TeV.

The LHC beams have a very high design luminosity, $10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. This opens the possibility of understanding the physics of phenomena with a cross-section as small as a few femtobarns, but it also presents technical challenges. The high bunch crossing frequency (one bunch crossing every 25 nanoseconds, at design luminosity) means that all detector components must be able to read out their data in a short time. Because of this high bunch crossing rate coupled with the large size of each event of ATLAS data (about 1.6 Mb of raw data), it is not possible with existing technology to record

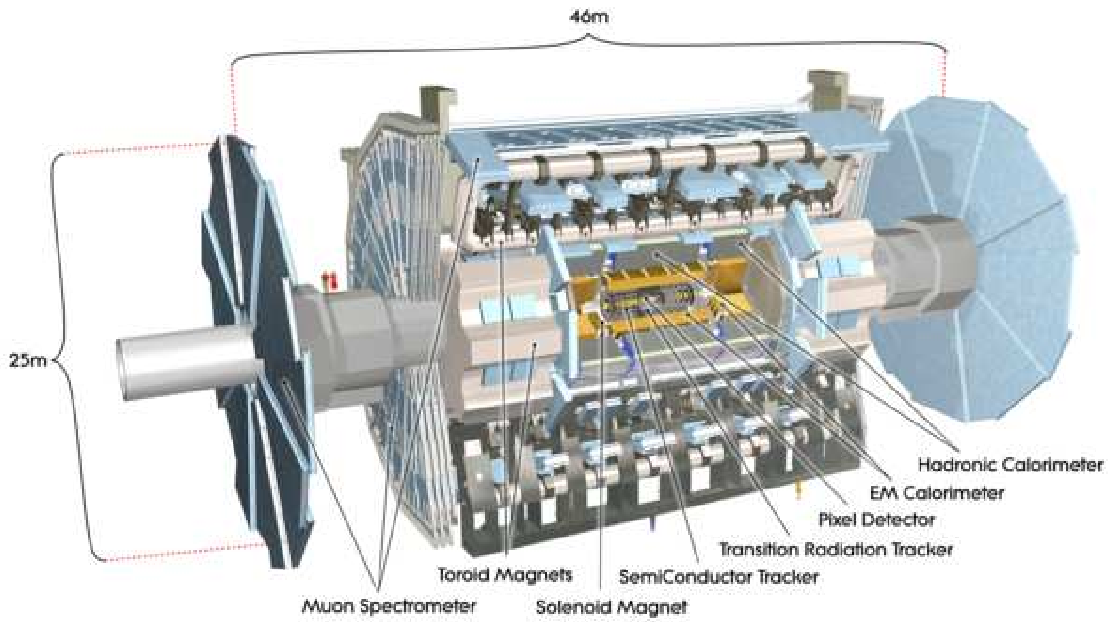

Figure 3.1: A schematic view of the ATLAS detector, reproduced from Ref. [15]

every event that takes place in ATLAS. This constraint is, in fact, quite dramatic; only about one in 200,000 events can actually be recorded to tape. Moreover, the presence of more than one proton-proton collision in some bunch crossings will degrade the quality of some reconstructed quantities, notably the missing E_T; to make matters worse, this effect will vary with time and running conditions.

These considerations, together with physics requirements from performance-driven channels like $H \to 4l$ and $H \to \gamma\gamma$, have had a strong influence on the design of many components of the ATLAS detector. This section will give a brief overview of the various subdetectors and their expected performance.

A schematic view of the ATLAS detector is shown in Figure 3.1. The innermost

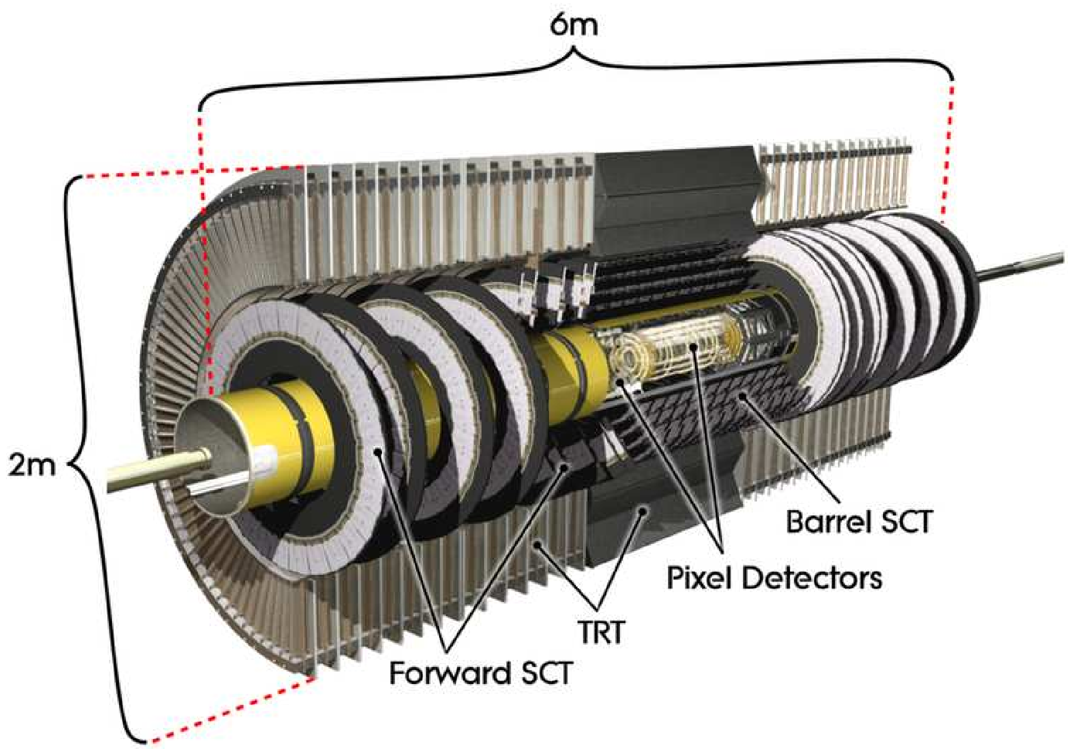

Figure 3.2: A schematic view of the ATLAS Inner Detector, reproduced from Ref. [15].

subdetectors, the Silicon Pixel Detector, the Semiconductor Tracker, and the Transition Radiation tracker, are surrounded by a 2 Tesla solenoid magnet and are described in Section 3.1. Calorimeter systems surround the Inner Detector; these subsystems are described in Section 3.2. The most distinctive subsystem is the muon spectrometer, with its toroidal magnetic field spread over a large volume; this subsystem is described in Section 3.3. Finally, a few important details about the trigger system are discussed in Section 3.4.

3.1 Inner Detector

The ATLAS inner detector, pictured in Figure 3.2, provides precise measurements of the trajectories of charged particles that pass through the region with pseudorapidity $|\eta| < 2.5$. The precise track measurements made by the Inner Detector are vital for b-tagging, which is very useful for suppressing backgrounds from, for example, top decay. The complete detector consists of three subsystems.

The innermost layer is a precision tracker comprised of three layers of silicon pixel detectors, arranged in a cylindrical geometry, with an inner radius of about 50 mm, an outer radus of about 150 mm, and a total length of about 1.3 m. It provides position measurements with an accuracy of about 10 μm in $R - \phi$ and about 115 μm in z in the barrel, with similar precision for the disks at either end.

The middle layer, the Semiconductor Tracker (SCT), uses four stereo layers of silicon strip detectors with one set of strips in each layer offset by a small angle with respect to the other, to provide additional position measurements with a precision of 17 μm in $R - \phi$ and 580 μm in z, again with similar precision for the disks at either end. The SCT has an inner radius of about 250 mm, an outer radius of about 600 mm, and a length of about 5.4 m.

The outer layer, the Transition Radiation Tracker (TRT), is a straw tube tracker that provides a large number (about 36 per track) of $R - \phi$ measurements in the region with R from about 550 mm to about 1100 mm. The straws are interleaved with polypropylene (fibers in the barrel region, foil in the endcaps) to provide transition radiation to help with electron identification. The total length of the TRT is about 5.4 m; this corresponds to pseudorapidities with $|\eta| < 2$. The precision of the $R - \phi$

measurements from the TRT is 130 μm per straw.

The material budget of the ATLAS Inner Detector ranges from less than 0.5 radiation lengths (X_0) and 0.2 interaction lengths (λ) near $|\eta| = 0$ to almost 2.5 X_0 and 0.7 λ in the region around $|\eta| = 1.6 - 1.7$. This large material budget can lead to degraded performance for some reconstruction tasks; an important example for this analysis is electron identification.

On the other hand, the large number of precise position measurement gives good performance for tracking and vertexing. The resolution of the transverse impact parameter d_0 is expected to be (10 μm)$\times 1 \oplus 14\text{GeV}/p_T$ in the barrel region and (12 μm)$\times 1 \oplus 20\text{GeV}/p_T$ in the endcaps. The resolution of the longitudinal impact parameter, $z_0 \times \sin\theta$ is 91 $\mu\text{m}\times 1 \oplus 2.3\text{GeV}/p_T$ in the barrel and 71 $\mu\text{m}\times 1 \oplus 3.7\text{GeV}/p_T$.

An important background in the $H \to WW$ searches is $t\bar{t}$ production. It is possible to suppress this background with b-tagging cuts, which rely heavily on the performance of the Inner Detector. The b-tagging algorithm used in ATLAS is based on a Likelihood discriminator which combines information about individual track impact parameters and reconstructed secondary vertices, if they are present. With the present b-tagging algorithm, it is possible to achieve a b-tagging efficiency of 60% with a rejection of > 100 against light jets. However, the specific cuts applied to the b-tagging weights are analysis dependent; see the discussions in Chapters 4, 5, and 6 for details.

3.2 Calorimetry

Figure 3.3 shows a schematic view of the ATLAS calorimeter systems. The innermost layer, comprised of the Liquid Argon (LAr) barrel and endcap electromagnetic

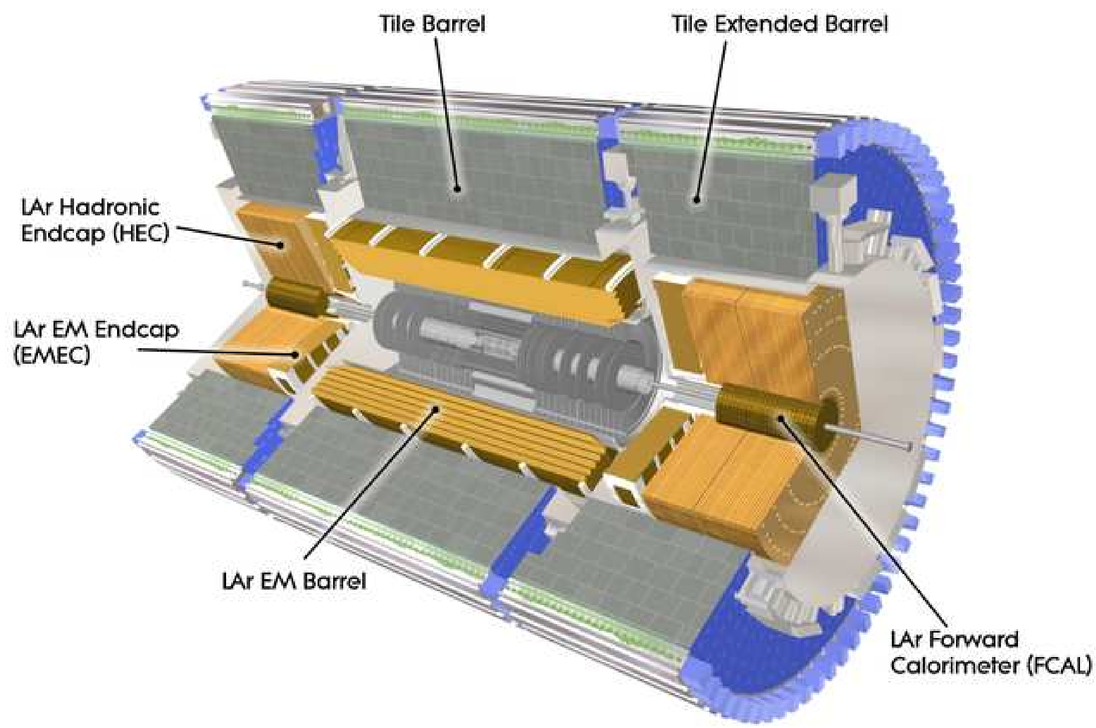

Figure 3.3: A schematic view of the ATLAS calorimeter systems, reproduced from Ref. [15].

calorimeters, is a lead-LAr detector with accordion-shaped absorbers and electrodes. This unusual geometry is intended to improve the readout time of the detector by avoiding current loops. It has a thickness of $> 22\ X_0$ in the barrel region and a thickness of $> 24\ X_0$ in the endcap. The electromagnetic calorimeters are arranged in three layers, plus a presampler in the region with $|\eta| < 1.8$. The granularity of the presampler in $\eta \times \phi$ is 0.025×0.1. The first sampling is segmented into strips of $\eta \times \phi = 0.025/8 \times 0.1$ in the region with $|\eta| < 1.4$; the granularity decreases in the forward regions. Similarly, the second sampling has a granularity of 0.025×0.025 in $|\eta| < 1.4$, with similar or lower granularity in other regions. Finally, the third sampling is segmented into cells of 0.05×0.025 in $|\eta| < 1.35$ and in $1.5 < |\eta| < 2.5$. Because of this fine granularity, a typical electron cluster consists of 50 cells or more; consequently, it is possible to use information about the shower shape to enhance the discrimination between electrons and jets.

The tile calorimeter, placed directly outside the LAr calorimeter, is a hadronic sampling calorimeter using steel as the absorber and scintillating tiles as the active material. It is separated into two regions, the barrel region with $|\eta| < 1$ and the extended barrel region with $0.8 < |\eta| < 1.7$. Like the electromagnetic calorimeter, the tile calorimeter is segmented longitudinally into three layers. In the barrel, the first layer has a thickness of about 1.5 λ, the second about 4.1 λ, and the thickness of the third sampling is 1.8 λ. In the extended barrel, these figures are 1.5 λ, 2.6 λ, and 3.3 λ, respectively. The granularity of the first two layers of the tile calorimeter is 0.1×0.1; the granularity of the last layer is 0.1×0.2.

In the region with $|\eta| > 1.5$, liquid Argon is used as the active medium for the hadronic calorimetry. The Hadronic End Cap (HEC) uses copper plates as the absorber and covers the range up to $|\eta| < 3.2$ to a depth of about 10 λ. It has three layers, with a granularity of 0.1×0.1 in the region with $|\eta| < 2.5$ and 0.2×0.2 in the region with $|\eta| > 2.5$.

The forward calorimeter (FCal) covers the region with $3.1 < |\eta| < 4.9$; like the other calorimeters, it is also divided into three longitudinal layers. The FCal is split into one electromagnetic calorimeter and two hadronic calorimeters, each about 45cm deep. The absorber for the electromagnetic FCal is copper; for the hadronic FCal modules, it is tungsten. The granularity is dependent on η and is slightly different for the three different FCal modules, but in general it is lower for the FCal than for the other calorimeters.

The jet energy resolution can be parameterized as

$$\frac{\sigma}{E} = \sqrt{\frac{a^2}{E} + \frac{b^2}{E^2} + c^2} \tag{3.1}$$

The hadronic calorimeters provide very good jet energy resolution; for example, for jets in the region $0.2 < |\eta| < 0.4$, $a \sim 60\%\sqrt{\mathrm{GeV}}$, $b \sim 0.5$ GeV, and c=3%. This performance is obviously important for the reconstruction of the Higgs candidate invariant mass in channels like $H \to WW \to l\nu qq$, but it clearly also plays a crucial role in the reconstruction of missing E_T. The missing E_T resolution can be roughly parameterized as $a\sqrt{\Sigma E_T}$, where $a \sim 0.53 - 0.57$.

3.3 Muon System

The muon spectrometer consists of a large air-core toroidal magnet system instrumented with a variety of detectors that provide position measurements for muon tracks passing through its volume. The momentum resolution of this sort of system rests on the bending power of the magnetic field. The relevant measure of bending power is the integral of the component of the magnetic field normal to the trajectory of an infinite-momentum track, $\int Bdl$. The toroidal magnetic field geometry was chosen so that the momentum resolution will not be degraded significantly in the forward directions; in a solenoid field, for example, this quantity decreases with $\sin\theta$, where θ is the angle of the track with respect to the solenoid field axis. The bending power of the ATLAS toroidal magnetic field ranges from 1.5-5.5 Tm in the region with $|\eta| < 1.4$, and from 1 to 7.5 Tm in the region $1.6 < |\eta| < 2.7$, with a somewhat lower power in

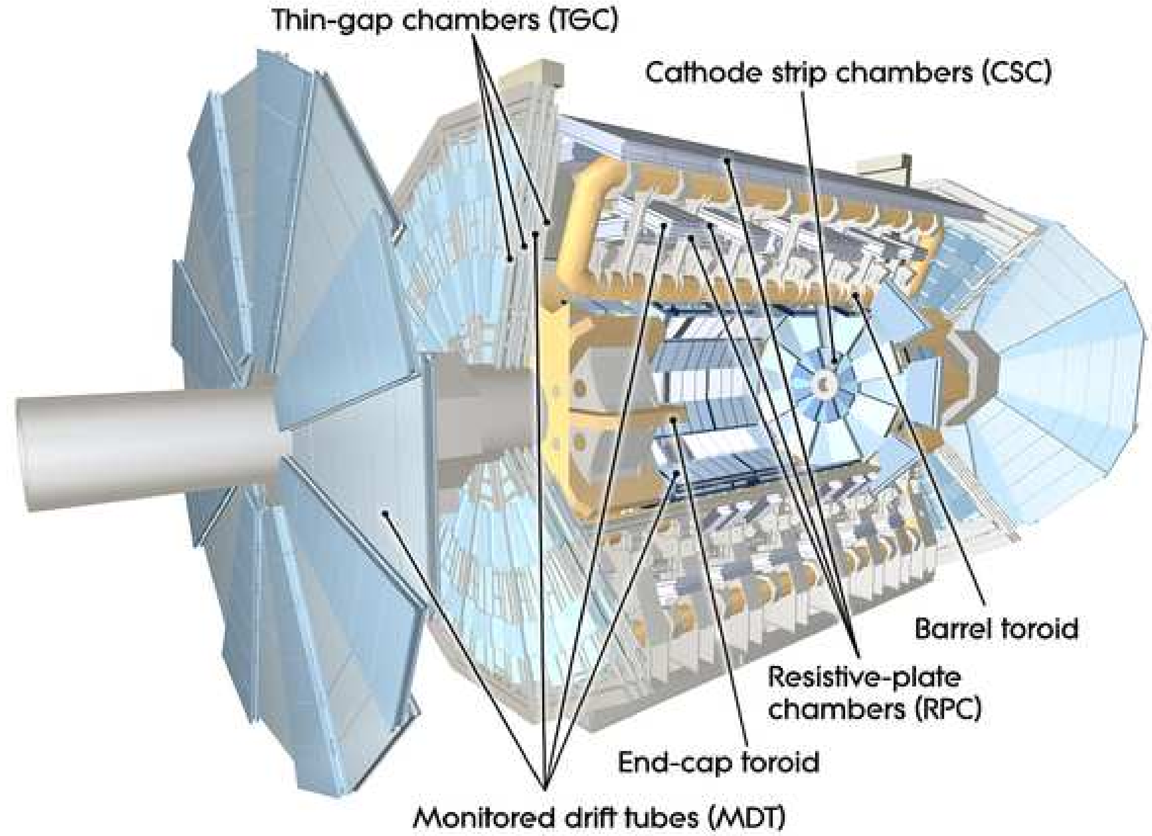

Figure 3.4: A schematic view of the ATLAS muon spectrometer, reproduced from Ref. [15].

between.

Muon trajectories are measured by three concentric cylindrical shells of detectors with radii of about 5, 7.5, and 10 m in the barrel region, and by four endcap "wheels" located at $|z|$ ~7.4, 10.8, 14, and 21.5 m. Because of a variety of performance requirements and the presence of varying radiation environments in different parts of the ATLAS cavern, four different detector technologies are used.

The Monitored Drift Tubes (MDTs) are aluminum tubes with a diameter of about 30 mm; they are filled with a gas mixture of 93% Argon and 7% CO_2, with a tungsten-rhenium wire running down the center. The wire and the outer edge of the tube are kept at a potential difference of 3080 V; the drift time in any given tube can be as long as 700 ns. Arrays of tubes are arranged into chambers consisting of two multilayers with three or four drift tube layers each. Each MDT chamber provides a measurement of the bending direction, z (in the barrel) or R (in the endcap), with a precision of $\sim 35\ \mu$m. The MDTs cover the full range of $|\eta| < 2.7$, except for the first layer which only covers $|\eta| < 2.0$.

Cathode Strip Chambers (CSCs) consist of thin anode wires suspended halfway between two parallel plates covered with cathode strips, with one layer oriented parallel to the wires and the other perpendicular; position measurements (in both the bending and azimuthal directions) are taken from these cathode strips and not the anode wires. The space between the cathode plates is filled with a gas mixture consisting of 80% Argon and 20% CO_2. The wires are kept at a potential of 1900 V with respect to the cathode strips. Four wire planes per chamber provide position measurements with a precision of about 60 μm per CSC plane in the bending direction (for a total of 45-65 μm, depending on background radiation conditions) and about 0.5 cm in the azimuthal direction. The electron drift time in the CSC is less than 40 ns; an OR of the four wire planes in a chamber yields a time-of-arrival distribution with a width of about 3.6 ns. Sixteen Cathode Strip Chambers form an endcap disk covering the region with $2.0 < |\eta| < 2.7$.

Resistive Plate Chambers (RPCs) cover the region with $|\eta| < 1.05$. RPCs are parallel-plate electrodes filled with a gaseous mixture of $C_2H_2F_4$, Iso-C_4H_10, and SF_6. The plates are made from a plastic with high electrical resistance; the gap between them is 2 mm, and a potential difference of 9.8 kV between the plates is maintained. The outside surfaces of the plates are painted with a graphite coating that is connected to the high-voltage power supply, and an insulating plastic coating is glued on to this layer. Metallic strips attached to the outside of this assembly provide position measurements; the strips on one side of the gas volume are orthogonal to the strips on the other, so the RPCs measure the track position in both the bending and azimuthal directions. Because the readout of the RPCs does not rely on the charge collected after ionized electrons drift to an electrode, the time resolution of these devices is quite good; the RPCs provide a space-time resolution of about 1 cm$\times$1 ns.

Thin Gap Chambers (TGCs) cover the region with $1.05 < |\eta| < 2.4$. These are multi-wire proportional chambers similar to the Cathode Strip Chambers, but here the position measurement in the bending direction is taken from the anode wires, the wire-to-wire distance is slightly larger than the wire-to-cathode distance (1.8 mm compared to 1.4 mm), and only the azimuthal coordinate is taken from the strips. The gas mixture is 55% CO_2 and 45% n-C_5H_{12}. The TGCs have an azimuthal granularity of about 2-3 mrad; the position measurement in the bending direction is obtained by reading out between 4 and 20 wires, for a precision of 7.2-36 mm.

The overall performance of the muon system is quite good; the P_T resolution, for example, is between 2% and 5% over the P_T and η range of interest. This resolution is very small compared to other effects that limit the sensitivity of the analysis.

3.4 Trigger

The large bunch-crossing rate of the LHC necessitates a trigger system that can reduce the data rate by a factor of 200,000 while still efficiently selecting interesting events. The ATLAS trigger system is broken down into three levels, denoted as Level 1 (L1), Level 2 (L2), and the Event Filter (EF). The L1 trigger is based on a reduced granularity readout of a subset of the detector systems with a fast response, such as the muon RPCs; the L1 decision is made by custom-built electronics located on the detector and off. It reduces the data rate by a factor of about 500, and it must make its decision within about 2.5 μs. In addition to an accept/reject decision, the L1 trigger identifies one or more Regions of Interest (ROIs) to be processed by the L2 trigger. Typically, the ROI will be a cone around, for example, an electron or muon candidate.

Detector information is read out with full granularity in the ROI; this constitutes about 2% of the total event size. A L2 decision is made by a processing farm comprised of commercially available computers; the decision is made within about 40 ms, and the output rate from L2 is about 3.5 kHz, or roughly a factor of 20 smaller than the L1 accept rate. The EF also runs on commercially available computers, but it is based on a full readout of the entire event and offline reconstruction algorithms. Its output rate is roughly 200 Hz, about another factor of 20 smaller than the L2 accept rate.

The analyses of $H \to WW$ discussed in this study use triggers based on the presence of electrons and muons. In particular, events are triggered if they contain one electron with $P_T > 60$ GeV, one isolated electron with $P_T > 25$ GeV, one isolated muon with $P_T > 20$ GeV, or two electrons with $P_T > 15$ GeV.[1]

---

[1] Two-muon and electron-muon triggers are forseen for ATLAS, but these triggers are omitted from the present study for technical reasons.

Chapter 4

Analysis of $H + 0j$, $H \to WW \to e\nu\mu\nu$

The $H + 0j$, $H \to WW \to \ell\nu\ell\nu$ channel is one of the most promising Higgs search channels available at the LHC. [24, 25, 26, 27] It is expected to be the first channel to be sensitive to a Standard Model Higgs, and thanks to its large event rate, it is the dominant discovery mode for a fairly broad range of masses around the WW threshold. However, because the final state contains two neutrinos, one cannot simply plot the invariant mass of Higgs candidates and expect to claim discovery based on a sharp peak. The $H + 0j$, $H \to WW$ channel is therefore an ostensibly attractive use case for multivariate techniques like Neural Networks. This chapter investigates the prospects for a multivariate analysis of this channel, and studies methods to understand the background shape and normalization using data.

Section 4.1 summarizes the Monte Carlo samples used in the analysis, and discusses the ways in which these samples are altered to compute theoretical uncertainties and simulate scenarios with degraded detector performance. Then, in Section 4.2, the loose set of cuts used to select events on which to perform a more detailed multivariate analysis is described. Section 4.3 introduces a Neural Network analysis of these events, including a strategy to normalize the dominant backgrounds using data. However, the systematic uncertainties present in this approach lead to a potentially large degradation in the sensitivity due to uncertainties in the P_T distribution of Higgs candidates. Moreover, even in an idealized case where the Higgs P_T is well-modelled by Monte Carlo, the Neural Network approach requires a few small but significant theoretical inputs in order to produce a result. Therefore, in Section 4.4, a fit-based strategy is introduced. This approach is better able to deal with uncertainties in the Higgs P_T distribution and requires fewer theoretical inputs than the Neural Network analysis or even the cut-and-count approach studied in previous works. The performance of the fitter is characterized with the help of toy Monte Carlo.

4.1 Monte Carlo

The Monte Carlo samples used in the present study are largely the same ones used in Ref. [28], augmented and updated with official fully simulated samples from the CSC production. In particular, the physics processes considered here and the generators used to model them are:

- $gg \to H \to WW$ signal. This process is modeled with MC@NLO [29, 30], both in fast simulation and in full simulation (ATLAS central production dataset 6533).

- VBF $H \to WW$. Pythia 6.2 [31, 32] is used to model this process in fast simulation; Pythia 6.4 and Sherpa [33] are used to model it in full simulation (ATLAS central production datasets 5322 and 6325, respectively).

- QCD $qq/qg \to WW$: MC@NLO is used to model this process in both fast simulation and full simulation (ATLAS central production datasets 5921-5929).

The fast simulation analysis makes use of large samples of events with altered Q^2 scales and with the CTEQ6 error PDF sets [34] to assess the uncertainty in the prediction of this process.

- $gg \to WW$ via quark box: The gg2WW matrix element calculation from Reference [35], interfaced to Herwig [36] for hadronization, is used here. The version of the program used for fast simulation does not include heavy flavor loops, but the version used for full simulation (datasets 2821-2829, from private production [37] but validated against samples from central production) does include them.

- $t\bar{t}$: The prediction of MC@NLO is taken as a baseline estimate of the top background in both fast simulation and full simulation (dataset 5200). Samples of $pp \to WWb\bar{b}$ and $pp \to t\bar{t} \to WWb\bar{b}$ generated with MadGraph are also studied (in fast simulation) in order to assess the impact of the Wt background.

- Drell Yan production: MC@NLO is used to model this process in fast simulation. The analysis based on full simulation only deals with the $H \to WW \to e\nu\mu\nu$ channel, so only the $Z \to \tau\tau$ background is considered there. The $Z \to \tau\tau$ Monte Carlo sample used for the full simulation analysis was ATLAS central production dataset 5146, generated with Pythia.

- $b\bar{b}$: The prediction of ALPGEN [38, 39, 40] is used to model this background in fast simulation. Due to the large computational requirements, this background is not studied in full simulation.

- W+jets. The prediction of ALPGEN is used to model this process in full simulation. Since fake electrons arising from jets are not modeled by the fast simulation, this process is studied only in full simulation. The Monte Carlo samples used are ATLAS datasets 6101-6118.

The fast simulation samples used in this study were generated using the last version of Atlfast-Fortran [41]. The studies presented here use an electron identification efficiency (not including the isolation efficiency) of 60.3%, and a muon identification efficiency of 92.5%, as suggested by recent full simulation studies [42]. Full simulation samples were generated using Athena versions 12.0.6 and 12.0.7.

In addition to the theoretical uncertainties discussed for some of the samples above, two scenarios involving degraded detector performance are considered:

- Raised/lowered jet energy scale: all components of jet four-momenta are multiplied by 1.05 or 0.95 for jets in the region with tracking; they are multiplied by 1.1 or 0.9 for jets elsewhere. Since this alteration is intended primarily to study the uncertainty in the jet veto, the changes in the jet P_T measurements are not propagated to the missing P_T.

- Smeared missing P_T: For each event, independent random numbers sampled from a gaussian centered at 0 with a width of 5 GeV are added to the x and y components of the missing P_T. This emulates certain degradations in the performance of the missing P_T reconstruction, but a similar effect could also arise from pileup or additional underlying event activity.

Cut	$gg \to H$	VBF	$t\bar{t}$	WW	$Z \to \tau\tau$	$b\bar{b}$
Lepton Selection	178.10	20.60	7707	875.97	4888	2828.75
$P_T^{miss} > 30$ GeV	139.54	18.20	6655	617.72	432.9	152.00
$Z \to \tau\tau$ Rej.	136.07	17.56	6200	593.98	104.0	130.16
Jet Veto	64.11	1.00	42.08	329.59	21.62	3.68
b-veto	63.83	0.99	30.67	328.33	21.21	3.68
$\Delta\phi_{ll} < 1.575$, $M_T < 600$ GeV	44.47	-	7.99	104.9	-	-
$\Delta\phi_{ll} > 1.575$, $M_T < 600$ GeV	20.07	-	21.37	171.6	21.16	-
b-tagged, $\Delta\phi_{ll} < 1.575$	-	-	2.93	-	-	-
b-tagged, $\Delta\phi_{ll} > 1.575$	-	-	7.94	-	-	-

Table 4.1: Cut flows (in fb) for $M_H = 170$ GeV in the $H + 0j$, $H \to WW \to e\nu\mu\nu$ channel in fast simulation. A '-' indicates that the corresponding contribution is ignored in the fit.

Cut	$gg \to H$	$t\bar{t}$	WW	$Z \to \tau\tau$	$W + jets$
Lepton Selection+M_{ll}	169.0	6501	718.12	4171	209.1
$P_T^{miss} > 30$ GeV	133.2	5617	505.25	526.3	181.6
$Z \to \tau\tau$ Rej.	129.8	5215	485.12	164.2	150.4
Jet Veto	52.85	14.84	238.35	31.91	76.12
b-veto	52.62	6.85	237.87	30.76	76.12
$\Delta\phi_{ll} < 1.575$, $M_T < 600$ GeV	35.7±1.1	2.3±1.6	85.4±2.7	<1.7	38±38
$\Delta\phi_{ll} > 1.575$, $M_T < 600$ GeV	16.9±0.7	4.6±2.3	151.9±3.6	30.8±4.2	38±38
b-tagged, $\Delta\phi_{ll} < 1.575$	-	1.14±1.14	-	-	-
b-tagged, $\Delta\phi_{ll} > 1.575$	-	5.71±2.55	-	-	-

Table 4.2: Cut flows (in fb) for $M_H = 170$ GeV in the $H + 0j$, $H \to WW \to e\nu\mu\nu$ channel in full simulation. A '-' indicates that the corresponding contribution is ignored in the fit. The WW background contains the two processes $q\bar{q} \to WW$ and $gg \to WW$.

4.2 Event Selection

The event selection consists of only a few simple cuts to select events with two leptonically decaying W bosons:

- Require that the event has exactly two isolated leptons (electron or muon) with $P_T > 15$ GeV. In the fast simulation, leptons have been required to pass an isolation criteria by default. In the full simulation, electrons are taken to be electromagnetic clusters that have isEM&0xfff==false and a matched track, with the energy measurement taken from the cluster and the η and ϕ measurements taken from the track. Muons are reconstructed using the STACO package, and for all leptons a track isolation is computed as the scalar sum(P_T) of all tracks in a cone of 0.4 units of R, excluding tracks within a cone of 0.01 units of R frm any reconstructed lepton. In order for a lepton to be considered 'isolated', this sum(P_T) isolation variable must be less than 5 GeV if the lepton is an electron, or 3 GeV if it is a muon. A cut is also placed on calorimeter isolation; the $etcone20$ variable is required to be less than 5 GeV for electrons or 2.5 GeV for muons. A similar set of isolation criteria has been shown to suppress backgrounds from $W + c$ and $W + b$ production [43, 44]. To suppress backgrounds from $b\bar{b}$ resonances, $c\bar{c}$ resonances, and single-top production, the invariant mass of the leptons is required to be between 12 GeV and 300 GeV, and to prevent outliers from causing trouble in Neural Network training, events are rejected if either lepton has $P_T > 400$ GeV.

- Require that the event has $P_T^{miss} > 30$ GeV if the leptons have opposite flavor, or $P_T^{miss} > 40$ GeV if they have the same flavor. To exclude outliers, require that the transverse mass M_T is less than 1 TeV.

- To suppress backgrounds from $Z \to \tau\tau$, reconstruct the invariant mass of a hypothetical τ pair using the collinear approximation. If the energy fractions x_τ^1 and x_τ^2 are both positive and the invariant mass of the τ pair $M_{\tau\tau}$ is in the range $|M_{\tau\tau} - M_Z| < 25$ GeV, reject the event.

- Reject events with any hard jets. In the fast simulation, the hadron-to-parton correction of ATLFAST-B is applied, the jet veto P_T threshold is 30 GeV, and the fake veto correction of Ref. [45] is applied. In the full simulation, no corrections are applied, and the jet veto P_T threshold is 20 GeV.

- Reject events with any b-tagged jets. In the fast simulation the b-tagging efficiency is taken to be 60% for jets with $P_T > 20$ GeV (after hadron-to-parton correction) and the rejection against mistags is assumed to be 10 for c-jets and 100 for light jets. In the full simulation, events are rejected where any jet has a b-tagging weight greater than 4.

Table 4.1 shows the cross-sections for signal and background after these cuts, obtained using fast simulation. Table 4.2 shows the numbers obtained using full simulation. Events that pass this set of preselection cuts are used in the Neural Network analysis or the fit-based analysis described in the sections that follow.

The trigger efficiency for signal has been computed and it is sufficiently large for this analysis to be viable. For the $H \to WW \to e\nu\mu\nu$ channel, the trigger efficiency for L1 is 99.0%; for L2 it is 96.7%, and for the EF it is 95.2%. Figure 4.1 shows comparisons of two of the main discriminating variables used in the fit described in

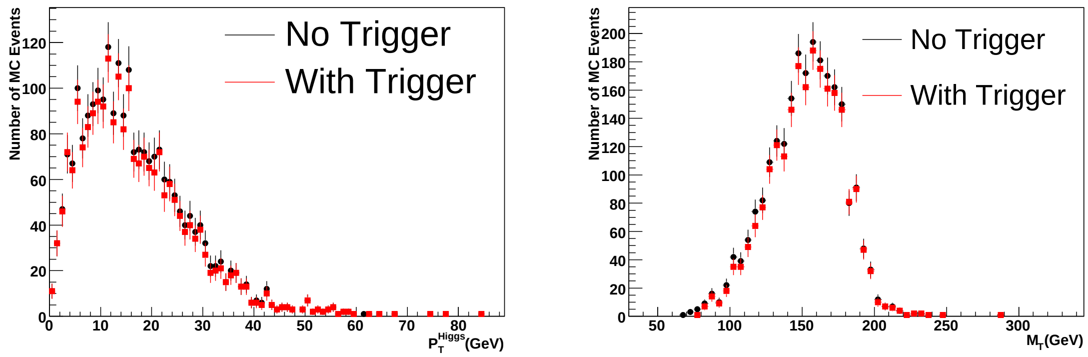

Figure 4.1: Left: the Higgs P_T distribution for the $H \to WW \to e\nu\mu\nu$ channel after preselection cuts, with and without the trigger efficiency. Right: the M_T distribution for the $H \to WW \to e\nu\mu\nu$ channel after preselection cuts, with and without the trigger efficiency.

Section 4.4, the transverse momentum of the Higgs candidate and the transverse mass. Neither is affected in a significant way by the trigger efficiency.

4.3 Neural Network Analysis

A tempting approach for the analysis of this channel is to use a Neural Network to discriminate between signal and background.[1] In this section, a 5-variable Neural Network is discussed. The Neural Network takes the following variables as input:

- P_T^{Higgs}, the P_T of the WW system
- P_T^{l1} and P_T^{l2}, the transverse momenta of the two leptons
- $\Delta\eta_{ll}$, the pseudorapidity gap between the two leptons
- $\Delta\phi_{ll}$, the dilepton opening angle in the transverse plane

[1]While this work was in preparation, another multivariate analysis using Boosted Decision Trees was published in Ref. [46].

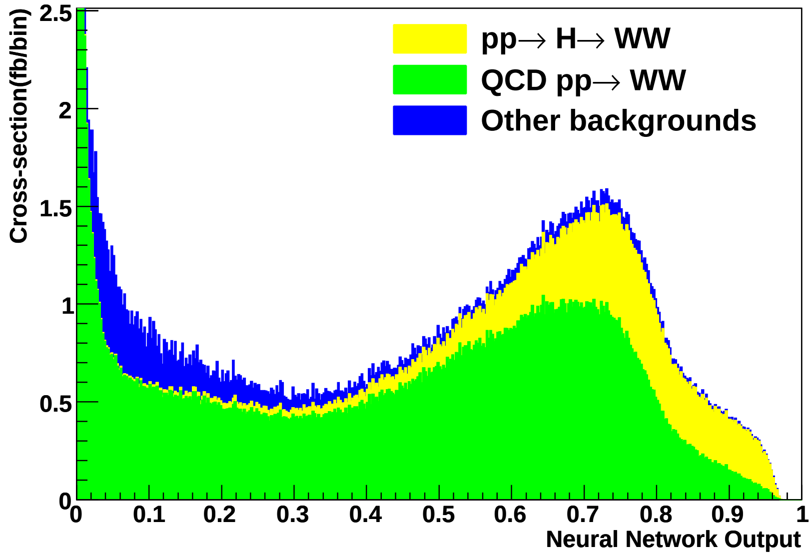

Figure 4.2: The Neural Network output distribution for the $H \to WW \to e\nu\mu\nu$ channel, for a representative Higgs mass of 160 GeV.

The Neural Network is a feed-forward Neural Network with two hidden layers of 6 neurons each; it is trained on a sample of events consisting of 25,000 signal events and 25,000 background events, with contributions from the various background processes according to cross-section. Independent Neural Networks are trained for Higgs mass hypotheses between 120 and 200 GeV, in steps of 5 GeV below 150 GeV and in steps of 10 GeV above, with the intent that in data one would select the mass hypothesis with the largest excess.

Figure 4.2 shows a plot of the Neural Network output for a representative Higgs mass hypothesis of 160 GeV, and Table 4.3 shows the cross-sections remaining after an optimal cut on the Neural Network output for a few different choices of the Higgs mass hypothesis. The exact shape of the Neural Network output varies somewhat between Higgs mass hypotheses, but in general the signal tends to be concentrated near 1. It

M_H (GeV)	Optimal Cut	VBF	$gg \to H$	QCD WW	$t\bar{t}$	$Z \to \tau\tau$
120	0.8475	0.05	3.49	22.87	1.45	0.34
125	0.8	0.11	7.55	39.81	2.65	0.58
130	0.7775	0.15	9.41	36.02	2.23	0.32
135	0.76	0.24	14.32	46.68	3.21	0.46
140	0.725	0.34	21.48	62.33	3.83	0.36
145	0.7175	0.44	24.68	60.61	4.59	0.38
150	0.67	0.54	29.74	63.77	4.88	0.44
160	0.7525	0.56	27.27	26.05	2.26	0.03
170	0.7225	0.69	31.67	38.75	3.78	0.08
180	0.6825	0.61	28.08	50.77	5.14	0.07
190	0.6575	0.44	20.54	55.32	6.62	0.08
200	0.67	0.35	16.02	48.66	7.56	0.14

Table 4.3: Optimal values for the cut on the Neural Network output and the cross-sections for signal and background in the signal-like region (in the $H + 0j$, $H \to WW \to e\nu\mu\nu$ channel) after cutting on the Neural Network output.

M_H (GeV)	Optimal Cut	VBF	$gg \to H$	QCD WW	$t\bar{t}$	$Z \to ll$
120	0.8475	0.01	0.73	4.96	0.27	0.31
125	0.8	0.02	1.54	8.71	0.47	0.62
130	0.7775	0.04	2.34	9.42	0.59	0.78
135	0.76	0.05	2.85	8.83	0.55	0.46
140	0.725	0.09	4.99	13.74	1.00	0.63
145	0.7175	0.10	5.38	11.81	0.83	0.31
150	0.67	0.15	8.36	18.01	1.26	0.47
160	0.7525	0.16	7.80	6.78	0.47	0.38
170	0.7225	0.18	8.39	8.91	0.77	0.69
180	0.6825	0.18	7.77	12.64	1.34	0.76
190	0.6575	0.12	5.16	12.74	1.64	0.92
200	0.67	0.07	3.08	7.96	1.21	1.07

Table 4.4: Optimal values for the cut on the Neural Network output and the cross-sections for signal and background in the signal-like region (in the $H + 0j$, $H \to WW \to e\nu e\nu$ channel) after cutting on the Neural Network output.

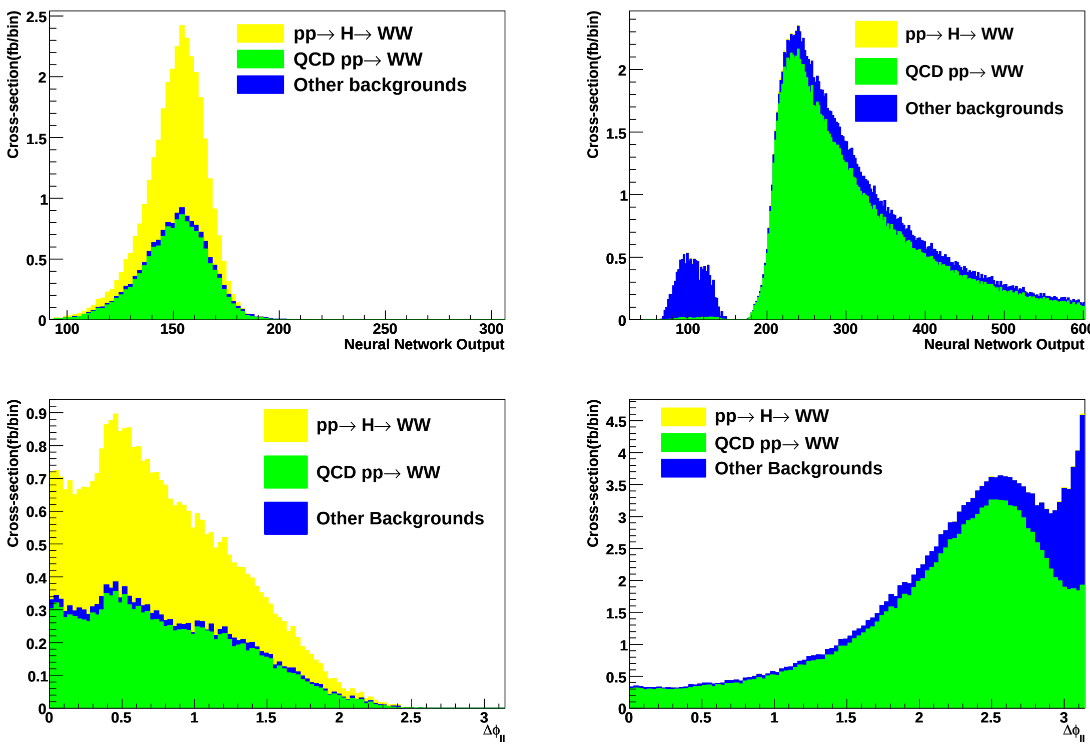

Figure 4.3: Upper left: The M_T distribution for events with Neural Network output larger than 0.8 for a representative Higgs mass of 160 GeV. Upper right: the corresponding distribution for events with Neural Network output less than 0.1. Similar distributions are shown for $\Delta\phi_{ll}$ in the second row.

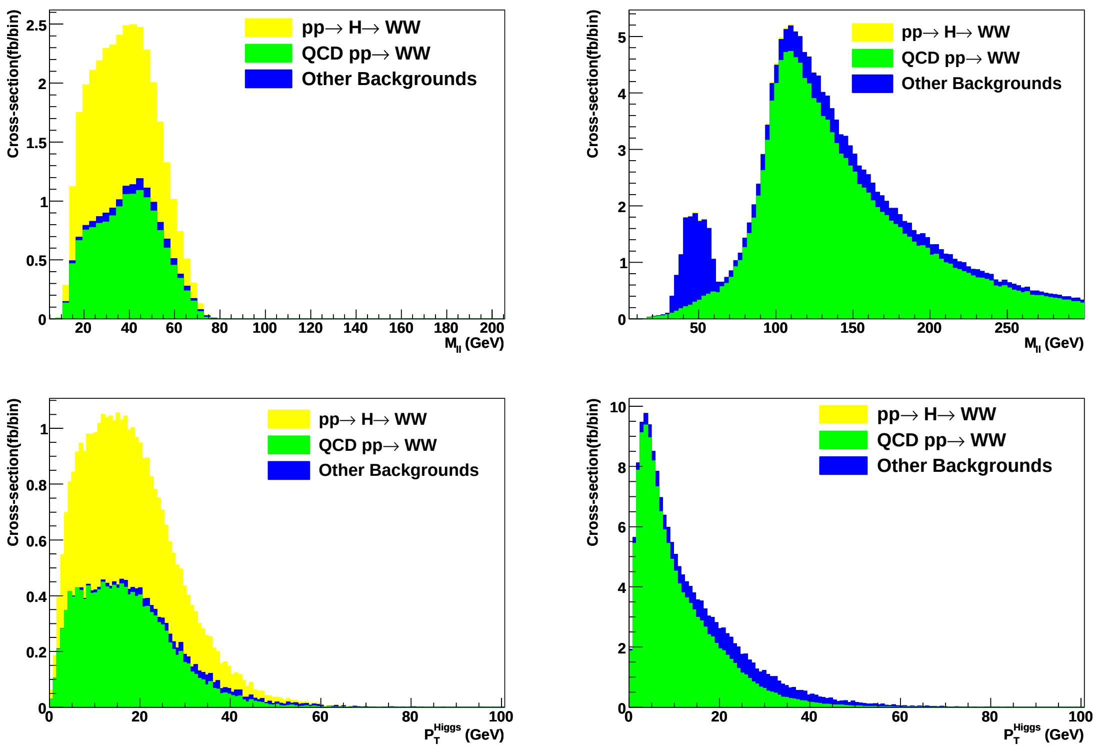

Figure 4.4: Upper left: The M_{ll} distribution for events with Neural Network output larger than 0.8 for a representative Higgs mass of 160 GeV. Upper right: the corresponding distribution for events with Neural Network output less than 0.1. Similar distributions are shown for P_T^{WW} in the bottom row.

M_H (GeV)	Optimal Cut	VBF	$gg \to H$	QCD/EW WW	$t\bar{t}$	$Z \to ll$
120	0.8475	0.02	1.35	8.62	0.30	0.24
125	0.8	0.06	4.16	22.32	0.94	0.70
130	0.7775	0.10	5.71	22.12	0.98	0.48
135	0.76	0.14	7.72	23.96	1.03	0.62
140	0.725	0.22	11.98	31.90	1.64	0.65
145	0.7175	0.26	11.88	24.25	1.62	0.59
150	0.67	0.39	20.37	42.05	2.26	1.01
160	0.7525	0.37	15.54	10.81	0.75	0
170	0.7225	0.40	15.99	13.30	1.10	0.39
180	0.6825	0.43	17.87	27.90	2.55	1.74
190	0.6575	0.31	13.02	32.60	3.54	1.97
200	0.67	0.19	8.66	24.72	3.18	1.93

Table 4.5: Optimal values for the cut on the Neural Network output and the cross-sections for signal and background in the signal-like region (in the $H + 0j$, $H \to WW \to \mu\nu\mu\nu$ channel) after cutting on the Neural Network output.

is easy to find out which regions of phase space are classified as signal-like by the Neural Network; one simply plots the distributions of important variables for events that survive a cut requiring that the Neural Network be large. The upper left plot in Figure 4.3 shows the transverse mass distribution for events with Neural Network output larger than 0.8; as one would expect, events classified as signal-like tend to have M_T close to 160 GeV. Similarly, the other plots on the left side of Figures 4.3 and 4.4 show that events with large Neural Network output have small dilepton opening angle in the transverse plane, a dilepton invariant mass less than about 60-70 GeV, and a Higgs P_T that is less concentrated at zero than the background. These trends are generally similar to the cuts one would apply in a number-counting analysis.

Background normalization is a crucial part of any analysis of $H + 0j$, $H \to WW \to l\nu l\nu$. In the Neural Nework approach, the background can be normalized by

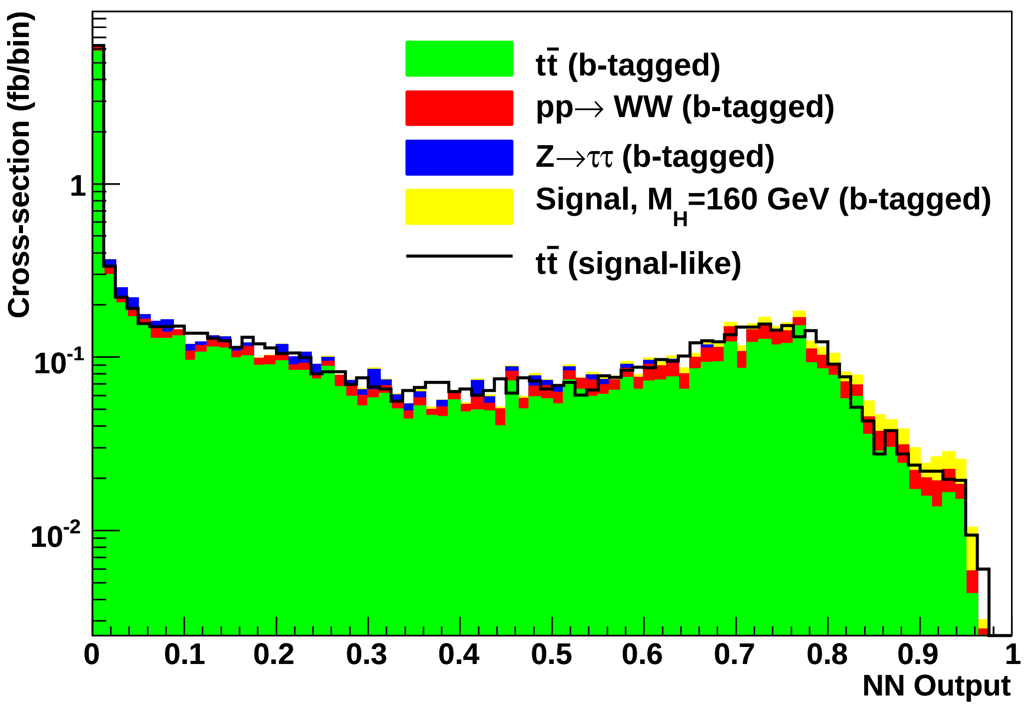

Figure 4.5: The Neural Network output distribution for events in the b-tagged region (shaded areas) and in the b-vetoed region (black line), for a representative Higgs mass hypothesis of 160 GeV.

counting events in the region with Neural Network output less than 0.1. The plots on the right side of Figures 4.3 and 4.4 show the same example variables (M_T, $\Delta\phi_{ll}$, and M_{ll}) for events that pass a cut requiring that the Neural Network output be less than 0.1. They show that the control region consists of events that have large transverse mass, large dilepton opening angle in the transverse plane, large dilepton invariant mass, and a P_T^{WW} distribution that peaks near zero.

The control sample strategy studied previously in Ref. [28] used the same cut on P_T^{Higgs} in the control region and the signal-like region. This allowed uncertainties in the Higgs P_T distribution to cancel out in that analysis. The bottom plots in Figure 4.4 highlight the fact that this is not the case in the Neural Network analysis. The signal-like region consists of events with large Higgs P_T, while the control region consists primarily of events with small Higgs P_T. As a consequence, the Neural Network analysis has better statistics in the control region but any uncertainty that affects the Higgs P_T distribution will no longer cancel out. The top background is normalized with the help of a b-tagged control sample. Table 4.1 shows the contributions of the various background processes to this region. A feature of this control sample is that the Neural Network output in this control region has roughly the same shape as the Neural Network output in the b-vetoed region; see Figure 4.5.

Because the background is normalized using a control sample, it is necessary to consider systematic uncertainties in the shape of the Neural Network output distribution for background. A ratio plot showing a systematically distorted histogram divided by the undistorted histogram (after normalizing both histograms to 1) can provide some intuition about how a systematic uncertainty affects the distribution.

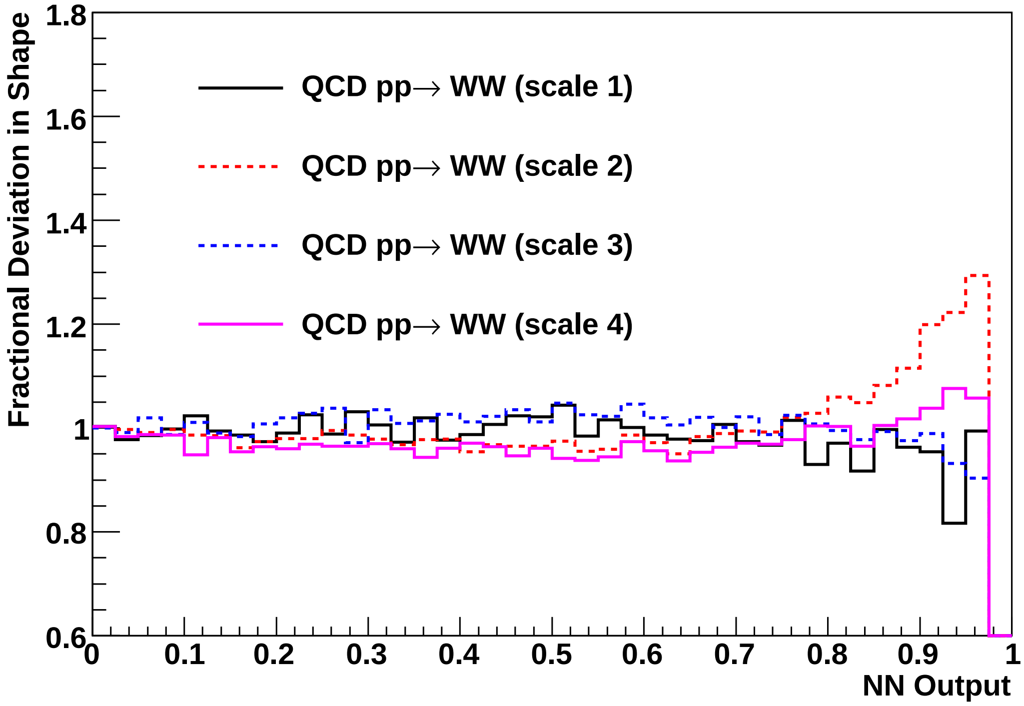

Figure 4.6: The ratio of the Neural Network Output distributions (for a Higgs mass hypothesis of 160 GeV) obtained by dividing each of the scale-varied QCD WW Monte Carlo samples by the distribution from the central-value QCD WW Monte Carlo sample. Here, the distributions have been normalized to equal area below a Neural Network output of 0.1 before taking the ratio.

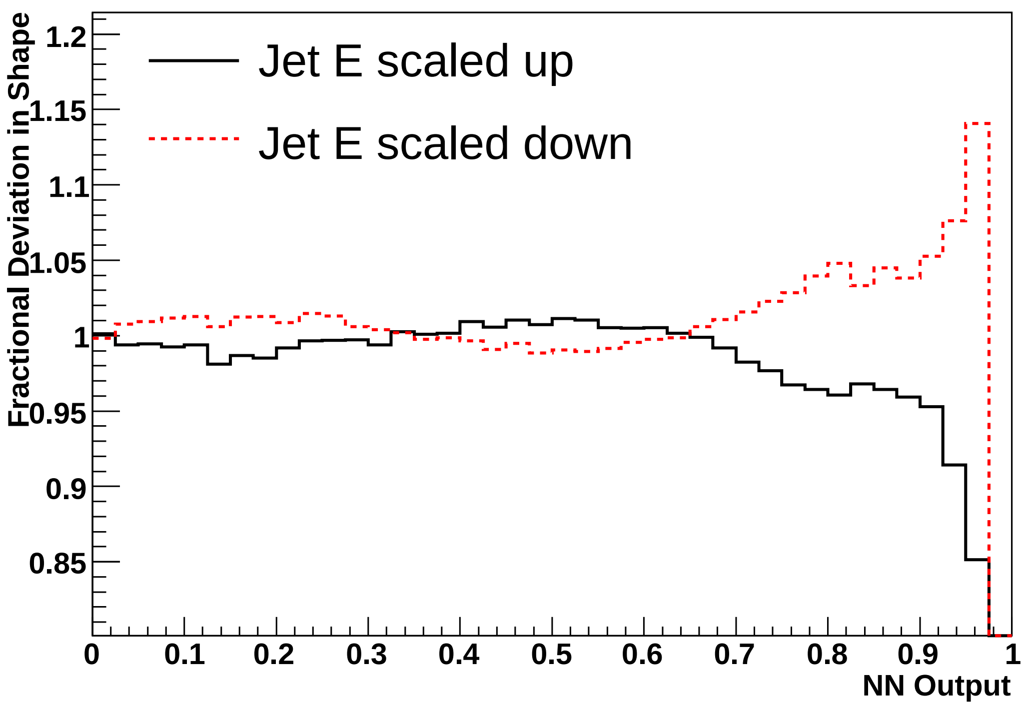

Figure 4.7: The ratio of the Neural Network Output distributions (for a Higgs mass hypothesis of 160 GeV) obtained by dividing each of the jet-energy-scale-varied QCD WW Monte Carlo samples by the distribution from the QCD WW Monte Carlo sample that assumes nominal detector performance. Here, the distributions are normalized to equal area below a Neural Network output of 0.1 before taking the ratio.

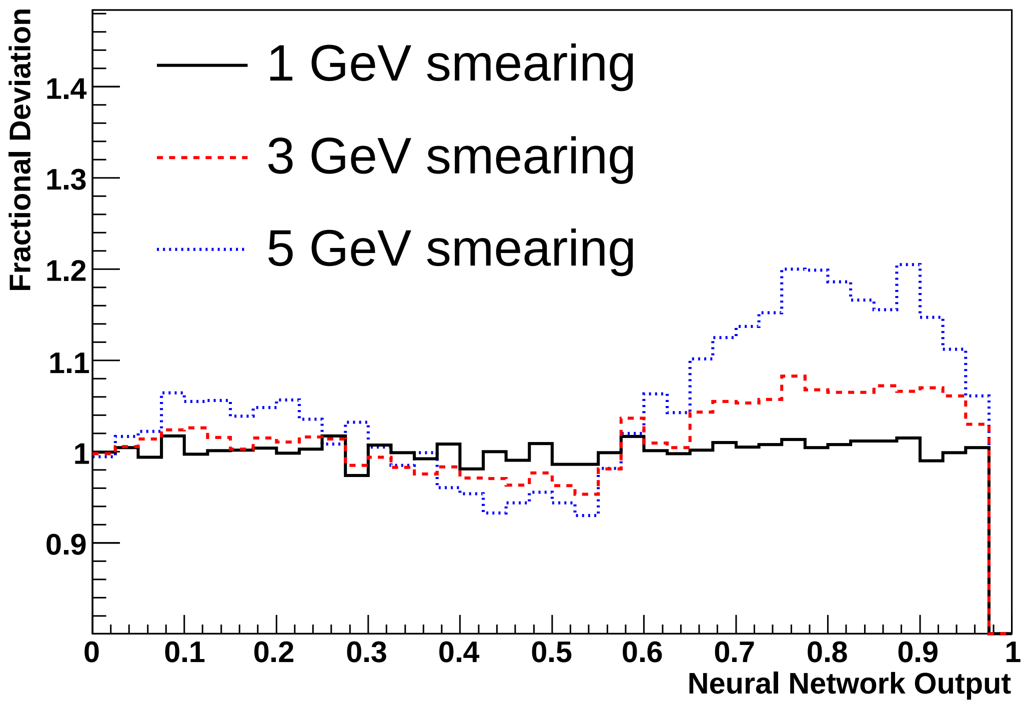

Figure 4.8: The ratio of the Neural Network Output distributions (for a Higgs mass hypothesis of 160 GeV) obtained by dividing the distribution from QCD WW Monte Carlo samples that have a degraded missing P_T by the distribution from the QCD WW Monte Carlo sample that assumes nominal detector performance. Here, the distributions are normalized to equal area below a Neural Network output of 0.1 before taking the ratio.

Figure 4.6 shows this ratio plot for each of the four altered Q^2 scale choices considered here, for a Higgs mass hypothesis of 160 GeV. The uncertainty becomes quite large in the region with Neural Network output close to 1, but it is worth noting that in this region the cross-section for background becomes quite small. Figure 4.7 shows similar ratio plots for the case of raised and lowered jet energy scales; the conclusion is roughly the same.

M_H (GeV)	WW	$t\bar{t}$	$Z \to \tau\tau$
120	215.27	23.18	3.57
125	198.90	22.22	3.21
130	189.49	21.46	4.34
135	174.68	19.83	3.14
140	165.53	19.95	6.76
145	152.33	19.31	8.19
150	145.99	17.74	10.34
160	136.47	17.40	10.55
170	132.04	17.64	13.04
180	103.66	14.88	10.45
190	83.30	13.04	12.46
200	66.37	10.64	11.65

Table 4.6: Cross-sections (in fb) in the $e\mu$ channel for various physics processes in the region with Neural Network output less than 0.1, for the full range of Higgs masses under consideration.

Figure 4.8 shows the change in the shape of the Neural Network output distribution in the cases where the missing P_T resolution has been degraded by smearing its x and y components by 1, 3, and 5 GeV. The change in the shape of the distribution is quite large, and it is not restricted to regions where the background cross-section is small. In other words, an inaccurate prediction of the Higgs P_T spectrum can cause a significant systematic uncertainty that will lead to a serious degradation of the sensitivity of the analysis. It is therefore highly desirable to constrain the Higgs P_T spectrum using data.

For each Higgs mass hypothesis, a cut is placed on the Neural Network output, requiring it to be larger than some value that depends on the Higgs mass hypothesis. The uncertainty on the ratio of the background cross-section in the region that passes the cut divided by the cross-section in the region with Neural Network output less than

M_H (GeV)	Optimal Cut	Jet E. scale(%)	Q^2(%)	PDF(%)	total
120	0.8475	1.66	2.17	1.74	3.24
125	0.8	1.91	2.03	1.82	3.33
130	0.7775	2.22	1.74	1.85	3.37
135	0.76	2.42	2.64	1.74	3.98
140	0.725	1.66	2.66	1.71	3.57
145	0.7175	2.30	2.73	1.91	4.05
150	0.67	2.68	2.76	1.98	4.33
160	0.7525	4.01	5.90	2.48	7.55
170	0.7225	4.06	6.17	2.32	7.74
180	0.6825	3.26	5.02	2.50	6.49
190	0.6575	3.20	4.35	2.42	5.92
200	0.67	4.31	3.78	2.64	6.31

Table 4.7: Optimal values for the cut on the Neural Network output and the uncertainties due to the jet energy scale, Q^2 scale, and PDF fit.

0.1 is computed. Table 4.3 shows the cross-section for the $e\mu$ channel in the signal-like region as a function of the Higgs mass hypothesis. Tables 4.4 and 4.5 show the results for the same-flavor channels, and Table 4.6 shows the cross-sections for the $e\mu$ channel in the control region. Table 4.7 shows the uncertainties in the ratio of background cross-sections in the signal-like and control regions arising from the jet energy scale, from the Q^2 scale, and from parton density functions.[2] The signal-to-background ratio is promising, and the systematic error in the background determination due to the Q^2 scale, jet energy scale, and parton density function uncertainties is at the level of less than 8%.

[2]These uncertainties are derived from studies of the $e\mu$ channel only.

4.4 Fit-based Analysis

The Neural Network analysis presented above relies heavily on accurate Monte Carlo. Especially during the early stages of data-taking, this is an undesirable feature. This section describes a fit-based approach to extract the background. It is shown that the fit is robust against the systematic uncertainties that were largest for the Neural Network analysis. The present study only considers a fit to those events with one electron and one muon, since an in-situ background extraction procedure has not been implemented for the $Z \to ee/\mu\mu$ background.

The fit is a 2-dimensional fit of transverse mass and Higgs P_T in two bins of the dilepton opening angle in the transverse plane, implemented using the RooFit toolkit [47]. After the preselection cuts of Section 4.2 are applied, the surviving events are separated into two subsamples, one with $\Delta\phi_{ll} < 1.575$ and the other with $\Delta\phi_{ll} > 1.575$. Table 4.2 shows the cross-sections for the various signals and backgrounds in these two regions. The region with large $\Delta\phi_{ll}$ is enriched in background and the region with small $\Delta\phi_{ll}$ is enriched in signal.

The details of the functional forms used in the fit are described in Appendix A; the discussion here will focus on conceptual aspects of the background extraction and the performance of the fitter on toy Monte Carlo. Section 4.4.1 discusses the procedure for normalizing the top background and how the fitter deals with the uncertainties in the top background prediction. Section 4.4.2 describes the procedure to normalize the $Z \to \tau\tau$ background, and Section 4.4.3 discusses the uncertainties in the QCD WW background and how the fitter deals with them. Finally, Section 4.4.4 evaluates the performance of the entire fitting algorithm using toy Monte Carlo.

4.4.1 Normalization of the Top Background

Upon inspection of the cross-sections in Table 4.1, one would be tempted to conclude that the normalization of the top background is a minor concern for the analysis. However, it is important to bear in mind that there are nontrivial uncertainties on the prediction of the top background. First and foremost, the Monte Carlo generator used as a baseline description of the top background ignores contributions from off-shell top quarks; however, one would expect that off-shell contributions to the top background would have a higher probability to survive the selection cuts (particularly the jet veto) than the doubly-resonant contribution would. One would therefore expect that the introduction of a complete and accurate description of the off-shell contributions to the top background would affect not only the overall normalization of the background, but also the ratio of the number of events in the b-tagged and b-vetoed regions as well as the distributions of important kinematic variables like Higgs P_T, M_T, and $\Delta\phi_{ll}$. Since the signal cross-section drops rapidly for Higgs masses that are not close to 160 GeV, a bias in the fitted distribution for the top background in these variables could in principle lead to significant problems in the fit for signals that are not close to the WW threshold. In order to be sensitive to Higgs signals at these masses, it is therefore necessary to ensure that the top background model is flexible enough to describe a reasonable range of distortions in the shapes of the key kinematic variables and in the relative normalization of the b-tagged and b-vetoed regions.

The b-tagged regions used to normalize the top background have the same kinematic cuts as the two main regions of the analysis, i.e. two leptons, missing transverse momentum, a $Z \rightarrow \tau\tau$ rejection cut, and a veto on events with hard jets. Table 4.1

Process	$\Delta\phi_{ll} < 1.575$	$\Delta\phi_{ll} > 1.575$
$pp \rightarrow t\bar{t} \rightarrow WWb\bar{b}$, b-vetoed	5.51	18.77
$pp \rightarrow t\bar{t} \rightarrow WWb\bar{b}$, b-tagged	1.87	7.03
$pp \rightarrow WWb\bar{b}$, b-vetoed	7.70	26.37
$pp \rightarrow WWb\bar{b}$, b-tagged	2.47	8.81

Table 4.8: The top background cross-sections (in fb) in several regions of the $H + 0j$, $H \rightarrow WW \rightarrow e\nu\mu\nu$ analysis, as predicted by the leading-order samples used for systematic studies.

shows the cross-sections in the b-tagged regions as obtained from fast simulation. In this study, the contamination from processes with only light jets in these regions is ignored. The cross-section estimates in these tables are based on Monte Carlo samples that use MC@NLO to model the top background; the results from calculations based on leading-order MadEvent samples with and without off-shell contributions are shown in Table 4.8.

One shortcoming of the b-tagged control samples used here is that the shape of the distribution of M_T and P_T^{Higgs} in the b-tagged region is not always exactly the same as the corresponding distribution in the b-vetoed region. Figure 4.9 shows the ratio of the transverse mass shape in the b-tagged and b-vetoed regions with $\Delta\phi_{ll} < 1.575$, for several bins of Higgs P_T. Figure 4.10 shows similar ratios for the Higgs P_T distribution in bins of M_T in the same region. There is a slight slope in all the distributions, but there is no significant difference among the results from the three models of the top background considered. However, Figures 4.11 and 4.12 show that this is not the case in the in the region with $\Delta\phi_{ll} > 1.575$. The inclusion of off-shell effects results in a significant change in the prediction of the ratio in the region with large M_T and small Higgs P_T, as well as a somewhat smaller change in the region with

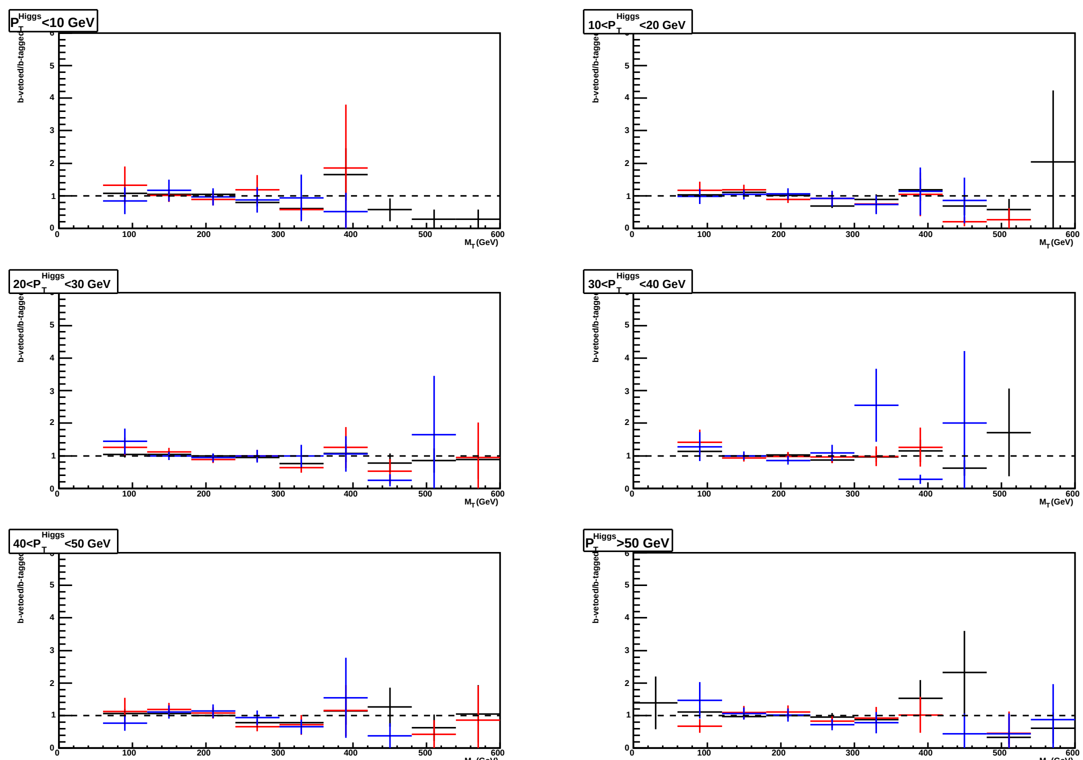

Figure 4.9: The ratio of the transverse mass shape for the top background in the b-vetoed region over the b-tagged region, in bins of Higgs P_T. Here, the distributions are shown for the region with small $\Delta\phi_{ll}$. Both numerator and denominator are normalized to 1 before taking the ratio.

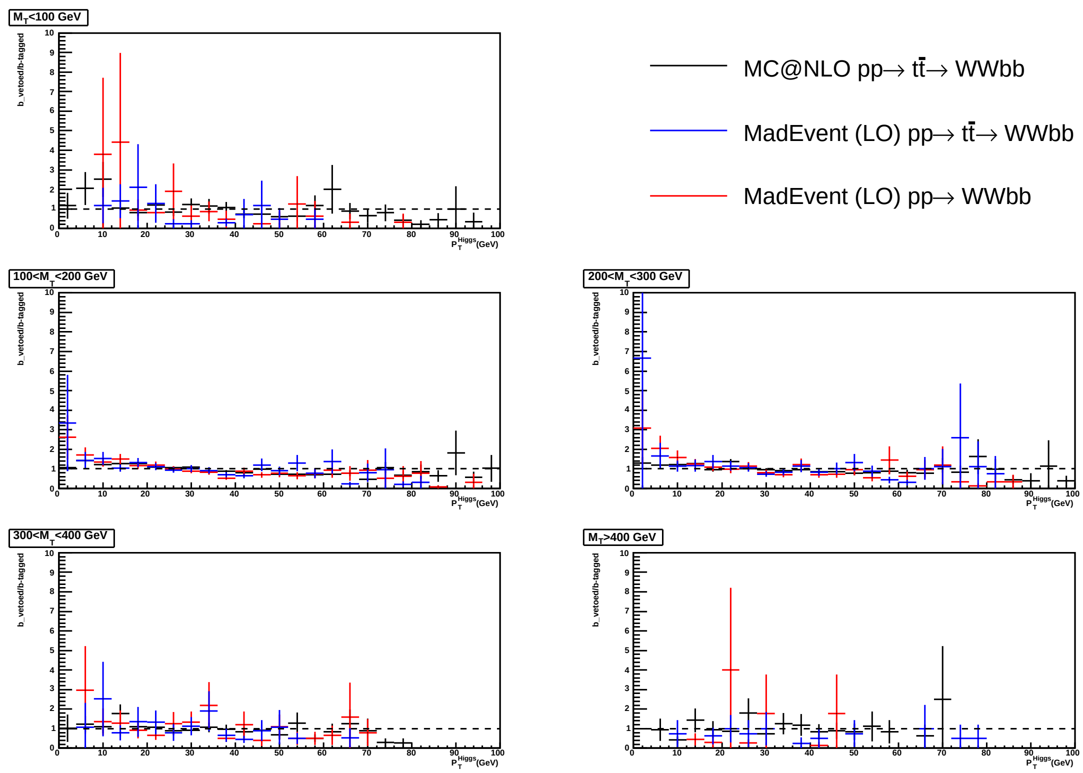

Figure 4.10: The ratio of the Higgs P_T shape for the top background in the b-vetoed region over the b-tagged region, in bins of transverse mass. Here, the distributions are shown for the region with small $\Delta\phi_{ll}$. Both numerator and denominator are normalized to 1 before taking the ratio.

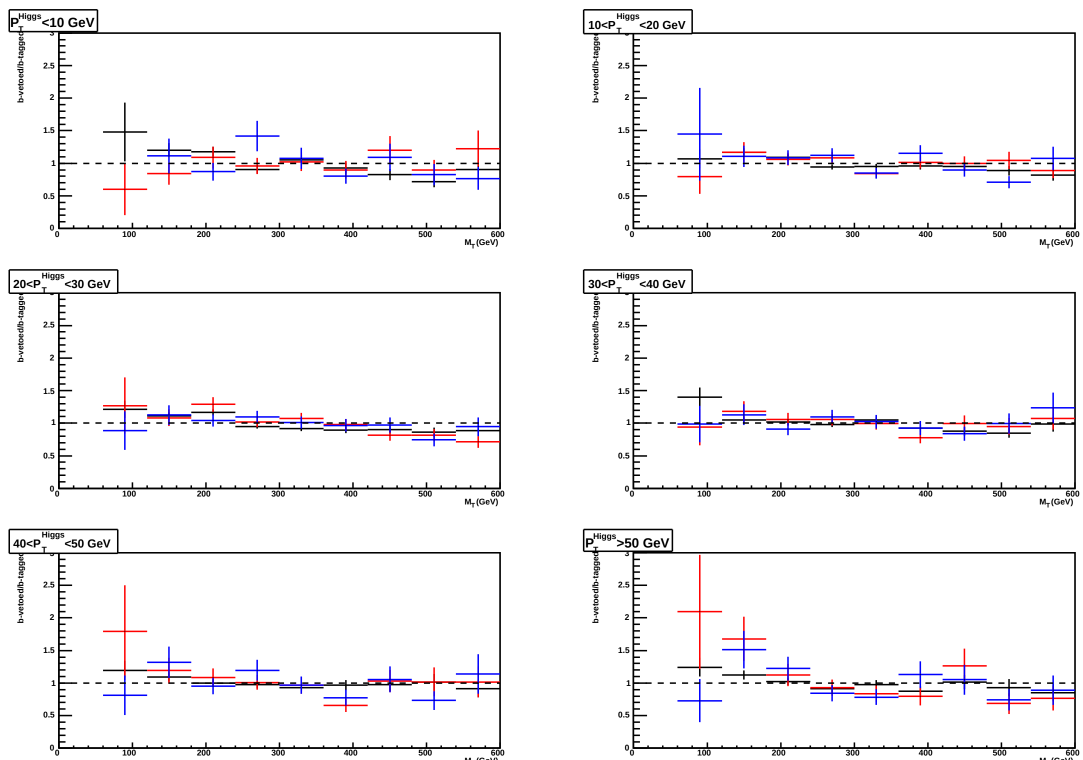

Figure 4.11: The ratio of the transverse mass shape for the top background in the b-vetoed region over the b-tagged region, in bins of Higgs P_T. Here, the distributions are shown for the region with large $\Delta\phi_{ll}$. Both numerator and denominator are normalized to 1 before taking the ratio.

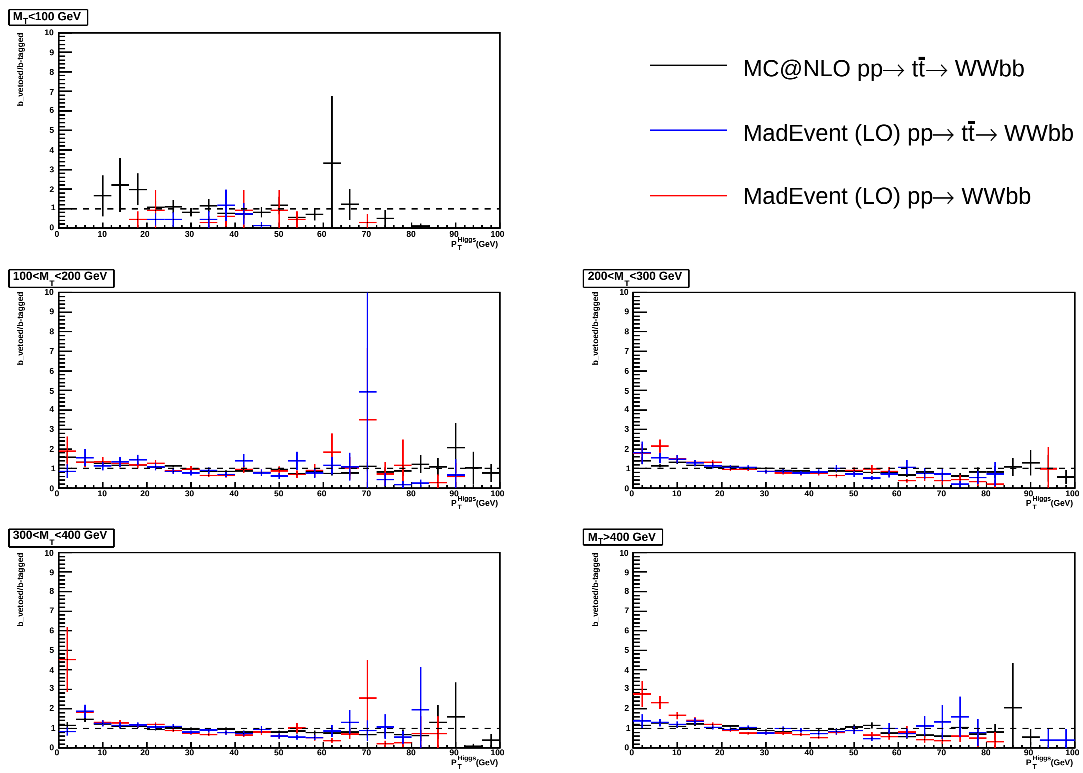

Figure 4.12: The ratio of the Higgs P_T shape for the top background in the b-vetoed region over the b-tagged region, in bins of transverse mass. Here, the distributions are shown for the region with large $\Delta\phi_{ll}$. Both numerator and denominator are normalized to 1 before taking the ratio.

large Higgs P_T and small M_T. In the fit, the shape of the M_T vs. P_T^{Higgs} distribution for the top background in the b-vetoed region is therefore distorted by multiplying the distribution obtained in the b-tagged region with a function that depends on several free parameters.

In practice, then, the procedure is as follows: perform an independent fit to the b-tagged region with small $\Delta\phi_{ll}$ and the b-tagged region with large $\Delta\phi_{ll}$. All but two of the parameters are fixed. The result obtained from these fits is the best-fit distributions and normalizations in the b-tagged regions. To extrapolate to the b-vetoed region, the distribution is distorted by multiplying it with an extrapolation function consisting of a product of exponentials in M_T and P_T^{Higgs}, with correction factors to model some small distortions in the distribution. In the main fit to the b-vetoed region, the parameters obtained from the fit to the b-tagged regions are fixed, but some of the parameters related to the extrapolation are allowed to float. In the present analysis, it is assumed that the b-tagging efficiency is well-understood.

4.4.2 Normalization of the $Z \rightarrow \tau\tau$ Background

Like the top background, the $Z \rightarrow \tau\tau$ background appears at first glance to be a minor concern: the cross-section in the signal-like region is so small that it is ignored it in the fit, and even in the large $\Delta\phi_{ll}$ region it is small compared to the QCD WW background. However, because of its sharply peaked structure, an accurate normalization of this background is necessary; if the peak is shifted to the wrong mass or normalized incorrectly, the fitter will try to explain away the excess by distorting the top and QCD WW background distributions and by misidentifying $Z \rightarrow \tau\tau$ events as signal.

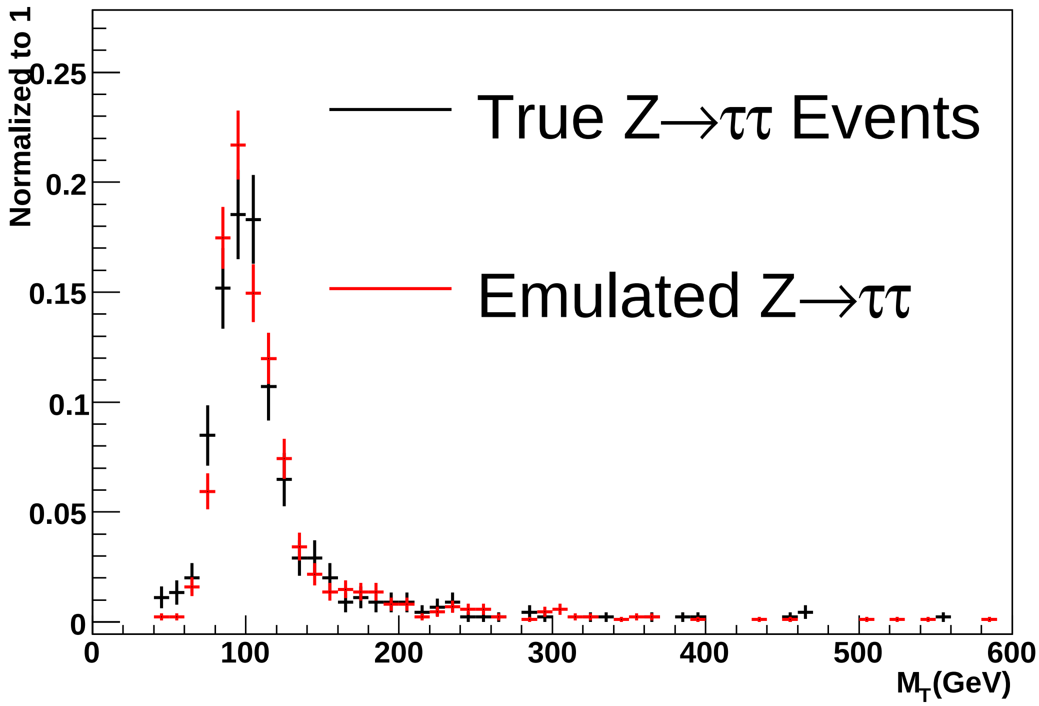

Figure 4.13: The M_T distribution for $Z \to \tau\tau$ Monte Carlo after all preselection cuts (except for the jet veto) and the approximation obtained using the Data-Monte-Carlo method. The distribution shown is integrated over the full range of the Higgs P_T.

The in-situ estimate of this background is based on the "Data-Monte-Carlo" method of Ref. [48]. Two-muon events with a dilepton invariant mass between 82 and 98 GeV are selected, and the same jet veto used in the rest of this analysis is applied. The effective cross-section after these cuts is roughly 360 pb. The real muons in these events are replaced with simulated tau leptons, and the remaining event selection cuts are applied. The efficiency of those cuts is about 0.07%, leaving an effective cross-section of roughly 250 fb. To be conservative, this study will assume that efficency factors will lower this figure to about 200 fb.

The Data-Monte-Carlo method produces a set of events that offers a reasonable

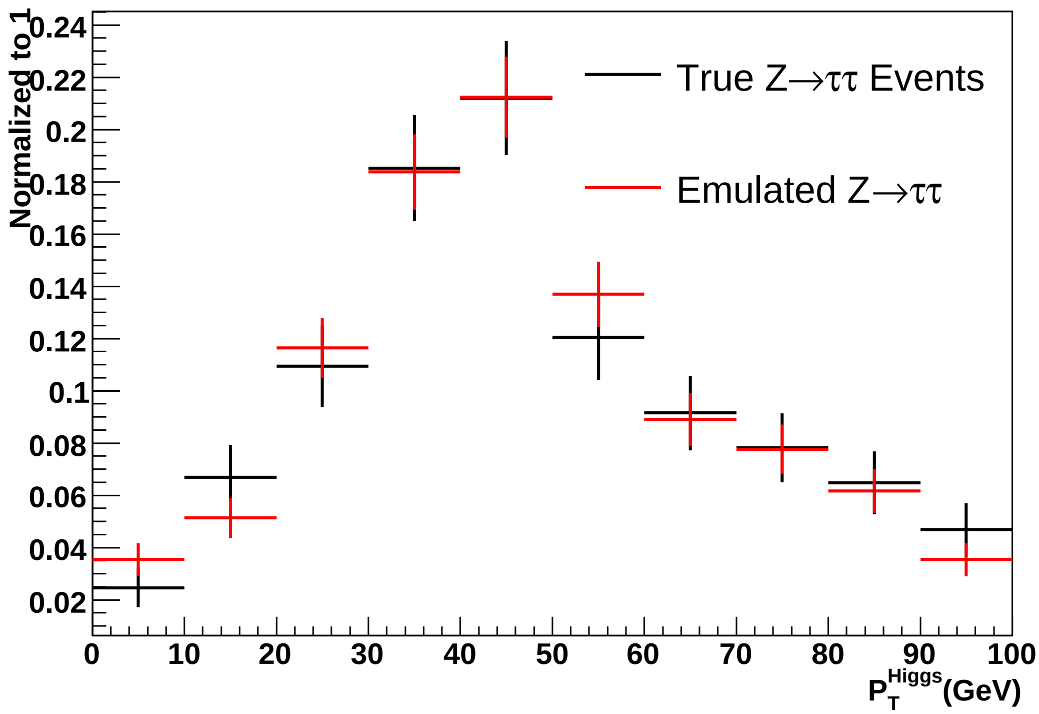

Figure 4.14: The P_T^{Higgs} distribution for $Z \to \tau\tau$ Monte Carlo after all preselection cuts (except for the jet veto) and the approximation obtained using the Data-Monte-Carlo method. The distribution shown is integrated over the full range of M_T.

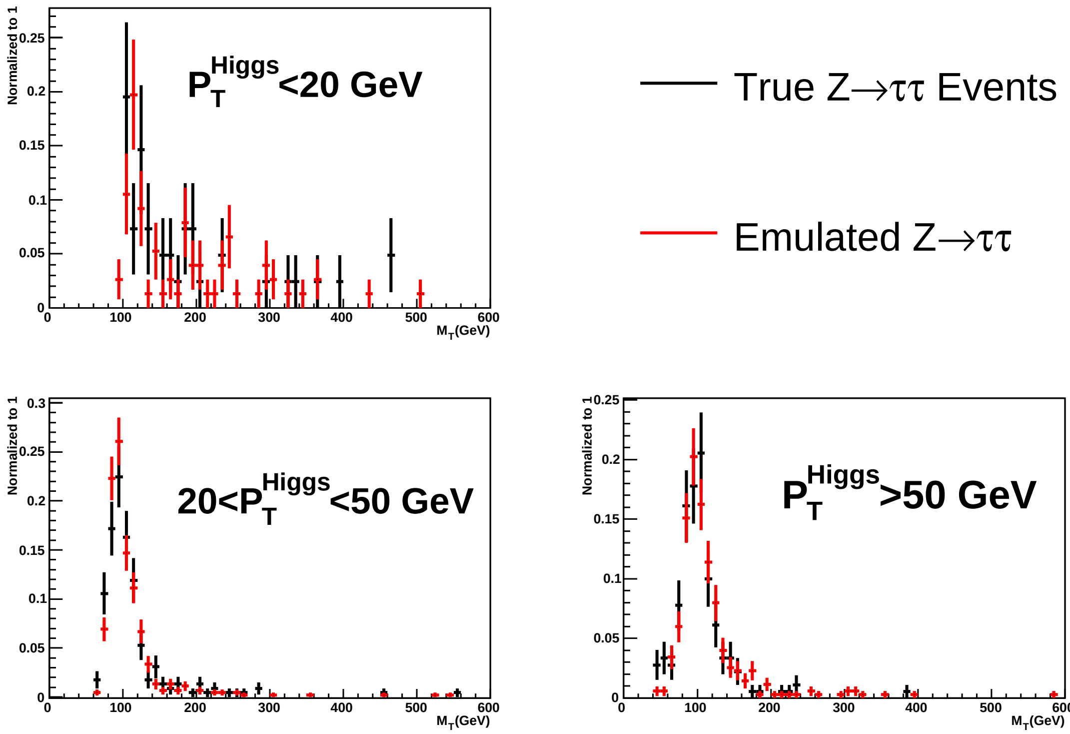

Figure 4.15: The M_T distribution for $Z \rightarrow \tau\tau$ Monte Carlo after all preselection cuts (except for the jet veto) and the approximation obtained using the Data-Monte-Carlo method. The distribution shown is integrated over several bins of Higgs P_T.

approximation of the real $Z \rightarrow \tau\tau$ background. Figure 4.14 shows the transverse mass distribution after preselection cuts,[3] integrated over Higgs P_T, as given by $Z \rightarrow \tau\tau$ Monte Carlo and by the Data-Monte-Carlo approximation; the agreement is quite good. Figure 4.14 shows that the agreement on the Higgs P_T distribution is very good as well. Figure 4.15 shows the transverse mass in several bins of the Higgs P_T and demonstrates that the approximation is accurately reproducing important correlations in the distribution of M_T vs. P_T^{Higgs}.

A fit is performed on the Data-Monte-Carlo events that pass the event selection

[3]Due to limited Monte Carlo statistics, the jet veto has not been applied in this comparison.

cuts. The shape parameters obtained from this fit are then fixed in the main fit. The relative normalization of the $Z \rightarrow \tau\tau$ background in the region of the analysis compared to the normalization of the Data-Monte-Carlo sample is assumed to be well-known.

4.4.3 QCD WW and the Hypothesis Test

Once the fits to the control samples are completed, two simultaneous fits to the events in the two $\Delta\phi_{ll}$ bins in the b-vetoed region are performed. In the first fit, it is assumed that the number of signal events is zero, allowing the normalizations for the QCD WW background in the two $\Delta\phi_{ll}$ bins to float independently. Two shape parameters for the QCD WW background are allowed to float, as are the extrapolation parameters associated with the top background.

For all luminosities, the nuisance parameter governing the Higgs P_T is free to float in the fit, but the Higgs P_T distributions are constrained to be similar in the two regions (i.e., the Higgs P_T distribution in the region with small $\Delta\phi_{ll}$ is given by the the Higgs P_T distribution in the region with large $\Delta\phi_{ll}$ times a second-order polynomial whose numerical value is close to 1). In this way, the operation of maximizing the likelihood "absorbs" the uncertainty in the prediction of the Higgs P_T distribution.

The second fit is the same as the first, except that the signal rate is no longer constrained to be zero. The Higgs mass is allowed to float as well. Using the Likelihoods obtained from the two fits, a hypothesis test is performed based on the widely used Likelihood Ratio estimator, defined as $LR = L_{s+b}/L_{bg-only}$. In the next section, the sampling distribution of this quantity will be studied using toy Monte Carlo.

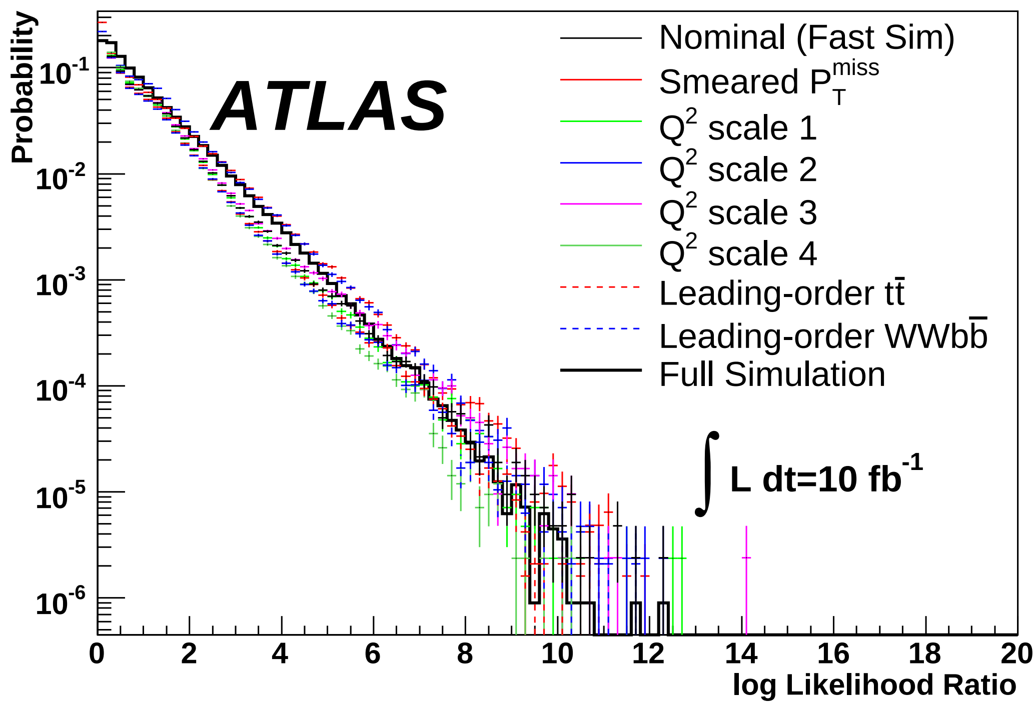

Figure 4.16: The log Likelihood Ratio distributions for background-only toy Monte Carlo outcomes corresponding to 10 fb^{-1}.

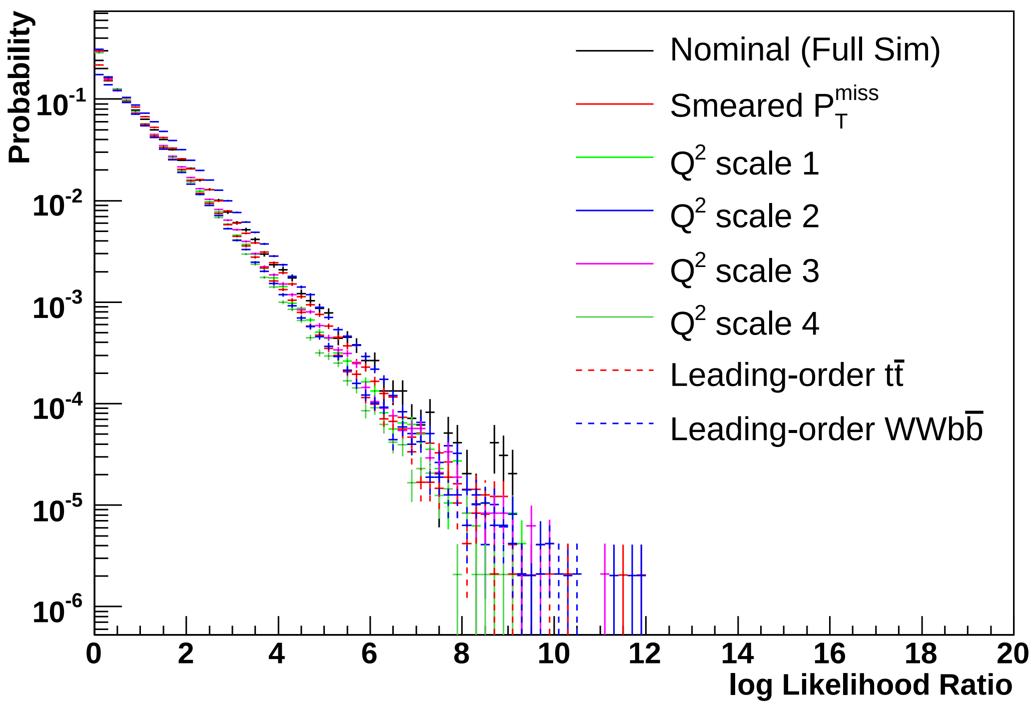

Figure 4.17: The log Likelihood Ratio distributions for background-only toy Monte Carlo outcomes corresponding to 3 fb^{-1}.

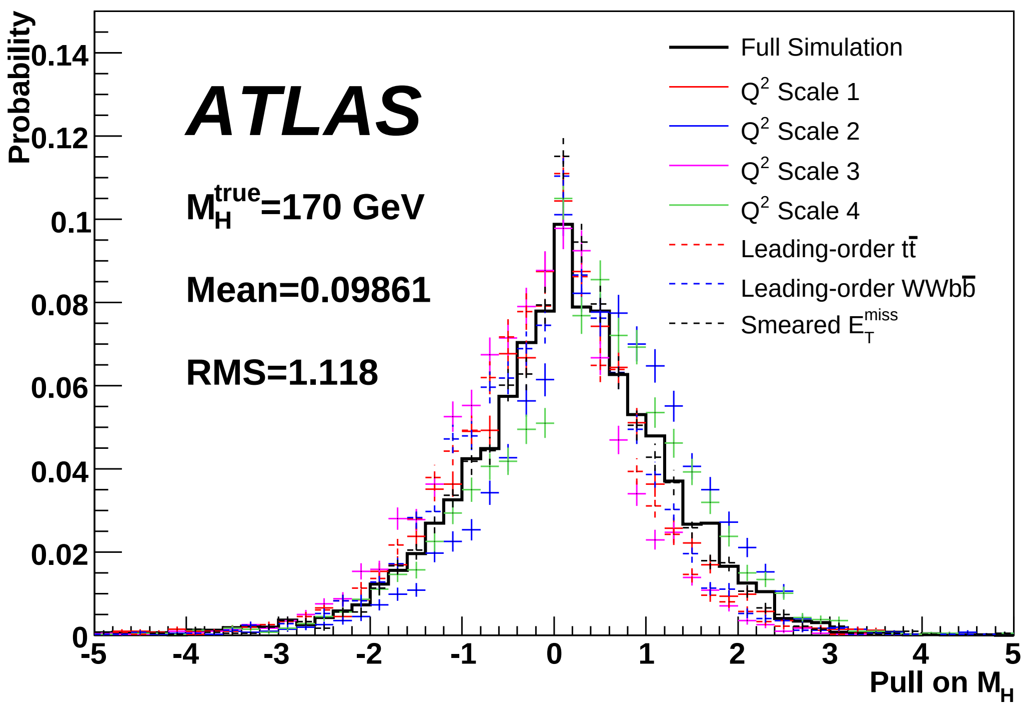

Figure 4.18: The pull distributions for M_H=170 GeV at a luminosity of 10 fb^{-1}, as obtained with various distorted background models..

4.4.4 Results

Toy Monte Carlo outcomes have been generated with and without a number of systematic distortions:

- Q^2 scale alterations. Four alternative truth PDFs are considered for the QCD WW background, obtained by altering the factorization and renormalization scales by factors of 8, as described in Section 4.1.

- Smeared missing P_T. One alternative truth PDF is considered for all backgrounds, obtained by smearing the x and y components of the missing P_T independently by 5 GeV each. The signal model does not include smearing; however, the most important check is the stability of the Likelihood Ratio in background-only outcomes.

- Alternative top background models. Two alternative truth PDFs are considered for the top background, one derived from leading-order MadEvent $pp \to t\bar{t} \to WWb\bar{b}$, and the other derived from leading-order MadEvent $pp \to WWb\bar{b}$, as described in Section 4.1.

If the Likelihood Ratio distributions for background-only toy Monte Carlo outcomes obtained from all these alternative truth PDFs are sufficiently similar to the distribution obtained from toy Monte Carlo generated with the baseline PDFs, then the confidence level calculation is trustworthy; otherwise, it may be desirable to use a free parameter tuning which reduces the biases and tails to an acceptable level.

Figures 4.16 and 4.17 show the results of the stability checks on the fitted Higgs mass for $M_H^{true} = 170$ GeV and on the background-only Likelihood Ratio distribution,

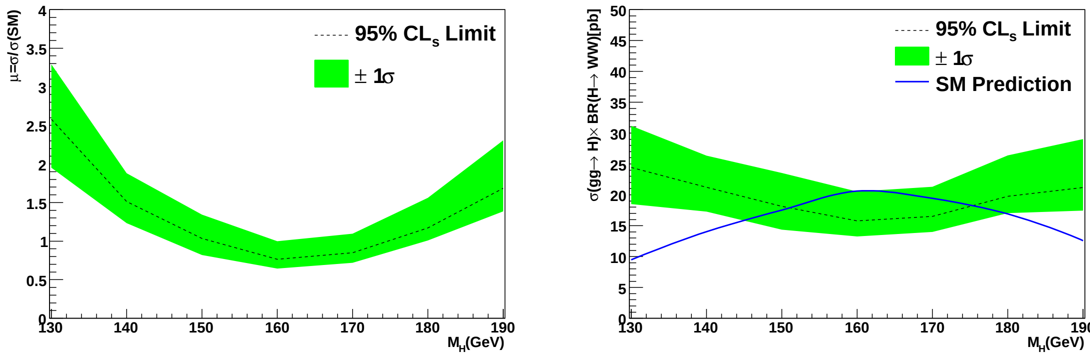

Figure 4.19: Left: the expected upper bound on the ratio of the signal cross-section to the Standard Model prediction as a function of the Higgs mass, after 1 fb^{-1} of integrated luminosity. Right: the corresponding upper bound on the absolute cross-section. In both plots, the green and yellow bands show the expected variability due to 1σ and 2σ fluctuations, respectively.

for integrated luminosities of 10 fb^{-1} and 3 fb^{-1}, respectively. From the plots, it is clear that the p-value of an excess is stable against systematics to within a factor on the order of 1. Figure 4.18 shows that the mass determination for a Standard Model Higgs with a mass in the neighborhood of 170 GeV is only stable against background systematics to within roughly 20% of the fit error. In a future revision of this analysis, it would be desirable to improve this feature, or at least to quote an error that is slightly larger than the fit error.

Finally, a set of toy Monte Carlo outcomes was generated with a signal model derived from a fit to a sample of fully simulated events with $M_H = 170$ GeV. The impact on the signal significance was not large.

Figures 4.19 and 4.20 show the upper limits on the Higgs cross-section one would expect to set after 1 fb^{-1} and 10 fb^{-1}, respectively, in the case that there is no Higgs observable in the mass range. To compute these limits, an exclusion Likelihood

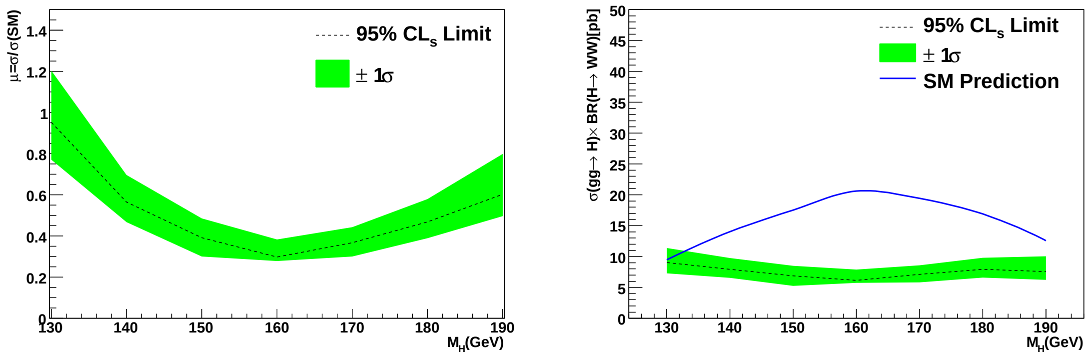

Figure 4.20: Left: the expected upper bound on the ratio of the signal cross-section to the Standard Model prediction as a function of the Higgs mass, after 10 fb^{-1} of integrated luminosity. Right: the corresponding upper bound on the absolute cross-section. In both plots, the green and yellow bands show the expected variability due to 1σ and 2σ fluctuations, respectively.

is constructed according to $LR = L_{s\ free}/L_{s\ fixed}$, where $L_{s\ free}$ is the maximum Likelihood obtained for a fit where the signal normalization and mass are free to float in the fit and $L_{s\ fixed}$ is the maximum Likelihood obtained for a fit where the signal normalization and mass are both fixed to the hypothesized values. A signal with a given mass and a given normalization is said to be excluded if the CL_s estimator [49] is less than 0.05. With only 1 fb^{-1} of integrated luminosity, the $H+0j$, $H \to WW \to e\nu\mu\nu$ channel is able to exclude a considerable range of Higgs masses; with 10 fb^{-1} of integrated luminosity, the Standard Model can be strongly excluded in the entire mass range if there is no Higgs boson.

The left plot in Figure 4.21 shows the expected significance after an integrated luminosity of 10 fb^{-1}; it is larger than 5σ for Higgs bosons in the mass range from roughly 145 GeV to roughly 180 GeV. The plot on the right shows the linearity of the mass determination; it is estimated by plotting the distribution of best-fit values of

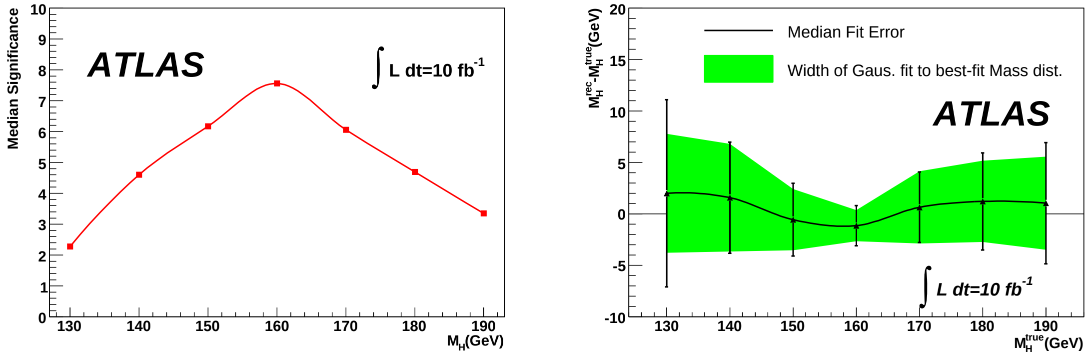

Figure 4.21: Left: the expected significance for an integrated luminosity of 10 fb^{-1}. Right: The linearity of the mass determination.

the Higgs mass parameter in signal-plus-background toy Monte Carlo outcomes and fitting a Gaussian function to the peak of the distribution. The black points show the difference between the mean of the Gaussian and the true Higgs mass; the green band shows the width of the Gaussian, and the black error bars indicate the median value of the error on the fitted Higgs mass. With the present fitting algorithm, a correction of less than 1σ would need to be applied to recover the true Higgs mass.

Chapter 5

Analysis of $H + 2j$, $H \to WW \to e\nu\mu\nu$

The $H + 2j$, $H \to WW \to l\nu l\nu$ channel has long been known to be a promising discovery mode for a Standard Model Higgs[8, 10, 12, 13] boson with a mass not too far from the WW threshold. [50, 51, 52, 53, 54, 55, 56, 57] This Chapter investigates the prospects for a fit-based approach to the $H \to WW \to e\nu\mu\nu$ channel. (For the moment, the same-flavor channels are not considered because a method for extracting the $Z \to ee/\mu\mu$ background has not yet been studied.)

Section 5.1 describes the Monte Carlo samples used to model the signal and the background. Section 5.2 lists the cuts applied to select the events to use in the fit. Section 5.3 describes the Neural Network used to characterize the jet activity in the event, and Section 5.4 describes the fit itself and evaluates its performance using toy Monte Carlo.

5.1 Monte Carlo

The Monte Carlo samples used in this study are a mixture of older samples made with Atlfast-Fortran [41] and newer samples made as part of the CSC production. The following physics processes are considered:

- $gg \to H \to WW$ signal. Alpgen [40, 39, 38] is used to model this process in fast simulation, and jet multiplicities as large as 3 are included at the matrix element level. This process is ignored in the full simulation analysis.

- VBF $H \to WW \to l\nu l\nu$. Pythia 6.2 [32, 31] is used to model this process in fast simulation, and Pythia 6.4 and Sherpa [58] is used to model it in full simulation. The official samples used for this signal process are datasets 5322 and 6325, respectively. In order to understand some differences between these two samples, signal samples have been generated with fast simulation using Pythia 6.2 and 6.4 with and without the Atlas underlying event tuning.

- $t\bar{t}$, the dominant background in this channel. In the fast simulation, MC@NLO version 3.3 [29, 30] is used to model this process (including spin correlations in the W decays). In full simulation, MC@NLO is used as well (but with spin correlations ignored). The official sample used for this process is dataset 5200.

- QCD/Electroweak WW background. Alpgen is used to model this process, and jet multiplicities as large as 3 are included at the matrix element level. The official samples for this process are datasets 6591-6594.

- $Z \to \tau\tau$. This minor background is modelled with MC@NLO in the fast simulation and Alpgen in full simulation. The official samples for this process are datasets 8154-8159.

- $b\bar{b}$. This potentially dangerous background is modelled with Alpgen, with up to 3 light jets at the matrix element level and using Herwig for hadronization.

Cut	$gg \to H$	VBF	$t\bar{t}$	WW	$Z \to \tau\tau$	$b\bar{b}$
No Cut	1993.38	292.952	85512.58	11457.37	210718.94	1.404×10^8
Lep. Selection	218.79	28.91	9449.73	1011.14	5877.74	3356.40
Fwd. Jet Tag	11.53	14.85	1034.98	33.97	56.91	46.36
Lep. Btw. Jets	9.25	14.00	657.84	22.85	36.69	35.16
$Z \to \tau\tau$ Rej.	8.33	13.25	598.90	20.92	9.33	28.98
P_T^{miss}, M_T	6.37	11.10	454.86	14.92	0.69	0.31
b-veto	6.05	10.76	158.66	14.30	0.60	0
Signal Box	3.92	8.10	46.52	4.30	0.11	0
Control Reg.	2.12	2.66	112.13	10.00	0.49	0

Table 5.1: Cut flows (in fb, for a true Higgs boson mass of 170 GeV) in the Vector Boson Fusion $H \to WW \to e\nu\mu\nu$ channel, using fast simulation.

There are very large uncertainties in the prediction of this background; it is included here only as a crude check.

The fully simulated samples used in this analysis were simulated and reconstructed with Athena version 12.

5.2 Event Selection

The events that are used in the fit are selected by applying the following cuts:

- Require that there be exactly two leptons (one electron and one muon) with $P_T > 15$ GeV. In the full simulation, electrons are taken to be electromagnetic clusters that satisfy isEM&0x3ff=false and have a matched track, with the energy measurement taken from the cluster and the η and ϕ measurements taken from the track. Muons are reconstructed using the STACO package, and for all leptons, a track isolation is computed as the scalar sum(P_T) of tracks in a cone of $\Delta R < 0.4$ around the lepton, excluding tracks within a cone of $\Delta R < 0.01$

Cut	Signal (170 GeV)	$t\bar{t}$	WW+jets	$Z \to \tau\tau$	W+jets
Lep. Selection	30.20	8317	838.96	(2096)	1323
Fwd. Jet Tag	17.27	946.6	32.77	79.30	31.83
Lep. Betw. Jets	16.47	617.8	22.92	55.13	27.91
$Z \to \tau\tau$ Rej.	15.68	561.8	21.20	39.03	27.91
P_T^{miss}, M_T	12.78	425.9	15.28	0	13.96
b-veto	12.67	206.72	-	-	-
sigbox, b-jet Veto	9.28±0.27	28.5±5.7	4.75±0.30	-	4.3±4.3
no b-jet Veto	9.65	114.2	4.99	-	6.07
Control, b-jet Veto	3.02±0.15	89±10	9.78±0.43	-	7.9±5.0
no b-jet Veto	3.13	311.7	10.28	-	7.89

Table 5.2: Cut flows (in fb) for $M_H = 170$ GeV in the $H + 2j$, $H \to WW \to e\nu\mu\nu$ channel, using full simulation. Numbers in parentheses are affected by generator-level filter cuts.

around any reconstructed lepton. Leptons are required to be isolated using a cut on this variable and on the *etcone*20 calorimeter isolation variable. Electrons are required to have both track and calorimeter isolation less than 5 GeV; muons are required to have track isolation less than 3 GeV and calorimeter isolation less than 2.5 GeV. In the fast simulation, the electron identification efficiency is taken to be 72.5% for electrons and 92.5% for muons [42].

- Reject events that have any b-tagged jets. In the fast simulation, the b-tagging efficiency is taken to be 60% for jets above 20 GeV; in the full simulation, events are rejected using the cuts described in [59].

- Require that the event has at least two jets with $P_T > 20$ GeV. The two leading jets are labelled as the "tag" jets. They are required to be in opposite hemispheres, $\eta_{j1} \cdot \eta_{j2}$, and they are required to be reasonably well-separated in pseudorapidity, $\Delta\eta_{jj} > 3$.

- The leptons are required to lie between the tagjets in pseudorapidity. The jets used here are reconstructed from topological clusters using a cone algorithm with $R = 0.4$. It has been shown that such jets have high efficiency in the forward regions of the detector and good resolution. [60, 61]

- To reject backgrounds from $Z \rightarrow \tau\tau$, assume that the leptons are coming from τ decays and compute x_τ^1 and x_τ^2, the fractions of the tau energies carried by the visible leptons in the collinear approximation. If x_τ^1 and x_τ^2 are positive and $M_{\tau\tau}$ is close to the Z mass ($|M_{\tau\tau} - M_Z| < 25$ GeV), the event is rejected.

- To prevent outliers from interfering with the Neural Network training, reject events with $M_{jj} > 6$ TeV.

- Require that the missing P_T is greater than 30 GeV, that the transverse mass M_T is between 50 GeV and 600 GeV, and that the transverse mass $M_T^{ll\nu}$ is at least 30 GeV.

Events surviving these cuts are used in the fit; they are partitioned into a signal box and a control region defined by additional cuts on the pseudorapidity gap between the leptons and the dilepton opening angle in the transverse plane. An event lies in the signal box if it has $\Delta\phi_{ll} < 1.5$ and $\Delta\eta_{ll} < 1.4$; otherwise, it lies in the control region. Table 5.1 shows the cross-sections obtained from fast simulation after the various preselection cuts listed above; Table 5.2 shows the corresponding result from full simulation.

The trigger efficiency has been studied in the context of this analysis; the trigger efficiency for this channel is 99.0% after Level 1, 96.8% after Level 2, and 94.5% after

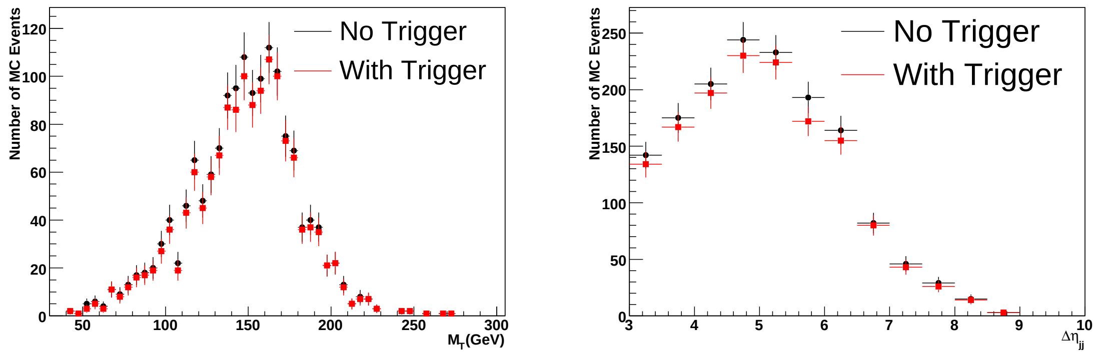

Figure 5.1: Left: the M_T distribution for the $H+2j$, $H \rightarrow WW \rightarrow e\nu\mu\nu$ channel after preselection cuts, with and without the trigger efficiency. Right: the $\Delta\eta_{jj}$ distribution for the $H+2j$, $H \rightarrow WW \rightarrow e\nu\mu\nu$ channel after preselection cuts, with and without the trigger efficiency.

the Event Filter. The plot on the left of Figure 5.1 shows the transverse mass for signal after the preselection cuts only and after the preselection and trigger cuts. There is no appreciable distortion in the shape of this variable due to the trigger efficiency. The plot on the right shows that the distribution of the pseudorapidity gap between the tagjets is also quite stable against the trigger efficiency.

5.3 Neural Network

After the event selection in the previous section is applied, a four-variable Neural Network is used to further enhance the separation between the signal and the background. The inputs to the Neural Network are:

- $\Delta\eta_{jj}$, the pseudorapidity gap between the tagjets
- M_{jj}, the invariant mass of the tagjets
- P_T^{veto}, the transverse momentum of the leading non-tag jet in the region $|\eta| < 3.2$,

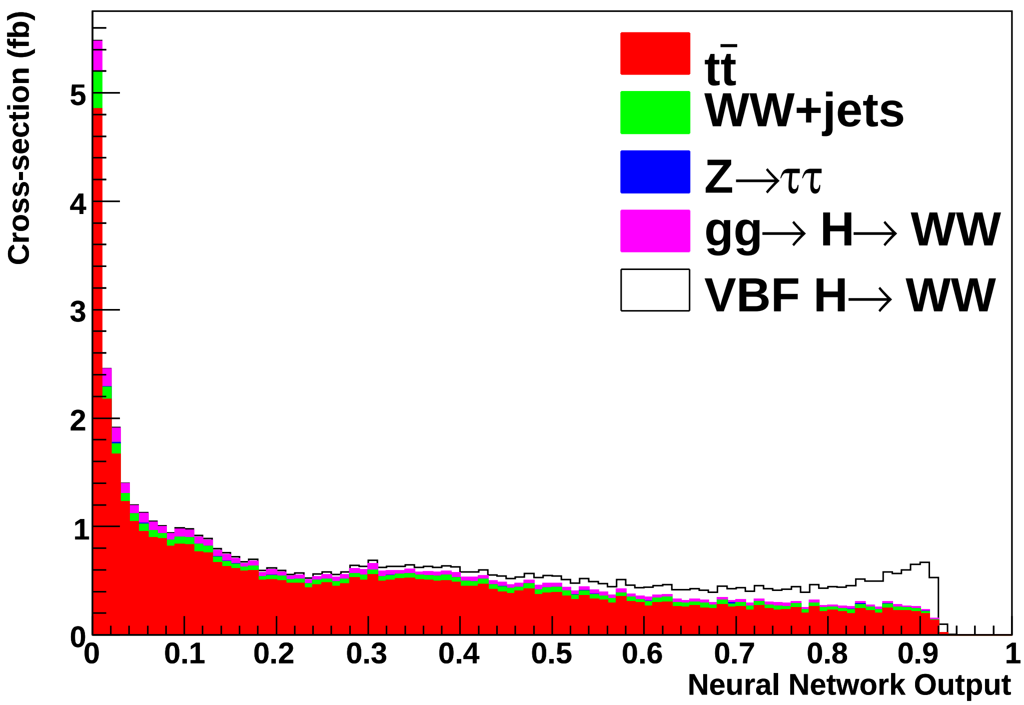

Figure 5.2: The Neural Network output distribution (based on fast simulation) for the $H + 2j$, $H \to WW \to e\nu\mu\nu$ channel in the signal box.

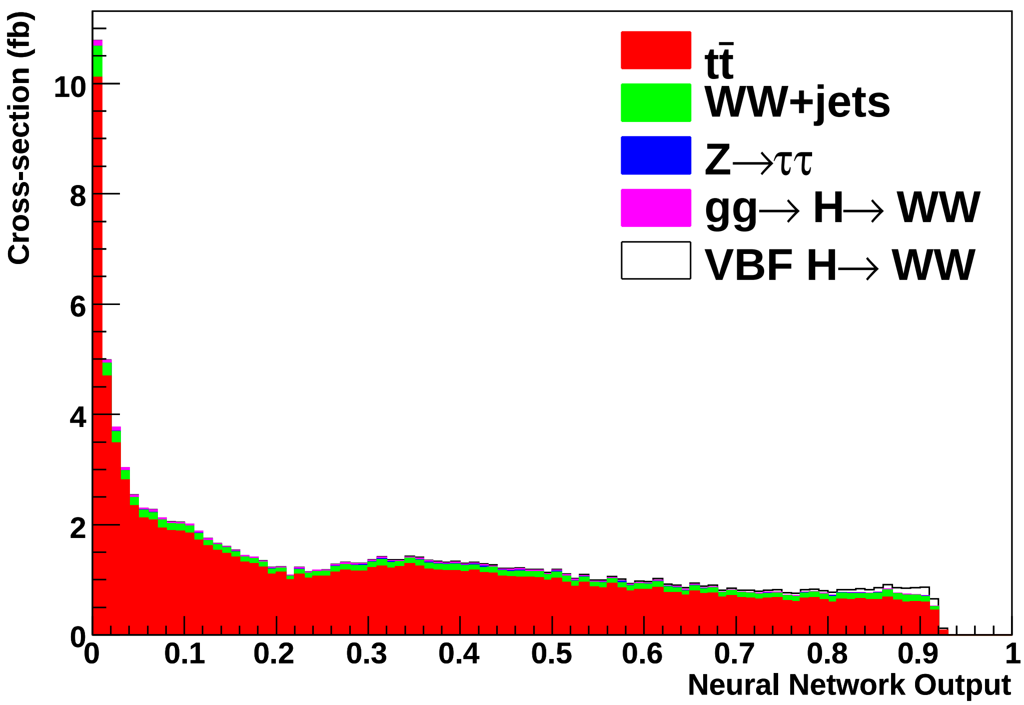

Figure 5.3: The Neural Network output distribution (based on fast simulation) for the $H + 2j$, $H \to WW \to e\nu\mu\nu$ channel in the control region.

and

- η^*, the pseudorapidity gap between the tag-tag system and the third jet. It is set to -9 if no third jet is present.

The Neural Network is a feed-forward network with two hidden layers of 4 variables each. It is trained it on a sample of 10000 signal events ($M_H = 160$ GeV) and 10000 background events ($t\bar{t}$ and $WW + 2/3j$ only, represented according to cross-section). Figures 5.2 and 5.3 show the Neural Network output distribution in the signal box and control region (integrated over M_T). The shape of the distributions in the two regions is quite similar; this will be used as part of the background determination strategy discussed in the next section.

Figure 5.4 shows the distributions of several important kinematic variables in bins of Neural Network output. The two top plots show that, as one would expect, events with small M_{jj} and small $\Delta\eta_{jj}$ are classified as background-like, and events with large values of these observables are classified as signal-like. The middle row of plots shows of P_T^{veto} and η^*. Events with any third jet are unlikely to be classified as very signal-like, but events where η^* is small and P_T^{veto} is large are classified as very background-like. The bottom row shows the distributions the transverse mass and the dilepton opening angle in the transverse plane. The plots indicate that the correlation between the Neural Network output and the lepton angular variables is very small.

The impact of altered jet energy scales on the Neural Network output distributions for signal and background has been checked by scaling the jet four-vectors up and down by 5% in the region with $|\eta| < 2.5$ and 10% elsewhere; the altered distributions are compared to the nominal distributions in Figure 5.5. The jet energy scale

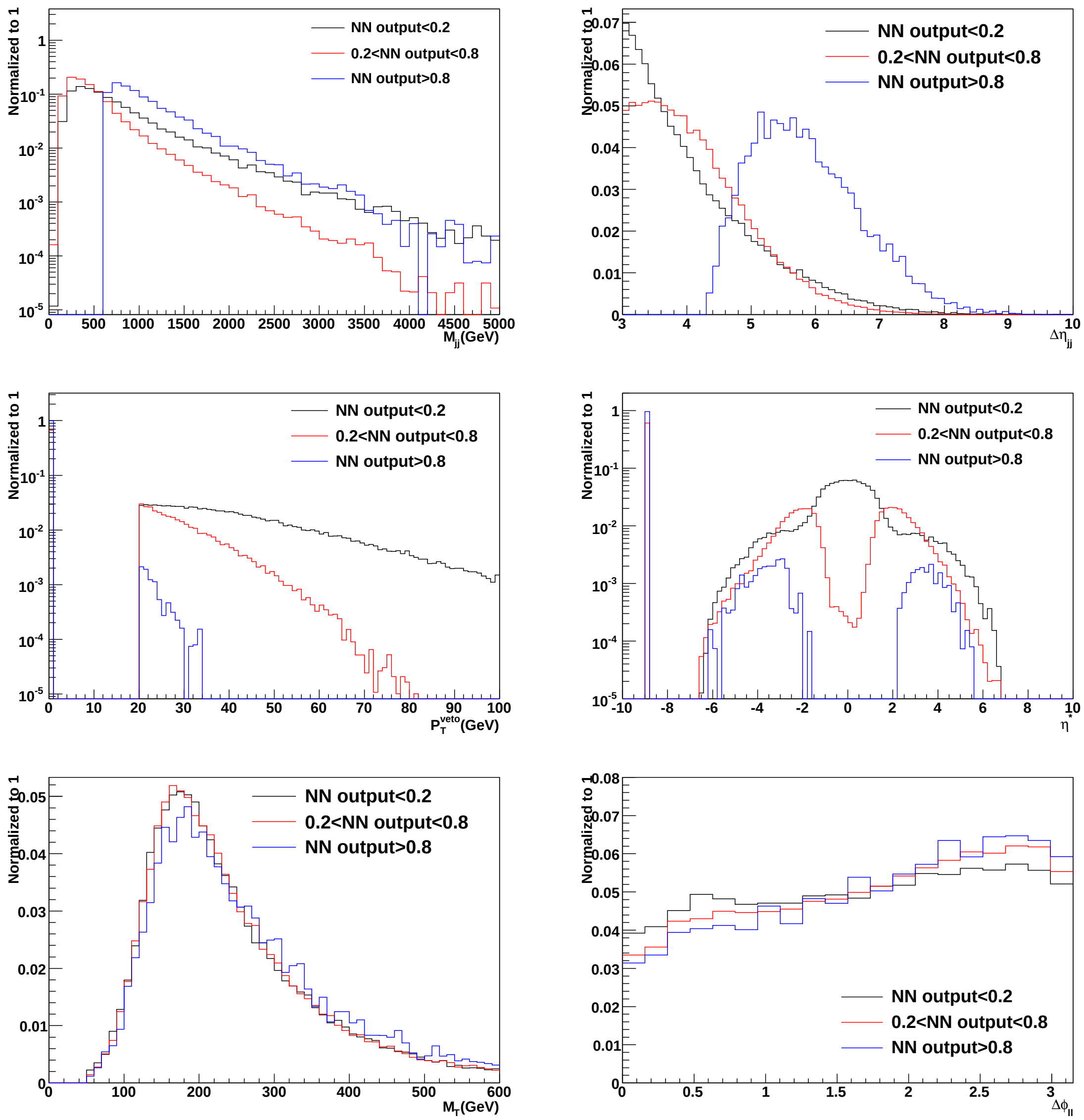

Figure 5.4: Correlation plots for the $H + 2j$, $H \rightarrow WW \rightarrow e\nu\mu\nu$ analysis, showing kinematic variables for the top background in bins of Neural Network output. Upper left: the invariant mass of the tagjets. Upper right: the pseudorapidity gap between the tagjets. Middle left: P_T^{veto}. Middle right: η^*. Lower left: the transverse mass. Lower right: the pseudorapidity gap between the tagjets.

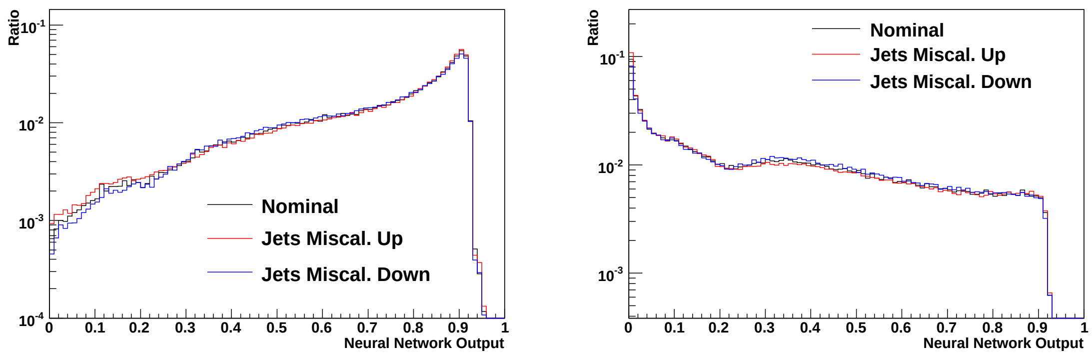

Figure 5.5: The Neural Network output distribution for $H \to WW \to e\nu\mu\nu$ signal with $M_H = 170$ GeV (left) and for ttbar background (right), with the jet energy scale shifted up and down.

uncertainty has a negligibly small effect on the shape of the Neural Network output distribution.

A lowered jet reconstruction efficiency for low-P_T jets will distort the shape of the Neural Network output distribution, largely due to the fact that the presence of a third jet, even at low P_T, will cause the Neural Network to classify the event as background-like. Figure 5.6 shows the ratio of the Neural Network output distribution for top background in the signal box and the control region in scenarios where a degraded jet reconstruction efficiency is modelled by increasing the minimum P_T threshold for jets from 20 GeV to 30 or 40 GeV. The ratio is quite stable against this sort of distortion; this feature is important for the fitting algorithm described in the next section.

The shape of the Neural Network output for signal can be distorted due to altered parton shower and underlying event models. Figure 5.7 shows the Neural Network output distribution obtained with two different models of the underlying event (Pythia default and ATLAS default) and with two different parton showering models (the “new” model used by Pythia 6.4 and the “old” model used by Pythia 6.2;

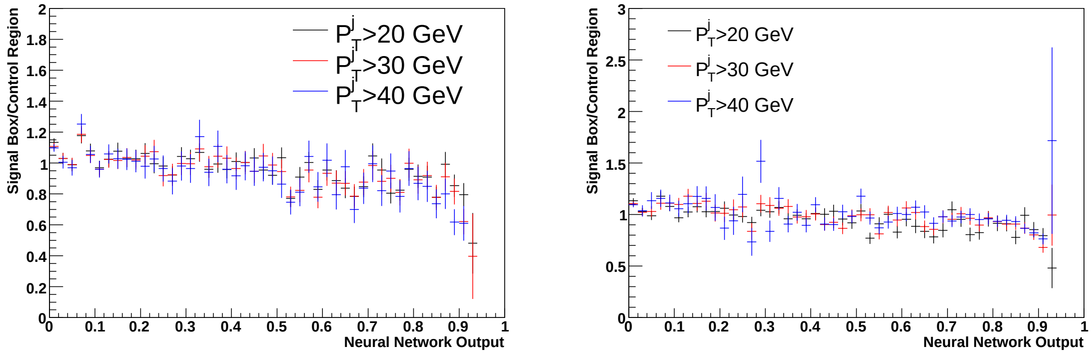

Figure 5.6: The ratio of the Neural Network output distribution in the signal box and the control region in several scenarios where a degraded jet reconstruction efficiency is modelled by raising the P_T threshold for jets. Left: only the P_T threshold for the tagjets is raised. Right: the P_T threshold for the jet used to calculate P_T^{veto} is also raised.

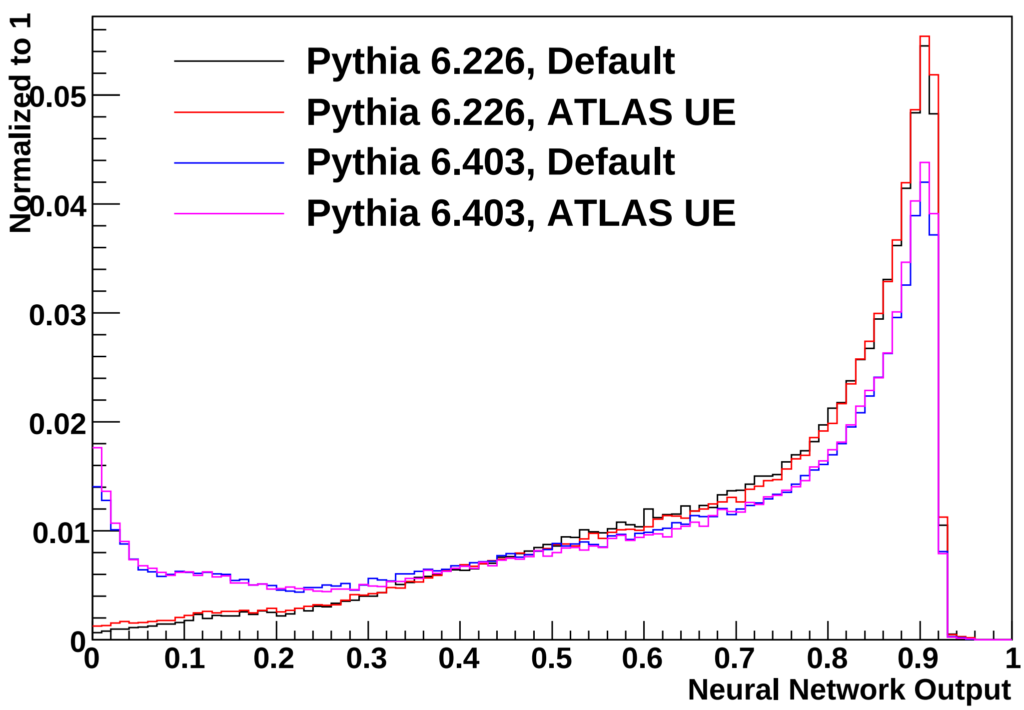

Figure 5.7: The Neural Network output distribution for signal (for a Higgs boson mass of 170 GeV, in the $H \to WW \to e\nu\mu\nu$ channel), with different models for the parton shower and underlying event.

the latter was used to train the Neural Network.) The underlying event is a small effect, but the showering model leads to a nontrivial difference in the distributions in the region with low Neural Network output. The large spike at 0 in the Pythia 6.4 distributions is caused by an excess of events with a third high-P_T jet; this feature is not reproduced by Sherpa.

5.4 Fitting Algorithm and Results

The signal is extracted with a two dimensional fit on the Neural Network output and the transverse mass M_T. The background is treated as a two-component background, comprised of an "irreducible" component consisting mostly of $t\bar{t}$ and WW+jets, and a "reducible" component, consisting mostly of W+jets background. The shape and normalization of the "irreducible" component are allowed to float in the fit; however, in the present study the shape and normalization of the "reducible" component of the background are taken from Monte Carlo. The functional forms used for signal, "irreducible" background, and "reducible" background are all documented in detail in Appendix B; this section will focus on conceptual aspects of the fit and on performance tests using toy Monte Carlo.

The in-situ background determination is based on the fact that the Neural Network output distribution in the signal-like region is very similar to the Neural Network output distribution in the control region. This approach has the advantage that a detailed understanding of the b-tagging efficiency is not required. Figure 5.8 shows the two distributions normalized to equal area on the left, and the ratio of the two distributions on the right. The plots show that the distributions are similar but not exactly the same. The fit assumes them to be identical up to a linear extrapolation factor.

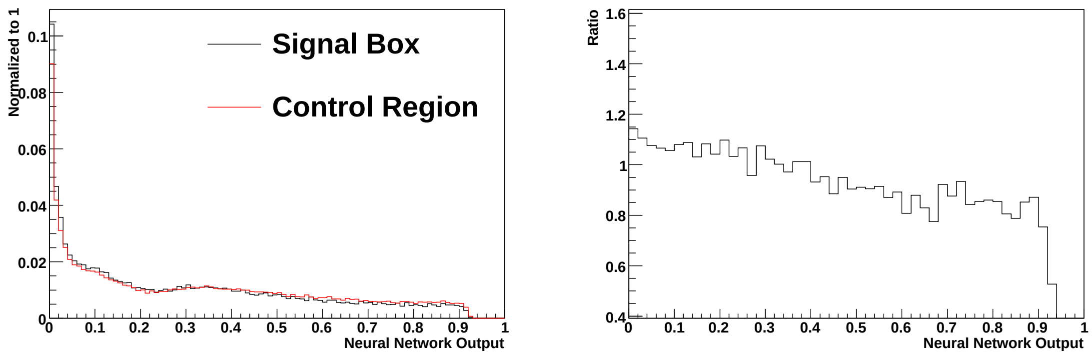

Figure 5.8: The distributions (normalized to 1) of the Neural Network output for the $H + 2j$, $H \to e\nu\mu\nu$ analysis in the signal box and the control region (left), and the ratio of the two distributions (right).

That is, the functional form in the signal box is given by the same functional form that is used in the control region, with the same values of the parameters, but multiplied by an extrapolation factor $1 + p_{12}X$ where X is the Neural Network output and p_{12} is an additional freeable fit parameter.

Given the fact that the Neural Network output shape is insensitive to the jet energy scale, one would expect that the ratio would be insensitive to the jet energy scale as well. Other possible sources of uncertainty in ratio of Neural Network output distributions are underlying event activity and the Q^2 scale uncertainty; these have not been explicitly checked these yet, but it is reasonable to think they might be small. If the uncertainty in the ratio of distributions is sufficiently small, then one can fix the parameter governing the extrapolation factor; if it is too large, one must allow it to float in the fit. In the latter case, it may be possible to place some constraints on it by looking into a b-tagged control sample [62]. All of these possibilities require further study; the present analysis will simply consider both scenarios: allowing p_{12} to float

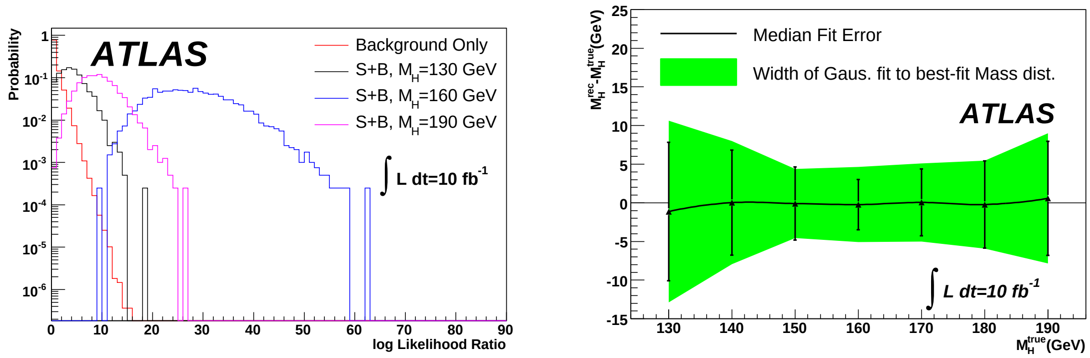

Figure 5.9: Left: the distribution of the Likelihood Ratio for background-only outcomes (red) and for signal-plus-background outcomes with several different Higgs masses, for an integrated luminosity of 10 fb^{-1}. Right: the linearity of the mass determination at the same luminosity.

in the fit and fixing it to a predetermined value taken from Monte Carlo.

The performance of the fitting algorithm has been studied using toy Monte Carlo. Background-only outcomes and signal-plus-background outcomes with masses in the range of 130 GeV to 190 GeV are generated, and two fits are performed on each outcome. The first one constrains the number of signal events to be zero, and the second allowing both the signal normalization and mass to float in the fit. These two fits yield (among other things) an estimate of the mass, the error on that estimate, and a value of the Likelihood Ratio, defined as $LR = L_{S+B}/L_{BG-only}$, where L_{S+B} is the maximum value of the Likelihood function obtained in the fit with a floating signal normalization and $L_{BG-only}$ is the maximum value of the Likelihood function obtained in the fit with the signal normalization constrained to be zero.

The left plot in Figure 5.9 shows the distribution of the Likelihood Ratio as obtained with the background-only toy Monte Carlo outcomes and the signal-plus-background outcomes with several different values of M_H. The background-only dis-

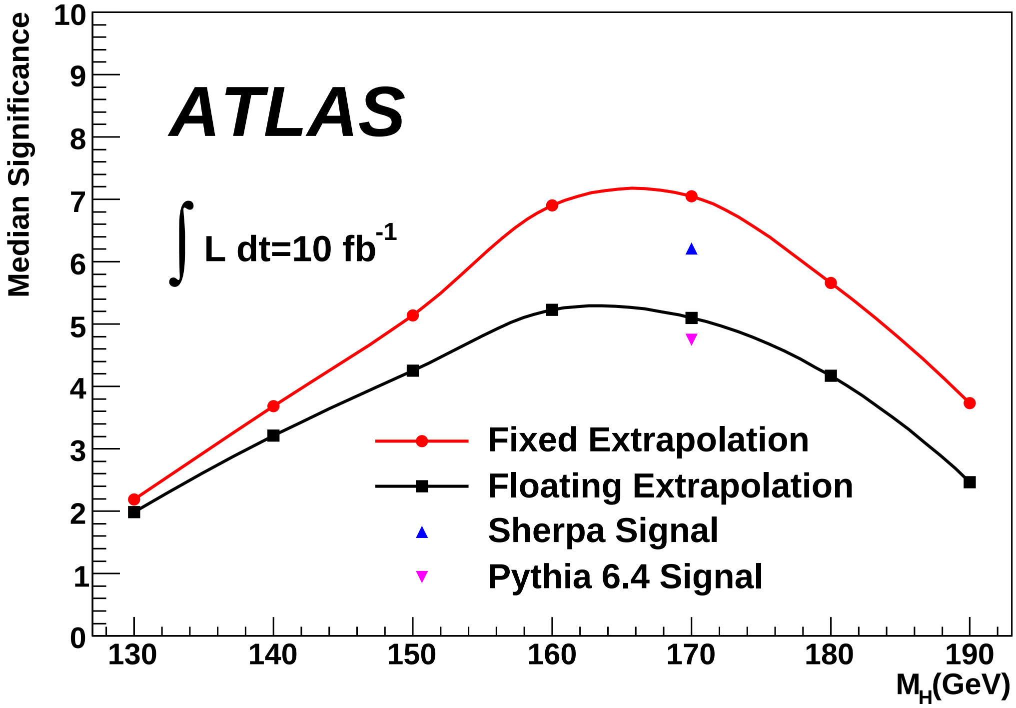

Figure 5.10: The expected significance for the $H + 2j$, $H \to WW \to e\nu\mu\nu$ analysis at an integrated luminosity of 10 fb^{-1}.

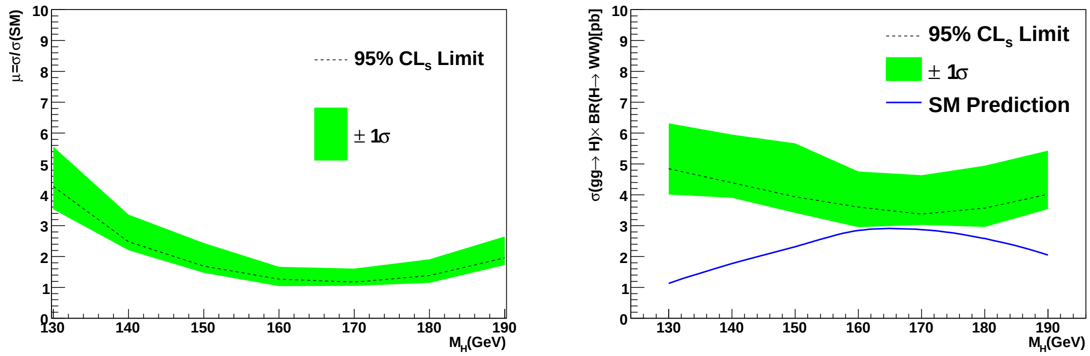

Figure 5.11: Left: the expected upper bound on the signal normalization with 1 fb^{-1} of integrated luminosity. Right: the corresponding upper bound on the absolute cross-section.

tribution is well-behaved over 6 decades in probability. The right plot in Figure 5.9 shows the linearity of the mass determination. The y-axis shows the difference (in GeV) between the fitted and true Higgs masses. The error bars show the median fit error, and the green band shows the width of a Gaussian fit to the region around the peak of the distribution of fitted Higgs masses.

The expected confidence level, CL_b, for a given mass is the integral of the background-only Likelihood Ratio distribution from the median of the signal-plus-background distribution for that mass up to infinity. Since the significance is sometimes rather large, the tail of the background-only distribution is extrapolated using an exponential function. The expected significance Z reported here is one-sided, i.e., $Z = \sqrt{2}\mathrm{erfc}^{-1}(2\ CL_b)$. The expected significance is shown as a function of the true Higgs mass in Figure 5.10 for the two treatments of the p_{12} parameter considered in this study. Although the present study uses a much more conservative treatment of systematics than previous studies (by assuming background normalization and the shapes of the jet variables are all more or less unknown), the sensitivity is comparable

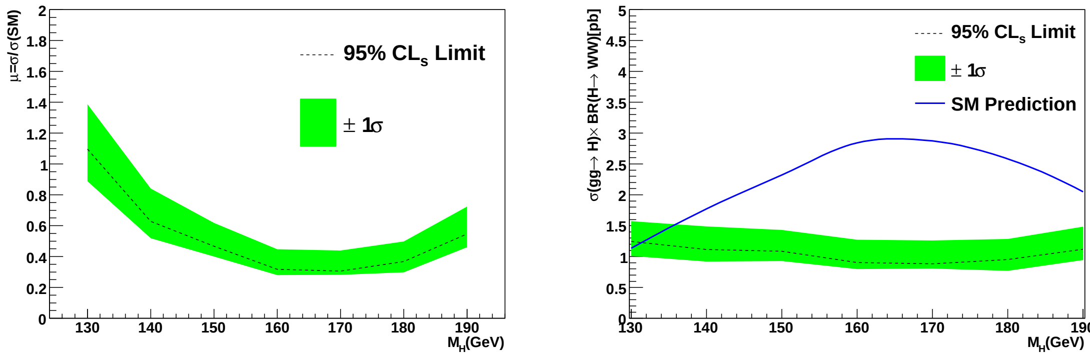

Figure 5.12: Left: the expected upper bound on the signal normalization with 10 fb^{-1} of integrated luminosity. Right: the corresponding upper bound on the absolute cross-section.

to what was quoted for the cut analysis in the previous study of this channel [57], at least for the case where the p_{12} parameter is constrained to a fixed value.

Figure 5.10 also shows the expected signal significance obtained based on two alternative models of the signal shape. The blue and purple triangles, corresponding to signal events generated with Sherpa and with Pythia 6.4, respectively, should be compared with the fixed extrapolation results presented in the red curve. The other signal models lead to a reduced expectation of the significance; the main difference between these and the default signal model is that the Pythia and Sherpa samples tend to have somewhat more jet activity in the central region. This causes the peak in the signal at high Neural Network output to become less sharp, leading to a lower signal significance.

In the absence of a convincing signal excess, it is important to set upper bounds on the production rate of a potential WW excess in this channel. Limits are set using the Likelihood Ratio $LR = L_{free\ s}/L_{fixed\ s}$, where in the numerator the signal normalization is allowed to float, and in the denominator, it is fixed to the hypothesized value. In the latter fit, a signal normalization uncertainty of 12% is taken into account by including in the fit an extra free parameter (corresponding to the signal efficiency) and constraining it to the predicted value by adding a Gaussian term to the Likelihood function. (This uncertainty corresponds to the jet energy scale uncertainty in the signal efficiency, evaluated using the procedure described in Section 5.3.) In both fits, the Higgs mass is allowed to float in the fit. A signal normalization hypothesis is said to be excluded if the $CL_s < 0.05$, where $CL_s = CL_{s+b}/CL_b$, as described in Ref. [49]. Because the CL_{s+b} and CL_b values of interest are generally not as small in the limit-setting case as they are in the discovery case, both CL_{s+b} and CL_b are computed numerically in this calculation, i.e., no extrapolation is needed to compute CL_b. Figures 5.11 and 5.12 show the expected exclusions as a function of mass, for integrated luminosities of 1 fb^{-1} and 10 fb^{-1}, respectively. The dashed black curve shows the median expectation; the green band shows the size of the expected $\pm 1\sigma$ fluctuations around this median. With an integrated luminosity of 10 fb^{-1}, one would expect to exclude Standard Model Higgs bosons with Higgs masses larger than about 135 GeV using only the $H + 2j$, $H \to WW \to e\nu\mu\nu$ channel.

Chapter 6

Analysis of $H + 2j$, $H \to WW \to \ell\nu qq$

For Higgs masses significantly above the WW threshold, it is possible to use a W mass constraint to reconstruct the invariant mass of the WW system in events where one W decays leptonically and the other decays to jets. This chapter will discuss the analysis of events where Higgs bosons are produced in association with two jets and decay via $H \to WW \to \ell\nu qq$. A Higgs search in this channel was first proposed in Refs. [51] and [63], and later adapted to the ATLAS environment in Refs. [64, 65, 66, 56, 67] and [68]. The analysis presented in this Chapter includes control samples that can be used to determine from data both the shape and the normalization of the dominant backgrounds. A fitting algorithm that constitutes a practical implementation of this idea is discussed as well, and its performance is studied with toy Monte Carlo.

This Chapter is organized as follows: Section 6.1 discusses the signal production mechanism and the dominant backgrounds. Section 6.2 present the event selection, as well as its efficiency in selecting signal and rejecting background. Section 6.3 describes a set of control samples that can be used to normalize the background in data and summarizes the results of a number of systematic checks that demonstrate the robustness of this method. Section 6.4 discusses the fitting algorithm and a toy Monte

Carlo study of its performance.

6.1 Signal and Background Processes

The most important physics processes relevant to the $H \rightarrow WW \rightarrow \ell\nu qq$ channel are the Vector Boson Fusion signal and the W+jets and $t\bar{t}$ backgrounds. The response of the ATLAS detector has been simulated using the last version of ATLFAST-Fortran [41], as well as with the full simulation package in Athena release 12.0.6. In both cases, the hadronization corrections of ATLFAST-B have been included. The following physics processes are considered:

- Vector Boson Fusion signal, $qq \rightarrow qqH \rightarrow qqWW$. This process is modeled with Herwig, and Higgs boson masses ranging from 160 GeV to 600 GeV have been considered.

- QCD W+jets production. This is very likely to be the dominant background in this analysis (although because of the large uncertainty in the prediction of the cross-section, it is not clear that this is definitely the case). This backgroud process is modeled with Alpgen version 2.0.6, using Herwig to handle the hadronization of the final state. In the present calculation, matrix elements are used to model final-state jet multiplicities up to and including 5 jets.

- $t\bar{t}$ production. This background is modeled using MC@NLO.

- QCD and Electroweak WW+jets. This process is modeled using Alpgen version 2.0.5 with Herwig hadronization, and matrix elements are used for final-state jet multiplicities up to and including three jets.

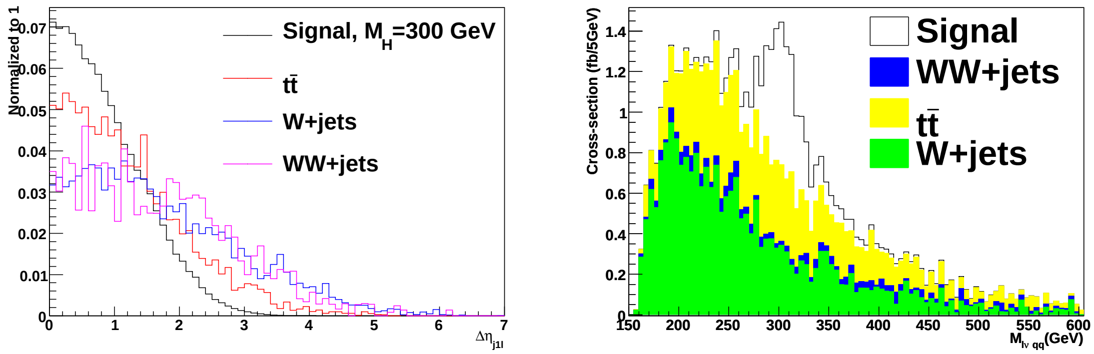

Figure 6.1: Left: the pseudorapidity gap between the lepton and the leading jet from W decay. Right: the invariant mass of the Higgs candidates after cuts. For signal, $M_H^{true} = 300$ GeV. Both of these plots are based on fast simulation.

In principle, there are other minor backgrounds to this process, such as QCD and Electroweak $WZ+2$ jets, with $WZ \to l\nu qq$. However, these backgrounds are expected to be small compared to the large uncertainty in the W+jets background.

6.2 Event Selection

The event selection consists of the following cuts:

- Leptonic W selection. The event must contain exactly one hard lepton, with transverse momentum $P_T > 25$ GeV for electrons and $P_T > 20$ GeV for muons. Lepton identification efficiencies in fast simulation and the lepton identification cuts applied in full simulation follow the treatment in Chapter 4. It is also required that the missing transverse momentum, P_T^{miss}, be larger than 30 GeV.

- Hadronic W selection. Out of all jets with $P_T > 30$ GeV, select the two whose invariant mass is closest to the known value of the W mass. It is required that the reconstructed invariant mass of these two jets be between 64 GeV and 90 GeV.

Cut	W+jets	ttbar	WW+jets	Signal ($M_H = 300$ GeV)
Leptonic W Selection	199875*	99618.2	10864	168.22
Hadronic W Selection	60483*	76780.1	3913	78.19
Forward Jet Tagging	1061.11	1924.41	45.36	25.27
Lepton Between Jets	863.85	1669.30	40.50	24.92
M_{jj}	200.18	404.09	12.77	11.21
Central Jet Veto	50.01	68.67	6.72	10.14
b-jet Veto	46.58	27.39	6.29	9.46
$\Delta\eta_{j1,l}$	23.65	18.80	3.08	7.84
Scale Varied Samples				
Scale 1	7.77	-	-	-
Scale 2	124.08	-	-	-
Control Regions				
b-tagged	1.79	30.81	0.29	0.56
Control Region(b-veto)	407.98	215.39	27.82	20.17
Control Region(b-tag)	40.16	517.28	2.57	1.68

Table 6.1: Cross-sections (in fb) for the signal and various backgrounds in fast simulation after successive cuts in the $H \to WW \to l\nu qq$ channel. A '*' indicates a number that is biased by generator-level cuts.

Cut	W+jets	ttbar	Signal ($M_H = 300$ GeV)
Leptonic W Selection	2353291*	128654	174.27
Hadronic W Selection	134483	70872	73.26
Forward Jet Tagging	1076.8	1929	23.16
Lepton Between Jets	867.0	1679	22.93
M_{jj}	131.0	367.7	9.16
Central Jet Veto	57.98	58.24	8.43
$\Delta\eta_{j1,l}$	16.07	47.96	6.93
b-jet Veto	16.07	14.84	6.06
Trigger Selection	13.06	12.40	5.08
$167 < M_{l\nu qq} < 1000$ GeV	13.1±4.7	12.4±3.4	5.08±0.29
Control Regions			
b-tagged	0	26±5	0.75±0.12
Control Region	500.16	778.76	20.79
Control Region(b-veto)	441.64	186.13	19.95
(b-veto and Trigger)	369.38	157.47	15.06
($167 < M_{l\nu qq} < 1000$ GeV)	358±29	154±12	15.1±0.5
Control Region(b-tag)	37±14	493±22	2.3±0.2

Table 6.2: Cross-sections (in fb) for the signal and various backgrounds after successive cuts in the $H \to WW \to l\nu qq$ channel. A '*' indicates a number that is biased by generator-level cuts.

- Forward Jet Tagging. Make a list of all the jets in the event except for the two jets that comprise the W candidate, and select the two that have the highest transverse momentum. These are labelled as the "tagging" jets. They are required to be high-P_T ($P_T^{j1} > 50$ GeV, $P_T^{j2} > 30$ GeV), to lie in opposite hemispheres ($\eta_{j1} \cdot \eta_{j2} < 0$), and to be well-separated in pseudorapidity ($|\eta_{j1} - \eta_{j2}| > 4.4$).

- The lepton is required to be between the tagjets in pseudorapidity.

- M_{jj}: require that the invariant mass of the two tagging jets is greater than 1500 GeV.

- Central Jet Veto: reject the event if it contains any extra jets (in addition to the two jets from the decay of the W and the two tagging jets) with $P_T > 30$ GeV and $|\eta| < 3.2$.

- b-jet veto: in fast simulation, reject the event if it contains any b-tagged jets with $P_T > 20$ GeV and $|\eta| < 2.5$. In full simulation, apply the b-tagging cuts described in Ref. [59].

- $\Delta\eta_{j1,l}$: require that the pseudorapidity gap between the leading jet from the decay of a W and the lepton be no larger than 1.5. The left plot in Figure 6.1 shows the distribution of this quantity for a representative Higgs mass of 170 GeV after all the above cuts have been applied.

- Reconstruct the mass of the Higgs candidate, $M_{l\nu qq}$, by applying the W mass constraint to the lepton-P_T^{miss} system. The resolution can be enhanced by performing a kinematic fit to the mass of the hadronically decaying W, minimizing the function $\chi^2 = (M_W^{rec} - M_W^{true})/\Gamma_W + (\Delta E_{j1})^2/\sigma_{j1}^2 + (\Delta E_{j2})^2/\sigma_{j2}$, where σ, the energy resolution for the jets, is given by $\sigma/E_j = 0.03 \oplus 0.5/\sqrt{E_j}$. The plot on the right side of Figure 6.1 shows the distributions of the Higgs candidate invariant mass after all cuts for a representative Higgs mass of 300 GeV.

Table 6.1 shows the cut flow in fast simulation for background and signal with a representative Higgs mass of 300 GeV, and Table 6.2 shows the cross-sections obtained from full simulation. It is worth noting that the row labelled "Control Region(b-tag)" in Table 6.2 includes the trigger efficiency and the cut $167 < M_{l\nu qq} < 1000$ GeV; the cross-sections in this row and the row labelled "Control Region(b-veto)" are not complementary.

As one would expect from a Weak Boson Fusion search, the QCD backgrounds are substantially reduced by the Forward Jet Tagging and Central Jet Veto cuts; the resulting signal-to-background ratio is close to 1 in the region of the mass peak.

6.3 Control Samples

The dominant background in this channel most likely comes from the QCD W+4 jets process. This process is $O(\alpha_s^4)$, and it is modeled using a tree-level matrix element. As a result, the scale uncertainty alone on this process is quite large; for example, simply multiplying or dividing the Q^2 scale by 4, the predicted cross-section changes by a factor of three or more. Moreover, it is not guaranteed that the shape of the $M_{l\nu qq}$ distribution is even accurately predicted by Monte Carlo; in particular, an imperfect jet calibration or a degraded missing P_T performance could distort the shape of the distribution. It is vital to understand such effects using data; to this end,

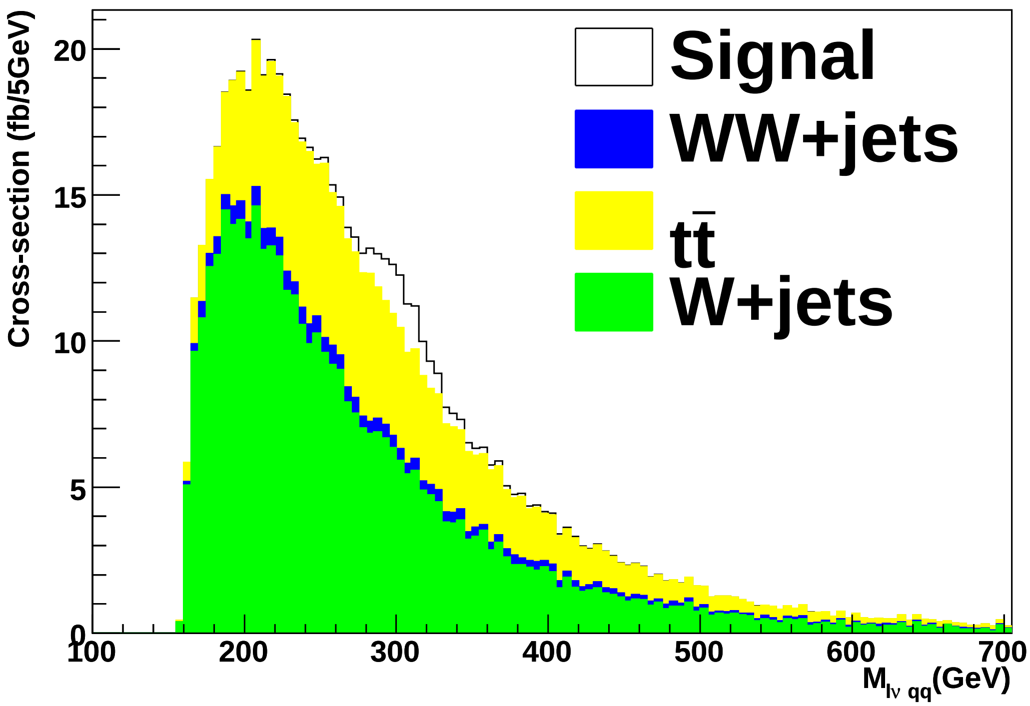

Figure 6.2: The $M_{l\nu qq}$ distributions in the b-vetoed control region of the $H+2j$, $H \to WW \to \ell\nu qq$ analysis.

a set of control samples have been studied.

The basis of the control sample strategy is the well-known observation that the both pseudorapidity gap between the tagjets and the invariant mass of the two tagjets tend to be larger in signal events than in the W+4 jet background events. The main control sample is defined as a region with loosened cuts on $\Delta\eta_{jj}$ and M_{jj}. In particular, the event selection is the same as in the signal-like region, except:

- The pseudorapidity gap between the tagjets is required to be less than 4 instead of larger than 4.4.
- The lower bound on the invariant mass of the tagjets is lowered to 500 GeV.

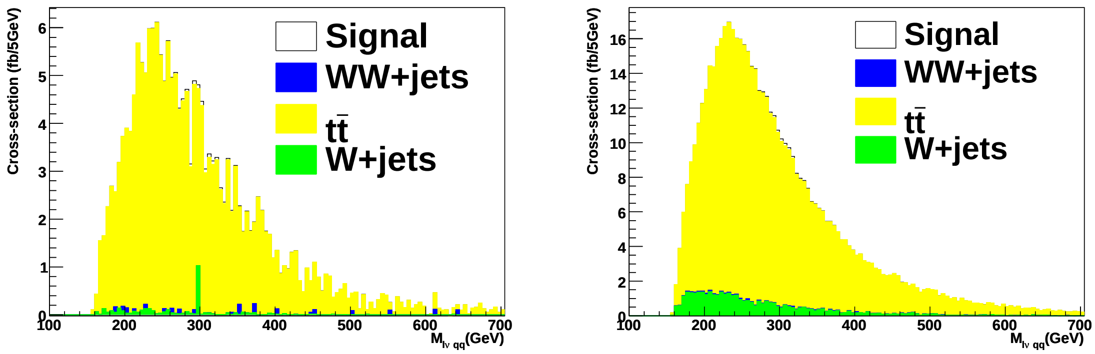

Figure 6.3: Left: the $M_{l\nu qq}$ distribution in the region with the same kinematic cuts as the signal-like region, but with the requirement of a b-tagged jet. Right: the $M_{l\nu qq}$ distribution in the region with a b-tag and altered kinematic cuts on the tagging jets.

Tables 6.1 and 6.2 include estimates of the signal and background cross-sections after these cuts. Figure 6.2 shows the $M_{l\nu qq}$ distribution in this region. Clearly, the shape of the W+jets background in this region is similar to the distribution in the signal-like region; however, both the signal-like region and this control sample include a nontrivial contamination from top background. It is therefore necessary to introduce two more control regions, with the same kinematic cuts as the existing regions but with the b-jet veto reversed. The cross-sections in these control samples are also shown in Tables 6.1 and 6.2, and the mass distributions are shown in Figure 6.3.

The following sources of systematic uncertainty have been studied to check that they do not introduce distortions that cause the background shapes in the signal-like and control regions to differ:

- the theoretical uncertainty as estimated by varying the relevant Q^2 scales in the Monte Carlo generator. The Q^2 scale for the W+jets background is varied by a factor of 4.

- the instrumental uncertainty due to the uncertainty in the jet energy scale. To model this, all jet four-momenta are scaled up by 5% in the region with $|\eta| < 2.5$ and by 10% elsewhere and the cross-sections are recomputed by applying the event selection cuts to the altered jet momenta; a similar exercise is performed by decreasing the energy scale by the same amount.

- the instrumental uncertainty that might result from a degraded jet energy resolution. To model a degraded resolution, the jet energies are smeared by a gaussian function with width given by $\frac{\delta E}{E} = \frac{a}{\sqrt{E}} \oplus b$. The values $a = 86.6\%$ and $b = 5.2\%$ are used for jets with $|\eta| < 3$, and $a = 173.2\%$ and $b = 12.12\%$ are used for jets with $|\eta| > 3$. These values cause the energy-dependent and constant terms of the smeared jet energy resolution to be approximately double the values applied in ATLFAST and used for the kinematic fit.

- the instrumental uncertainty that would arise if a shift were to be present in the missing P_T resolution. To model this scenario, the x and y components of the missing P_T are multiplied by a factor of 1.03.

- the instrumental uncertainty that might result from a degraded missing P_T resolution. To model such a degradation, the x and y components of the missing P_T are smeared by 5 GeV each.

It is worth noting that systematic errors arising from the uncertainty in the b-tagging efficiency have not yet been studied.

Figure 6.4 shows the ratio of the W+jets background shapes in the signal-like and control regions, with the systematic distortions discussed above. The plots are all

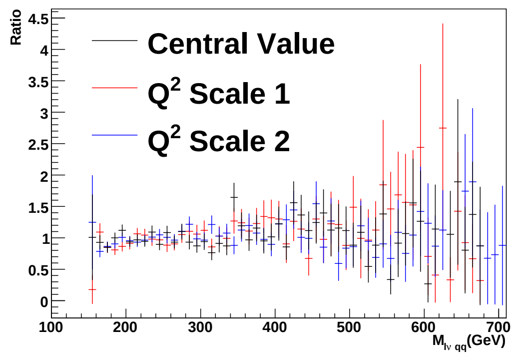

Figure 6.4: Left: the ratio of the $M_{l\nu qq}$ distribution in the signal-like and control regions, for the default choice of Q^2 scale, and with the Q^2 scale raised and lowered by a factor of 4. Right: the same ratio, but for several scenarios with instrumenal systematic distortions.

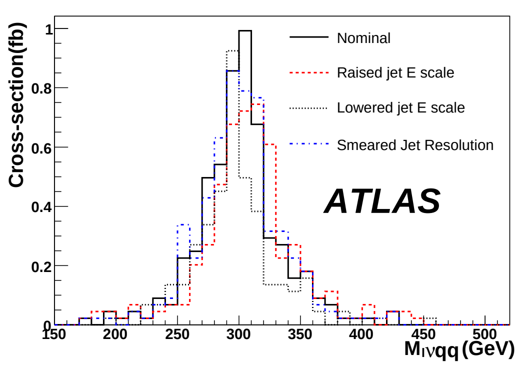

Figure 6.5: Left: the $M_{l\nu qq}$ distribution for the fully simulated signal sample, with and without several sources of instrumental systematic uncertainty. Right: the distribution of the Likelihood Ratio in background-only toy Monte Carlo outcomes and in outcomes which include a true Standard Model Higgs boson with several different choices of the M_H parameter, for an integrated luminosity of 10 fb^{-1}.

quite well-behaved to within statistics. The left plot in Figure 6.5 shows the effect of systematic distortions of the jet energy scale and resolution on the signal peak in full simulation.

6.4 Fitting Algorithm and Results

To turn the ideas presented in the previous section into a practical hypothesis test, a fitting algorithm which implements a simultaneous binned fit to the $M_{l\nu qq}$ distributions in the signal-like region and the main control region has been designed. The functional forms used are documented in Ref. [69]. The present discussion will give a brief overview of the fit procedure and then focus on performance evaluations using toy Monte Carlo.

The fit is a binned fit with a bin size of 5 GeV. As a first step, standalone fits of the top background model to the b-tagged control region are performed. The parameters are fixed, and the distributions are rescaled by the ratio of cross-sections in the b-tagged and b-vetoed regions to obtain estimates of the top background normalization and shape in both b-vetoed regions. Then, a simultaneous fit is performed, maximizing the Likelihood $L = L_{signal-like} \times L_{control}$ where $L_{signal-like}$ is the Likelihood in the signal-like region and $L_{control}$ is the Likelihood in the control region. The top background parameters are not allowed to float in the fit, but the parameters governing the shape and normalization of the W+jets background are. The W+jets background shape is assumed to be the same in the signal-like region and the control region, but the normalizations in the two regions are independent. Similarly, the signal shape is assumed to be the same in both regions, but the ratio of signal cross-sections is parameterized as a function of Higgs mass based on Monte Carlo. Two fits are

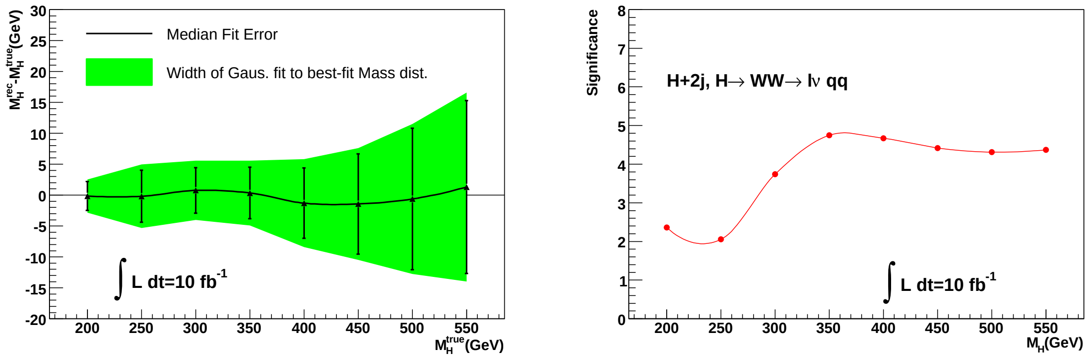

Figure 6.6: Left: the performance of the mass determination for the $H \to WW \to \ell\nu qq$ channel as a function of the Standard Model M_H parameter, after accumulating an integrated luminosity of 10 fb^{-1}. Right: the expected statistical significance at this luminosity.

performed, one fixing the signal normalization to zero and the other allowing both the mass and normalization to float. A Likelihood Ratio is defined as $L_{S+B}/L_{BG-only}$, where L_{S+B} is the value of the Likelihood function at the maximum of the fit with the signal normalization free and $L_{BG-only}$ is the value of the Likelihood function at the maximum of the fit with the signal normalization fixed to zero.

The performance of the fitting algorithm has been evaluated using two sets of toy Monte Carlo consisting of pseudo-experimental outcomes corresponding to a luminosity of 10 fb^{-1} with and without true signal events. The cross-sections for signal and background are taken from the full simulation results in Table 6.2; it is worth noting that the results presented here ignore any contribution from QCD multijet backgrounds and that the actual normalization of W+jets background may be quite different from the normalization assumed here. The right plot in Figure 6.5 shows the distribution of the Likelihood Ratio in background-only toy Monte Carlo outcomes and in several sets of signal-plus-background outcomes with various values of the

M_H parameter. The performance of the mass determination and the expected signal significance with an integrated luminosity of 10 fb^{-1} are shown in Figure 6.6. At this luminosity, the probability of a 5σ discovery approaches 50% if there is a Standard Model Higgs boson with a mass of 350 GeV. For Higgs masses below about 400 GeV, the expected error on the determination of M_H at this luminosity is less than about 6 GeV.

Chapter 7

Combined Sensitivity

The $H \to WW$ searches will of course not be the only Higgs boson searches performed at LHC; it is important to consider the sensitivity of the $H \to WW$ searches in the context of other Higgs boson searches. This Chapter will discuss a combined fit of the $H + 0j$ and $H + 2j$ searches described in Chapters 4 and 5 with each other and with the $H \to ZZ \to 4l$ search presented in Ref. [70].

The $H \to ZZ \to 4l$ channel has a much smaller signal cross-section than the $H \to WW$ channel because the branching ratio for $Z \to ll$ is considerably smaller than the branching ratio for $W \to l\nu$. However, it has the advantage that all the decay products of the Higgs are visible, so the mass of Higgs candidates can be fully reconstructed. The main discriminating variables in any $H \to 4l$ analysis are therefore M_{4l}, the invariant mass of the Higgs candidate, plus M_{12} and M_{34}, the invariant masses of the two Z bosons. In the analysis of Ref. [70], M_{12} was defined to be the mass of the Z candidate with mass closest to the true Z mass, and M_{34} was defined to be the mass of the more off-shell Z candidate. The $H \to ZZ \to 4l$ fit used here is a two-dimensional fit to the distribution of M_{34} and M_{4l}, with all correlations included. The main backgrounds in the $H \to ZZ \to 4l$ analysis are QCD ZZ and $Zb\bar{b}$ production.

The normalizations of these backgrounds are allowed to float freely in the fit, but the shapes of the distributions are taken to be fixed. The analysis of Ref. [70] studied the use of free parameters to incorporate shape uncertainties, and the loss in sensitivity was small (less than 1σ at 10 fb^{-1}) for most values of M_H.

The combined fits presented here follow the same formalism used in the previous chapters of this document. A Likelihood function is computed as the sum of the Likelihood functions of the individual channels, including separate free parameters for the normalization of each channel. The Likelihood function is maximized once allowing the signal normalizations to be free, and once subject to the constraint that they are all equal to zero. The Likelihood Ratio is computed as the ratio of the maximum values of the Likelihood function obtained for these two fits, $LR = L_{s+b}/L_{bg-only}$. Its sampling distribution is studied with toy Monte Carlo generated for a background-only scenario and for signal-plus-background scenarios with different Higgs masses. In the fits where the signal normalizations are not constrained to be zero, the Higgs mass is allowed to float in the range $120 < M_H < 200$ GeV.

Significances are computed using toy Monte Carlo. Since the available toy Monte Carlo statistics are insufficient to compute very small p-values numerically, the probability distribution for background-only outcomes is extrapolated to large Likelihood Ratios by fitting the tail of the distribution to an exponential function. Figure 7.1 shows the combined significances for the $H+0/2j$ ($H \to WW \to e\nu\mu\nu$) channels, and Figure 7.2 shows the impact of adding the $H \to ZZ \to 4l$ channel to the combined fit. Because the combined fits in these figures are more computationally expensive than the fits for individual channels, the number of toy Monte Carlo outcomes used

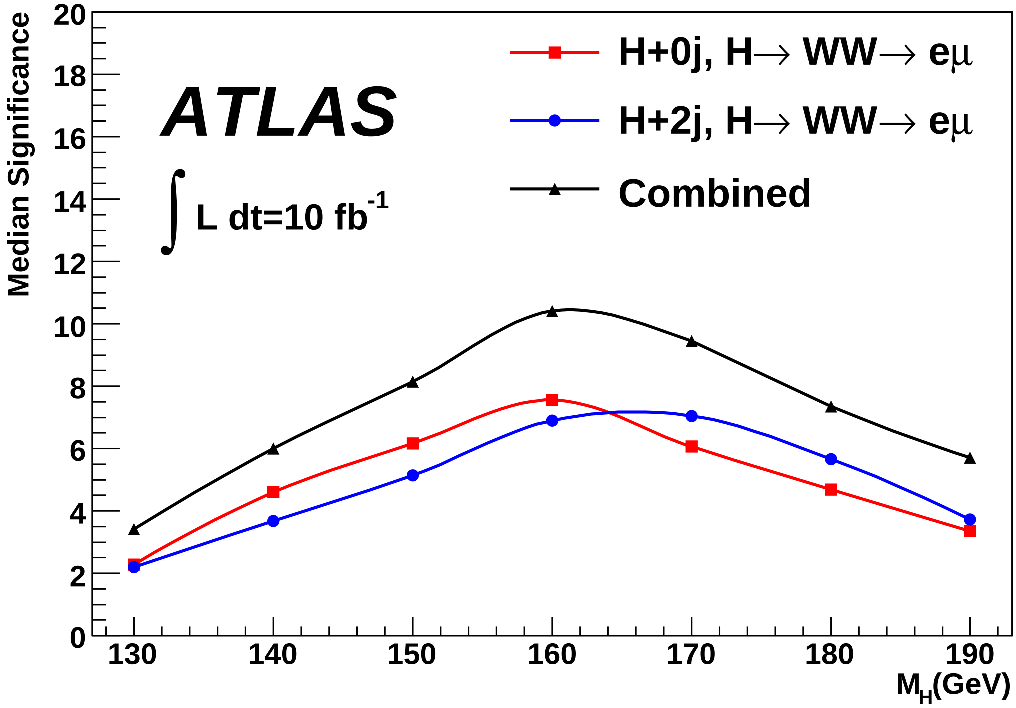

Figure 7.1: The combined significance for the $H + 0j$, $H \to WW \to e\nu\mu\nu$ channel and the $H + 2j$, $H \to WW \to e\nu\mu\nu$ channel after an integrated luminosity of 10 fb^{-1}.

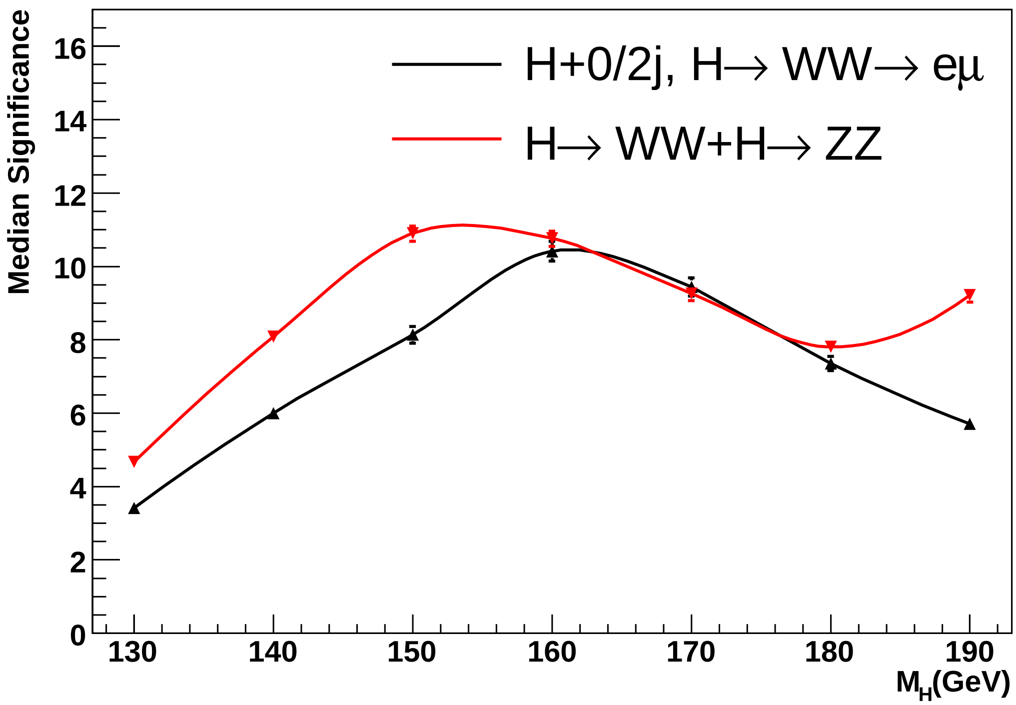

Figure 7.2: The combined significance for the $H + 0j$ ($H \to WW \to e\nu\mu\nu$), $H + 2j$ ($H \to WW \to e\nu\mu\nu$), and $H \to ZZ \to 4l$ channels after an integrated luminosity of 10 fb^{-1}.

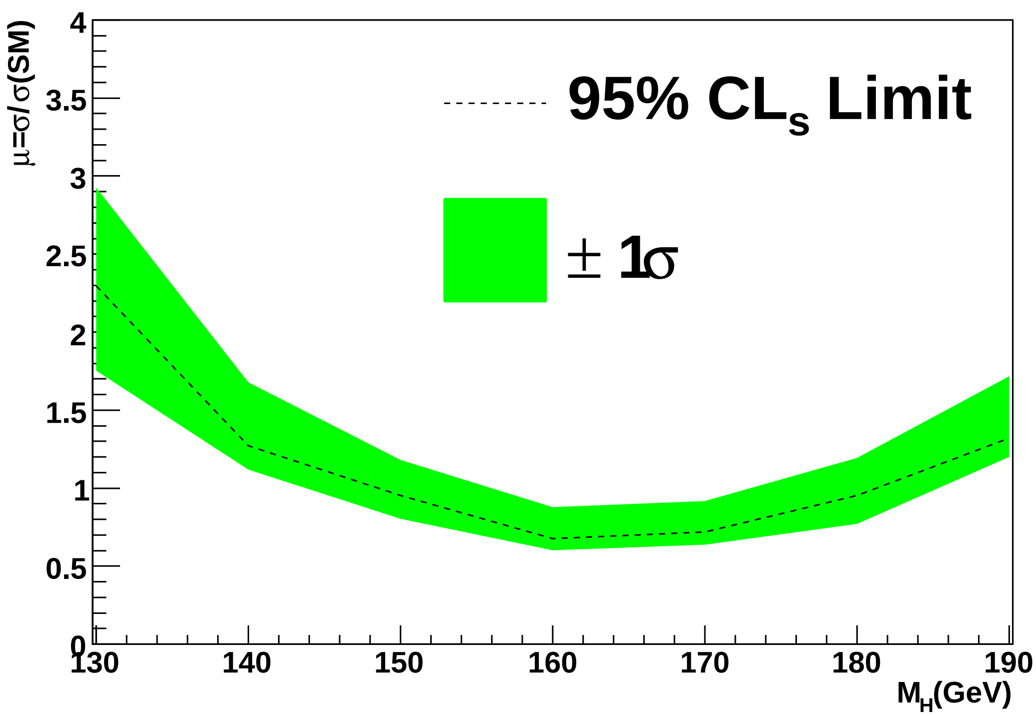

Figure 7.3: The upper bounds one would expect to set on the Higgs boson production cross-section (in units of the Standard Model prediction) using the combined fit of the $H \to WW \to e\nu\mu\nu$ channels at an integrated luminosity of 1 fb^{-1}.

to produce them is smaller than the statistics used for the individual channels, and the statistical error on the extrapolation of the Likelihood Ratio distribution to small p-values is important. The error bars in Figure 7.2 indicate the variation in the estimate of the significance due to the fit error in the slope of the tail of the Likelihood Ratio distribution. In the final analysis on real data, it would be desirable to use a larger toy Monte Carlo sample to evaluate the significance.

Figures 7.3 and 7.4 show the expected upper bound on the Higgs boson production cross-section in units of the Standard Model prediction. The former shows the limits for a combined fit of the $H \to WW \to e\nu\mu\nu$ channels; the latter includes the

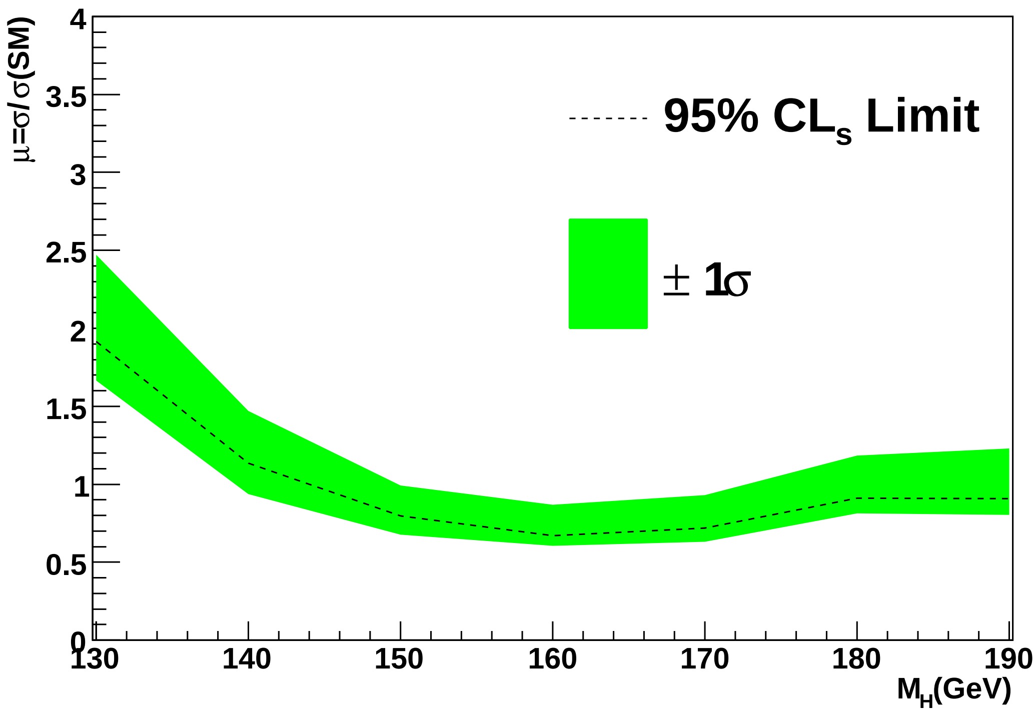

Figure 7.4: The upper bounds one would expect to set on the Higgs boson production cross-section (in units of the Standard Model prediction) using the combined fit of the $H \to WW \to e\nu\mu\nu$ and $H \to 4l$ channels at an integrated luminosity of 1 fb^{-1}.

$H \to 4l$ channel as well. The plots show that with 1 fb^{-1} of integrated luminosity, one would expect to be able to exclude a Standard Model Higgs boson in the range $150 < M_H < 180$ GeV using only the $H \to WW \to e\nu\mu\nu$ channels; if the $H \to 4l$ channel is included, the excluded region covers $M_H > 145$ GeV.

Bibliography

[1] Yang, C. N. and Mills, R. L., Phys. Rev. **96**, 191 (1954).

[2] Glashow, S. L., Nucl. Phys. **22**, 579 (1961).

[3] Salam, Abdus and Ward, John Clive, Nuovo Cim. **19**, 165 (1961).

[4] Salam, Abdus and Ward, John Clive, Phys. Lett. **13**, 168 (1964).

[5] Weinberg, Steven, Phys. Rev. Lett. **19**, 1264 (1967).

[6] Fritzsch, H. and Gell-Mann, Murray and Leutwyler, H., Phys. Lett. **B47**, 365 (1973).

[7] Phys. Rept. **427**, 257 (2006).

[8] Englert, F. and Brout, R., Phys. Rev. Lett. **13**, 321 (1964).

[9] Higgs, Peter W., Phys. Lett. **12**, 132 (1964).

[10] Higgs, Peter W., Phys. Rev. Lett. **13**, 508 (1964).

[11] Higgs, Peter W., Phys. Rev. **145**, 1156 (1966).

[12] Guralnik, G. S. and Hagen, C. R. and Kibble, T. W. B., Phys. Rev. Lett. **13**, 585 (1964).

[13] Kibble, T. W. B., Phys. Rev. **155**, 1554 (1967).

[14] Bruning, O. , (Ed.) and others. CERN-2004-003.

[15] The ATLAS Collaboration, *The ATLAS Experiment at the Large Hadron Collider*. Submitted to JINST.

[16] Wu, C. S. and Ambler, E. and Hayward, R. W. and Hoppes, D. D. and Hudson, R. P., Phys. Rev. **105**, 1413 (1957).

[17] Garwin, R. L. and Lederman, L. M. and Weinrich, Marcel, Phys. Rev. **105**, 1415 (1957).

[18] Goobar, Ariel and Hannestad, Steen and Mortsell, Edvard and Tu, Huitzu, JCAP **0606**, 019 (2006).

[19] Danby, G. and Gaillard, J-M. and Goulianos, K. and Lederman, L. M. and Mistry, N. and Schwartz, M. and Steinberger, J., Phys. Rev. Lett. **9**, 36 (1962).

[20] J. F. Gunion, H. E. Haber, and G. Kane, *The Higgs Hunters' Guide*. Perseus Publishing, Cambridge, Mass., 1990.

[21] Barate, R. and others, Phys. Lett. **B565**, 61 (2003).

[22] B. Mellado, A. Nisati, D. Rebuzzi, S. Rosati, G. Unal, and Sau Lan Wu, *Higgs Production Cross-sections and Branching Ratios for the ATLAS Higgs Working Group*. ATLAS Internal Note, in preparation.

[23] Djouadi, A. and Kalinowski, J. and Spira, M., Comput. Phys. Commun. **108**, 56 (1998).

[24] Barger, V. and Bhattacharya, G. and Han, T. and Kniehl, B. A., Phys. Rev. D **43**, 779 (1991).

[25] M. Dittmar and H. Dreiner, Phys. Rev. **D55**, 167 (1997).

[26] ATLAS Collaboration, *ATLAS Detector and Physics Performance*. CERN/LHCC 99-15 (1999).

[27] The CMS Collaboration, *CMS Techical Design Report*. J. Phys. G: Nucl. Part. Phys. 34.

[28] B. Mellado, W. Quayle, Sau Lan Wu, *Prospects for a Higgs discovery in the channel $H \to WW \to l\nu l\nu$ with no hard jets*. ATLAS communication ATL-COM-PHYS-2005-074(2005).

[29] S. Frixione and B.R. Webber, JHEP **0206**, 029 (2002).

[30] S. Frixione, P. Nason and B.R. Webber, JHEP **0308**, 007 (2003).

[31] Torbjörn Sjöstrand, *High-Energy Physics Event Generation with PYTHIA 5.7 and JETSET 7.4*. Comp. Phys. Comm. **82** (1994) 74.

[32] Torbjörn Sjöstrand et al., *High-Energy-Physics Event Generation with PYTHIA 6.1*. Comp. Phys. Comm. **135** (2001) 238.

[33] T. Gleisberg, S. Hoeche, F. Krauss, A. Schaelicke, S. Schumann, J. Winter, JHEP **0402**, 056 (2004).

[34] Pumplin, J. and others, JHEP **07**, 012 (2002).

[35] Binoth, T. and Ciccolini, M. and Kauer, N. and Kramer, M., JHEP **03**, 065 (2005).

[36] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour and B. R. Webber, JHEP **0101**, 010 (2001).

[37] T. Dai, *$gg \to WW \to l\nu l\nu$ Production with JIMMY/HERWIG*. Talk in CSC-HG4 meeting of 24 September, 2007.

[38] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, JHEP **07**, 001 (2003).

[39] Mangano, Michelangelo L. and Moretti, Mauro and Pittau, Roberto, Nucl. Phys. **B632**, 343 (2002).

[40] F. Caravaglios, M. L. Mangano, M. Moretti, and R. Pittau, Nucl. Phys. **B539**, 215 (1999).

[41] E. Richter-Was, D. Froidevaux and L. Poggioli, *ATLFAST2.0 a Fast Simulation Package for ATLAS*. ATLAS Internal Note ATL-PHYS-98-131.

[42] A. Armbruster, H. Liu, J. Purdham, J. Qian, J. Strandberg, R. Thun, *Lepton ID and Fake Rate Studies for the CSC $H \to WW$ Analysis*. ATLAS Communication ATL-COM-PHYS-2008-017 (2008).

[43] M. Rast, *Studie zum Entdeckungspotential im Prozess Higgs $\to WW \to l\nu l\nu$ mit dem ATLAS-Experiment unter Beruecksichtigung des W+Jets-Untergrundes*. Available as BONN-IB-2007-09.

[44] M. Rast, E. von Törne, M. Pleier, N. Wermes, *W+jets background to $H \to WW \to l\nu l\nu$ searches at the ATLAS experiment*. Available as ATL-INT-2008-016.

[45] V. Cavasinni, D. Costanzo, I. Vivarelli, *Forward Tagging and Jet Veto Studies for Higgs Events Produced via Vector Boson Fusion*. ATLAS internal note ATL-PHYS-2002-008 (2002).

[46] H. Yang, T. Dai, D. Levin, X. Li, A. Wilson, Z. Zhao, B. Zhou, *ATLAS $H \to WW$ Detection Sensitivity with Boosted Decision Trees*. ATLAS Communication ATL-COM-PHYS-2008-023 (2008).

[47] W. Verkerke and D. Kikrby, *The RooFit Toolkit for Data Modeling*. Proceedings of the PhyStat2005 workshop.

[48] T. Vickey *et al.*, *A Data-driven Control Sample for $Z \to \tau\tau$ + jets Backgrounds to Higgs Searches*. In preparation.

[49] Read, Alexander L., J. Phys. **G28**, 2693 (2002).

[50] Cahn, R. N. and Dawson, Sally, Phys. Lett. **B136**, 196 (1984).

[51] Cahn, Robert N. and Ellis, Stephen D. and Kleiss, Ronald and Stirling, W. J., Phys. Rev. D **35**, 1626 (1987).

[52] Rainwater, D. and Zeppenfeld, D., Phys. Rev. D **60**, 113004 (1999).

[53] Kauer, N. and Plehn, T. and Rainwater, David L. and Zeppenfeld, D., Phys. Lett. **B503**, 113 (2001).

[54] K. Cranmer, P. McNamara, B. Mellado, W. Quayle, S. L. Wu, *Search for Higgs Bosons Decay $H \to WW^{(*)} \to l^+l^- P_T^{miss}$ for $115 < M_H < 130$ GeV Using Vector Boson Fusion.* ATLAS Internal Note ATL-PHYS-2003-02 (2003).

[55] K. Cranmer, P. McNamara, B. Mellado, Y. Pan, W. Quayle, Sau Lan Wu, *Neural Network Based Search for Higgs Boson Produced via VBF with $H \to W^+W^- \to l^+l^- P_T^{miss}$ for $115 < 130$ GeV.* ATLAS Internal note ATL-PHYS-2003-007.

[56] S. Asai *et al.*, European Physical Journal C **32**, 19 (2004).

[57] K. Cranmer, Y. Fang, B. Mellado, S. Paganis, W. Quayle, Sau Lan Wu, *Analysis of VBF $H \to WW \to l\nu l\nu$.* ATLAS Internal Note ATL-PHYS-2004-019.

[58] Gleisberg, Tanju and others, JHEP **02**, 056 (2004).

[59] A. Armbruster, H. Liu, J. Purdham, J. Qian, J. Strandberg, R. Thun, *Using b-tagging to Reject ttbar Background in the CSC $H \to WW$ Analysis.* ATLAS Communication ATL-COM-PHYS-2008-018 (2008).

[60] I. Tsukerman, *Characteristics of jets and E_t^{miss} in VBF Higgs and minimum-bias CSC MC samples.* ATL-COM-PHYS-2007-084.

[61] I. Tsukerman, *Forward jet tagging and central jet veto in VBF Higgs CSC MC samples.* ATL-COM-PHYS-2007-083.

[62] M. Baak, *Note on 5D fit of H+2j, $H \to WW \to l\nu l\nu$.* ATLAS Internal Note, in preparation.

[63] Iordanidis, K. and Zeppenfeld, D., Phys. Rev. D **57**, 3072 (1998).

[64] A. Erdogan, D. Froidveaux, S. Klioukhine, and S. V. Zmushko, *Study of $H \to WW \to l\nu jj$ and $H \to ZZ \to lljj$ decays for M_H = 1 TeV.* ATLAS internal note ATL-PHYS-92-008 (1992).

[65] D. Froidevaux, Luc Poggioli, and S. V. Zmushko, *$H \to WW \to l\nu jj$ and $H \to ZZ \to lljj$ Particle level studies.* ATLAS internal note ATL-PHYS-97-103.

[66] V. Cavasinni, D. Costanzo, S. Lami, and F. Spano, *Search for H to WW to l nu jj with the ATLAS detector (mH = 300-600 GeV)*. ATLAS internal note ATL-PHYS-98-127.

[67] V. Cavasinni, D. Costanzo, E. Mazzoni, I. Vivarelli, *Search for a Intermediate Mass Higgs boson produced via Vector Boson Fusion in the channel* $H \rightarrow WW \rightarrow l\nu jj$ *with the ATLAS detector*. ATLAS internal note ATL-PHYS-2002-010 (2002).

[68] C. Le Maner, Luc Poggioli, H. Przysiezniak, Elzbieta Richter-Was. ATLAS internal note ATL-PHYS-2004-003.

[69] N. Goldschmidt, J. King, B. Mellado, S. Padhi, W. Quayle, Sau Lan Wu, *Fit-based Search for* $H + 2j$, $H \rightarrow \ell\nu qq$. In preparation.

[70] L. Flores *et al.*, *Search for* $H \rightarrow ZZ^{(*)} \rightarrow 4l$ *using a Global Fit Procedure*. In preparation.

Appendix A

Functional Forms for the $H + 0j$ Fit

This Appendix describes the functional forms that are used for signal and the various backgrounds of the $H{+}0j$, $H \to WW \to e\nu\mu\nu$ fit and documents the true values of the parameters for the nominal and systematically distorted cases considered. Section A.1 discusses the signal model and the parameterization of the expected ratio of cross-sections in the regions with large and small $\Delta\phi_{ll}$. Section A.2 describes the functional form used to model the QCD WW background, the parameterization of the nuisance parameters that govern the Higgs P_T distribution, and the parameters that govern the extrapolation of the Higgs P_T shape from the large-$\Delta\phi_{ll}$ region to the small-$\Delta\phi_{ll}$ region. Section A.3 deals primarily the extrapolation of the shape of the M_T vs. P_T^{Higgs} distribution from the b-tagged regions to the b-vetoed regions, since the functional form that is used to describe the top background in the b-tagged control regions is the same as the functional form for QCD WW. Section A.4 describes the functional form that is used to model the $Z \to \tau\tau$ background, and Section A.5 describes the functional form used to model the W+jets background.

A.1 Signal Model

The functional form used for the signal model is a 2-dimensional product PDF which assumes that the Higgs P_T and the transverse mass are uncorrelated. The function used to describe the transverse mass distribution is an exponentiated gradual step function with a Fermi-Dirac cutoff at low M_T:

$$e^{-\sqrt{(M_T-m_3)^2(m_1+m_2/(1+e^{10(M_T-m_3)}))^2+m_4}}/(1+e^{m_5(M_T-m_6)}) \tag{A.1}$$

The function used to describe the Higgs P_T is a falling exponential with a gradual cutoff near zero and an inverted Gaussian to model a small dip in the distribution:

$$e^{-|p_1|P_T^{Higgs}}e^{-e^{(p_2-P_T^{Higgs})/p_3}}[1-p_6\text{Gaus}(P_T^{Higgs},p_4,p_5)] \tag{A.2}$$

It would be impractical to allow all the m_i and p_i to float in the fit, so they are parameterized as functions of the Higgs mass based on fits to a set of fast simulation Monte Carlo samples. Separate parameterizations in the region with small $\Delta\phi_{ll}$ and the region with large $\Delta\phi_{ll}$ are necessary in order to achieve a good fit to the Monte Carlo. The m_i and p_i parameters for the region with small $\Delta\phi_{ll}$ are parameterized as follows:

- $m_1 = 27.92371 - 0.6688822M_H + (5.958503\times10^{-3})M_H^2 - (2.329679\times10^{-5})M_H^3 + (3.374844\times10^{-8})M_H^4$
- $m_2 = 60.84838 - 2.001317M_H + 0.02612993M_H^2 - (1.692307\times10^{-4})M_H^3 + (5.431399\times10^{-7})M_H^4 - (6.905989\times10^{-10})M_H^5$
- $m_3 = -121.9439 + 2.633943M_H - (5.661873\times10^{-3})M_H^2$

- $m_4 = 341.8834 - 4.607360 M_H + 0.01558272 M_H^2$ for $M_H < 160$ GeV, and $m_4 = -566.338808 + 3.562264 M_H$ for $M_H > 160$ GeV

- $m_5 = -0.2211633$

- $m_6 = 76.60894$

- $p_1 = -2.431103 + 0.06633037 M_H - (6.238295 \times 10^{-4}) M_H^2 + (2.543430 \times 10^{-6}) M_H^3 - (3.804727 \times 10^{-9}) M_H^4$

- $p_2 = 842.8998 - 26.80440 M_H + 0.3394536 M_H - (2.132685 \times 10^{-3}) M_H + (6.642227 \times 10^{-6}) M_H^4 - (8.199587 \times 10^{-9}) M_H^5$

- $p_3 = 479.1628 - 15.32836 M_H + 0.1954649 M_H^2 - (1.236101 \times 10^{-3}) M_H^3 + (3.873801 \times 10^{-6}) M_H^4 - (4.810455 \times 10^{-9}) M_H^5$

- $p_4 = 1694.568 - 54.00929 M_H + 0.6856918 M_H^2 - (4.322087 \times 10^{-3}) M_H^3 + (1.350899 \times 10^{-5}) M_H^4 - (1.673492 \times 10^{-8}) M_H^5$

- $p_5 = -1494.894 + 38.44581 M_H - 0.3580025 M_H^2 + (1.451681 \times 10^{-3}) M_H^3 - (2.163122 \times 10^{-6}) M_H^4$

- $p_6 = 52.46458 - 2.207557 M_H + 0.03826075 M_H^2 - (3.446074 \times 10^{-4}) M_H^3 + (1.706065 \times 10^{-6}) M_H^4 - (4.413346 \times 10^{-9}) M_H^5 + (4.671749 \times 10^{-12}) M_H^6$

The parameterization in the region with large $\Delta\phi_{ll}$ is:

- $m_1 = 4.599432 - 0.08268124 M_H + (5.089156 \times 10^{-4}) M_H^2 - (1.040934 \times 10^{-6}) M_H^3$

- $m_2 = -4.153717 + 0.08102855 M_H - (5.178389 \times 10^{-4}) M_H^2 + (1.079090 \times 10^{-6}) M_H^3$

- $m_3 = 138.2423 - 0.7532171 M_H + (5.531297 \times 10^{-3}) M_H^2$
- $m_4 = 93.30377$
- $m_5 = -0.3505511 + (1.223013 \times 10^{-3}) M_H$
- $m_6 = 28.89743 + 0.3718254 M_H$
- $p_1 = 0.6391961 - (8.421393 \times 10^{-3}) M_H + (4.639370 \times 10^{-5}) M_H^2 - (8.304164 \times 10^{-8}) M_H^3$
- $p_2 = 10.03146 - 0.08925018 M_H + (2.557221 \times 10^{-4}) M_H^2$
- $p_3 = 6.091894 - 0.04697794 M_H + (1.329996 \times 10^{-4}) M_H^2$
- $p_4 = 26.12350 - 0.2629622 M_H + (7.383113 \times 10^{-4}) M_H^2$
- $p_5 = 73.66466 - 0.6967552 M_H + (2.352613 \times 10^{-3}) M_H^2$
- $p_6 = 1.299686 - (4.964745 \times 10^{-3}) M_H + (1.633660 \times 10^{-5}) M_H^2$

Figures A.9 through A.48 show comparisons between the signal model and the signal Monte Carlo for Higgs masses ranging from 120 GeV to 200 GeV. Generally, the agreement is quite good.

The signal model discussed above was derived from privately generated Monte Carlo samples. As a cross-check, a special signal model for $M_H = 170$ GeV was generated based on the official fully simulated Monte Carlo sample of Dataset 6533 (MC@NLO). Figures A.49 through A.52 show the the fit to this Monte Carlo sample. The main difference between this model and the one used in the fit is that the Higgs P_T distribution is slightly harder in this model than in the one used in the fit.

For signal, it is assumed that the ratio of cross-sections in the two bins of $\Delta\phi_{ll}$ is well-known. This ratio is computed from Monte Carlo and the result is parameterized as a function of M_H. The function used to model the ratio R is a fifth order polynomial minus two gaussians:

$$\begin{gathered} R = 339.415 - 11.2164 M_H + 0.146685 M_H^2 - (9.46890 \times 10^{-4}) M_H^3 \\ +(3.01629 \times 10^{-6}) M_H^4 - (3.79195 \times 10^{-9}) M_H^5 - 0.150730 \text{ Gaus}(M_H, 160, 9.76027) \\ -0.0546370 \text{ Gaus}(M_H, 138.615, 5.86168) \end{gathered} \tag{A.3}$$

Figure A.53 shows a comparison between the values of R obtained from Monte Carlo and the parameterization; the agreement between the Monte Carlo and the model is good.

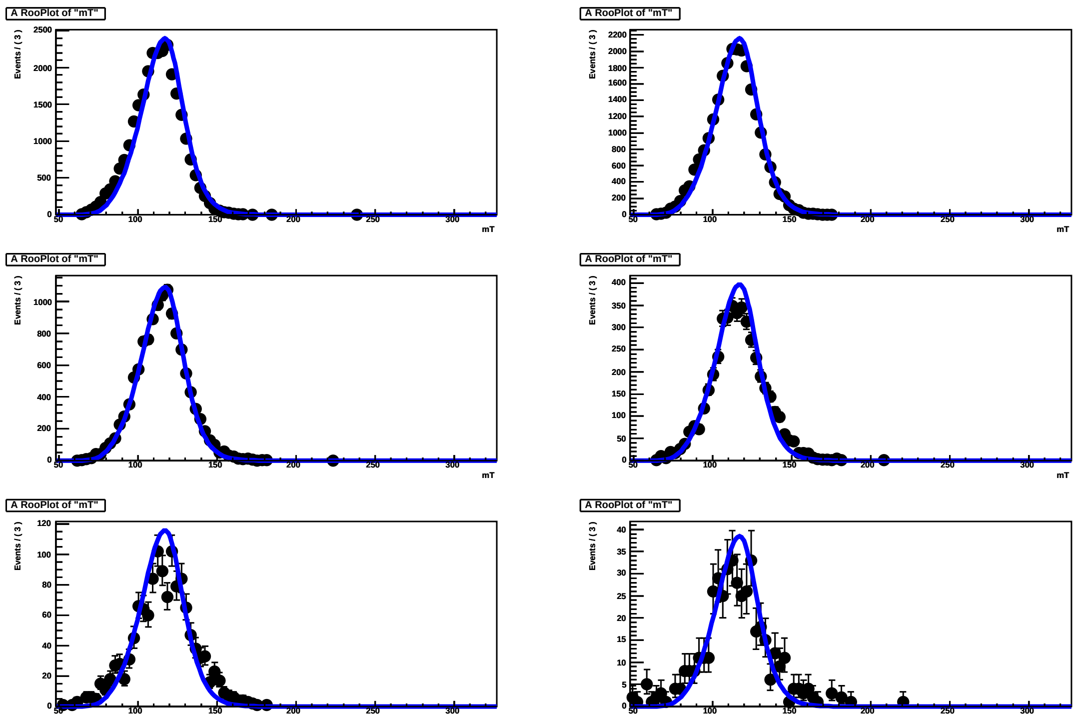

Figure A.1: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 120 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

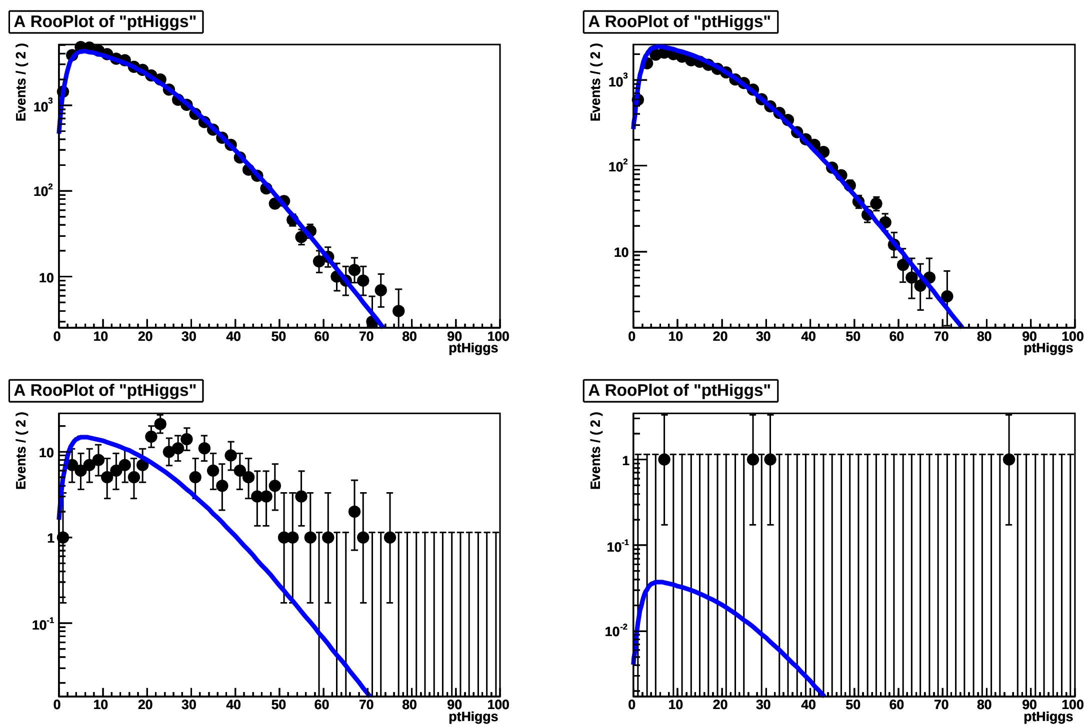

Figure A.2: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 120 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

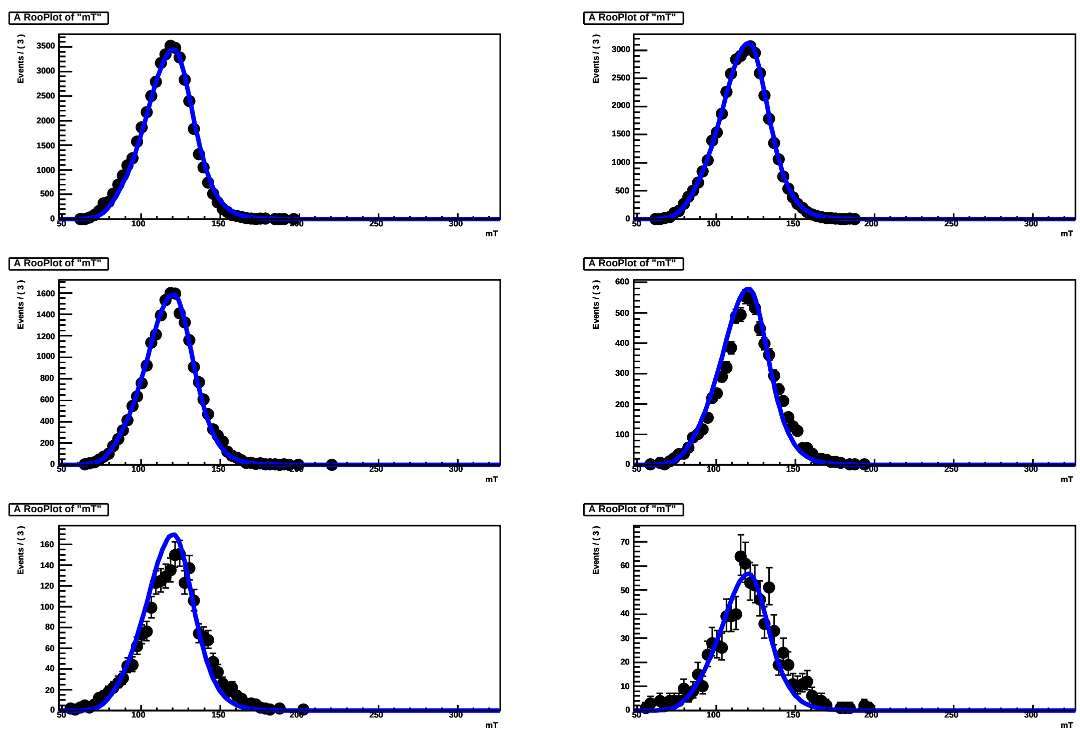

Figure A.3: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 125 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

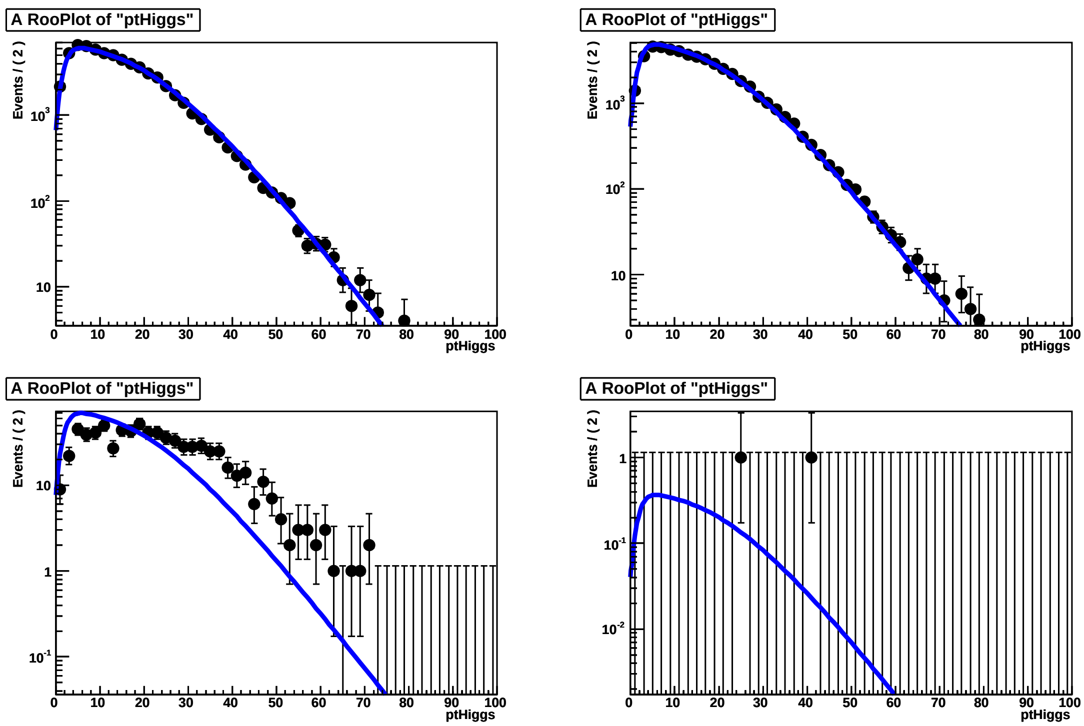

Figure A.4: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 125 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

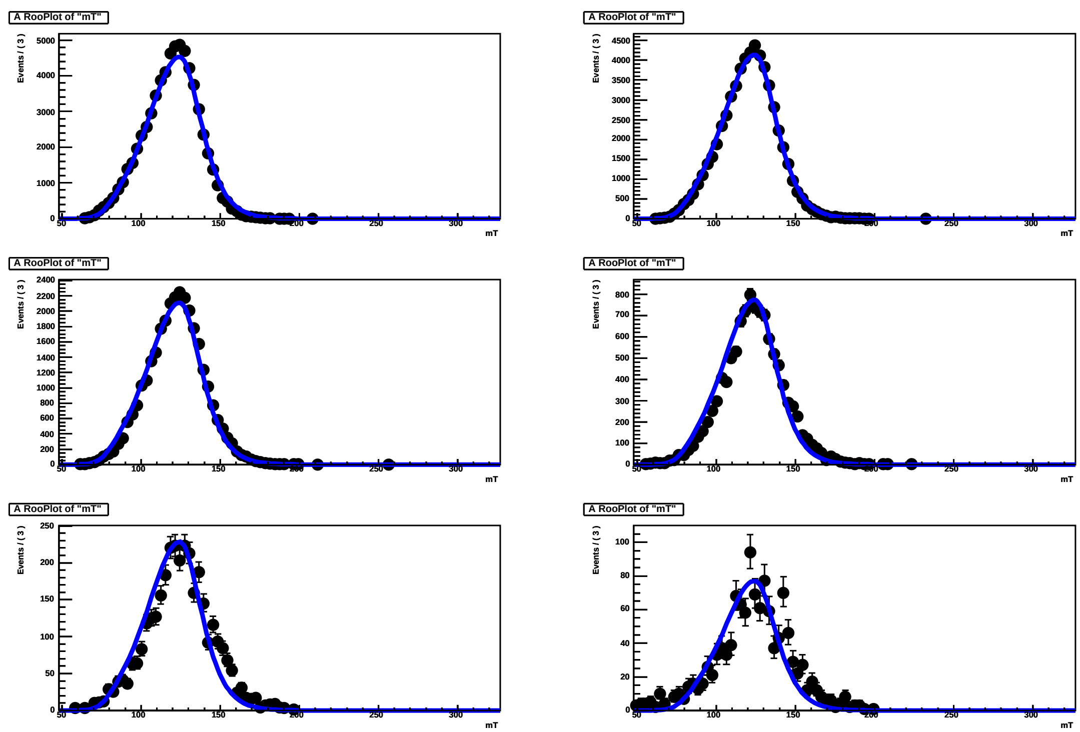

Figure A.5: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 130 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

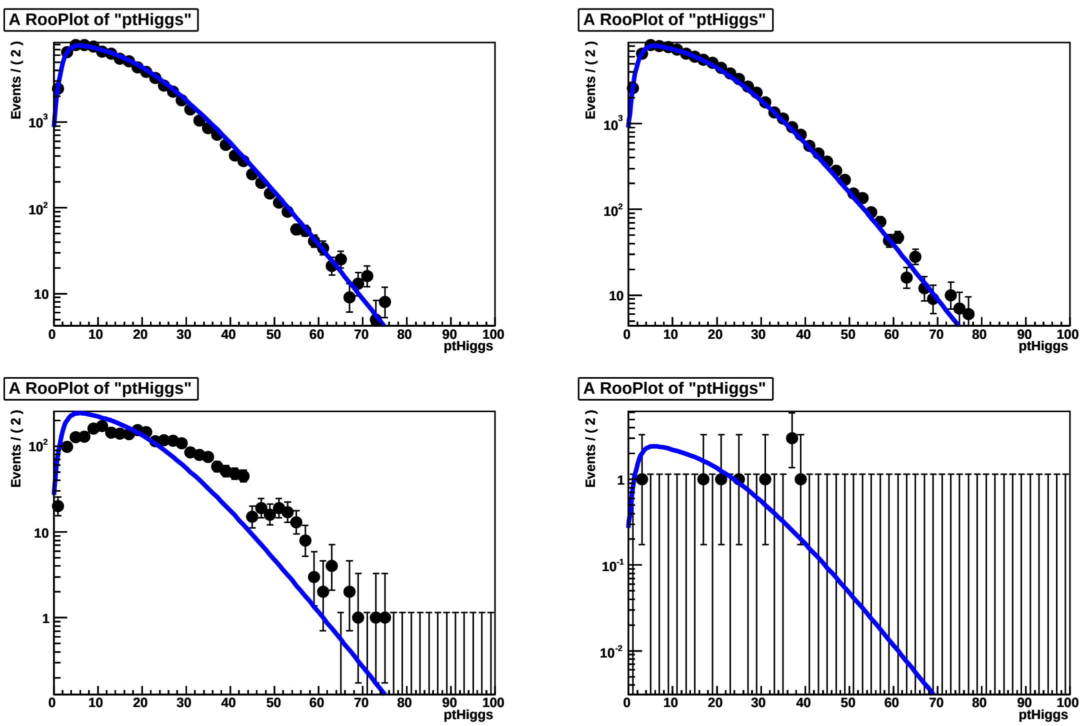

Figure A.6: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 130 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

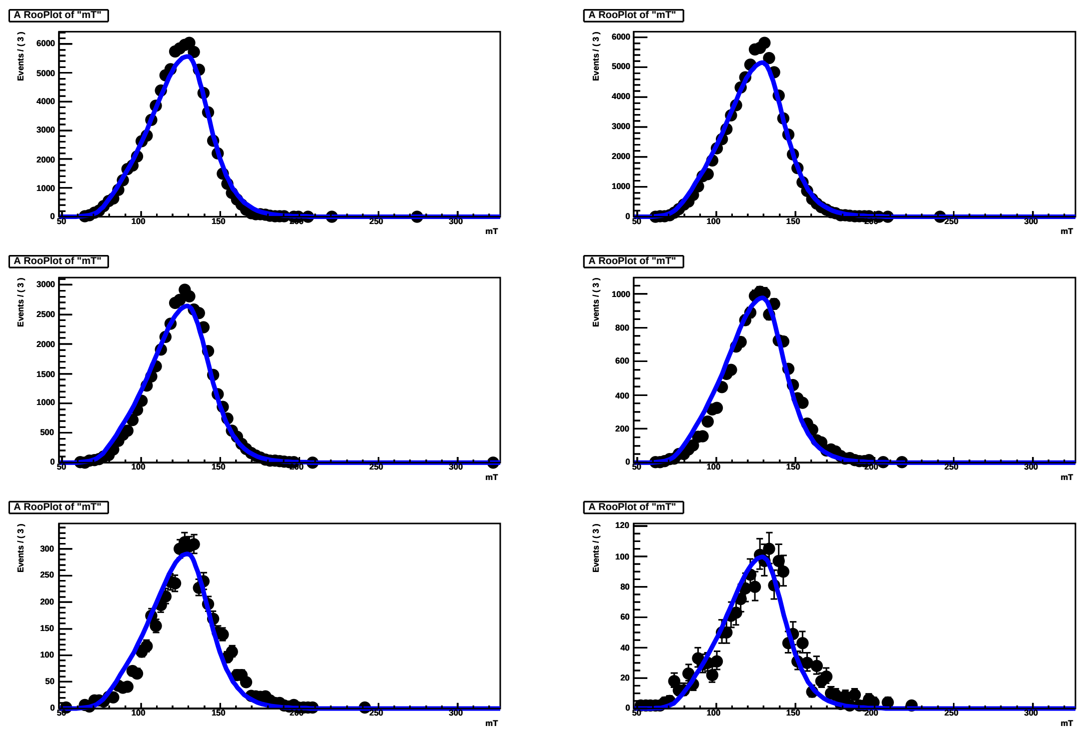

Figure A.7: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 135 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

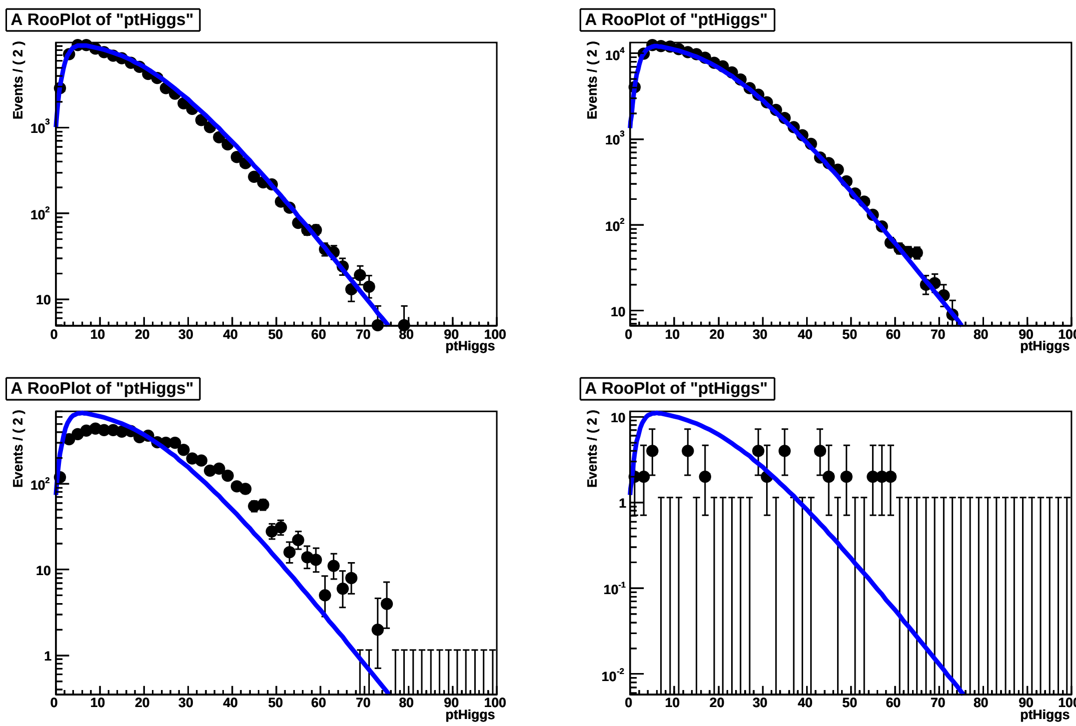

Figure A.8: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 135 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

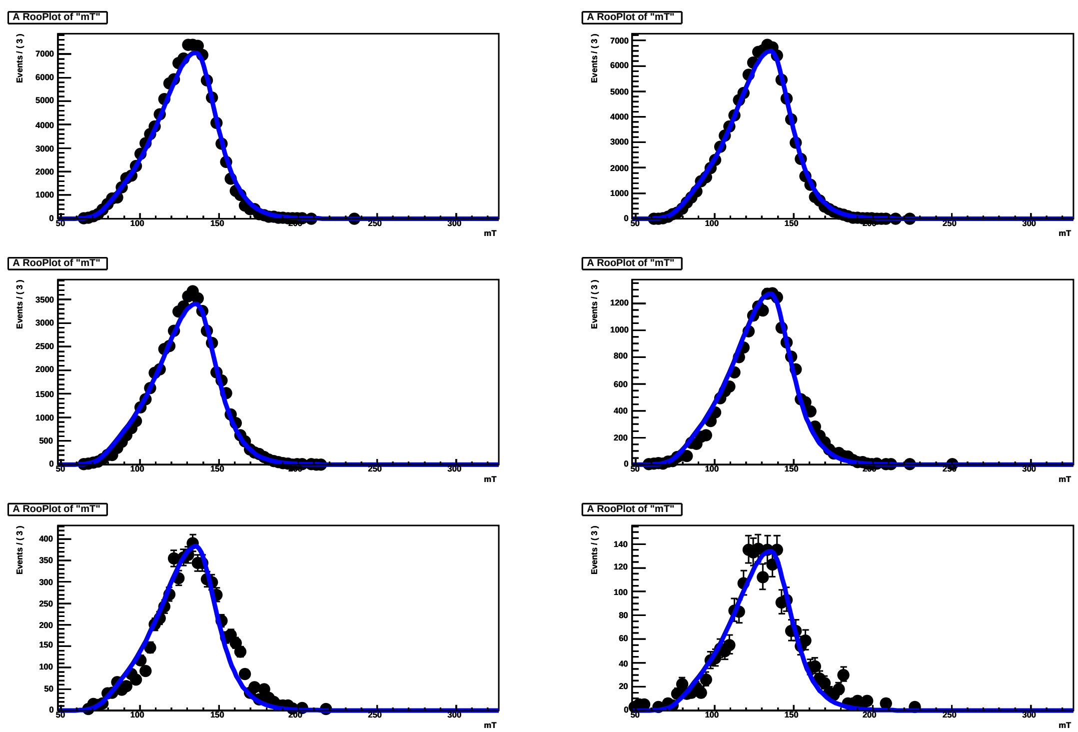

Figure A.9: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 140 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

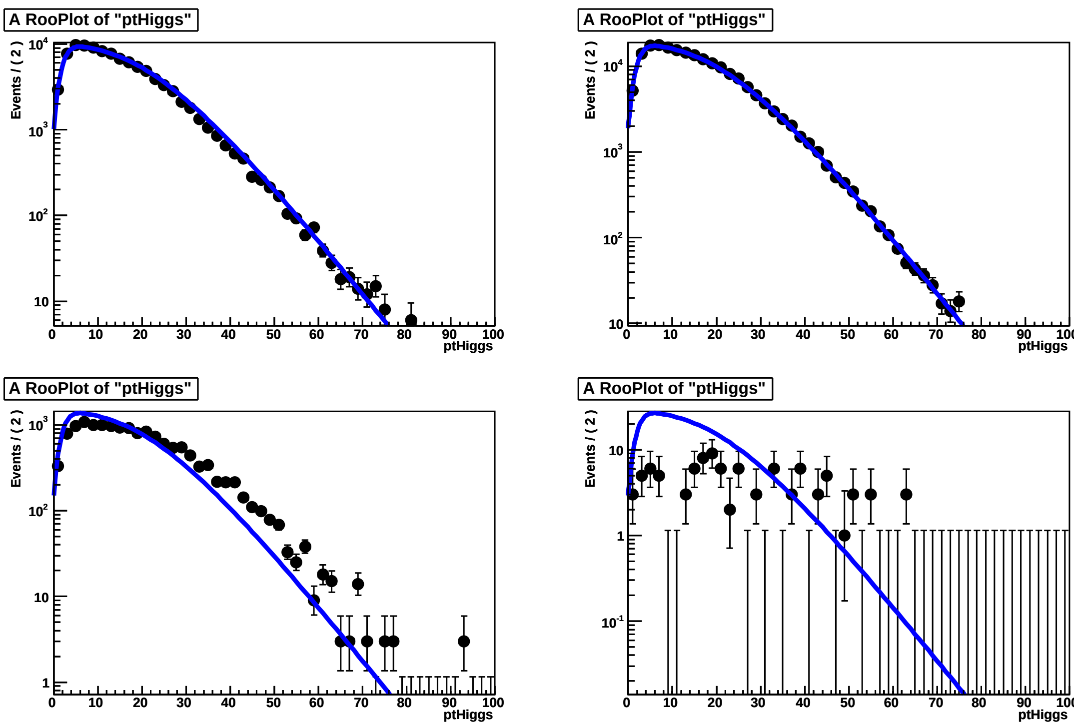

Figure A.10: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 140 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

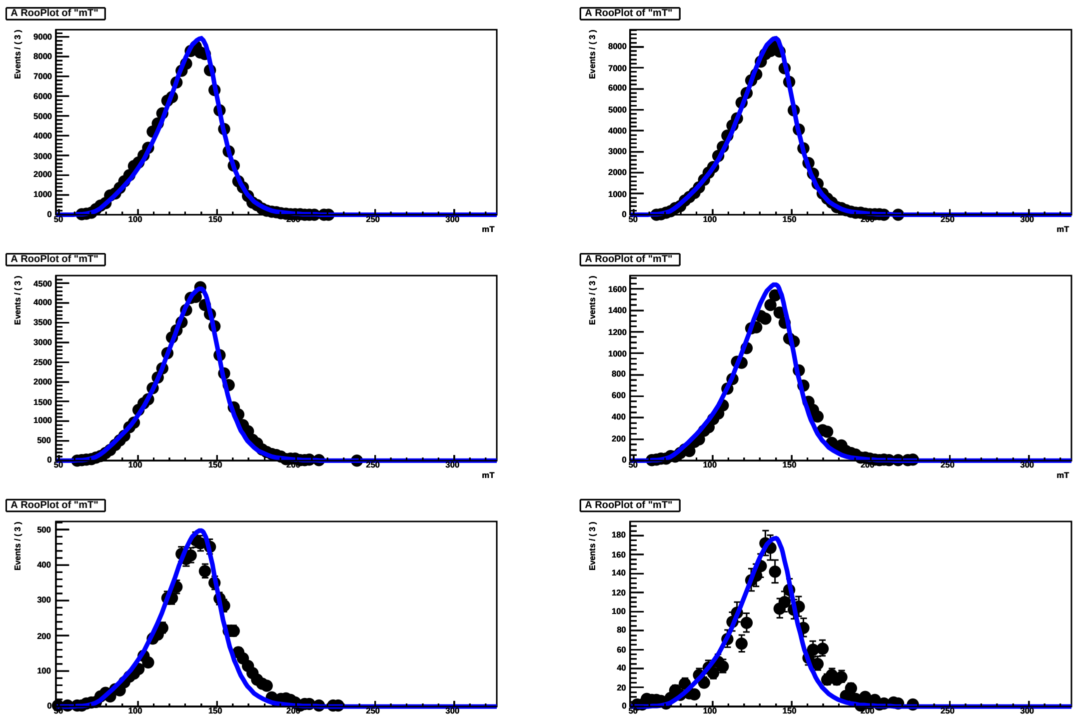

Figure A.11: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 145 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

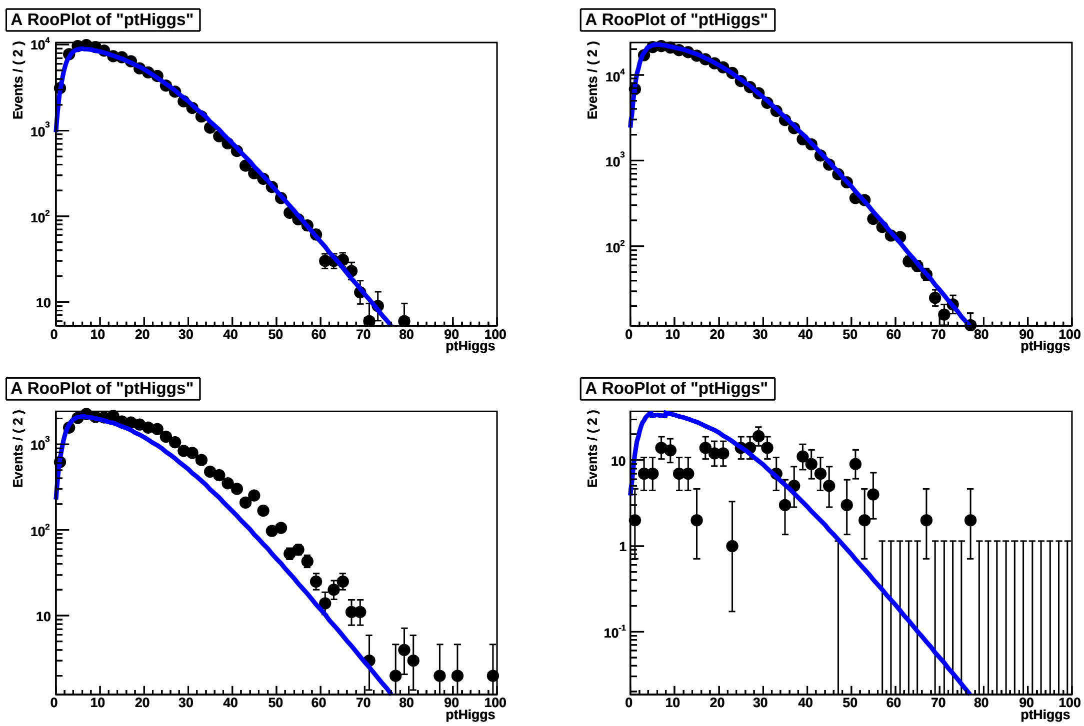

Figure A.12: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 145 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

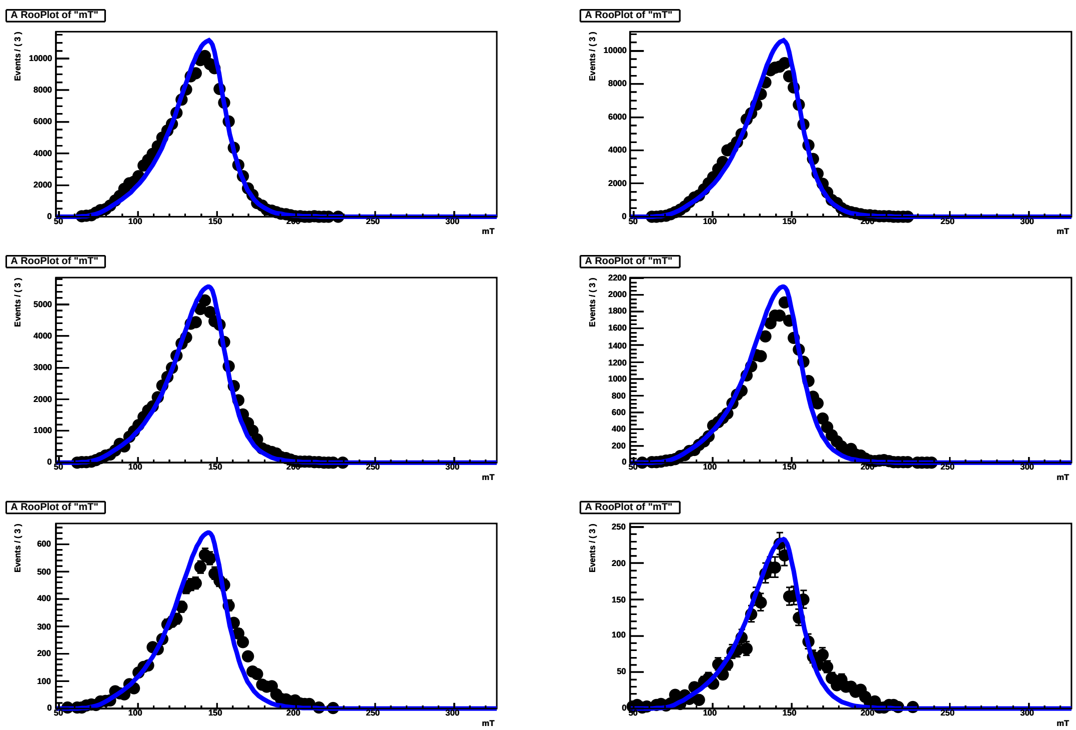

Figure A.13: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 150 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

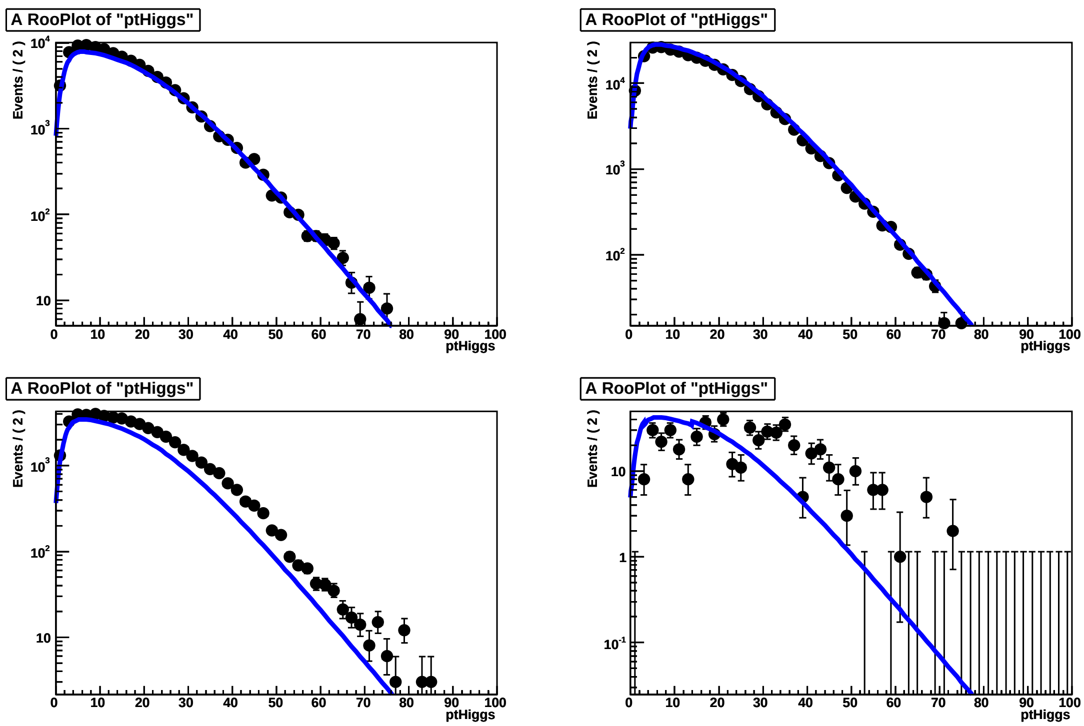

Figure A.14: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 150 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

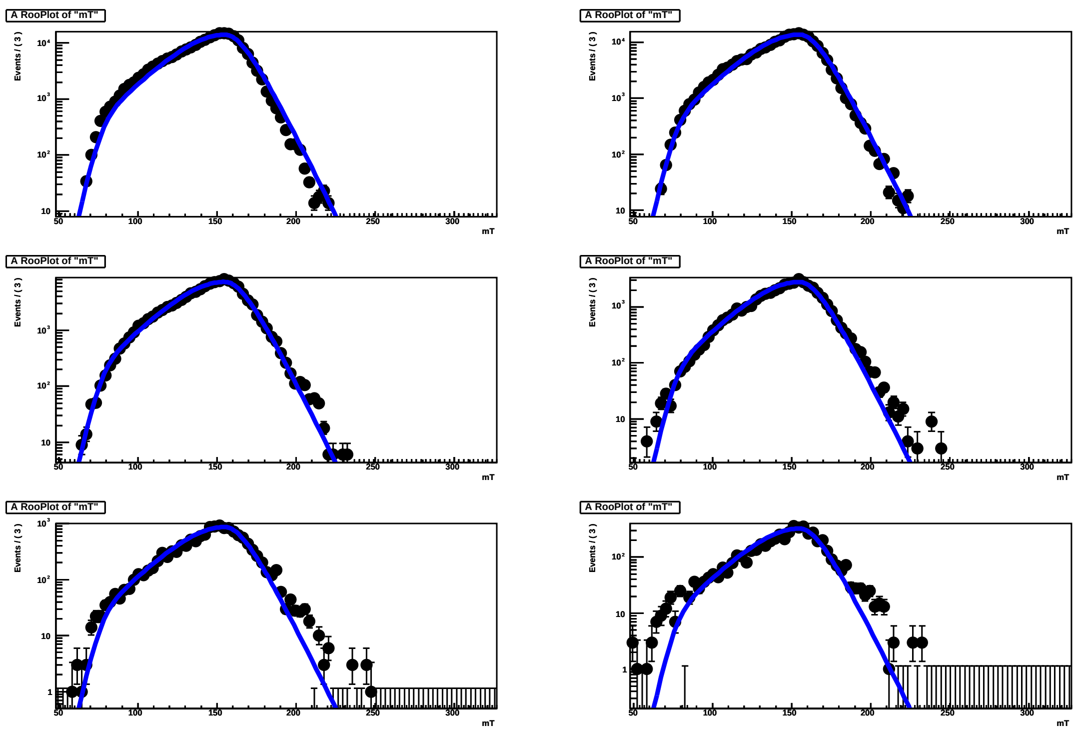

Figure A.15: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 160 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

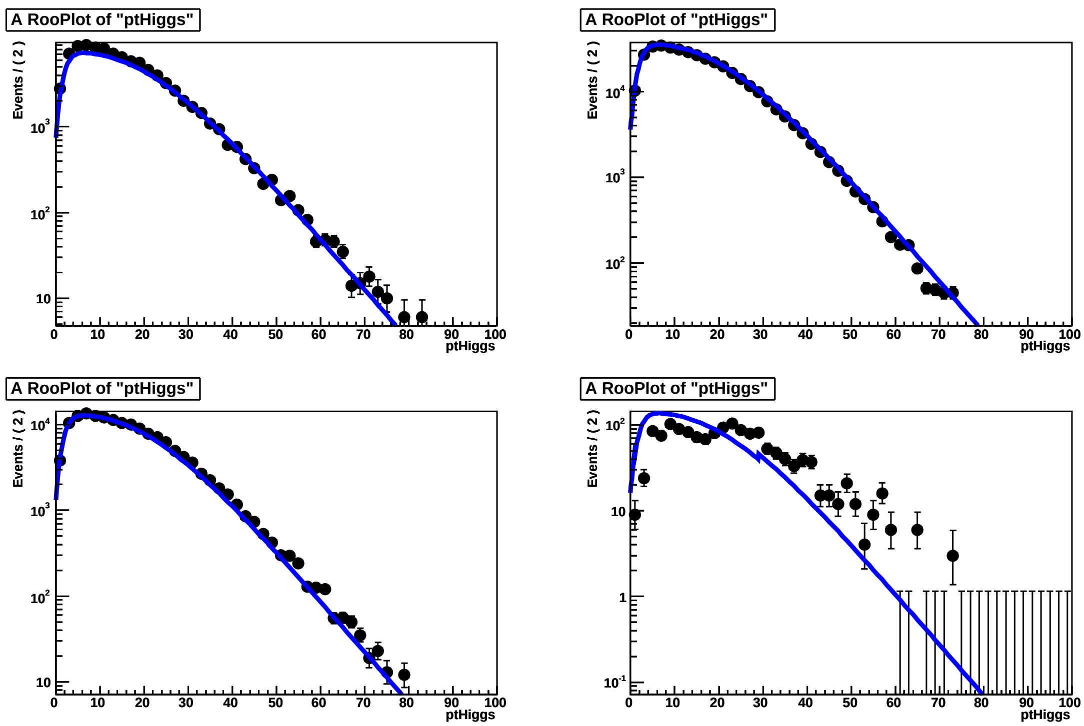

Figure A.16: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 160 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

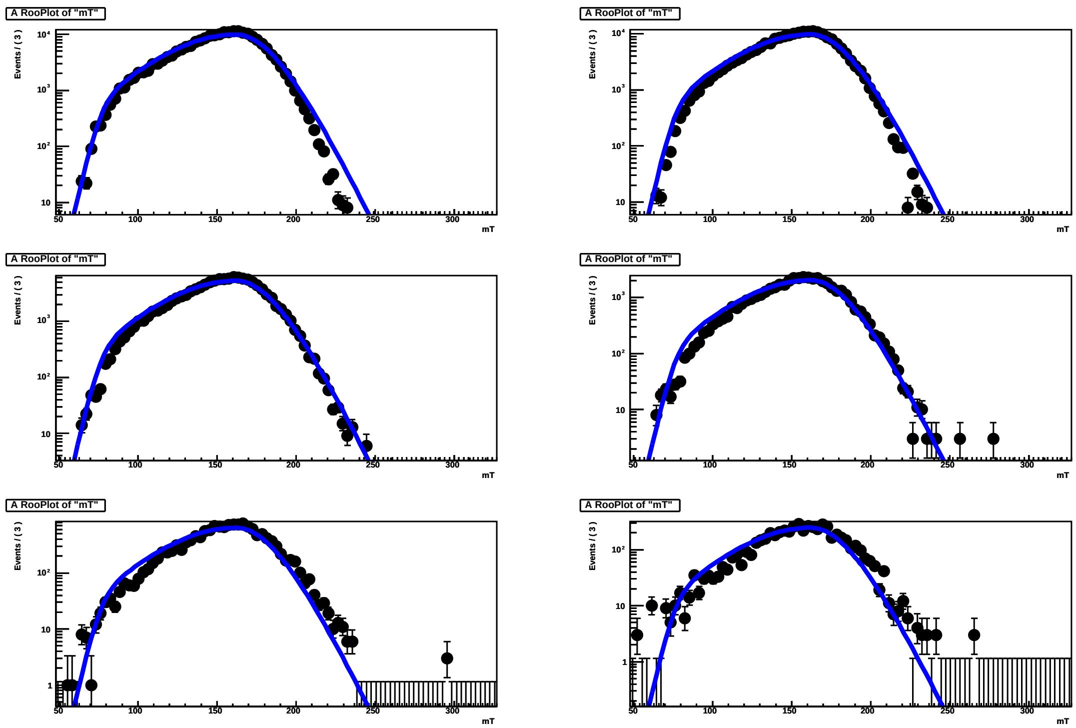

Figure A.17: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 170 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

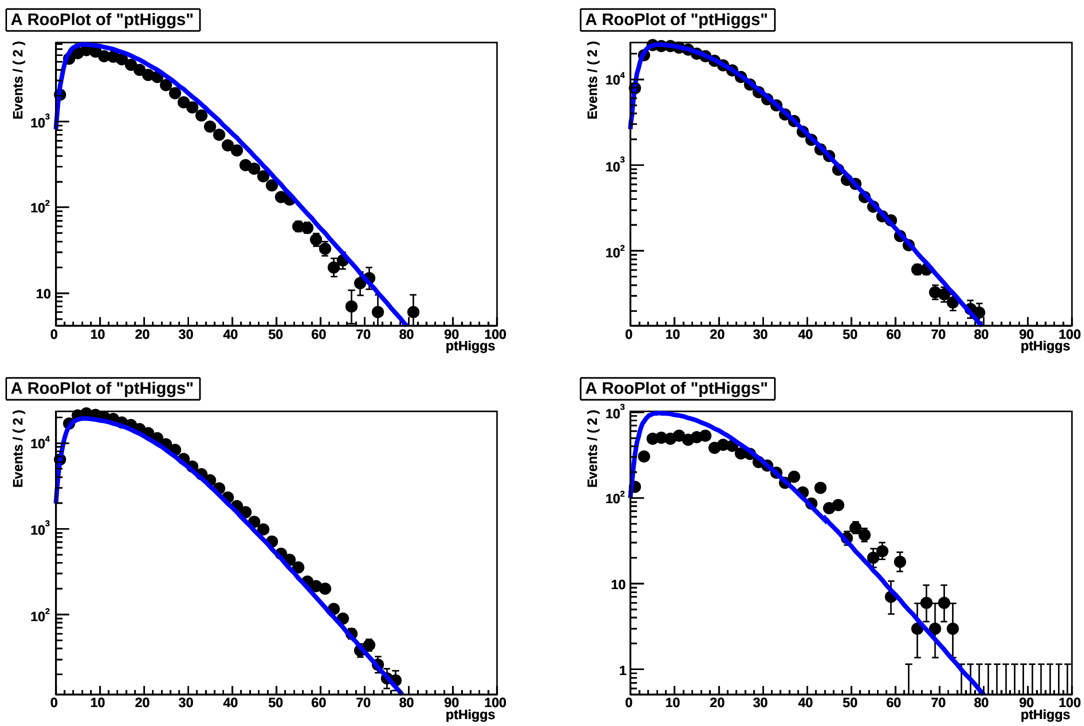

Figure A.18: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 170 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

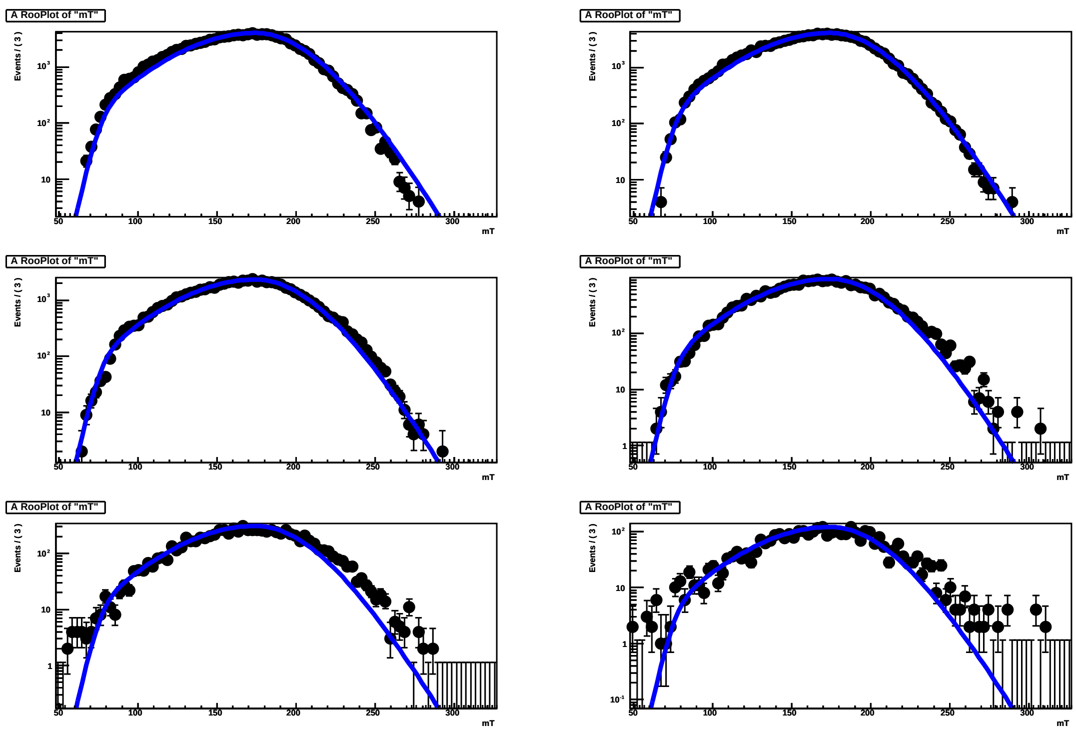

Figure A.19: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 180 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

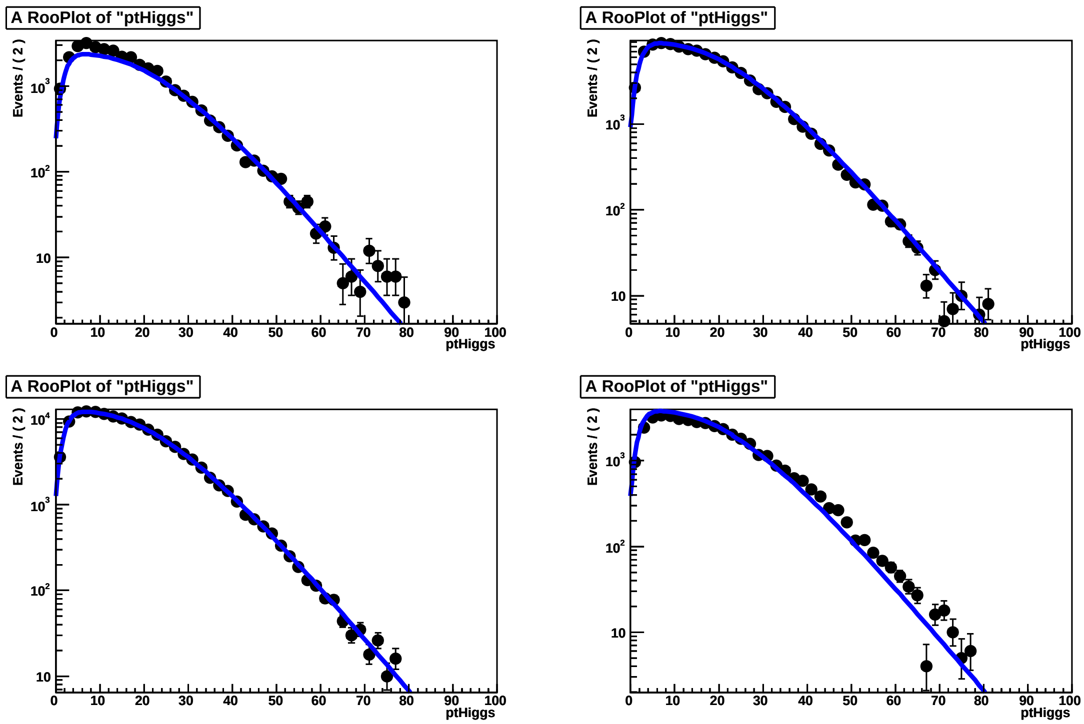

Figure A.20: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 180 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

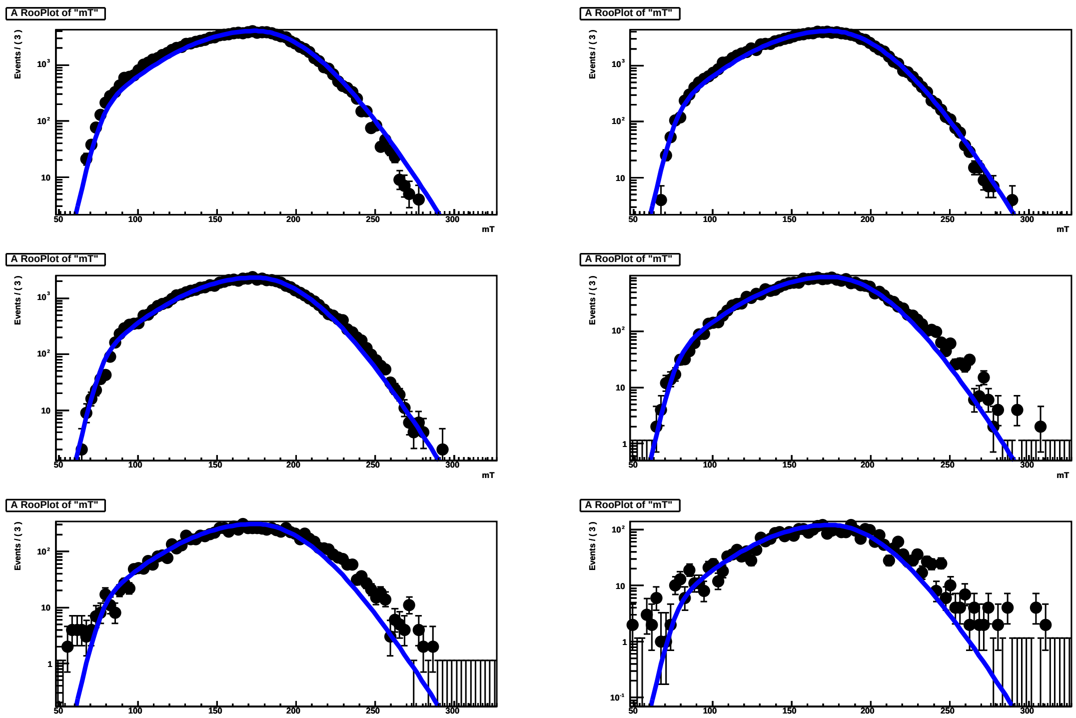

Figure A.21: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 190 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

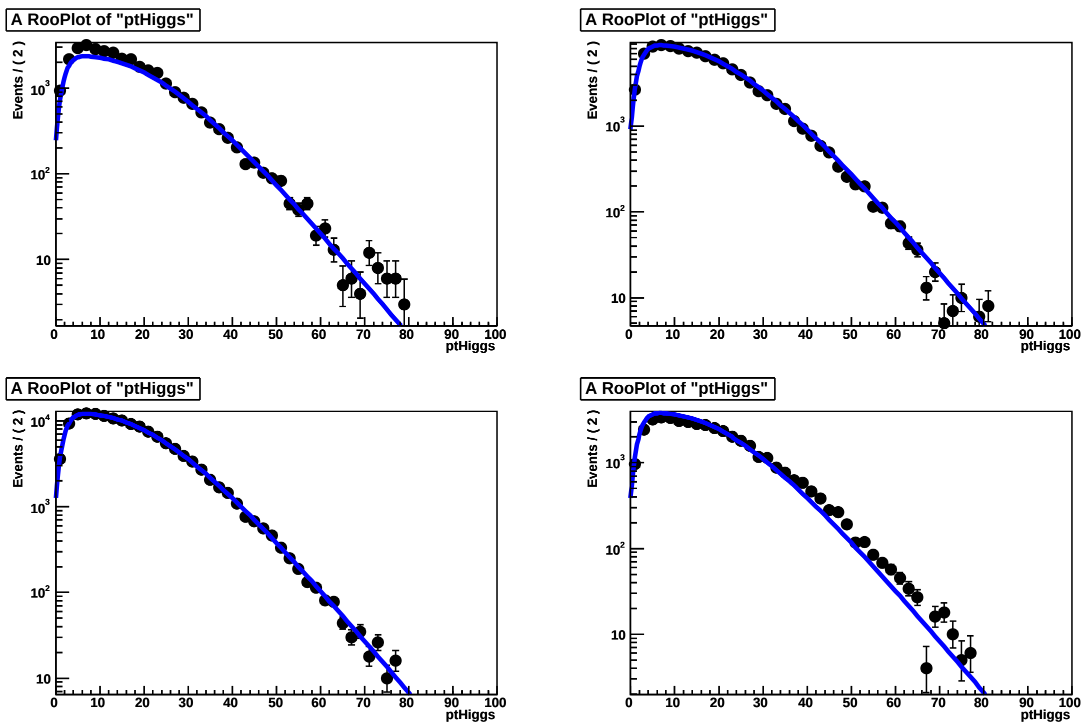

Figure A.22: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 190 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

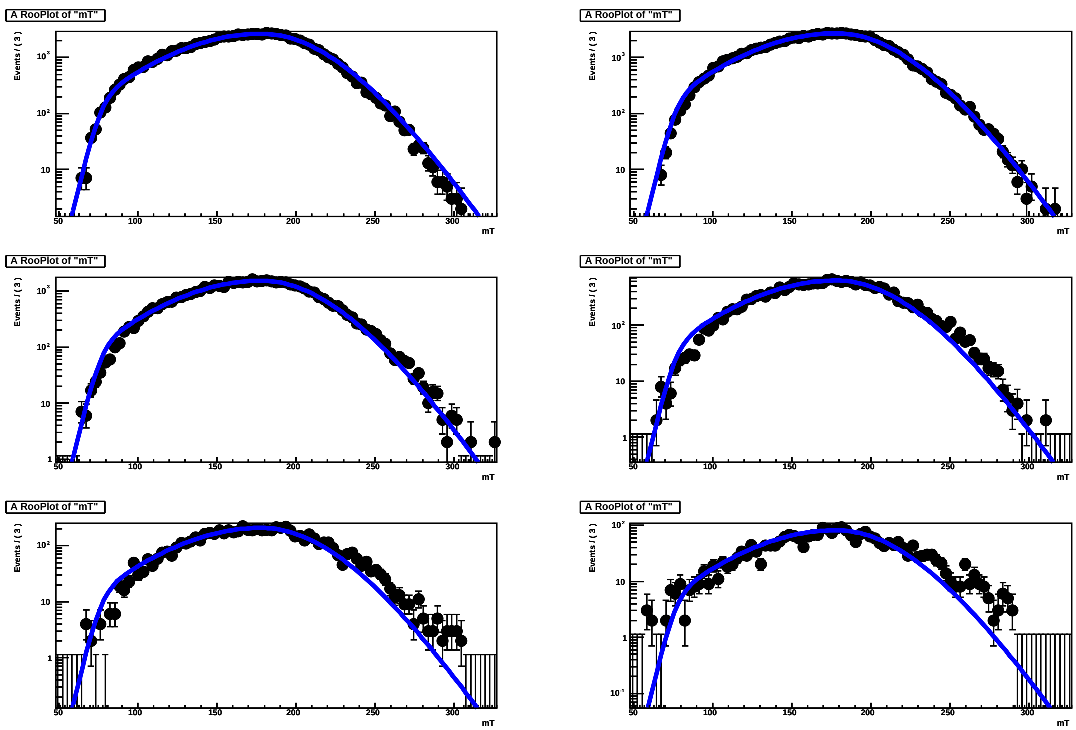

Figure A.23: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 200 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

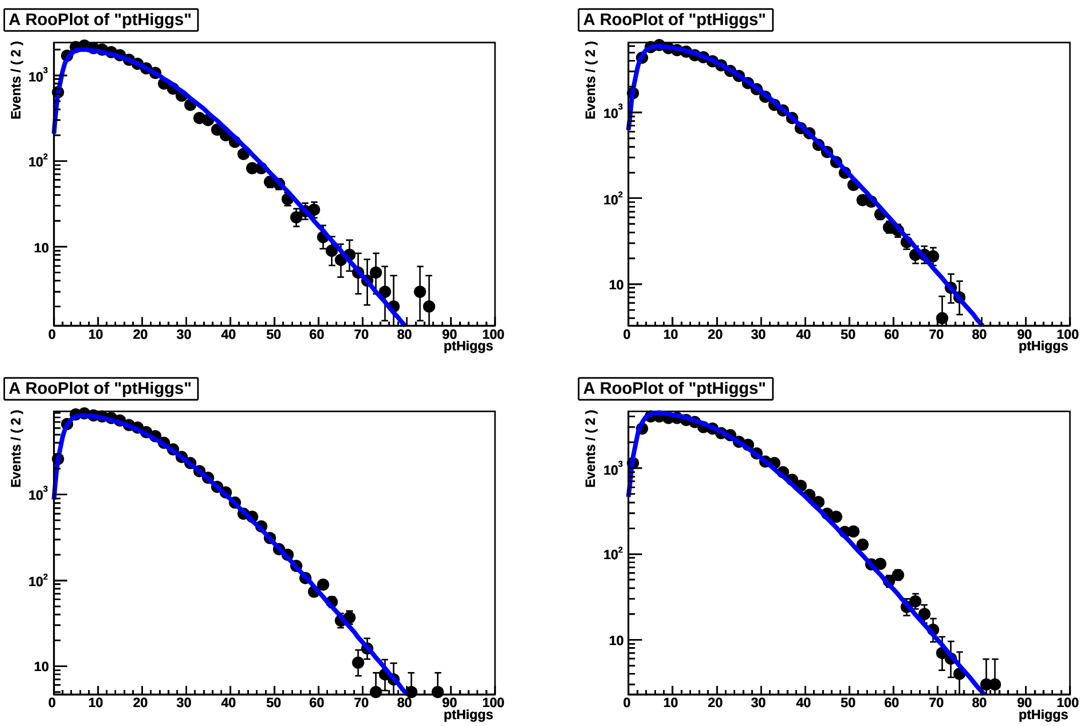

Figure A.24: A comparison between the parameterized signal model and the real Monte Carlo in the region with small $\Delta\phi_{ll}$, for a true Higgs mass of 200 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

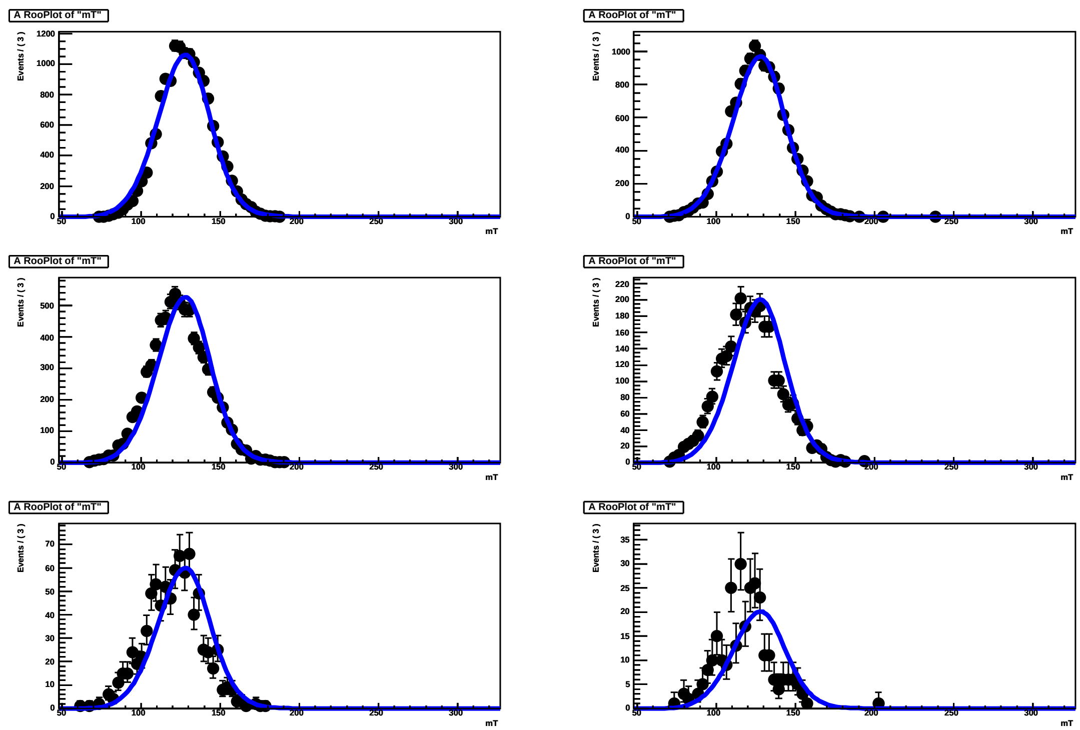

Figure A.25: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 120 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

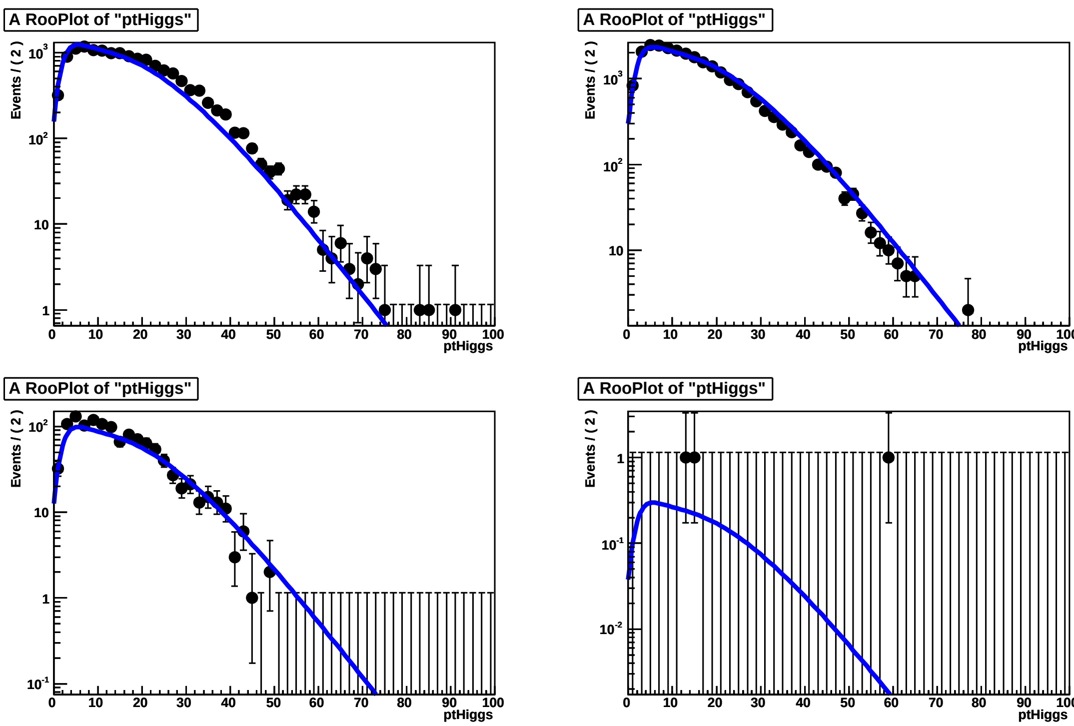

Figure A.26: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 120 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

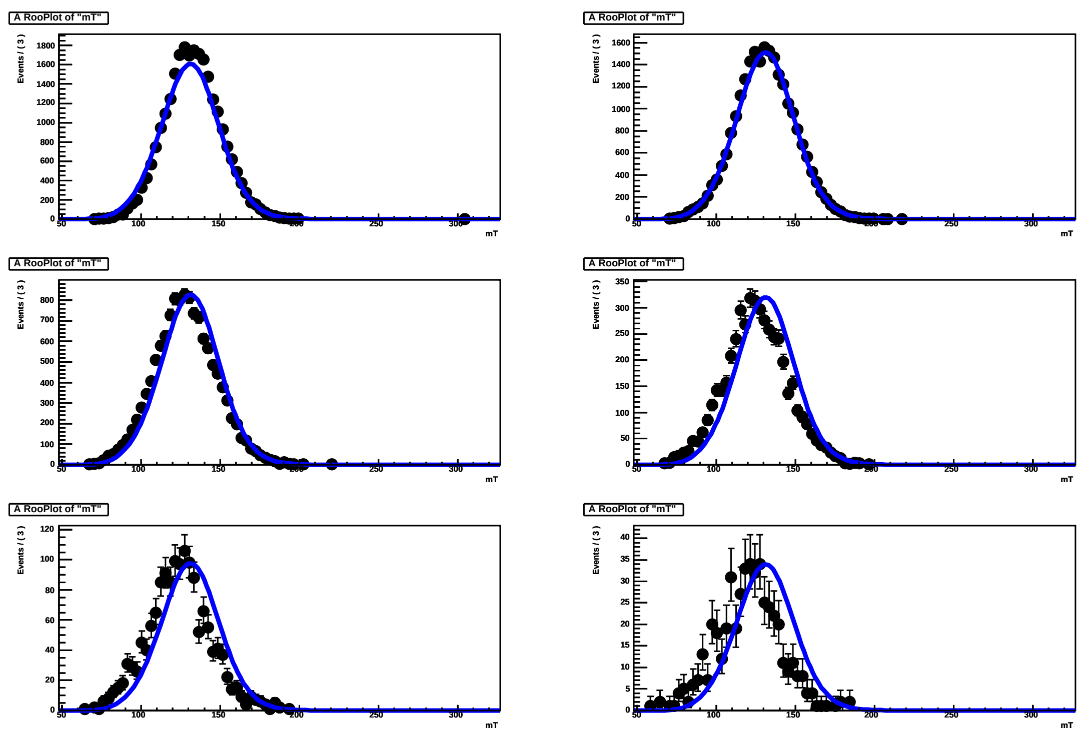

Figure A.27: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 125 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

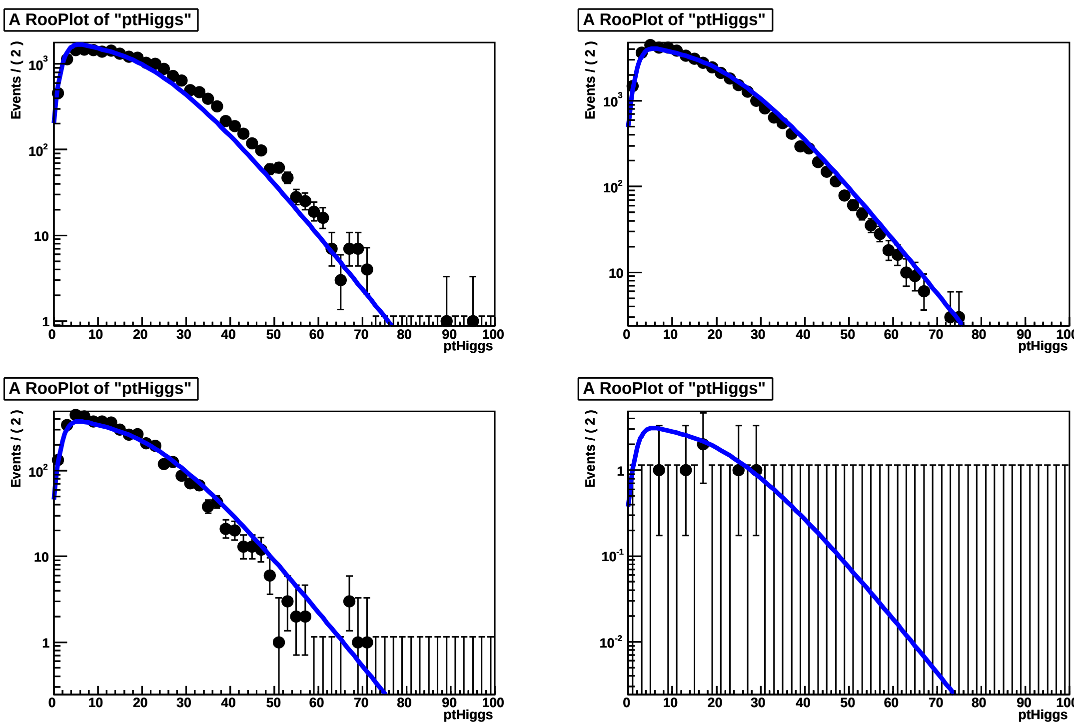

Figure A.28: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 125 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

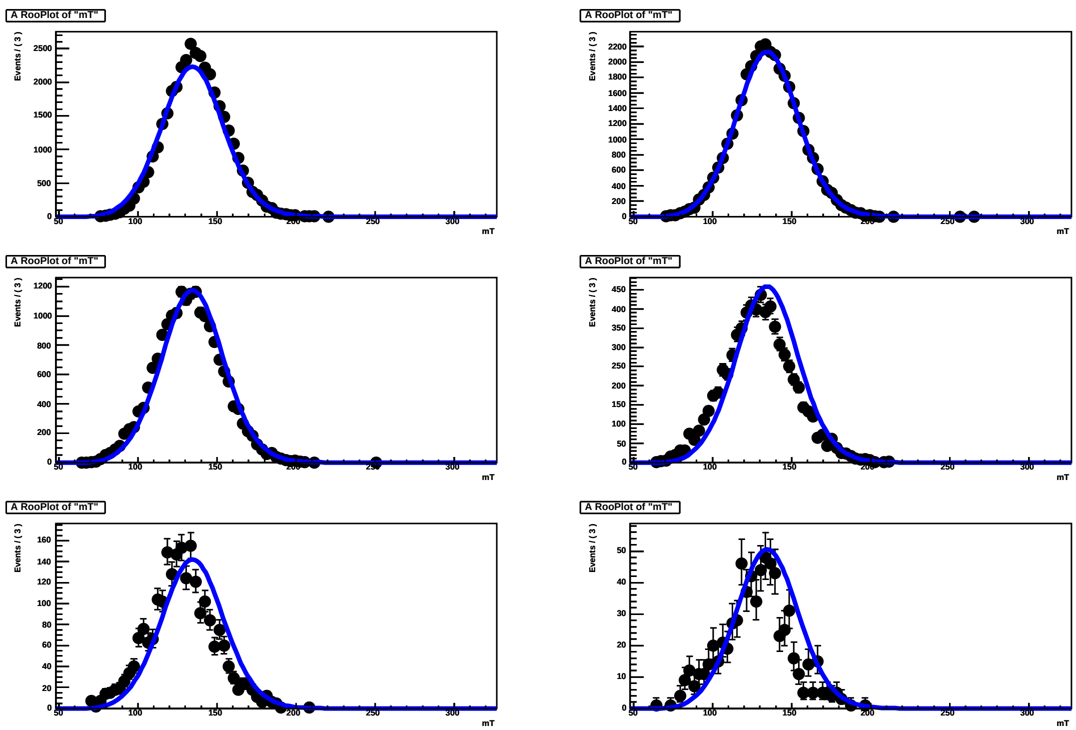

Figure A.29: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 130 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

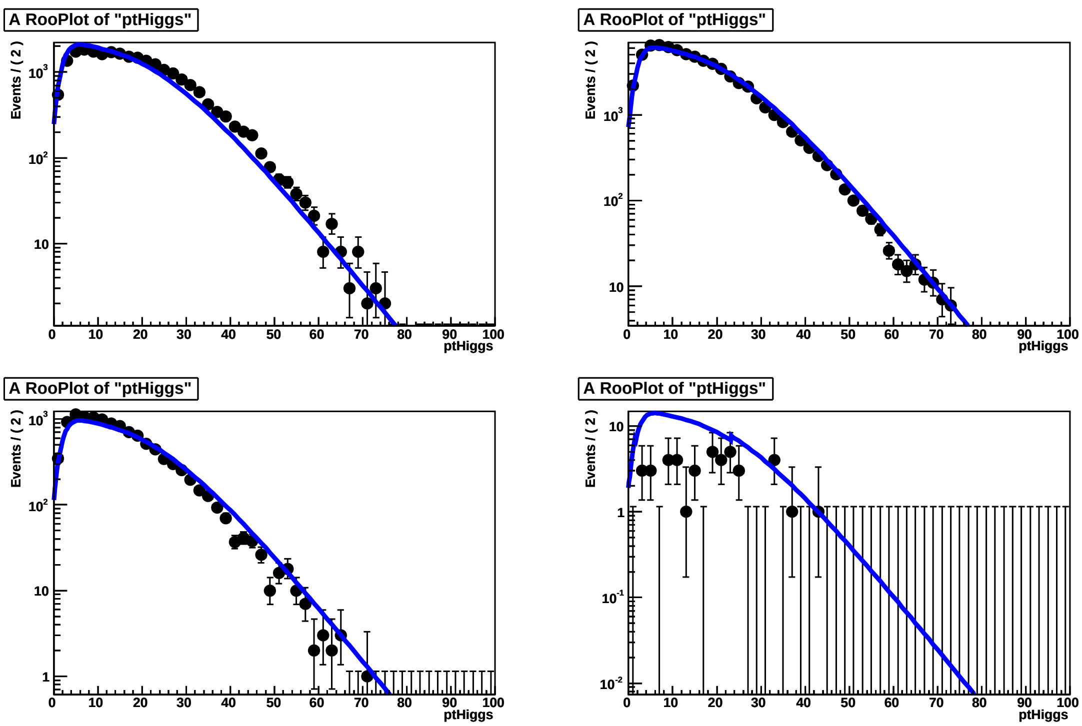

Figure A.30: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 130 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

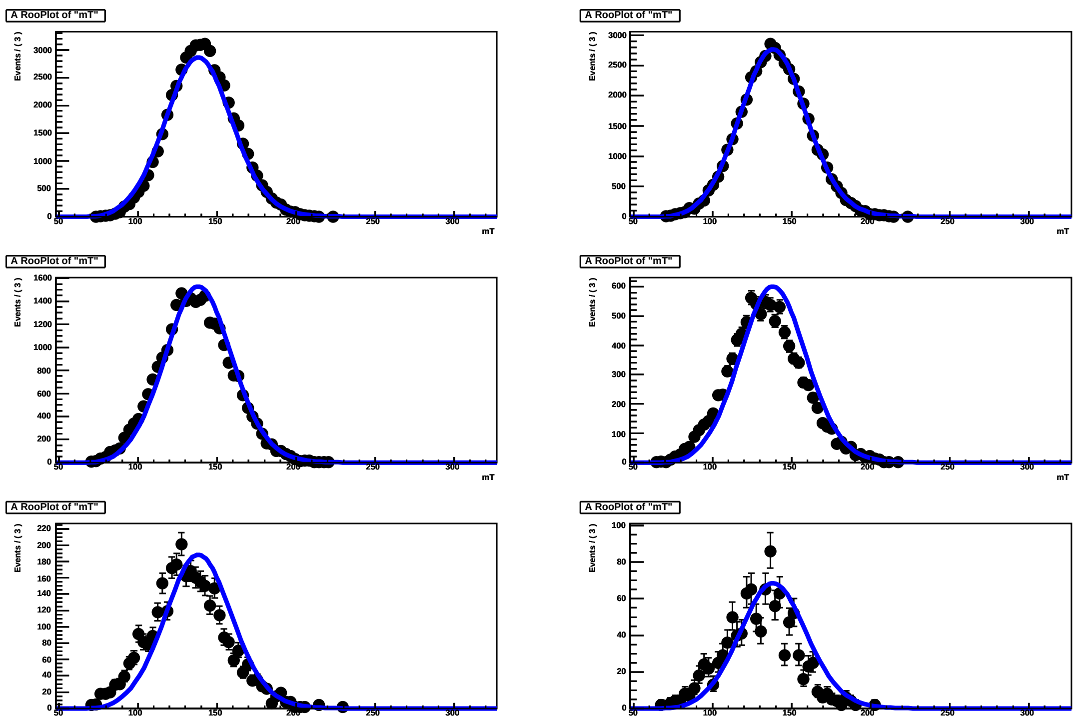

Figure A.31: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 135 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

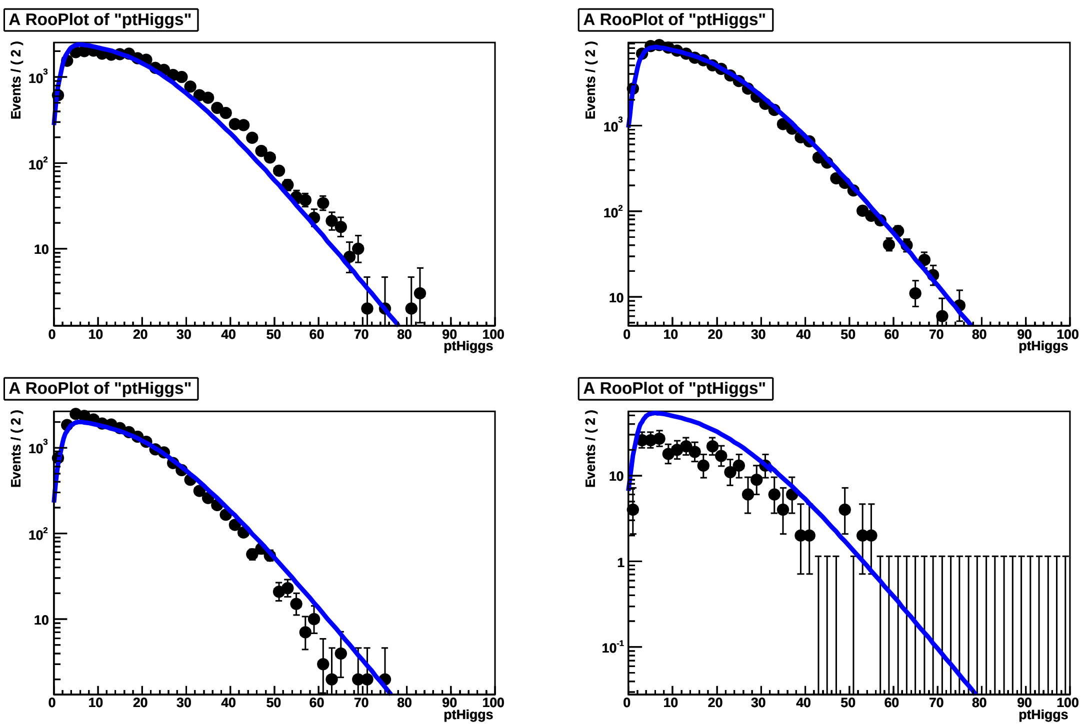

Figure A.32: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 135 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

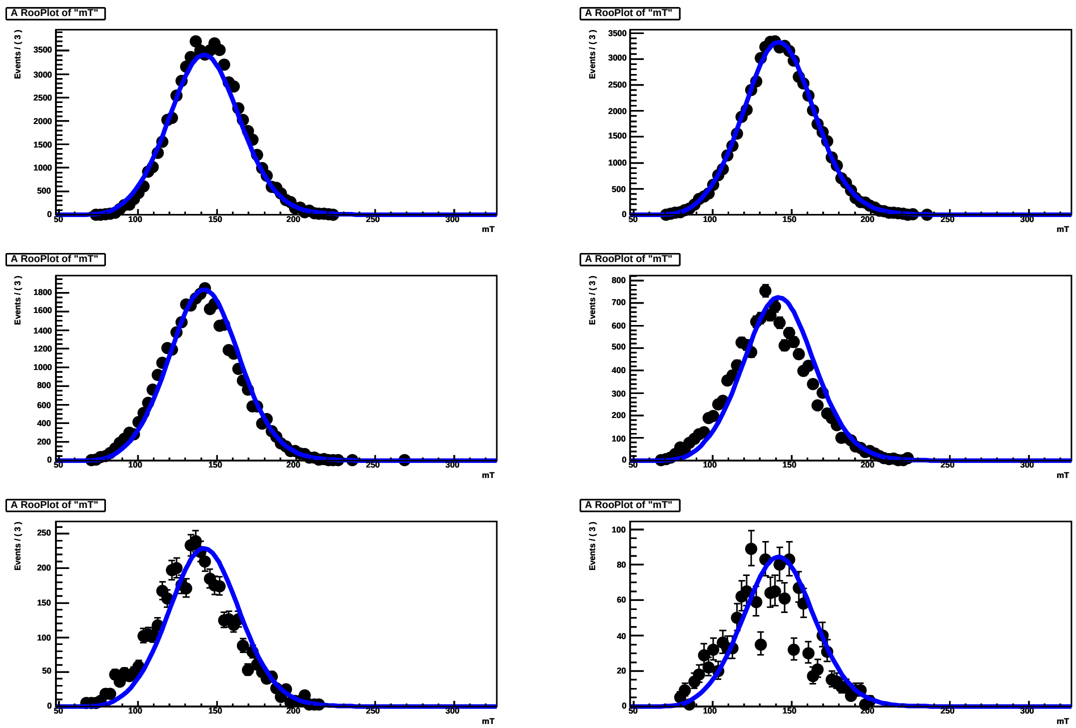

Figure A.33: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 140 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

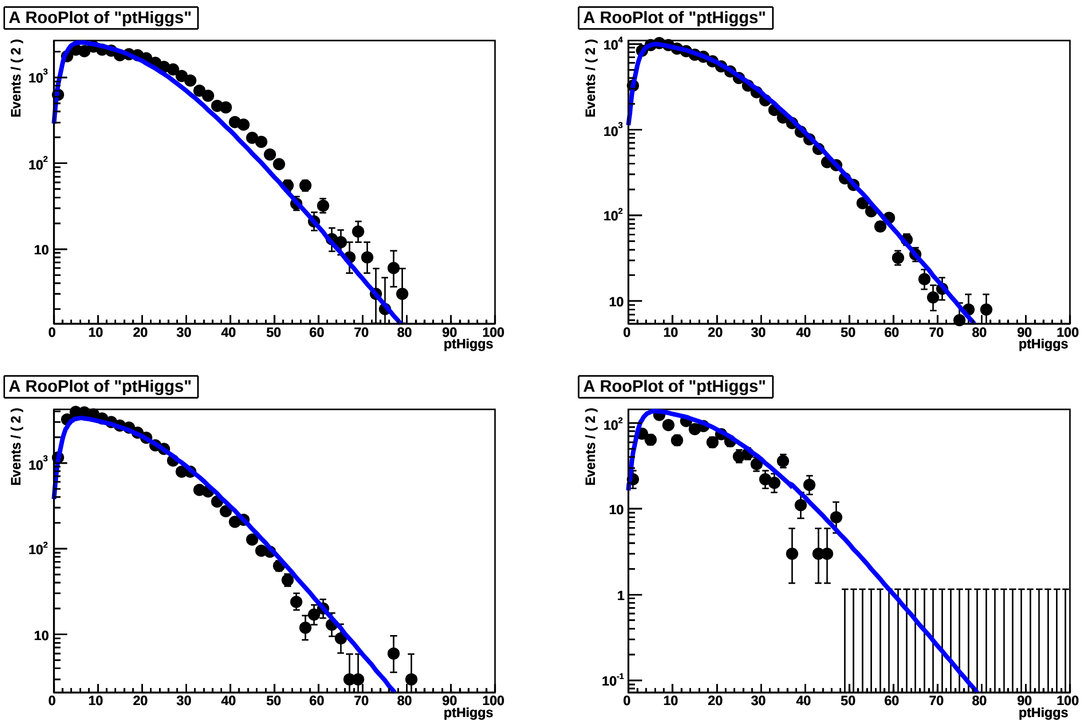

Figure A.34: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 140 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

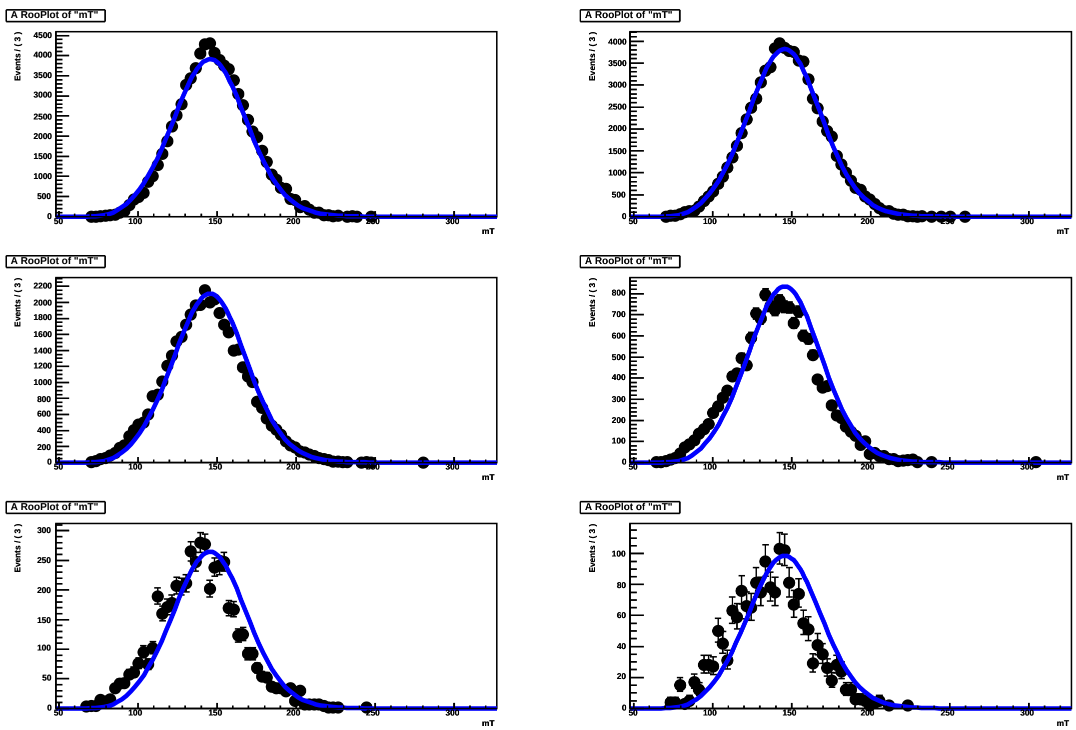

Figure A.35: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 145 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

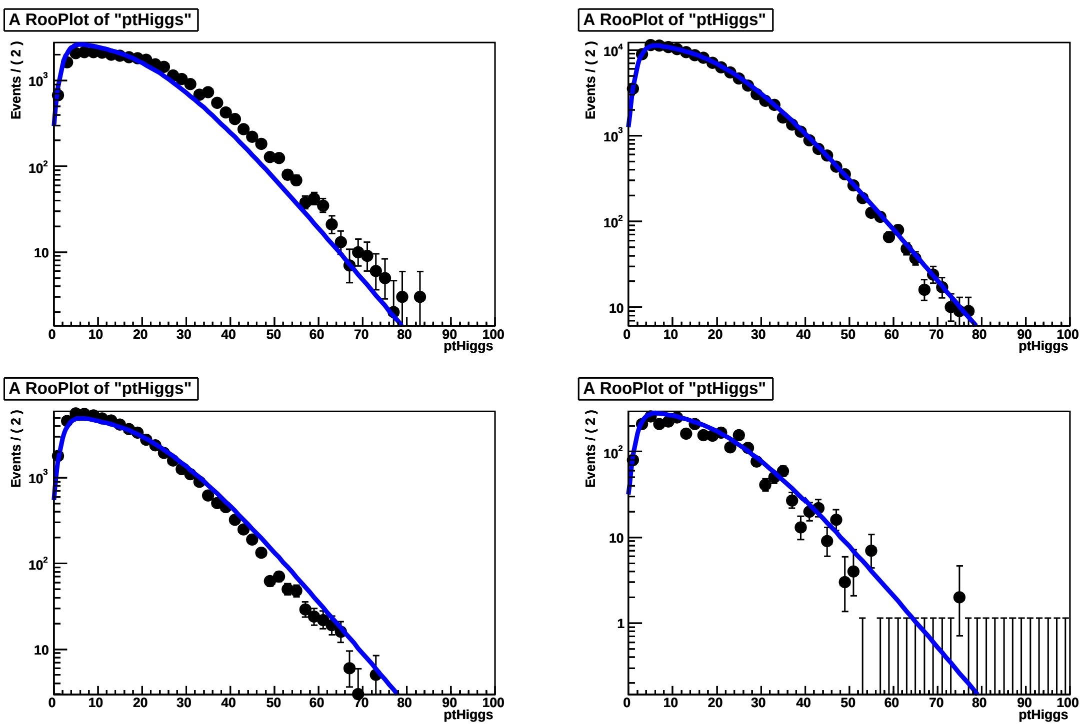

Figure A.36: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 145 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

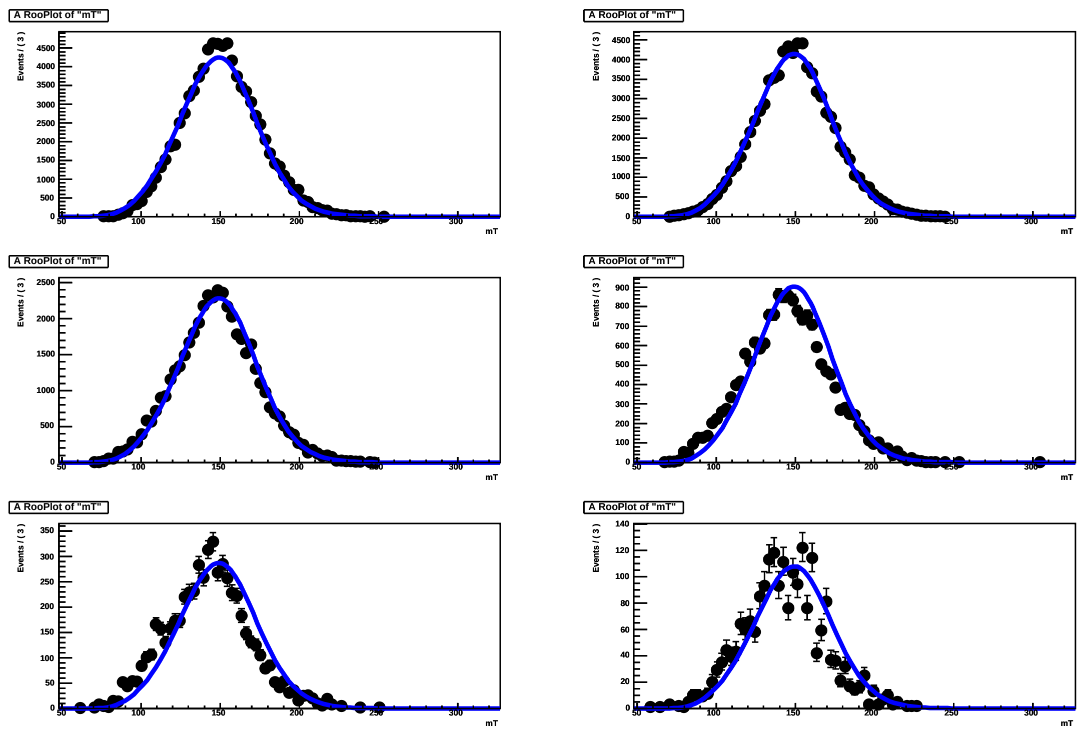

Figure A.37: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 150 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

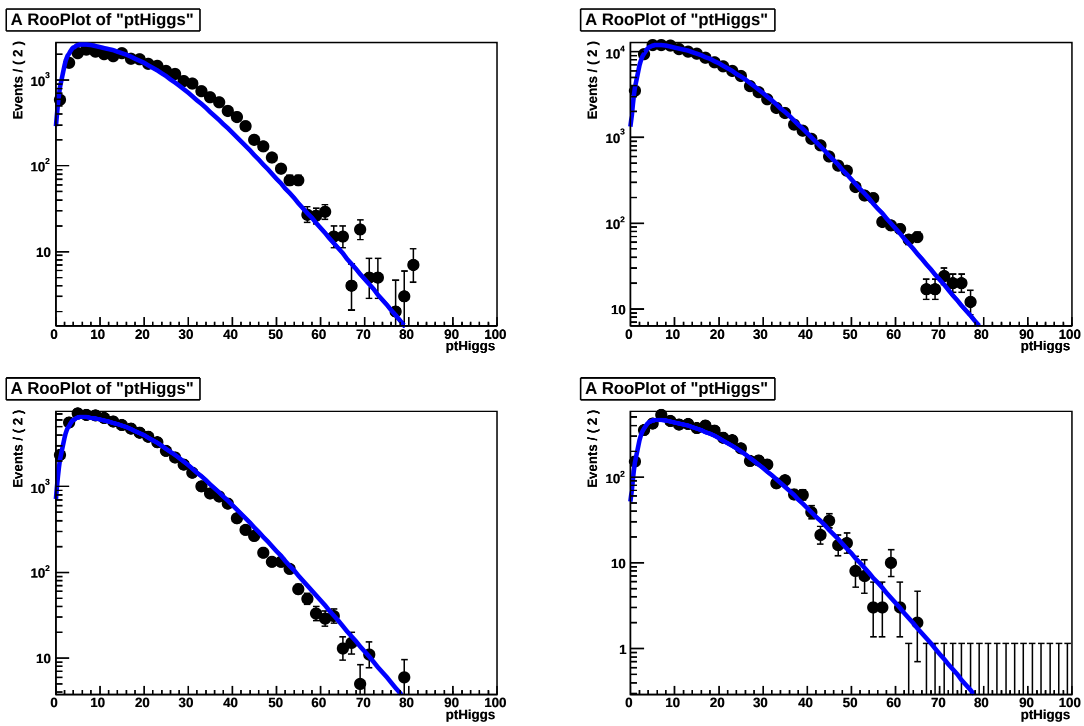

Figure A.38: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 150 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

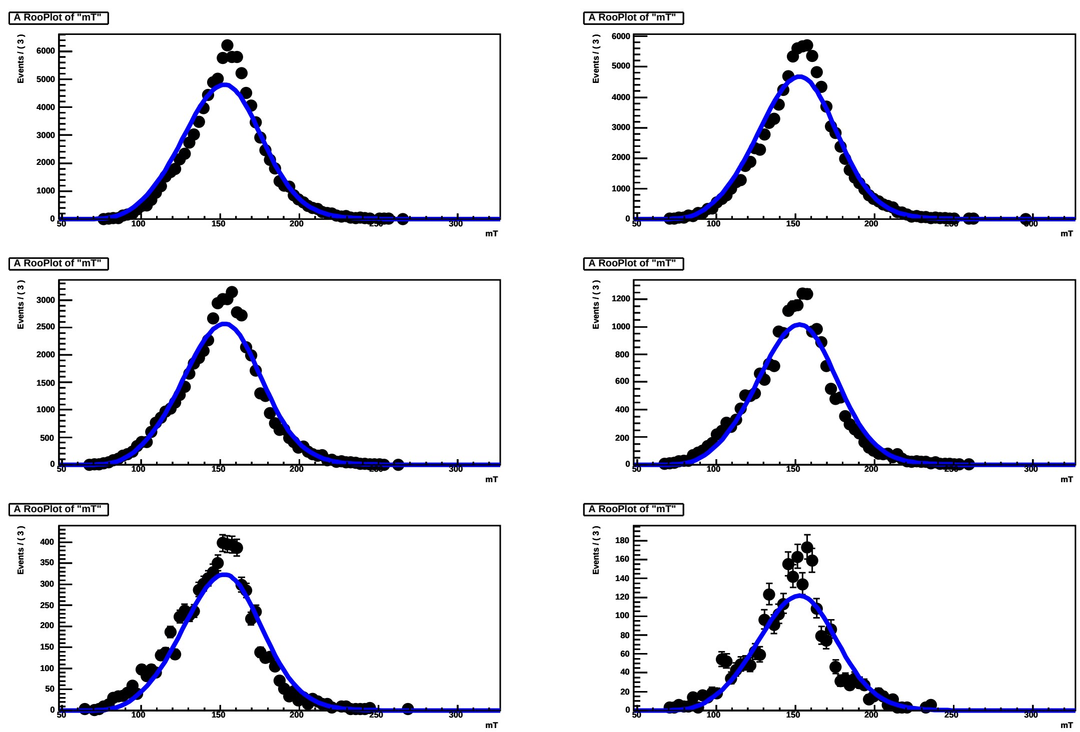

Figure A.39: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 160 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

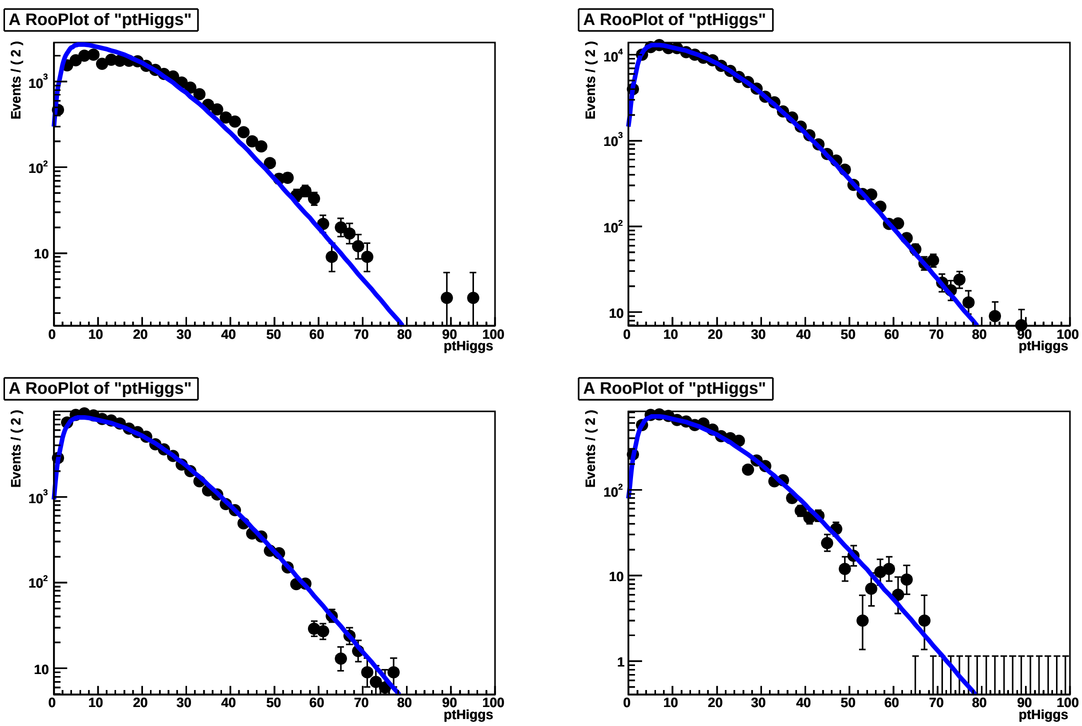

Figure A.40: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 160 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

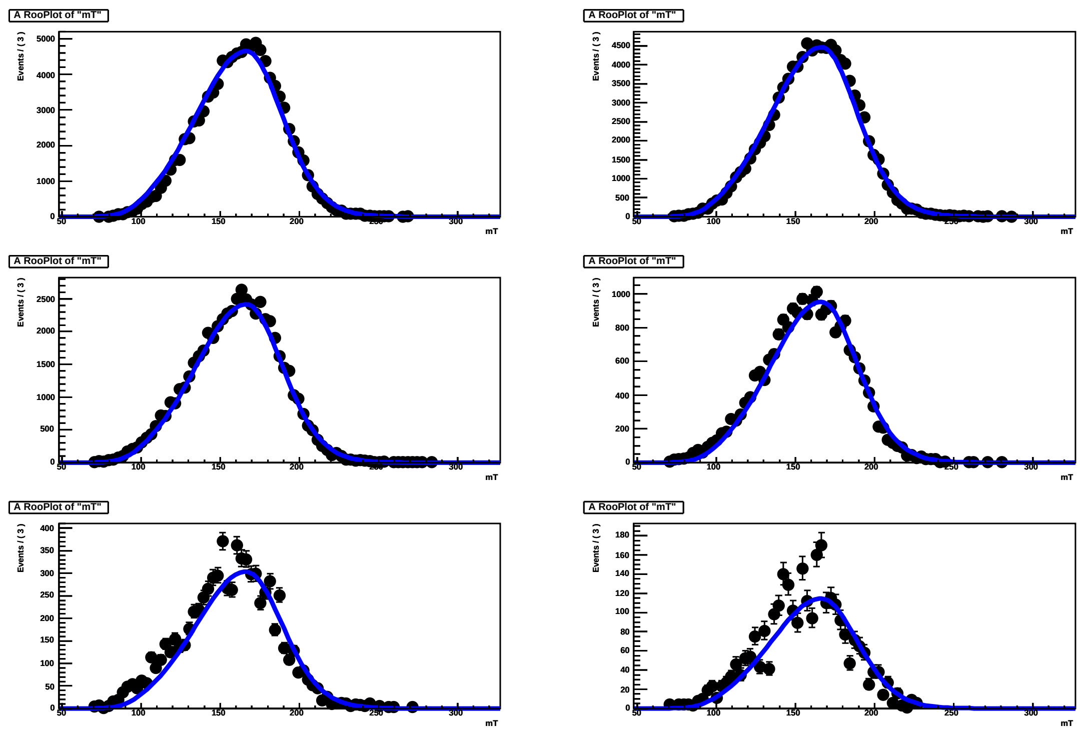

Figure A.41: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 170 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

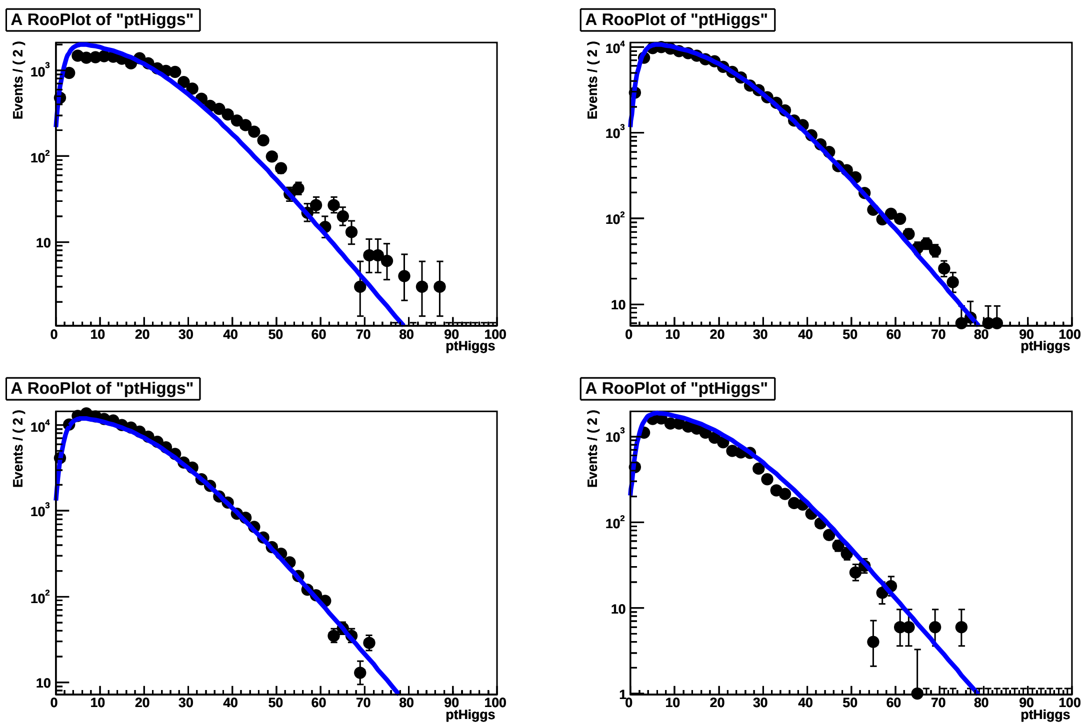

Figure A.42: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 170 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

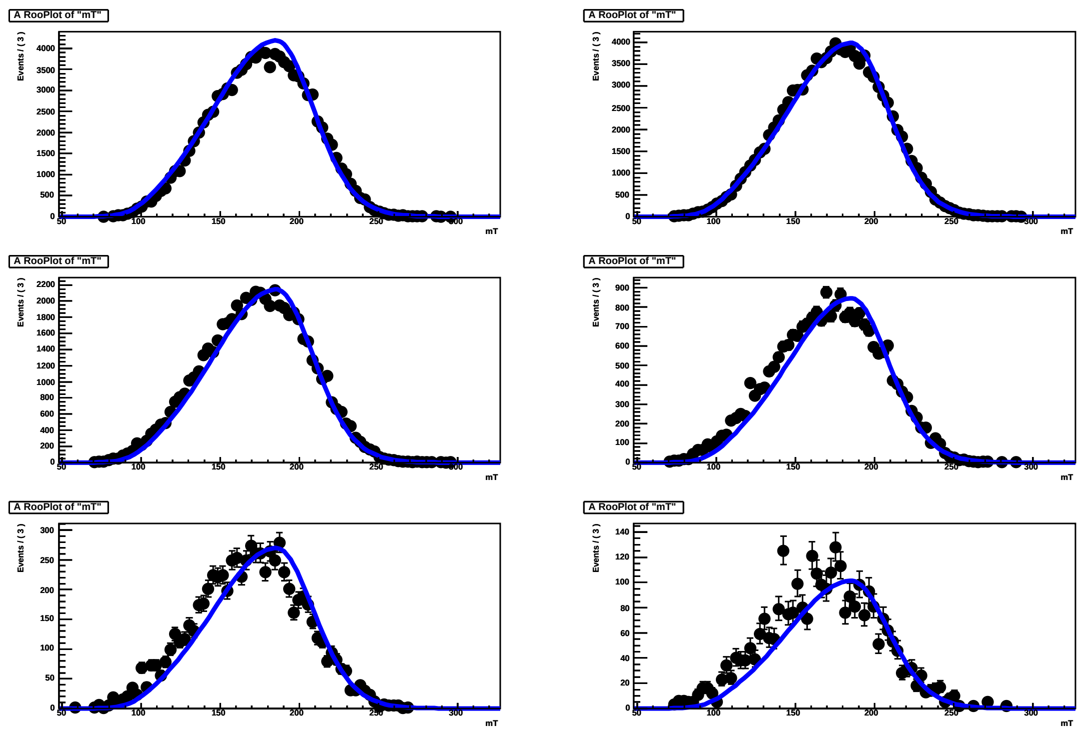

Figure A.43: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 180 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

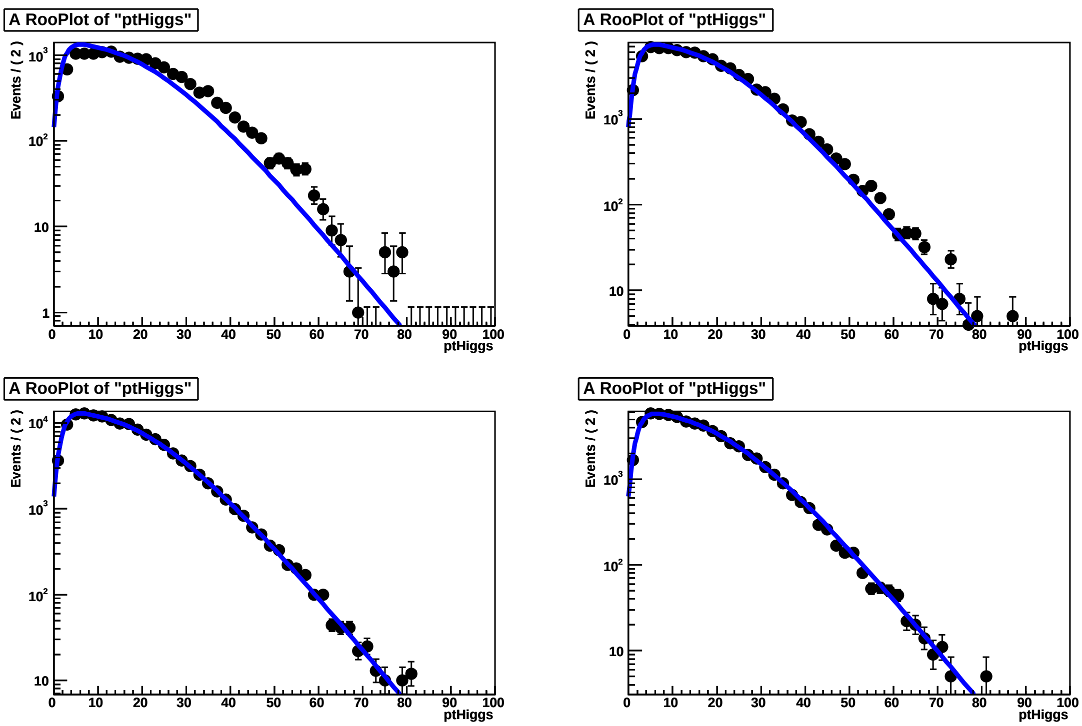

Figure A.44: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 180 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

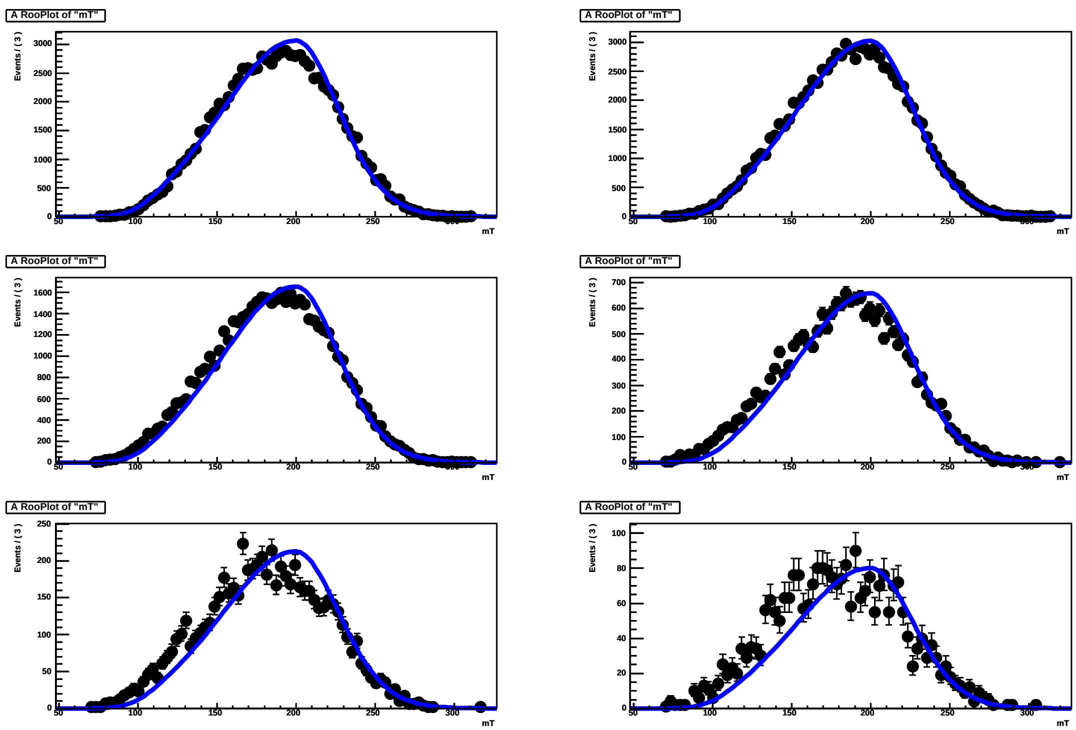

Figure A.45: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 190 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

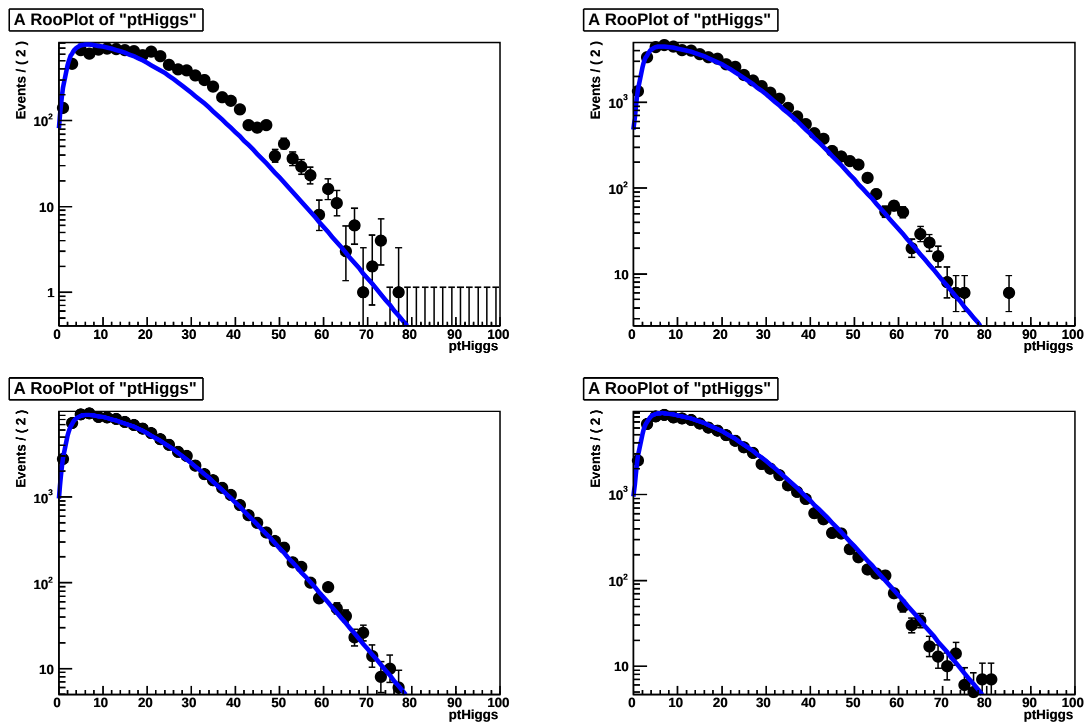

Figure A.46: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 190 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

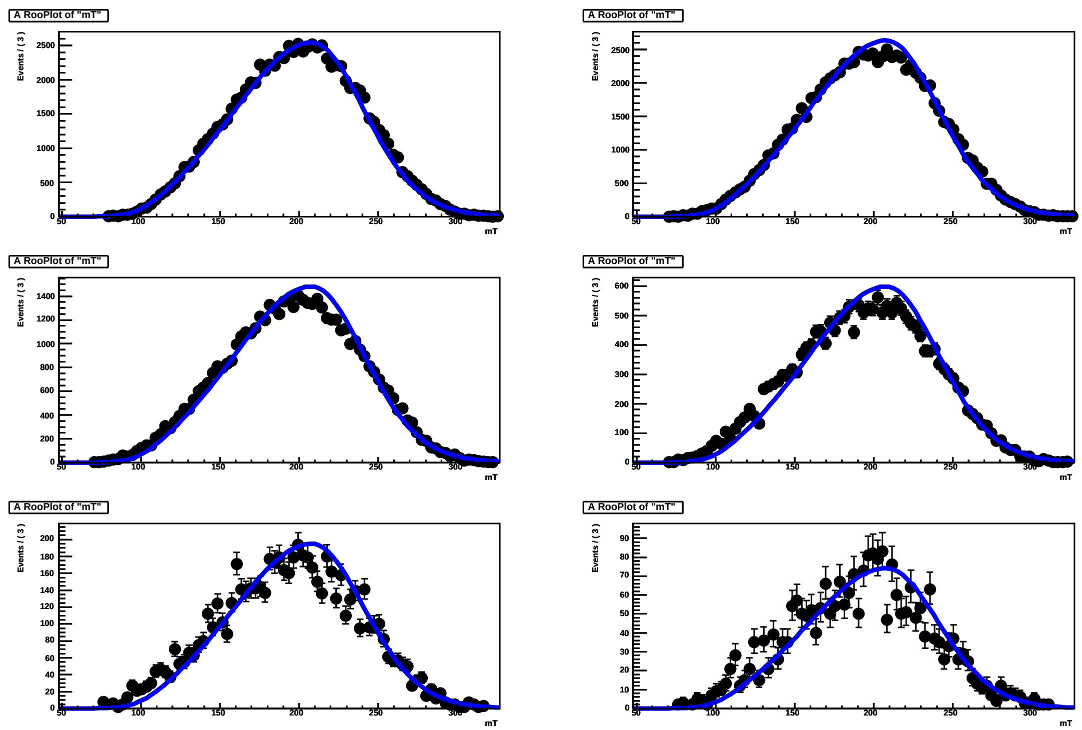

Figure A.47: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 200 GeV. Here, the transverse mass is shown in bins of Higgs P_T. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

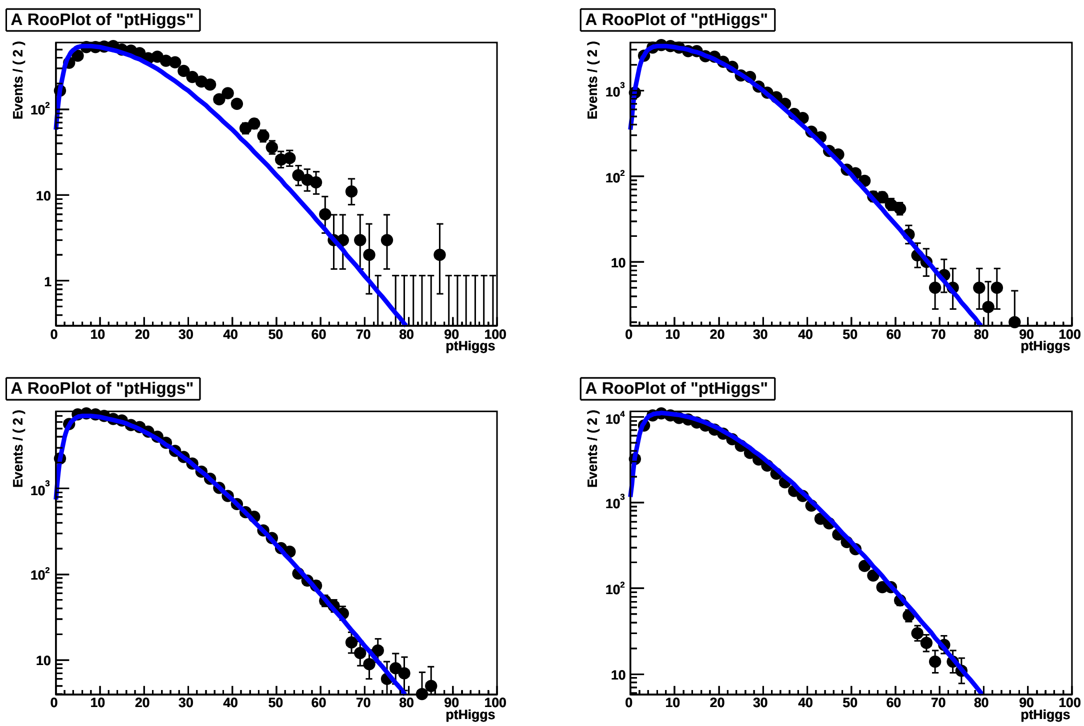

Figure A.48: A comparison between the parameterized signal model and the real Monte Carlo in the region with large $\Delta\phi_{ll}$, for a true Higgs mass of 200 GeV. Here, the Higgs P_T is shown in bins of transverse mass. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

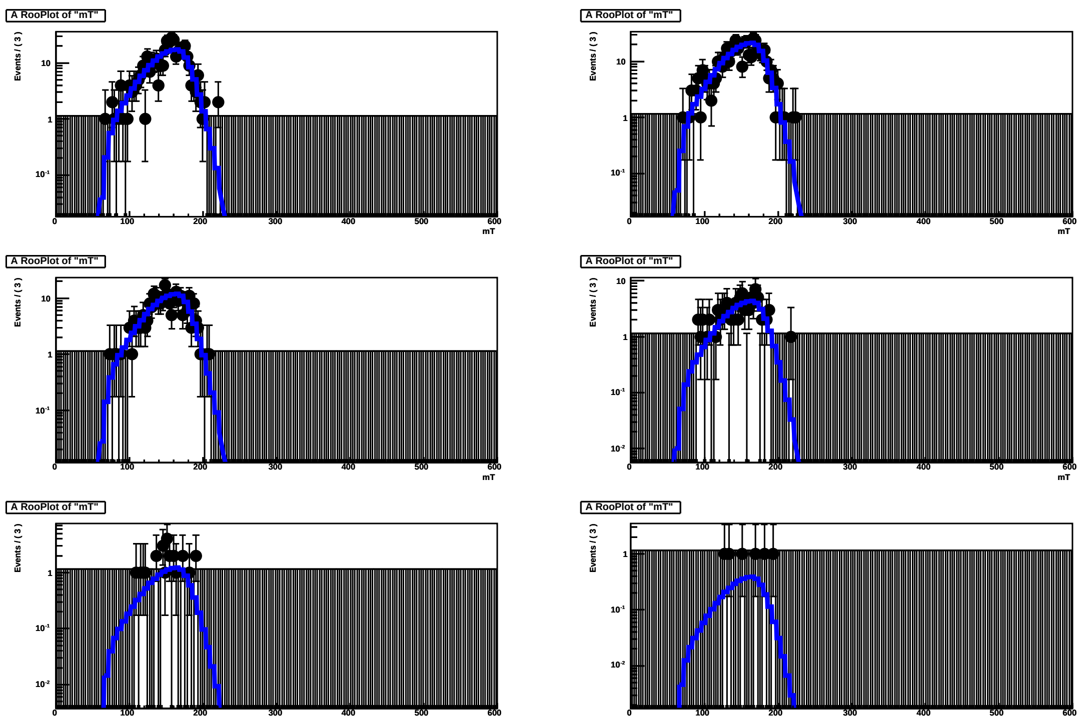

Figure A.49: The best fit of the functional form for signal to the fully simulated signal Monte Carlo, in the region with small $\Delta\phi_{ll}$. Here, M_T is shown in bins of P_T^{Higgs}. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

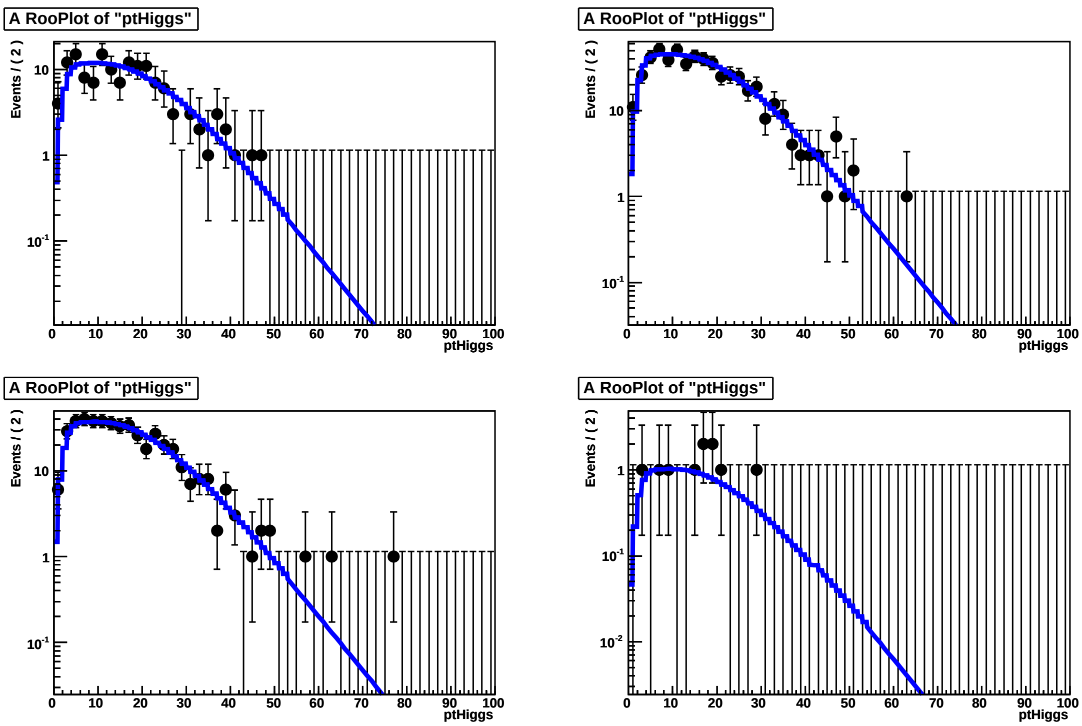

Figure A.50: The best fit of the functional form for signal to the fully simulated signal Monte Carlo, in the region with small $\Delta\phi_{ll}$. Here, P_T^{Higgs} is shown in bins of M_T. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

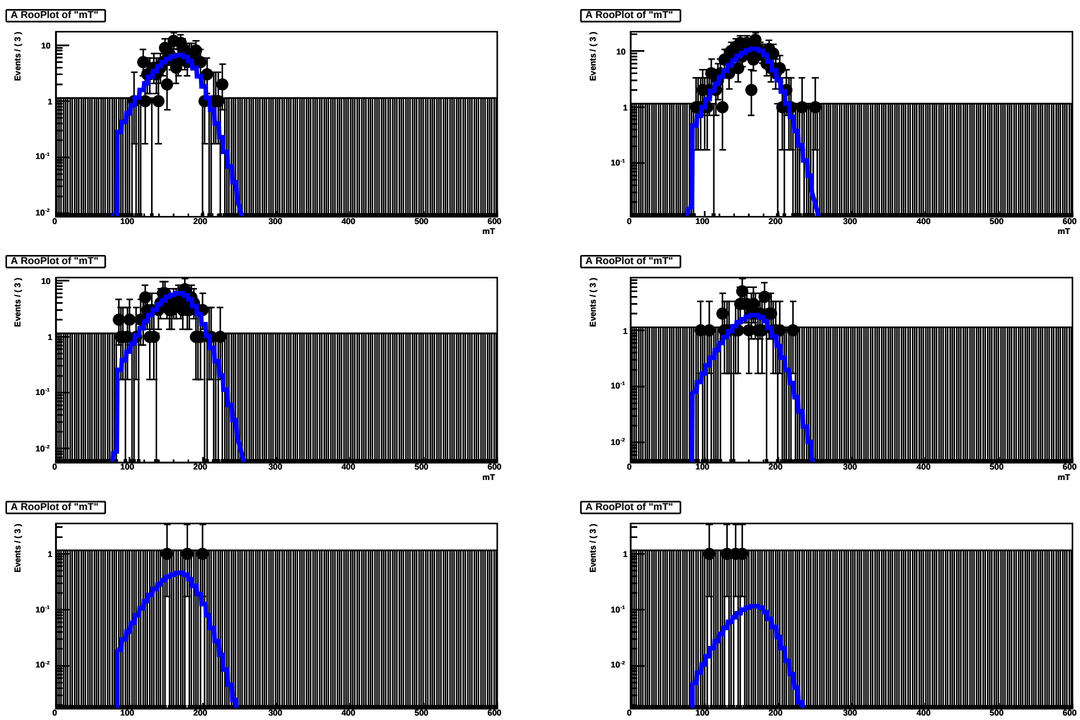

Figure A.51: The best fit of the functional form for signal to the fully simulated signal Monte Carlo, in the region with large $\Delta\phi_{ll}$. Here, M_T is shown in bins of P_T^{Higgs}. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

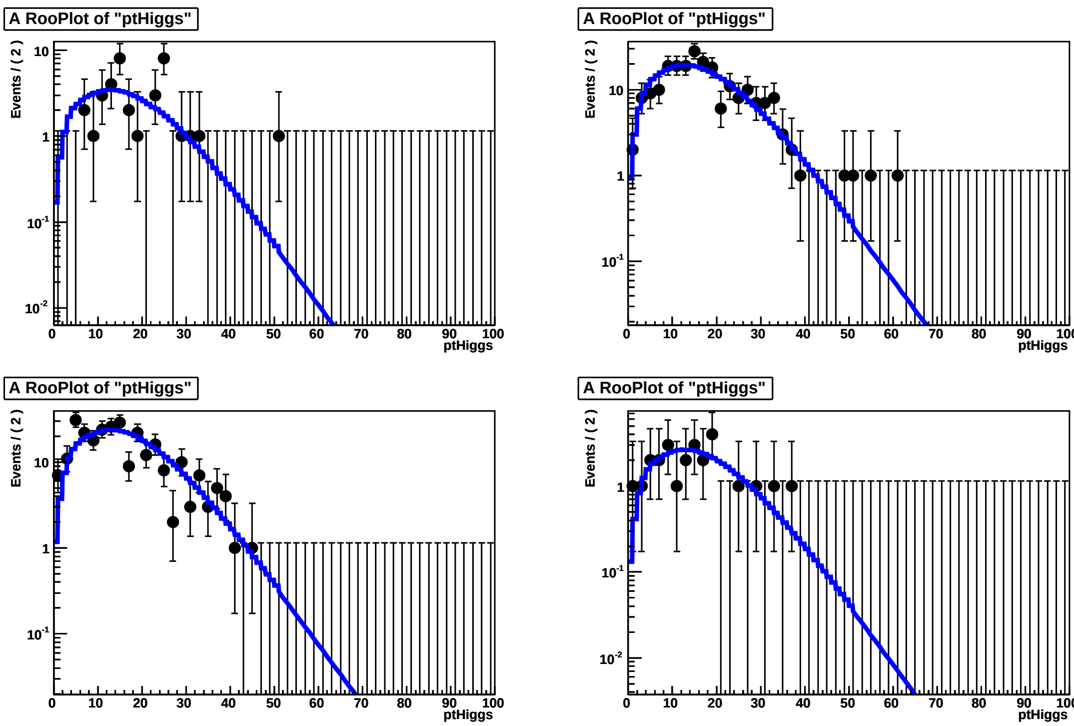

Figure A.52: The best fit of the functional form for signal to the fully simulated signal Monte Carlo, in the region with large $\Delta\phi_{ll}$. Here, P_T^{Higgs} is shown in bins of M_T. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

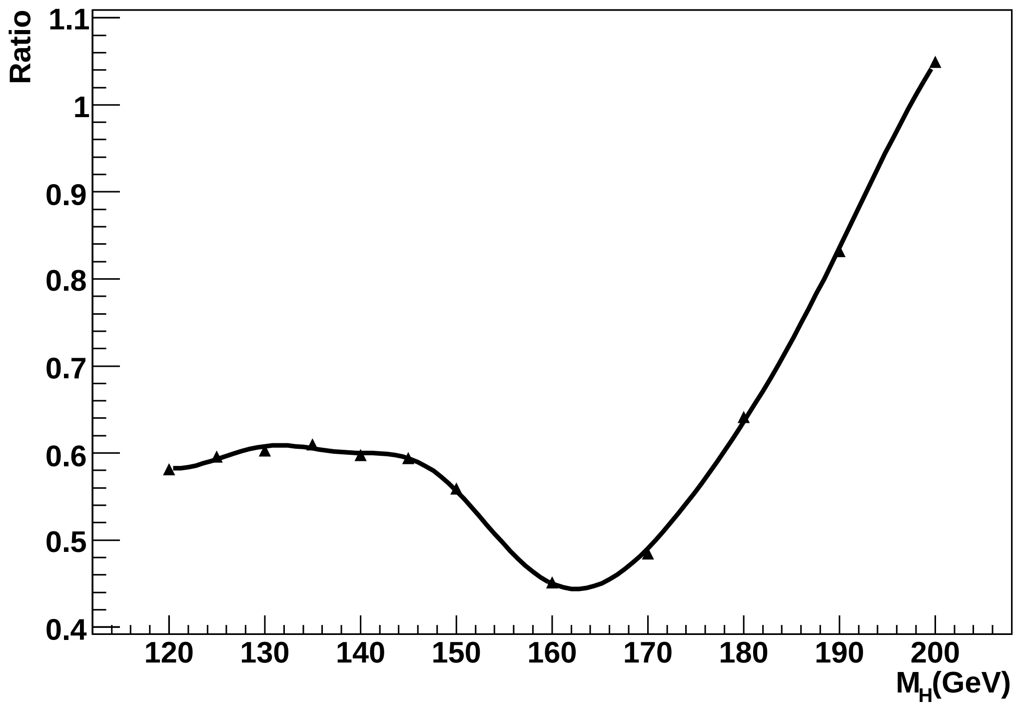

Figure A.53: The ratio of signal cross-sections in the small-$\Delta\phi_{ll}$ and large-$\Delta\phi_{ll}$ regions. The points show the results derived from fast simulation, and the curve shows the parameterization used in the fit.

A.2 QCD WW Model

The functional form used for the QCD WW background is also a product of two distributions, but it does not assume the two observables are uncorrelated. The Higgs P_T spectrum is modelled with the same function as for signal; the transverse mass is modelled as the sum of two exponentials modulated at low M_T by a product of two Fermi-Dirac cutoffs:

$$\frac{(e^{-|m_1|M_T} + |m_5|e^{-(|m_1|+|m_2|)M_T})}{(1+e^{-|m_3|(M_T-m_4)})(1+e^{-(|m_6|+|m_3|)(M_T-m_4+|m_7|)})} \tag{A.4}$$

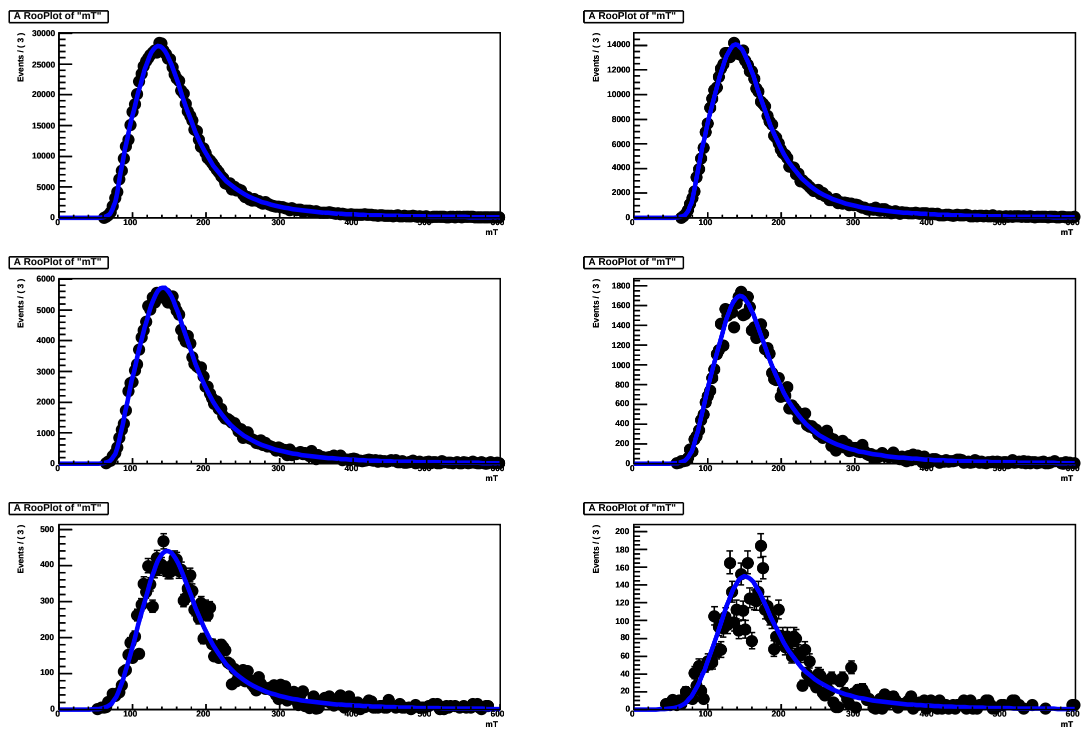

Figure A.54: The best fit of the functional form for QCD WW background to the central-value (fast simulation) Monte Carlo, in the region with small $\Delta\phi_{ll}$. Here, M_T is shown in bins of P_T^{Higgs}. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

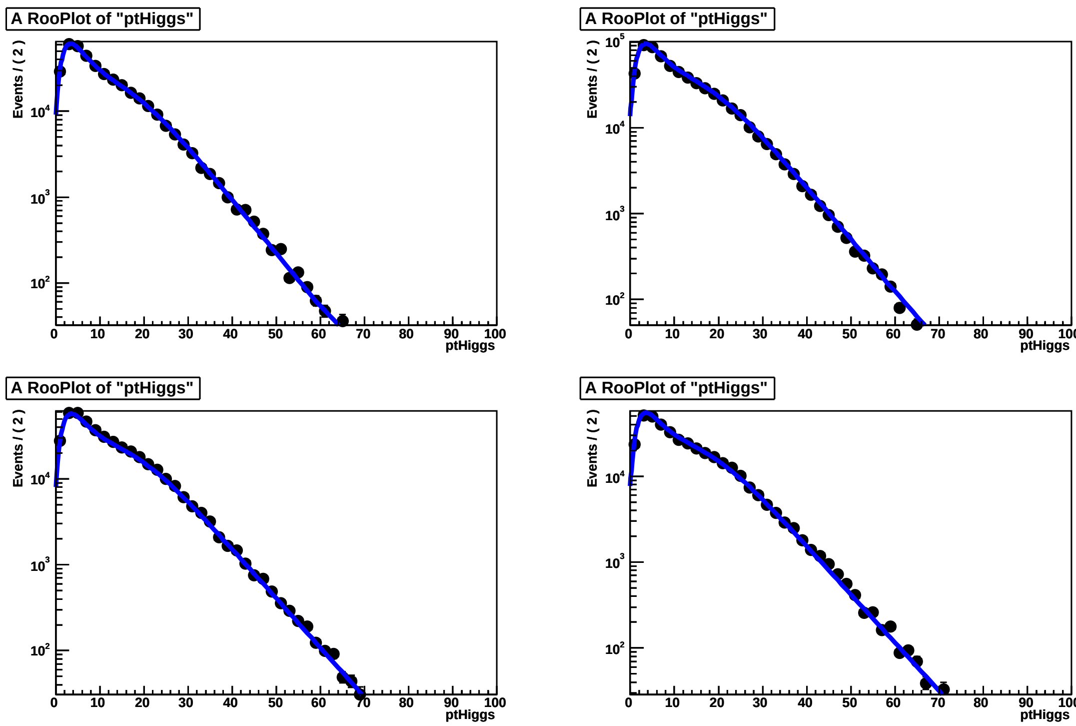

Figure A.55: The best fit of the functional form for QCD WW background to the central-value (fast simulation) Monte Carlo, in the region with small $\Delta\phi_{ll}$. Here, P_T^{Higgs} is shown in bins of M_T. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

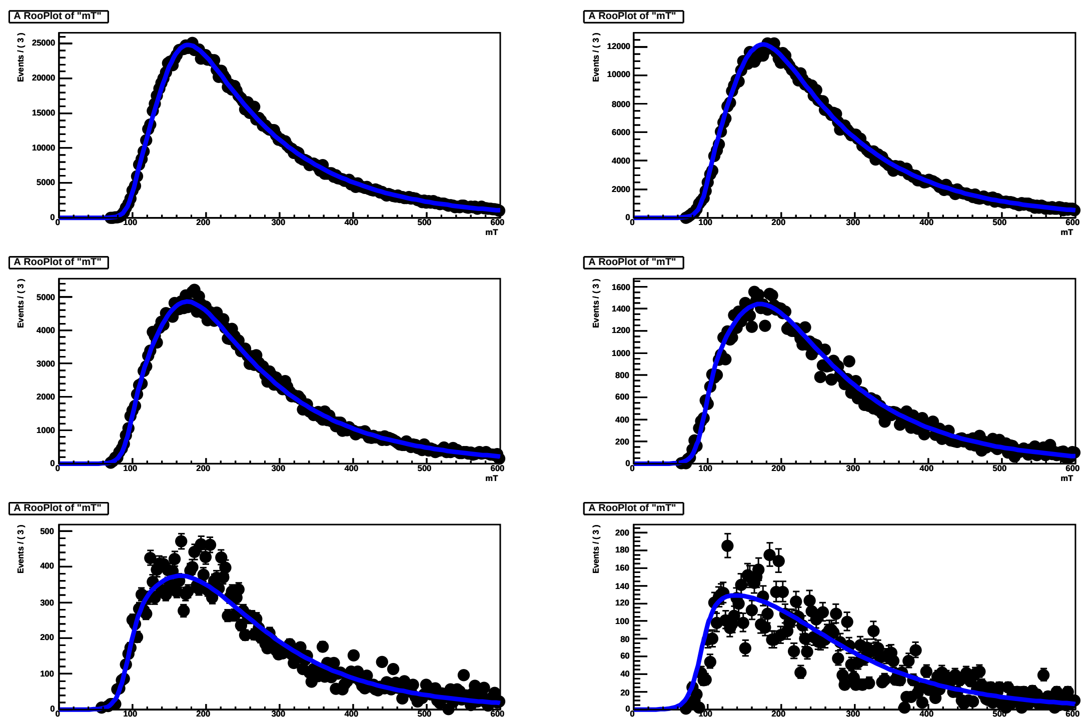

Figure A.56: The best fit of the functional form for QCD WW background to the central-value (fast simulation) Monte Carlo, in the region with large $\Delta\phi_{ll}$. Here, M_T is shown in bins of P_T^{Higgs}. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

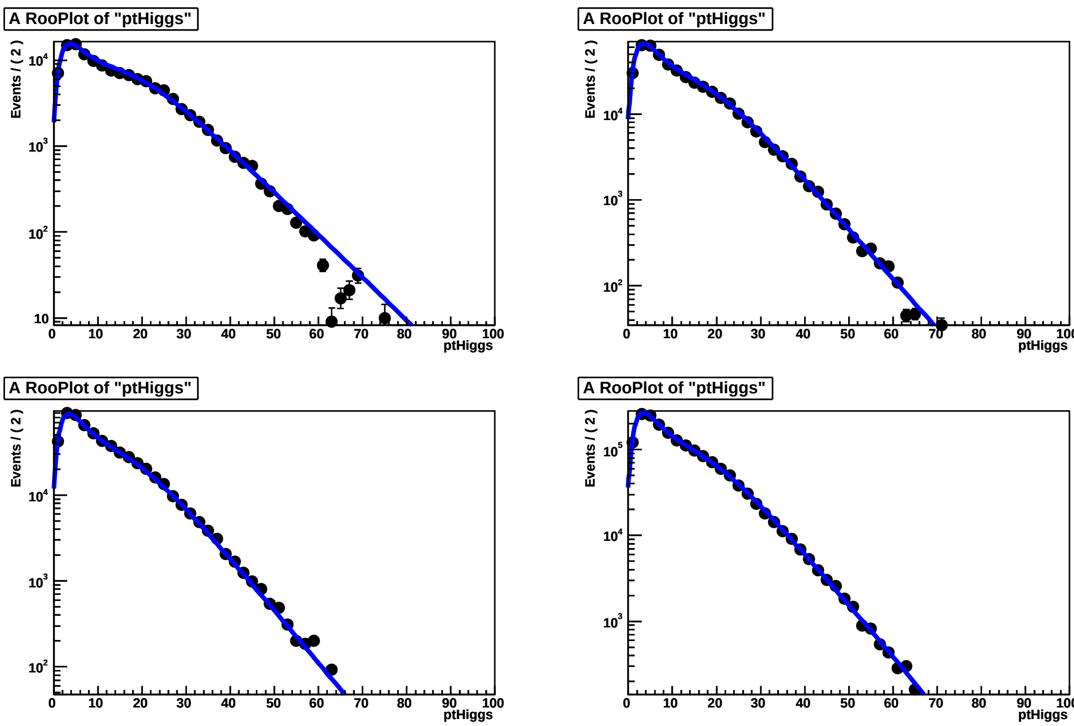

Figure A.57: The best fit of the functional form for QCD WW background to the central-value (fast simulation) Monte Carlo, in the region with large $\Delta\phi_{ll}$. Here, P_T^{Higgs} is shown in bins of M_T. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

Par.	Central	Scale 1	Scale 2	Scale 3	Scale 4
$m^{1,l}$	7.89399e-3	7.74915e-3	7.61075e-3	7.97652e-3	7.56339e-3
$m_1^{3,l}$	5.04709e-2	5.17819e-2	4.83655e-2	4.98330e-2	5.08593e-2
$m_2^{3,l}$	-6.14682e-4	-4.96490e-4	-3.76064e-4	-4.58609e-4	-5.44418e-4
$m_1^{4,l}$	142.836	142.463	143.474	142.964	142.916
$m_2^{4,l}$	-0.149756	-0.234516	-0.244052	-0.199722	-0.166972
$m_1^{6,l}$	9.84367e-2	0.108271	9.73652e-2	9.87036e-2	0.102740
$m_1^{7,l}$	43.6412	44.2309	43.3567	43.2286	44.3149
p_1^l	0.132675	0.132326	0.136403	0.142865	0.135606
p_2^l	1.41064	1.36847	1.35709	1.35533	1.29402
p_3^l	1.34420	1.31804	1.30269	1.28405	1.23852
p_4^l	6.41960	6.49235	5.32836	5.25156	5.00120
p_5^l	10.6345	10.0473	12.7063	13.5127	13.0080
p_6^l	0.601179	0.585548	0.675420	0.710809	0.673786
$m^{1,s}$	9.72549e-3	9.13931e-3	9.24093e-3	1.00824e-2	9.89887e-3
$m^{2,s}$	1.76511e-2	1.77102e-2	1.62585e-2	1.76904e-2	1.69364e-2
$m_1^{3,s}$	5.98103e-2	5.63175e-2	5.79895e-2	6.34147e-2	5.85427e-2
$m_2^{3,s}$	5.88854e-5	5.20478e-5	3.07535e-4	-2.60887e-4	1.97029e-4
$m_3^{3,s}$	-1.83741e-6	-1.29932e-6	-3.41183e-6	4.13317e-6	-2.65197e-6
$m_1^{4,s}$	128.272	129.237	128.839	127.453	128.484
$m_2^{4,s}$	0.3	0.299999	0.236827	0.299999	0.299959
$m^{5,s}$	67.3322	80.4070	71.0050	61.3279	56.5612
$m_1^{6,s}$	0.144568	0.157108	0.146928	0.147571	0.145820
$m_2^{6,s}$	-3.99664e-4	-3.99994e-4	2.81109e-4	3.99719e-4	-3.27887e-4
$m_3^{6,s}$	-3.56321e-5	-4.61194e-5	-5.55893e-5	-5.81854e-5	-3.45456e-5
$m_1^{7,s}$	48.8034	48.7930	48.7309	49.4631	48.8890
$m_2^{7,s}$	2.49471e-2	7.13783e-2	0.182497	-0.137196	4.24381e-2
$m_3^{7,s}$	3.98931e-3	2.78041e-3	-4.77115e-5	8.21856e-3	5.16798e-3
p_1^s	0.129192	0.130007	0.126903	0.130840	0.128203
p_2^s	1.54008	1.64142	1.70302	1.27166	1.59884
p_3^s	1.45654	1.57837	1.61372	1.21032	1.50391
p_4^s	6.73532	7.54258	7.03631	5.00263	6.71636
p_5^s	10.1610	9.68259	9.93167	11.7709	10.6771
p_6^s	0.615696	0.611256	0.626470	0.630411	0.643113

Table A.1: The best-fit values of the QCD WW background parameters for the various scale choices. The superscripts l and s denote the large-$\Delta\phi_{ll}$ and small-$\Delta\phi_{ll}$ regions, respectively.

Par.	Sm. P_T^{miss}
$m^{1,l}$	7.91755e-3
$m_1^{3,l}$	4.96175e-2
$m_2^{3,l}$	-4.64561e-4
$m_1^{4,l}$	140.534
$m_2^{4,l}$	-0.222328
$m_1^{6,l}$	9.66496e-2
$m_1^{7,l}$	44.0098
p_1^l	0.139104
p_2^l	3.24049
p_3^l	2.48369
p_4^l	1.55756e-6
p_5^l	19.5476
p_6^l	0.741383
$m^{1,s}$	9.51781e-3
$m^{2,s}$	1.72799e-2
$m_1^{3,s}$	6.33652e-2
$m_2^{3,s}$	-2.87795e-4
$m_3^{3,s}$	1.69285e-6
$m_1^{4,s}$	123.176
$m_2^{4,s}$	0.411895
$m^{5,s}$	65.3249
$m_1^{6,s}$	0.142402
$m_2^{6,s}$	-3.99958e-4
$m_3^{6,s}$	-3.06024e-5
$m_1^{7,s}$	48.7462
$m_2^{7,s}$	-7.99424e-2
$m_3^{7,s}$	7.77057e-3
p_1^s	0.133065
p_2^s	3.16860
p_3^s	2.43222
p_4^s	1.39549e-3
p_5^s	18.7282
p_6^s	0.747592

Table A.2: The best-fit values of the QCD WW background parameters for the scenario with smeared missing P_T. The superscripts l and s denote the large-$\Delta\phi_{ll}$ and small-$\Delta\phi_{ll}$ regions, respectively.

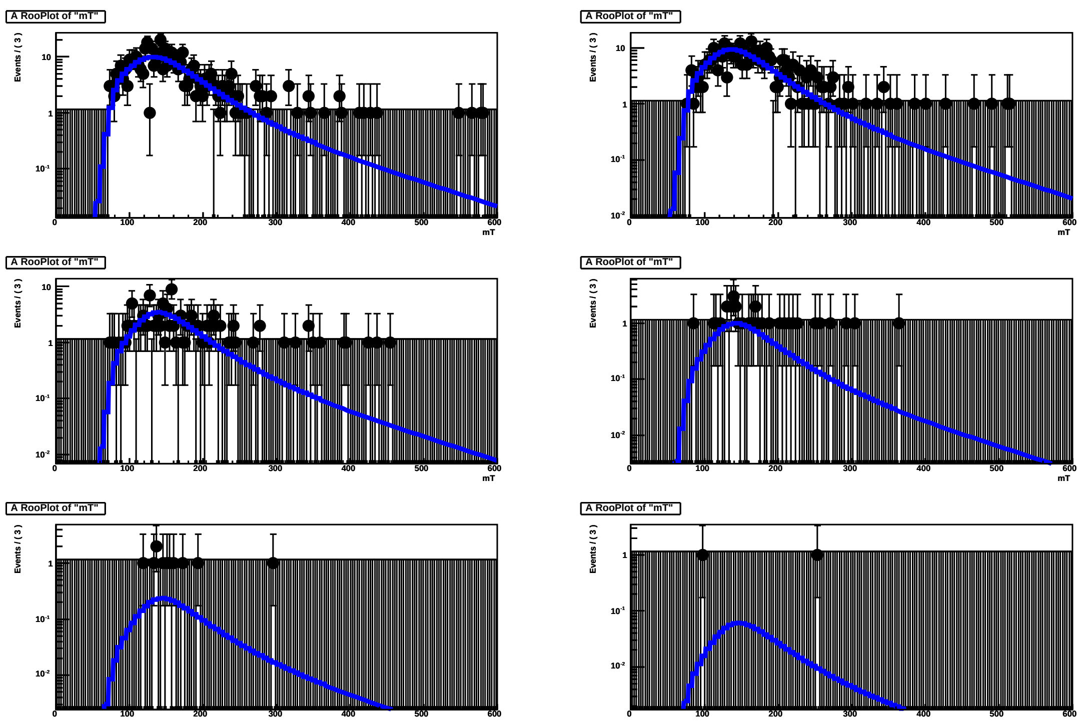

Figure A.58: The best fit of the functional form for QCD WW background to the central-value (full simulation) Monte Carlo, in the region with small $\Delta\phi_{ll}$. Here, M_T is shown in bins of P_T^{Higgs}. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

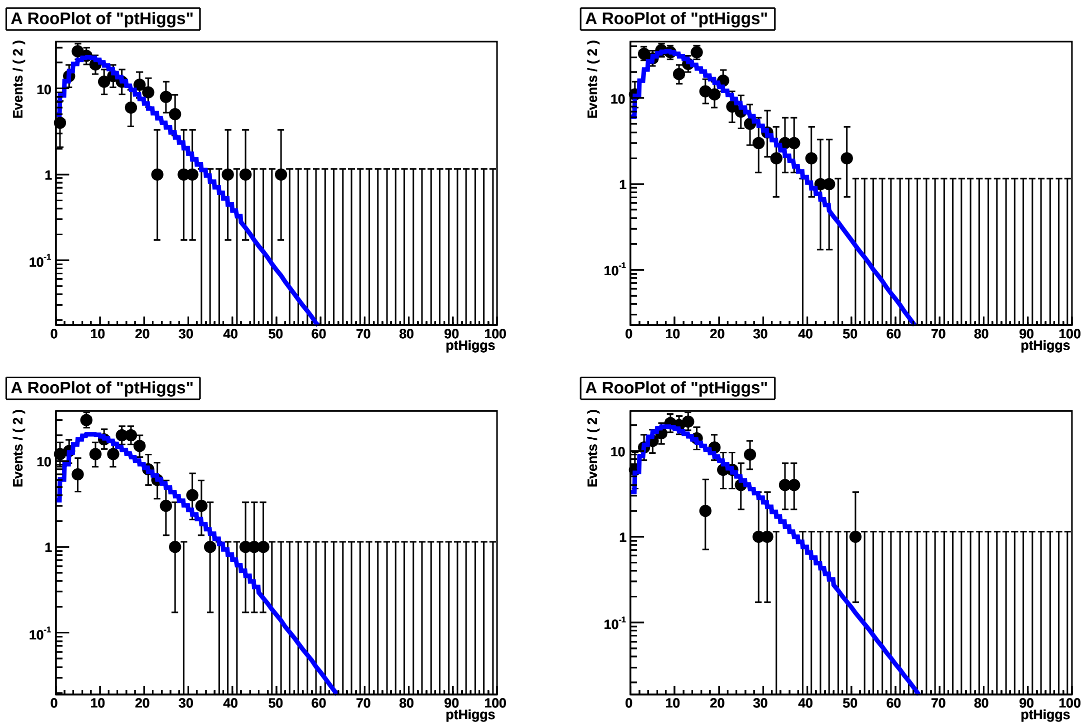

Figure A.59: The best fit of the functional form for QCD WW background to the central-value (full simulation) Monte Carlo, in the region with small $\Delta\phi_{ll}$. Here, P_T^{Higgs} is shown in bins of M_T. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

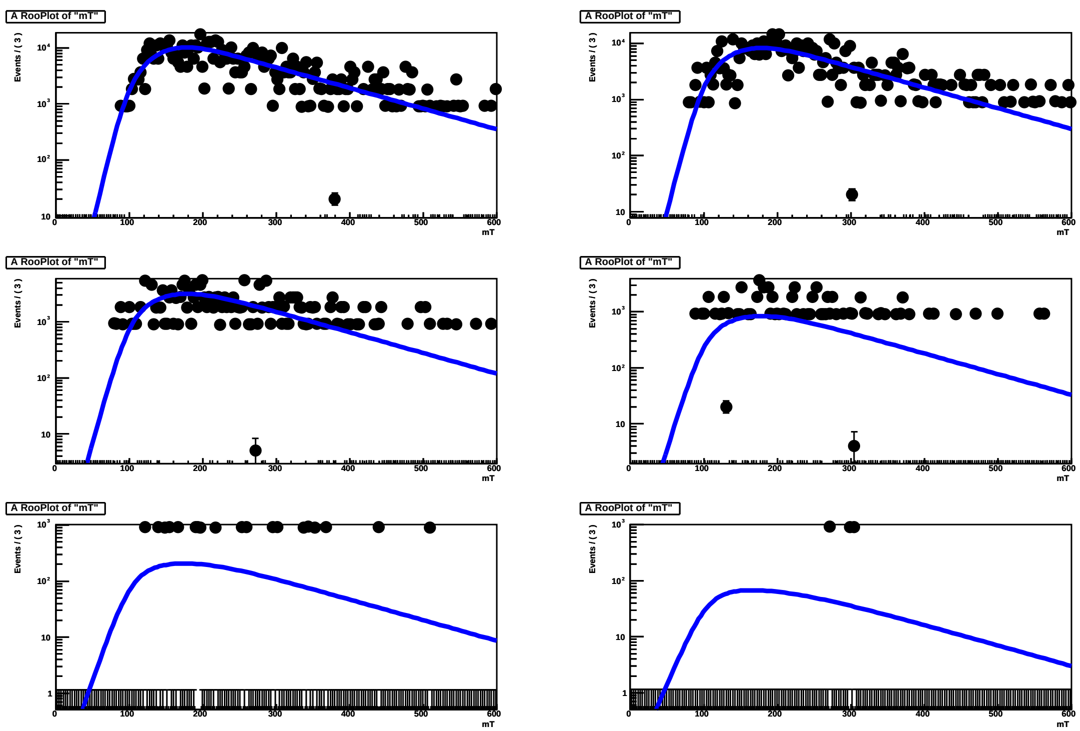

Figure A.60: The best fit of the functional form for QCD WW background to the central-value (full simulation) Monte Carlo, in the region with large $\Delta\phi_{ll}$. Here, M_T is shown in bins of P_T^{Higgs}. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

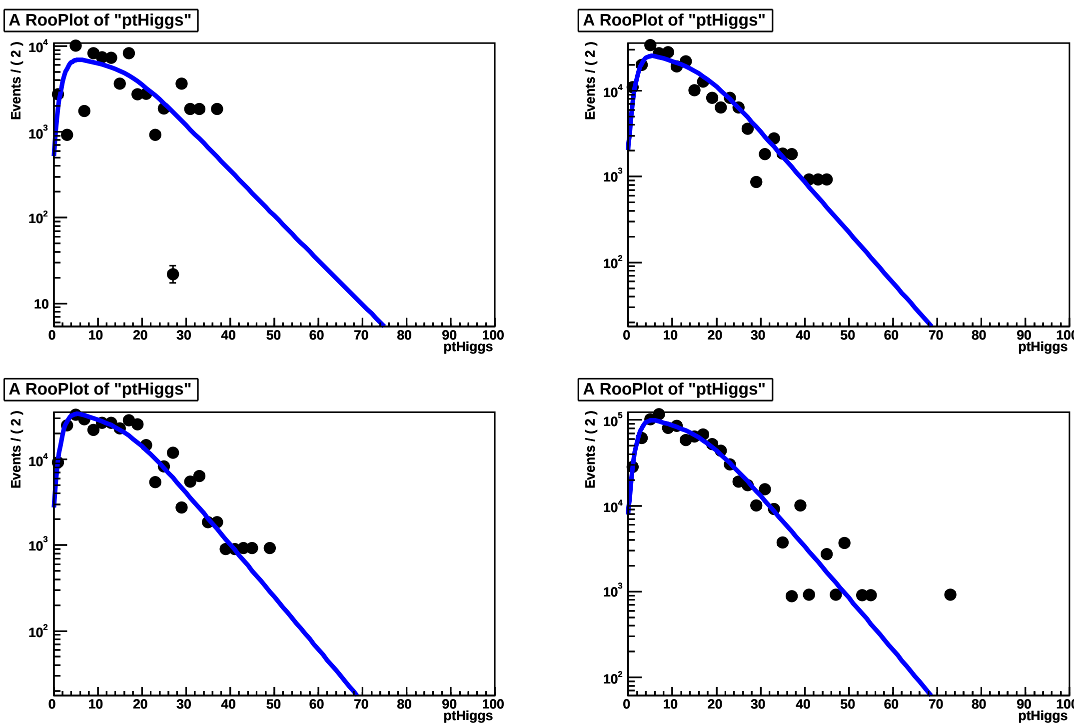

Figure A.61: The best fit of the functional form for QCD WW background to the central-value (full simulation) Monte Carlo, in the region with large $\Delta\phi_{ll}$. Here, P_T^{Higgs} is shown in bins of M_T. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

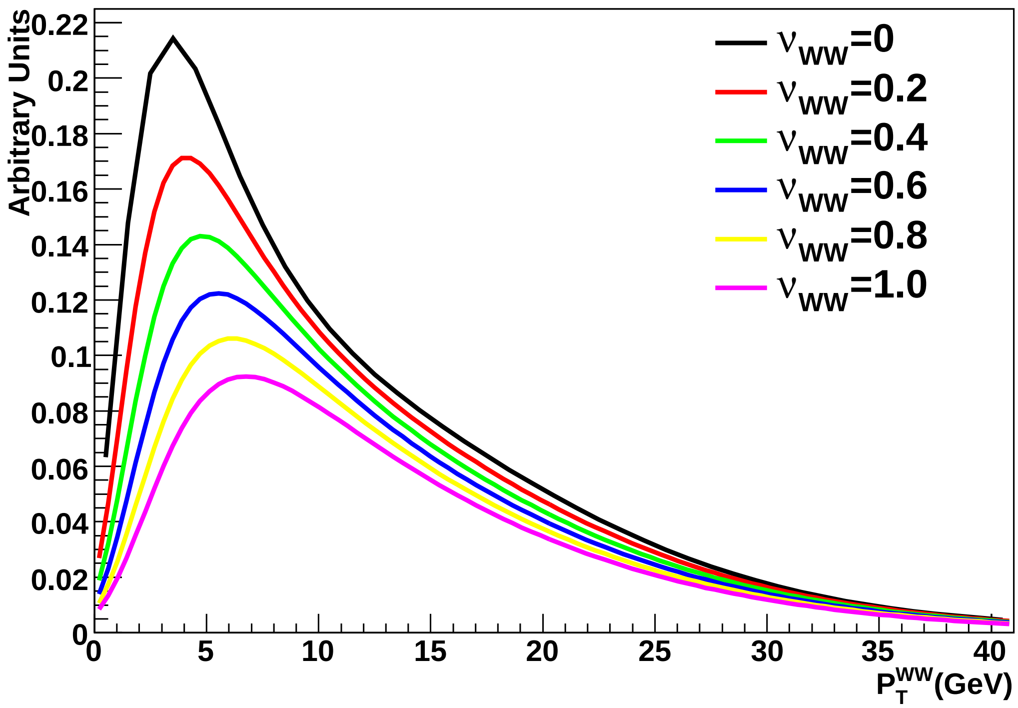

Figure A.62: The dependence of the P_T^{WW} component of the QCD WW M_T/P_T^{WW} fit model on ν_{WW}.

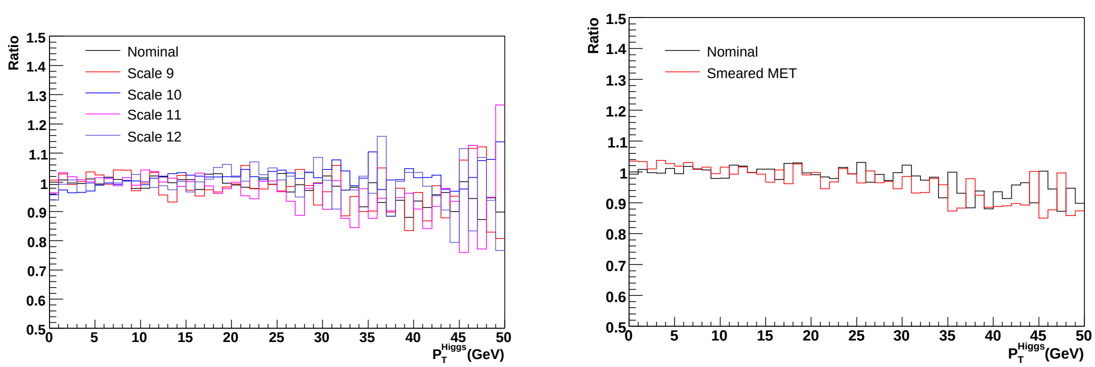

Figure A.63: Left: the ratio of the Higgs P_T shapes for the QCD WW background (normalized to 1) in the small-$\Delta\phi_{ll}$ and large-$\Delta\phi_{ll}$ regions, for the central-value case and for the various scale-varied cases. Right: the same ratio, for the central-value case and for the case with smeared P_T^{miss}.

Correlations are introduced by allowing some of the m_i to vary with Higgs P_T:

- $m_3 = m_3^1 + m_3^2 P_T^{Higgs} + m_3^3 (P_T^{Higgs})^2$
- $m_4 = m_4^1 + m_4^2 P_T^{Higgs}$
- $m_6 = m_6^1 + m_6^2 P_T^{Higgs} + m_6^3 (P_T^{Higgs})^2$
- $m_7 = m_7^1 + m_7^2 P_T^{Higgs} + m_7^3 (P_T^{Higgs})^2$

where the superscripts in m_3^1, m_3^2, etc. are indices, not exponents. Figures A.54 through A.57 show comparisons between the QCD WW Monte Carlo (from fast simulation, including the contribution from $gg \to WW$ via quark box) and the best-fit QCD WW model. The agreement is generally very good, and the correlation terms appear to be doing a good job describing the transverse mass distribution across a broad range of Higgs P_T.

To assess the systematic uncertainty due to the Q^2 scale, toy Monte Carlo samples have been generated from distributions obtained by fitting the same functional form to Monte Carlo samples with the Q^2 scales varied by 8, as described in Section 4.1. The parameters for both the central-value and scale-varied cases are shown in Table A.1.

A similar exercise has been performed to assess the systematic uncertainty due to a degraded Missing P_T resolution (or underlying event activity, etc.). Toy Monte Carlo outcomes have been sampled from a distribution obtained by fitting to Monte Carlo where the x and y components of the missing P_T have been smeared by 5 GeV. The best-fit values of the parameters for this scenario are also shown in Table A.1.

In the significance calculation presented in Figure 4.21, a fit to fully simulated QCD WW Monte Carlo samples was used as the generator for toy Monte Carlo outcomes. Figures A.58 through A.61 show the comparison between the fully simulated QCD WW events and the fit to them.

The purpose of the fits to “real” (fast or full simulation) Monte Carlo samples described above was to obtain parameterized models that could be used as generators for toy Monte Carlo outcomes. They were performed allowing all the parameters listed in Table A.1 to float (although the allowed ranges were restricted to improve convergence.)

An inspection of Tables A.1 and A.2 suggests that the transverse mass distribution is actually rather well-predicted. Therefore, in the fits to the toy Monte Carlo outcomes described in the main body of the text, most of the parameters governing the transverse mass distribution are fixed to the values in the “Central” column of Table A.1. There are two exceptions: $m^{1,l}$, which is allowed to float in the range 0.0074 to 0.0085, and $m_1^{3,s}$, which is allowed to float in the range 0.05 to 0.065. These parameters govern the slope of the tail of the M_T distribution in the large-$\Delta\phi_{ll}$ region and the steepness of the turnover of the low-M_T portion of the distribution in the small-$\Delta\phi_{ll}$ region, respectively.

The parameters p_1^l through p_6^l are combined into a single nuisance parameter ν_{WW} intended to absorb the shape uncertainty in the Higgs P_T spectrum:

- $p_1^l = 0.132675 + 0.006429\nu_{WW}$
- $p_2^l = 1.41064 + 1.82985\nu_{WW}$
- $p_3^l = 1.34420 + 1.13949\nu_{WW}$

- $p_4^l = 6.41960 - 6.41960\nu_{WW}$
- $p_5^l = 1.06345 + 8.91310\nu_{WW}$
- $p_6^l = 6.01179 + 0.140204\nu_{WW}$

The evolution of the Higgs P_T component of the distribution from $\nu_{WW} = 0$ to $\nu_{WW} = 1$ is shown in Figure A.62. Clearly, this sort of change in the shape can absorb a variety of systematic uncertainties, but in a future update of this analysis, it would nevertheless be desirable to explicitly check for biases in a more complete set of degraded scenarios. The nuisance parameter is used in the fits to pseudodata and (eventually) real data, but not in the fits that defined the truth models used to generate toy Monte Carlo outcomes.

To a good approximation, the Higgs P_T distribution in the small-$\Delta\phi_{ll}$ region is the same as the distribution in the large-$\Delta\phi_{ll}$ region. To the extent that this is true, it is possible to constrain the Higgs P_T shape parameters in the two regions to be the same in the fit, i.e., to use only one ν_{WW} to describe the Higgs P_T in both the large-$\Delta\phi_{ll}$ and small-$\Delta\phi_{ll}$ regions.

Figure A.63 shows the ratio of the Higgs P_T shapes in the small-$\Delta\phi_{ll}$ and large-$\Delta\phi_{ll}$ regions in the central-value and systematically distorted scenarios (using fast simulation). The plots show that the shapes are, in fact, slightly different at the level of a few percent. Moreover, the difference in the Higgs P_T shape between the two regions is not exactly the same in all systematically distorted cases. Therefore, in the fit, the distribution of the Higgs P_T for QCD WW background in the small-$\Delta\phi_{ll}$ region is taken to be $(1 + c_1 \times P_T^{WW} + c_2 \times (P_T^{WW})^2)f(P_T^{WW})$, where $f(P_T^{WW})$ is the Higgs P_T distribution in the large-$\Delta\phi_{ll}$ region, and the parameters c_1 and c_2 are allowed to

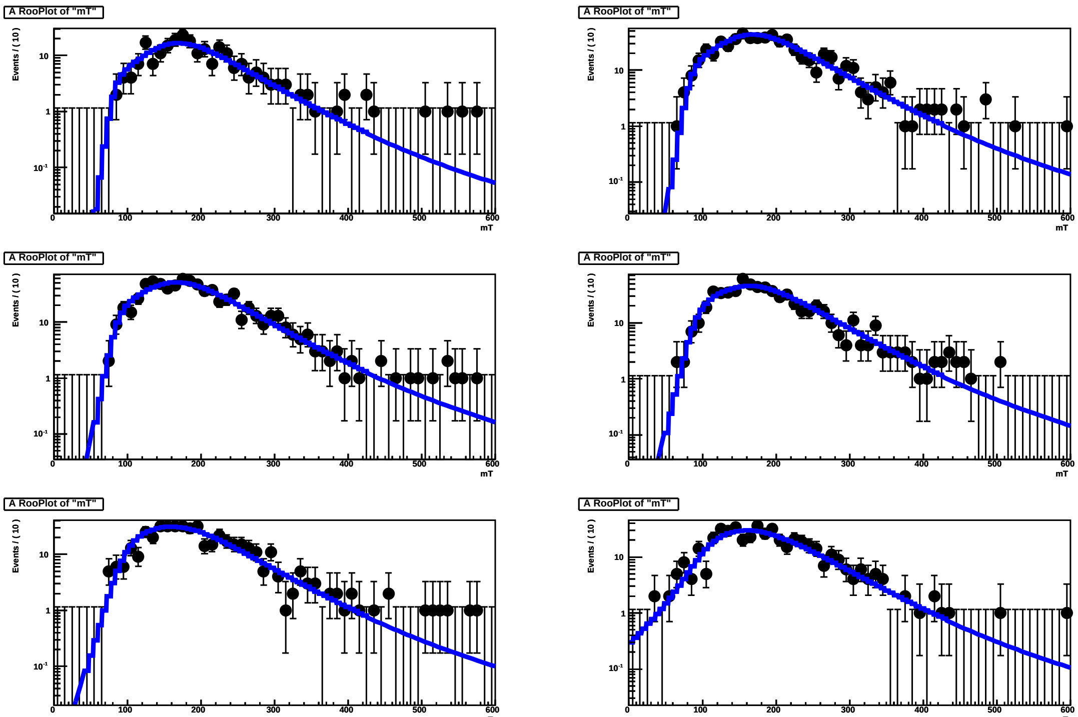

Figure A.64: The best fit of the functional form for the $t\bar{t}$ background to the (fast simulation) Monte Carlo, in the region with a b-tagged jet (with $P_T < 20$ GeV) and small $\Delta\phi_{ll}$. Here, M_T is shown in bins of P_T^{Higgs}. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

float in the fit.

A.3 Model for the Top Background

The functional form used for the top background in the b-tagged regions is the same correlated product PDF used to model the QCD WW background. To extrapolate the top background from the b-tagged regions to the signal-like regions, the shape of the distribution is distorted by multiplying it with a function that is the product of two uncorrelated exponentials in P_T^{Higgs} and M_T plus Gaussian ridges

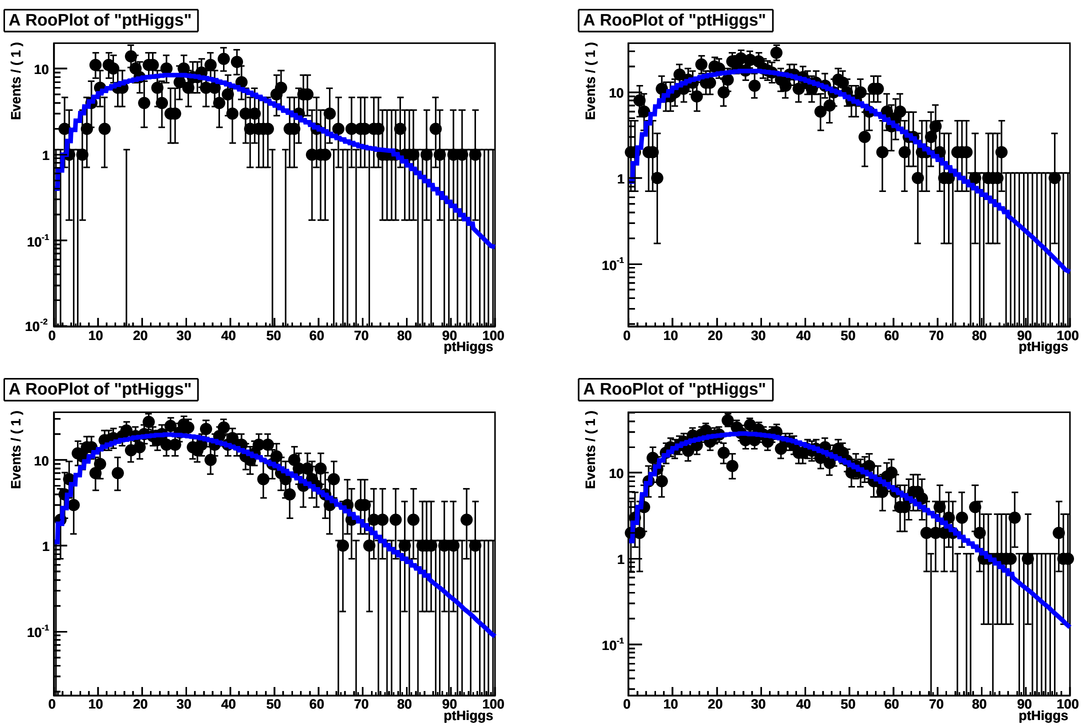

Figure A.65: The best fit of the functional form for the $t\bar{t}$ background to the (fast simulation) Monte Carlo, in the region with a b-tagged jet (with $P_T < 20$ GeV) and small $\Delta\phi_{ll}$. Here, P_T^{Higgs} is shown in bins of M_T. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

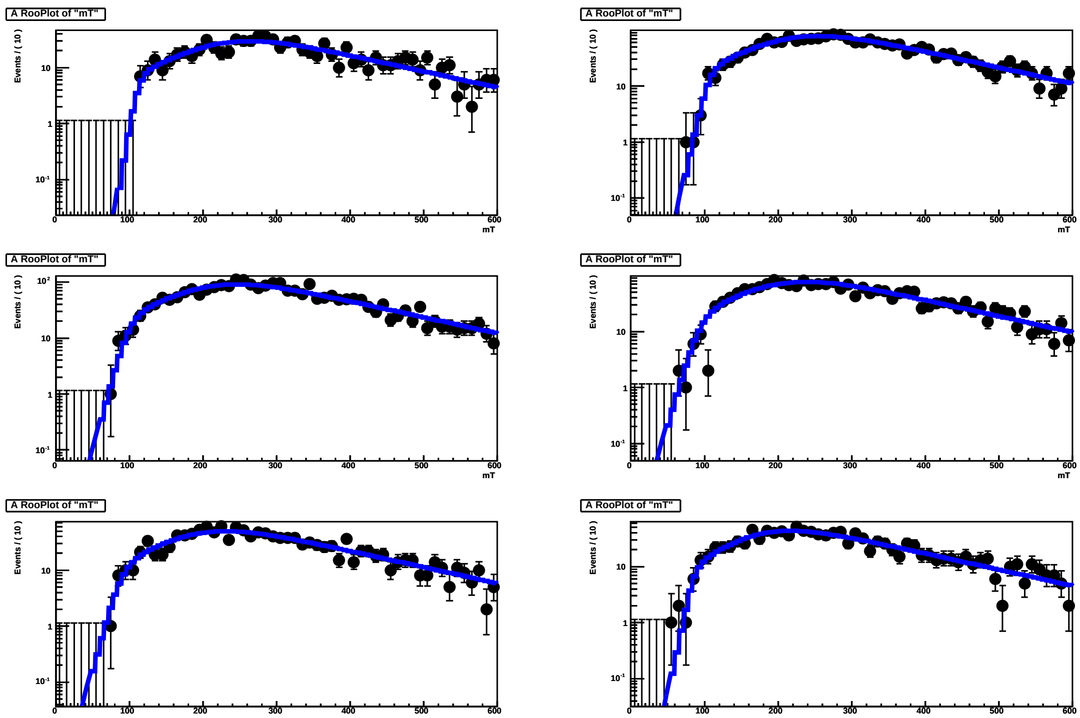

Figure A.66: The best fit of the functional form for the $t\bar{t}$ background to the (fast simulation) Monte Carlo, in the region with a b-tagged jet (with $P_T < 20$ GeV) and large $\Delta\phi_{ll}$. Here, M_T is shown in bins of P_T^{Higgs}. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

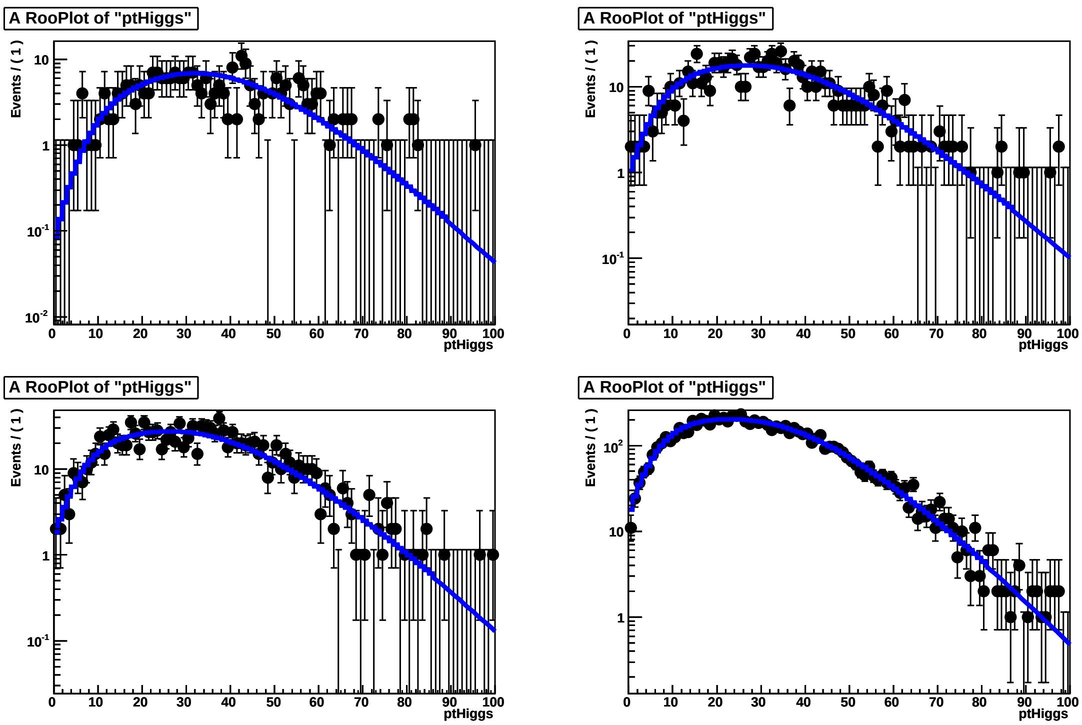

Figure A.67: The best fit of the functional form for $t\bar{t}$ background to the (fast simulation) Monte Carlo, in the region with a b-tagged jet (with $P_T < 20$ GeV) and large $\Delta\phi_{ll}$. Here, P_T^{Higgs} is shown in bins of M_T. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

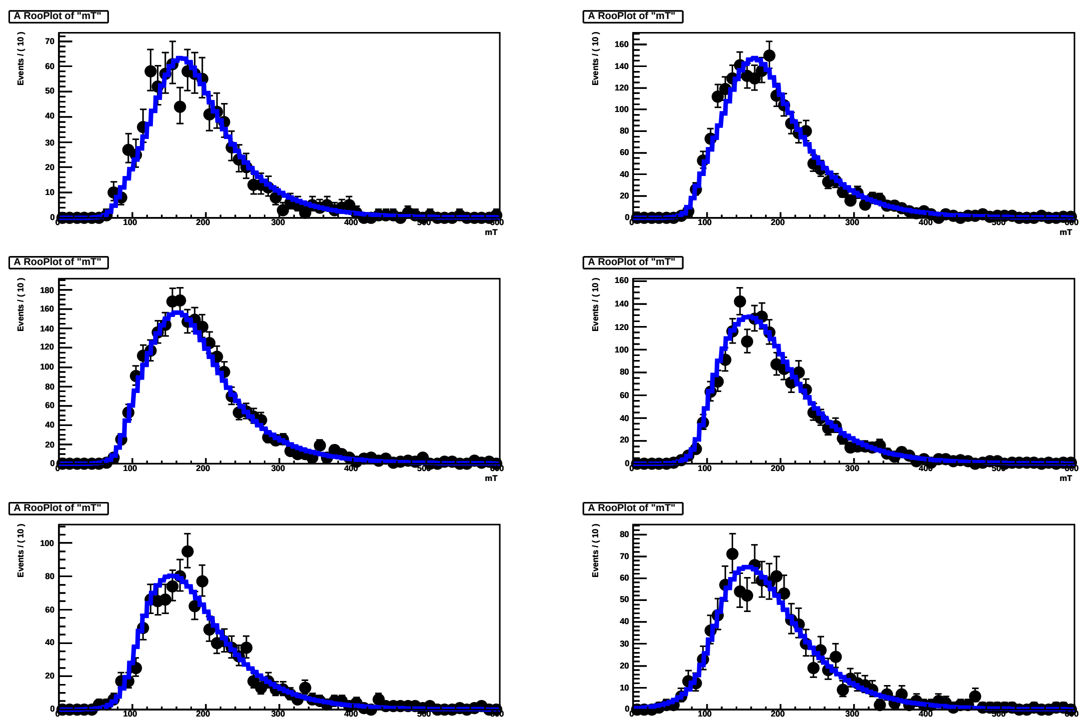

Figure A.68: The best fit of the functional form for the $t\bar{t}$ background to the (fast simulation) Monte Carlo, in the region with a b-veto and small $\Delta\phi_{ll}$. Here, M_T is shown in bins of P_T^{Higgs}. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

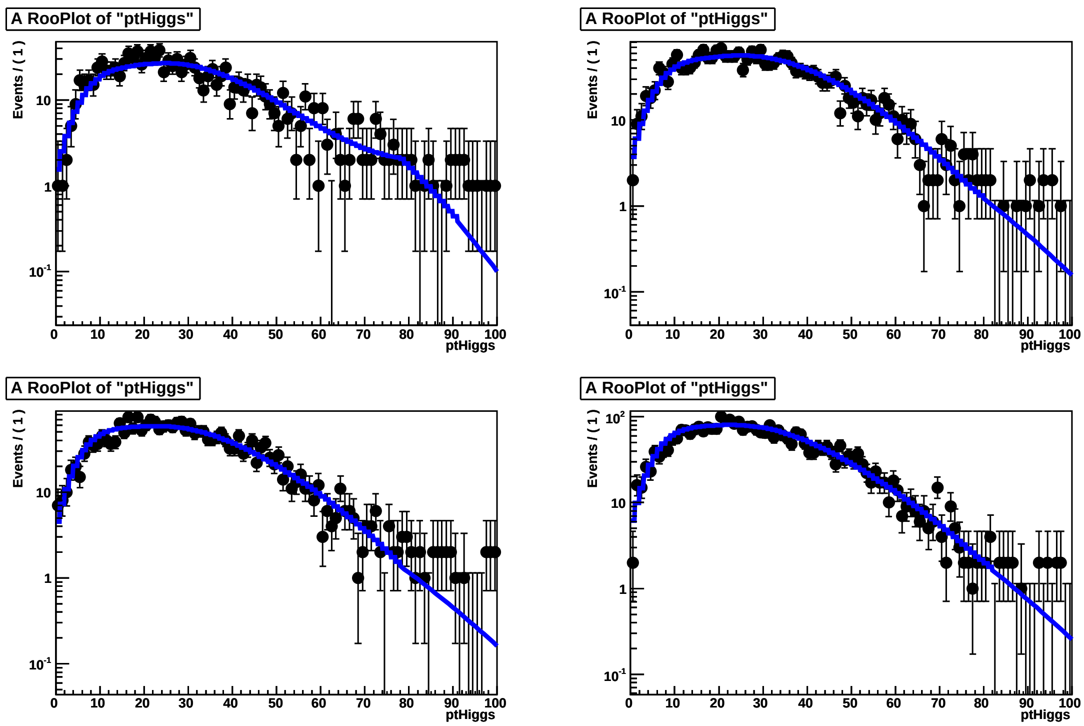

Figure A.69: The best fit of the functional form for the $t\bar{t}$ background to the (fast simulation) Monte Carlo, in the region with a b-veto and small $\Delta\phi_{ll}$. Here, P_T^{Higgs} is shown in bins of M_T. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

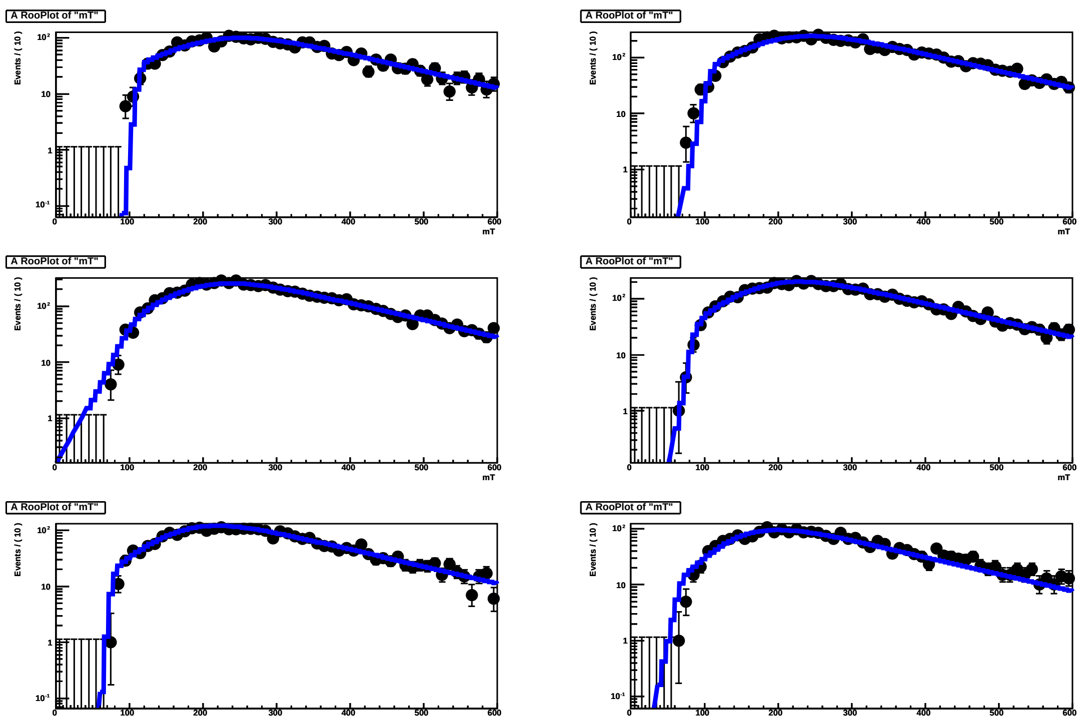

Figure A.70: The best fit of the functional form for the $t\bar{t}$ background to the (fast simulation) Monte Carlo, in the region with a b-veto and large $\Delta\phi_{ll}$. Here, M_T is shown in bins of P_T^{Higgs}. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

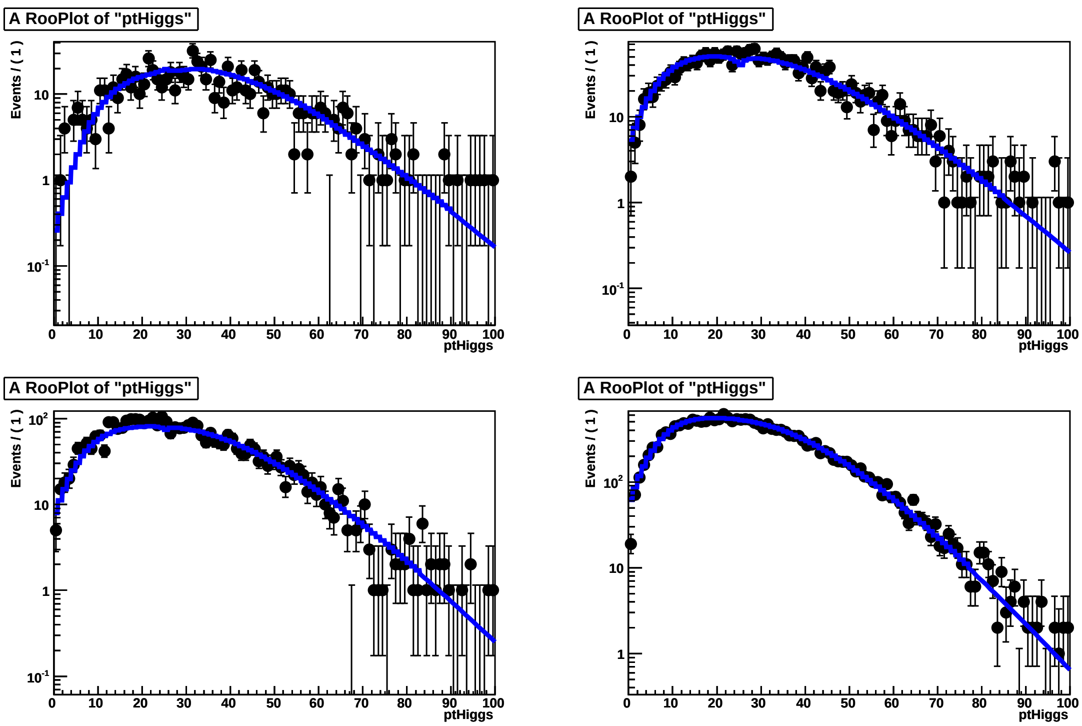

Figure A.71: The best fit of the functional form for $t\bar{t}$ background to the (fast simulation) Monte Carlo, in the region with a b-veto and large $\Delta\phi_{ll}$. Here, P_T^{Higgs} is shown in bins of M_T. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

Par.	NLO $t\bar{t}$	LO $t\bar{t}$	LO $WWbb$	Smeared P_T^{miss}
m_1	6.63892e-03	4.48110e-03	4.22416e-03	6.16086e-03
m_3^1	2.55281e-02	1.75013e-02	2.59352e-02	2.37549e-02
m_3^2	-1.31528e-04	3.08511e-04	-1.02382e-04	8.44752e-05
m_3^3	2.28065e-06	-2.23157e-06	3.17158e-06	0
m_4^1	2.20413e+02	2.31617e+02	2.13014e+02	2.11193e+02
m_4^2	-6.99873e-01	-8.91229e-01	-4.20417e-01	-6.66409e-01
m_5	0	0	0	0
m_6^1	1.95813e-01	4.83507e-01	1.94574e-01	9.09544e-02
m_6^2	-7.15139e-03	-9.01966e-03	-5.78590e-03	-3.32428e-03
m_6^3	9.97966e-05	-5.50480e-04	9.66482e-05	9.05324e-05
m_7^1	1.07565e+02	1.05967e+02	1.11119e+02	9.97263e+01
m_7^2	1.40394e-01	8.26276e-01	-6.99707e-01	-7.59743e-02
m_7^3	-5.27562e-03	-1.57606e-02	1.57741e-02	0
p_1	1.19194e-01	1.23987e-01	1.15350e-01	1.14223e-01
p_2	2.31874e+01	1.74193e+01	1.07577e+01	1.86785e+01
p_3	1.22939e+01	9.62367e+00	5.97721e+00	1.03284e+01
p_4	1.66758e+01	1.31288e+01	1.04125e+01	1.45740e+01
p_5	3.68996e+01	4.05542e+01	3.52405e+01	4.55179e+01
p_6	8.89877e-01	9.48201e-01	9.63429e-01	9.46228e-01
α	-6.22344e-04	-5.40450e-04	-5.56795e-04	-5.93258e-04
β	-8.53154e-03	-7.06149e-03	-7.87352e-03	-9.33475e-03
B	8.13035e-02	3.48034e-01	9.55527e-01	1.47569e-03
C	5.35996e-03	4.53368e-03	1.19551e-02	6.08304e-03
σ_1	10.5516	10.5516	10.5516	10.5516
σ_2	40	40	40	40
μ_2	126.441	126.441	126.441	126.441

Table A.3: The values of the $t\bar{t}$ background parameters used as the truth for the NLO scenario, the LO scenario, the $WWbb$ scenario, and the smeared P_T^{miss} scenario, in the region with large $\Delta\phi_{ll}$.

Par.	NLO $t\bar{t}$	LO $t\bar{t}$	LO $WWbb$	Smeared P_T^{miss}
m_1	7.64271e-03	5.39861e-03	1.37554e-02	7.02518e-03
m_2	1.05850e-02	1.65217e-02	2.14264e-03	1.03707e-02
m_3^1	4.58939e-02	2.62394e-02	4.09461e-02	3.77398e-02
m_3^2	-1.04324e-04	0	0	4.49588e-04
m_3^3	-2.77136e-06	0	0	-7.72104e-06
m_4^1	1.57590e+02	1.94669e+02	1.49260e+02	1.48632e+02
m_4^2	-2.22667e-02	4.39706e-01	3.28447e-01	1.03651e-01
m_5	1.30920e+02	9.85432e+02	1.22097e+02	1.11171e+02
m_6^1	1.69359e-01	2.82696e-02	1.00884e-01	2.16821e-01
m_6^2	-3.09060e-03	9.52507e-04	3.47790e-04	-3.99988e-03
m_6^3	1.13184e-05	0	0	1.13924e-05
m_7^1	7.97975e+01	8.43493e+01	4.87105e+01	6.99218e+01
m_7^2	6.20855e-03	4.98858e-01	9.46112e-01	4.10164e-04
m_7^3	-8.18529e-03	0	0	-6.19589e-04
p_1	9.29387e-02	1.09340e-01	8.37908e-02	1.05901e-01
p_2	1.40092e+01	1.82701e+01	2.39638e+00	2.35630e+01
p_3	8.12949e+00	8.87560e+00	1.00004e+00	1.23338e+01
p_4	1.35427e+01	1.60903e+01	3.89121e+00	1.83678e+01
p_5	3.30666e+01	4.48057e+01	3.20858e+01	3.51516e+01
p_6	9.11801e-01	9.63861e-01	9.88816e-01	8.91439e-01
α	-8.91892e-04	-1.07281e-03	-1.29808e-03	-4.90488e-04
β	-9.00708e-03	-2.56898e-03	-7.80195e-03	-9.34491e-03
B	5.21577e-02	8.58119e-01	7.28849e-01	4.26286e-03
C	0	0	0	0
σ_1	10.5516	10.5516	10.5516	10.5516
σ_2	40	40	40	40
μ_2	126.441	126.441	126.441	126.441

Table A.4: The values of the $t\bar{t}$ background parameters used as the truth for the NLO scenario, the LO scenario, the $WWbb$ scenario, and the smeared P_T^{miss} scenario, in the region with small $\Delta\phi_{ll}$.

centered at $P_T^{Higgs} = 0$ and $M_T = \mu_2$:

$$e^{\alpha M_T} e^{\beta P_T^{Higgs}} + B \text{ Gaus}(P_T^{Higgs}, 0, \sigma_1) + (C P_T^{Higgs})^2 \text{ Gaus}(M_T, \mu_2, \sigma_2) \tag{A.5}$$

Some parameters are always fixed:

- $C = 0$ in the region with small $\Delta\phi_{ll}$ (but it is free to float in the region with large $\Delta\phi_{ll}$)
- $\sigma_1 = 10.5516$ GeV
- $\sigma_2 = 40$ GeV
- $\mu_2 = 126.441$ GeV

All other parameters are freeable, but some are fixed at runtime to values specified in a jobOption file.

In the large-$\Delta\phi_{ll}$ region, the various parameters governing the Higgs P_T distribution are combined into a single nuisance parameter $\nu_{t\bar{t}}$, as follows:

- $p_1 = 0.119194 - 0.004971\nu_{t\bar{t}}$
- $p_2 = 23.1874 - 4.5089\nu_{t\bar{t}}$
- $p_3 = 12.2939 - 1.9655\nu_{t\bar{t}}$
- $p_4 = 16.6758 - 2.1018\nu_{t\bar{t}}$
- $p_5 = 36.8996 + 8.6183\nu_{t\bar{t}}$
- $p_6 = 0.889877 + 0.056351\nu_{t\bar{t}}$

In practice, the parameters governing the Higgs P_T distribution in the small-$\Delta\phi_{ll}$ region are not allowed to float in the fit, so there is no need for a quantity like $\nu_{t\bar{t}}$ in the small-$\Delta\phi_{ll}$ region.

Figures A.64 through A.67 show comparisons between the fastsim MC@NLO $t\bar{t}$ Monte Carlo in the b-tagged regions with large and small $\Delta\phi_{ll}$ and the corresponding fits. Figures A.68 through A.71 show similar comparisons for the b-tagged regions. The agreement between the model and the Monte Carlo is quite good.

To study the robustness of the fitting algorithm against differences between the description of the $t\bar{t}$ final state from MC@NLO and other treatments which include contributions from non-resonant production of $WWb\bar{b}$ events, models of the top background have been obtained using Monte Carlo samples generated with the MadEvent $pp \rightarrow t\bar{t} \rightarrow WWb\bar{b}$ and MadEvent $pp \rightarrow WWb\bar{b}$ generators. Also, to assess the systematic uncertainty due to a degraded Missing P_T resolution, a model has been obtained using a Monte Carlo sample (generated using MC@NLO) with the x and y components of the missing P_T smeared by 5 GeV. Table A.3 shows the parameters for the large-$\Delta\phi_{ll}$ region, and Table A.3 shows the parameters for the small-$\Delta\phi_{ll}$ region.

A.4 Model for the $Z \rightarrow \tau\tau$ Background

The functional form used for the $Z \rightarrow \tau\tau$ background is also a correlated product PDF. The Higgs P_T distribution is described by the same functional form as all the other processes; the transverse mass distribution is modelled by a bifurcated Gaussian $G_b(M_T, \mu, \sigma_L, \sigma_R)$. Correlations are introduced by allowing μ, σ_L, and σ_R to depend on the Higgs P_T:

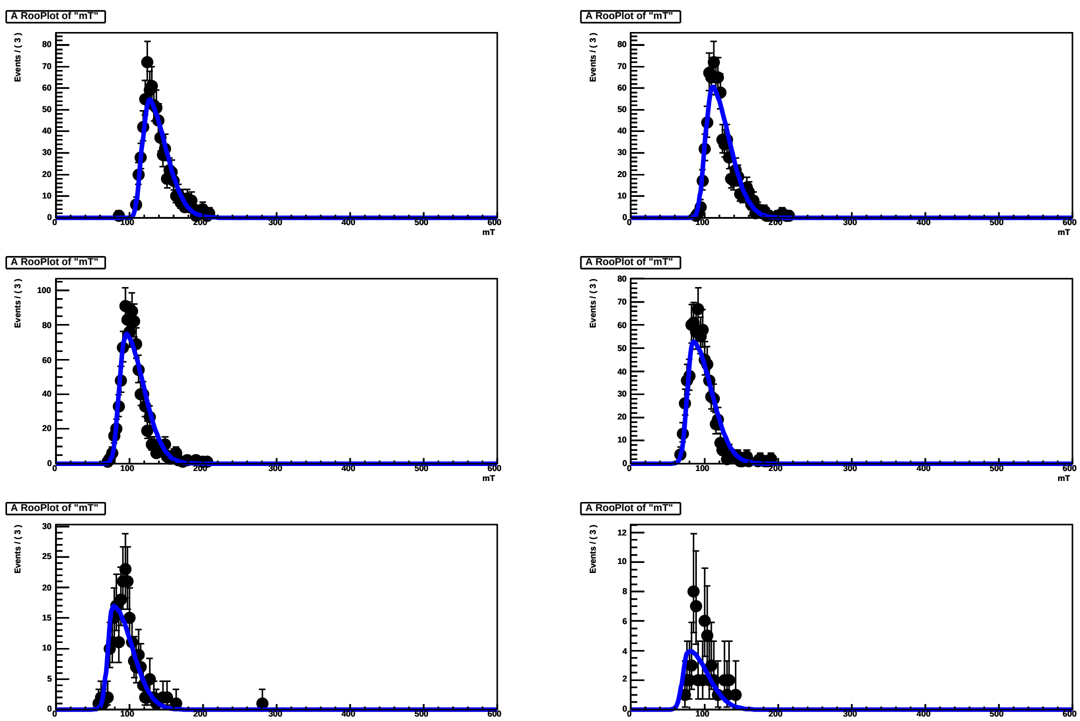

Figure A.72: The best fit of the functional form for the $Z \to \tau\tau$ background to the (fast simulation) Monte Carlo, in the region with large $\Delta\phi_{ll}$. Here, M_T is shown in bins of P_T^{Higgs}. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

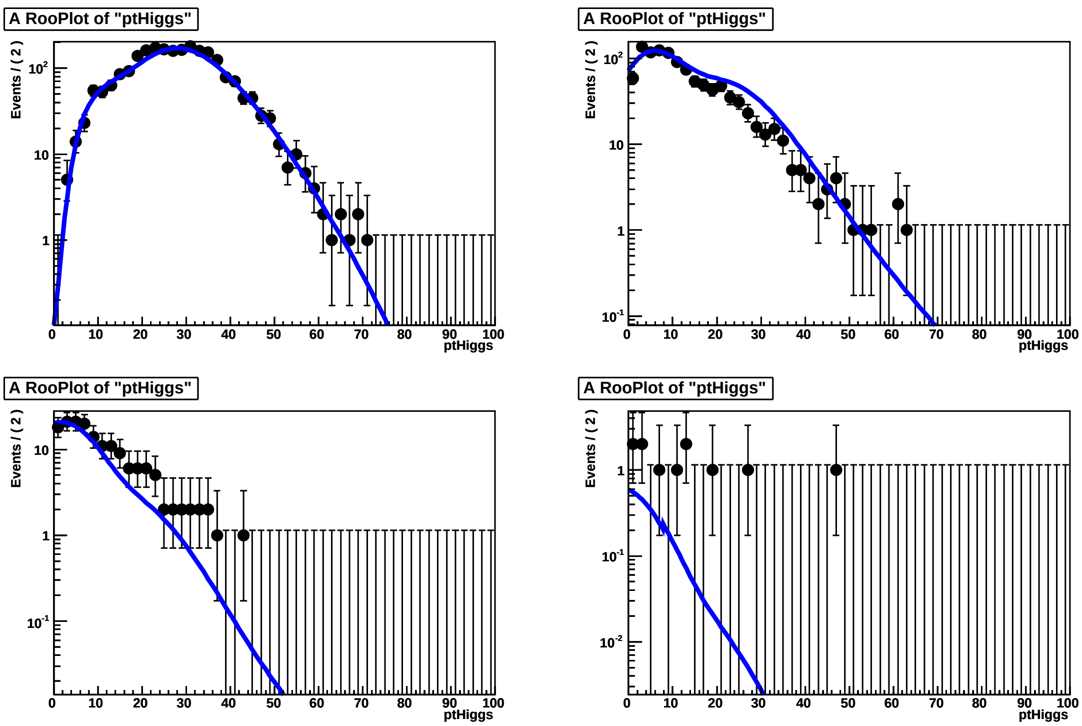

Figure A.73: The best fit of the functional form for $Z \rightarrow \tau\tau$ background to the (fast simulation) Monte Carlo, in the region with large $\Delta\phi_{ll}$. Here, P_T^{Higgs} is shown in bins of M_T. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

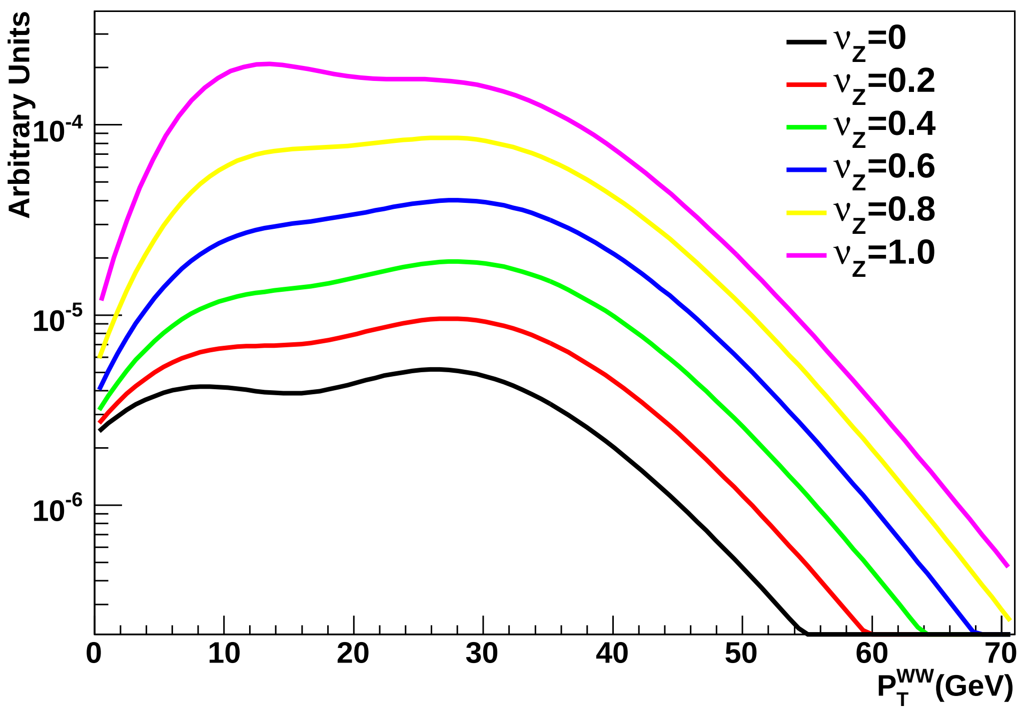

Figure A.74: The dependence of the Higgs P_T component of the $Z \rightarrow \tau\tau$ background model on the ν_Z parameter.

Par.	Nominal	Smeared P_T^{miss}
m_0	1.29903e+02	1.30997e+02
m_1	-5.28082e+01	-4.96372e+01
s_0	3.39644	3.34068
s_1	3.55946	5.05652
s_2	2.69603e+01	2.39765e+01
s_3	-1.90856e	-1.00920e-01
p_1	2.44788e-01	2.05242e-01
p_2	7.48873e+01	2.93603e+01
p_3	2.94392e+01	1.24034e+01
p_4	1.57873e+01	1.92211e+01
p_5	1.03804e+01	2.17654e+01
p_6	6.83593e-01	9.09709e-01

Table A.5: The values of the $Z \to \tau\tau$ background parameters used as the truth for the nominal and smeared P_T^{miss} scenarios. This table shows the values for the region with large $\Delta\phi_{ll}$; the $Z \to \tau\tau$ background is ignored in the small-$\Delta\phi_{ll}$ region.

- $\mu = \mu_0 + \mu_1 \sin(\pi P_T^{Higgs}/100.)$
- $\sigma_R = \sigma_R^0 + \sigma_R^1 \sin(\pi P_T^{Higgs}/100.)$
- $\sigma_L = \sigma_L^0 + \sigma_L^1 \sin(\pi P_T^{Higgs}/100.)$

where P_T^{Higgs} is measured in GeV. Figures A.72 and A.73 show comparisons between the fastsim $Z \to \tau\tau$ Monte Carlo and the best-fit parameterization in the region with large $\Delta\phi_{ll}$. (This background is negligible in the region with small $\Delta\phi_{ll}$.)

To study the robustness of the fitting algorithm against a degraded missing P_T resolution, best-fit values of the parameters have been obtained from a fit to a Monte Carlo sample with a 5 GeV smearing applied to the x and y components of the missing P_T. Table A.5 shows the best-fit parameters in the region with large $\Delta\phi_{ll}$ for the nominal and smeared cases.

As was done for the QCD WW background, the parameters governing the Higgs P_T distribution for this background are combined into one single nuisance parameter ν_Z in order to enhance the stability of the fit:

- $p_1 = 0.244788 - 0.039546\nu_Z$
- $p_2 = 74.8873 - 45.52700\nu_Z$
- $p_3 = 29.4392 - 17.03580\nu_Z$
- $p_4 = 15.7873 + 3.433800\nu_Z$
- $p_5 = 10.3804 + 11.38500\nu_Z$
- $p_6 = 0.683593 + 0.226116\nu_Z$

Figure A.74 shows how the Higgs P_T distribution for $Z \to \tau\tau$ evolves as one changes ν_Z from 0 (unsmeared case) to 1 (5 GeV smearing).

A.5 Model for the W+jets Background

The functional form used to model the W+jets background is an uncorrelated product PDF consisting of the Expression A.2 to describe the Higgs P_T times Expression A.4 to describe the transverse mass distribution. The functional form is the same in both the small-$\Delta\phi_{ll}$ region and the large-$\Delta\phi_{ll}$ region, but the values of the parameters are determined separately for the two regions.

After all analysis cuts have been applied, only one W+jets Monte Carlo event remains in each of the two b-vetoed regions of the analysis. In order to obtain an estimate of the shape of the distribution in P_T^{WW} and M_T, a loosened event selection

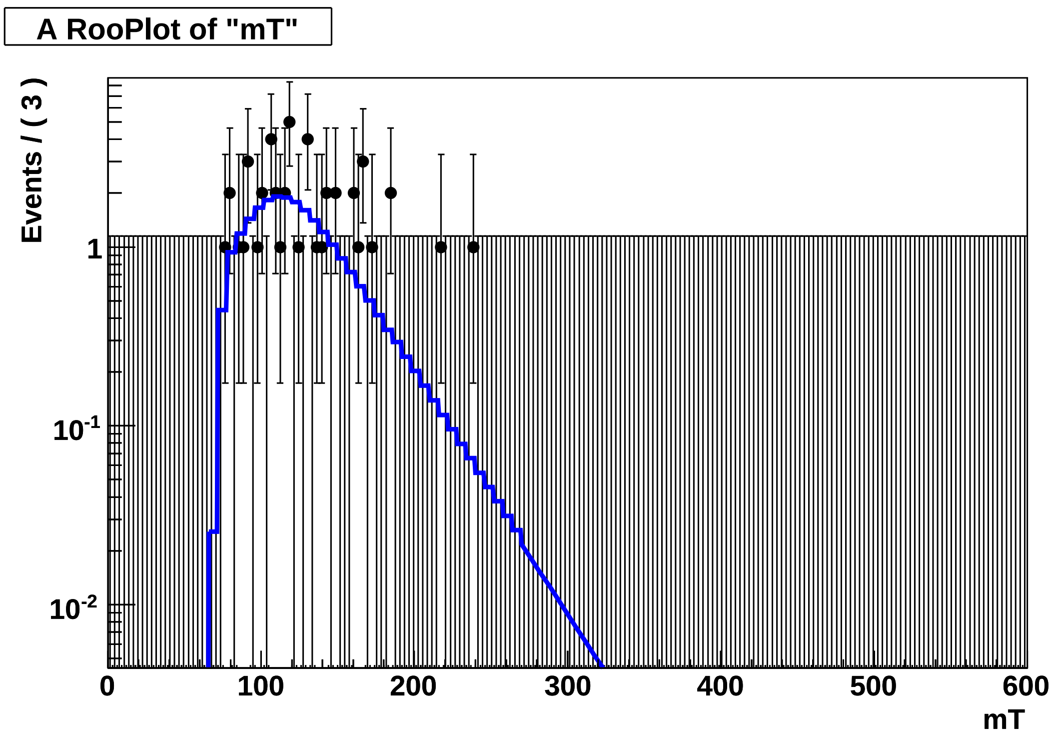

Figure A.75: The best fit of the functional form for the W+jets background to the (full simulation) Monte Carlo, in the region with small $\Delta\phi_{ll}$. Here, M_T is shown in bins of P_T^{Higgs}. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

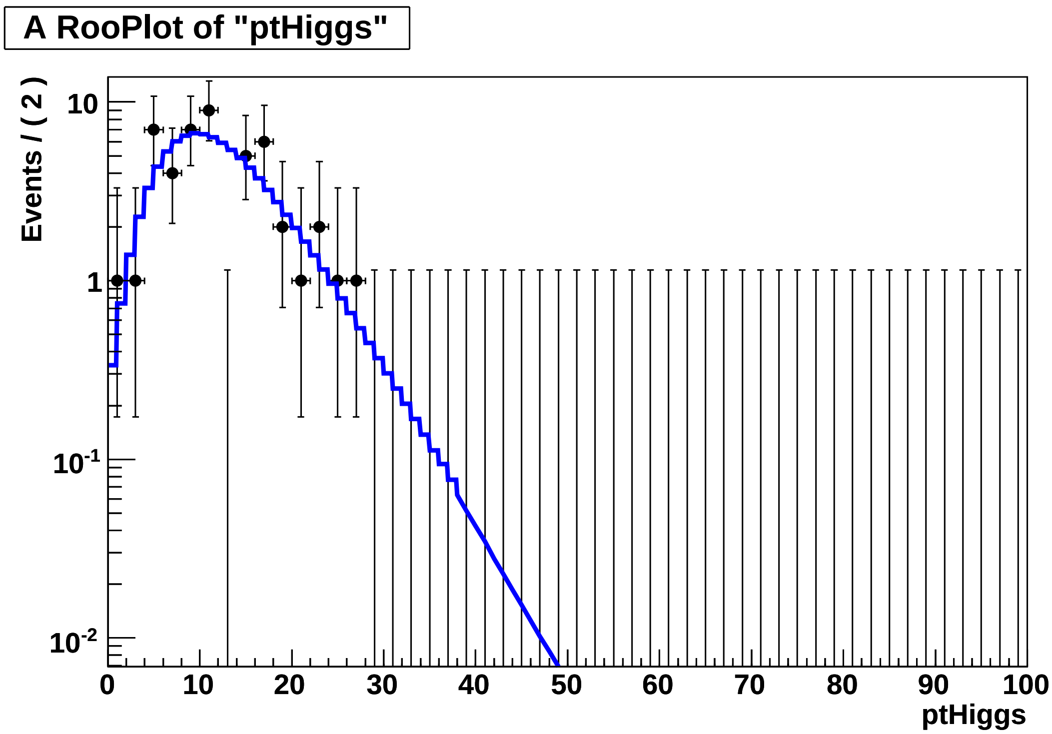

Figure A.76: The best fit of the functional form for the W+jets background to the (full simulation) Monte Carlo, in the region with small $\Delta\phi_{ll}$. Here, P_T^{Higgs} is shown in bins of M_T. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

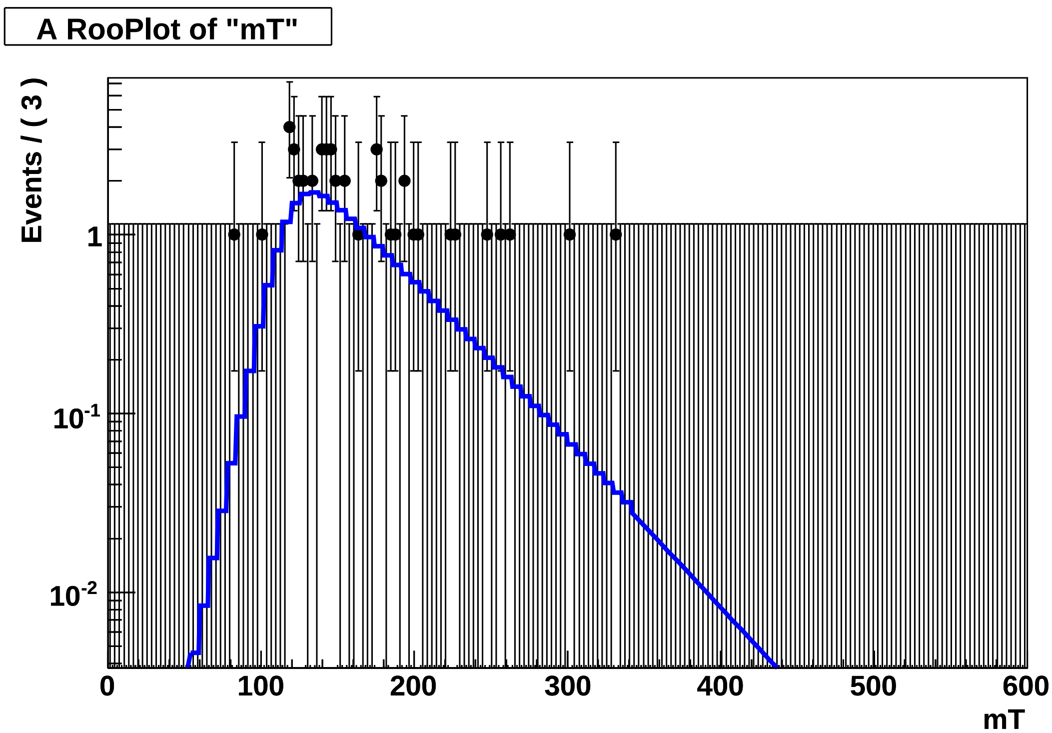

Figure A.77: The best fit of the functional form for the W+jets background to the (full simulation) Monte Carlo, in the region with large $\Delta\phi_{ll}$. Here, M_T is shown in bins of P_T^{Higgs}. Upper left: $0 < P_T < 10$ GeV. Upper right: $10 < P_T < 20$ GeV. Middle left: $20 < P_T < 30$ GeV. Middle right: $30 < P_T < 40$ GeV. Lower left: $40 < P_T < 50$ GeV. Lower right: $50 < P_T < 100$ GeV.

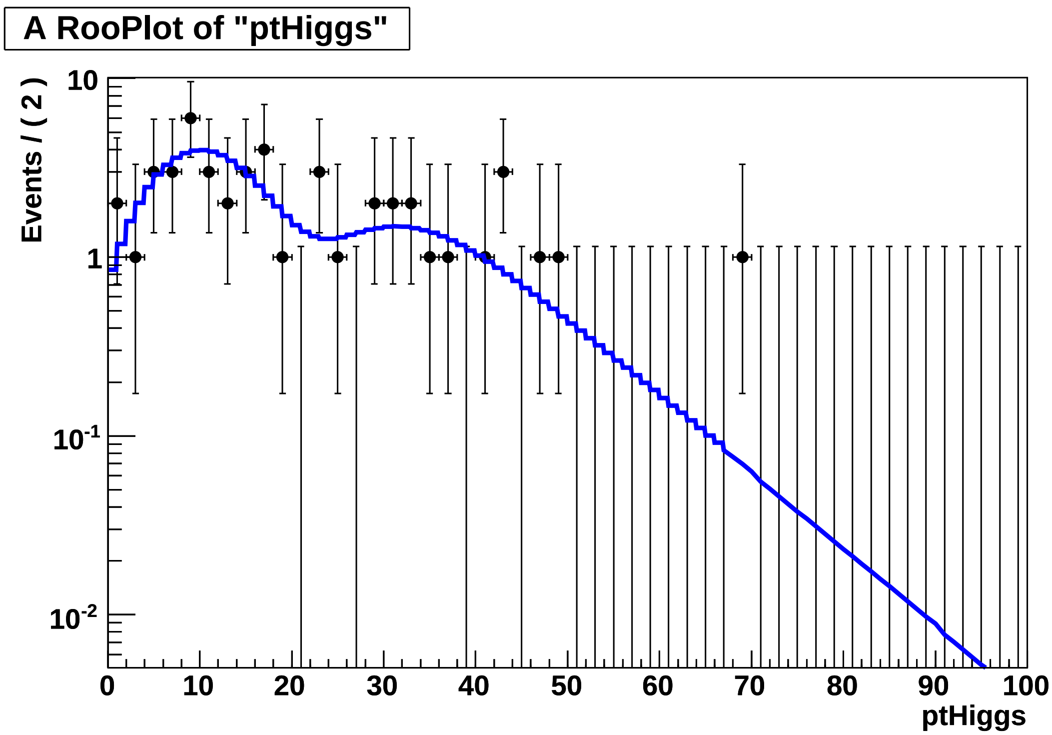

Figure A.78: The best fit of the functional form for W+jets background to the (full simulation) Monte Carlo, in the region with large $\Delta\phi_{ll}$. Here, P_T^{Higgs} is shown in bins of M_T. Upper left: $0 < M_T < 120$ GeV. Upper right: $120 < M_T < 160$ GeV. Lower left: $160 < M_T < 200$ GeV. Lower right: $200 < M_T < 1000$ GeV.

Par.	Small $\Delta\phi_{ll}$	Large $\Delta\phi_{ll}$
m_1	0.02	2.21926e-02
m_2	1.19050e-02	5.71598e-02
m_3	7.90040e-02	4.85430e-03
m_4	104.549	126.845
m_5	199.997	0
m_6	0.5	0.116366
m_7	33.1900	10.0333
p_1	0.2	9.66196e-02
p_2	9.67113e	12.5119
p_3	5.43865e	8.82230
p_4	1.65367e	21.5901
p_5	29.6031	8.35656e
p_6	8.37372e-06	0.717697

Table A.6: The values of the W+jets background parameters used as the truth in the region with small $\Delta\phi_{ll}$ and large $\Delta\phi_{ll}$.

was defined. The event selection is the same as the one used in the rest of the analysis, but electrons are only required to pass a loose (0x7) IsEM requirement instead of the usual "tight" (0xfff) IsEM cut. This selection was applied to a large sample of Pythia W+jets events from ATLAS central production, Datasets 5104-5105. The functional form was fit to the surviving events, and the resulting shapes were normalized to the cross-section obtained with the Alpgen W+jets samples. (Alpgen was used to determine the normalization for consistency with the WW+2j analysis.) Figures A.75 through A.78 show the resulting W+jets background shapes, and Table A.6 summarizes the values of the various parameters.

Appendix B

Functional Forms for the $H + 2j$ $(WW \to e\nu\mu\nu)$ Fit

This appendix documents the functional forms used in the Likelihood function. Section B.1 describes the signal model, and Section B.2 describes the background model. For simplicity, overall normalization factors are ignored. These are computed numerically from a sum; the distributions actually used are 2-dimensional histograms (100 bins from 0 to 1 in Neural Network output and 100 bins from 0 to 600 GeV in transverse mass) with the contents of each bin given by evaluating the functions described in this section at the "bottom" corner of the bin (i.e. the corner with small M_T and small Neural Network output).

B.1 Signal Model

The signal model is an uncorrelated product PDF. The transverse mass is modelled by a line times a gradual step function, exponentiated and cut off by a Fermi-dirac term at low M_T:

$$\frac{e^{-\sqrt{((M_T-m_3)(m_1+\frac{m_2}{1+e^{10(M_T-m_3)}}))^2+m_4}}}{1+e^{m_5(M_T-m_6)))}} \tag{B.1}$$

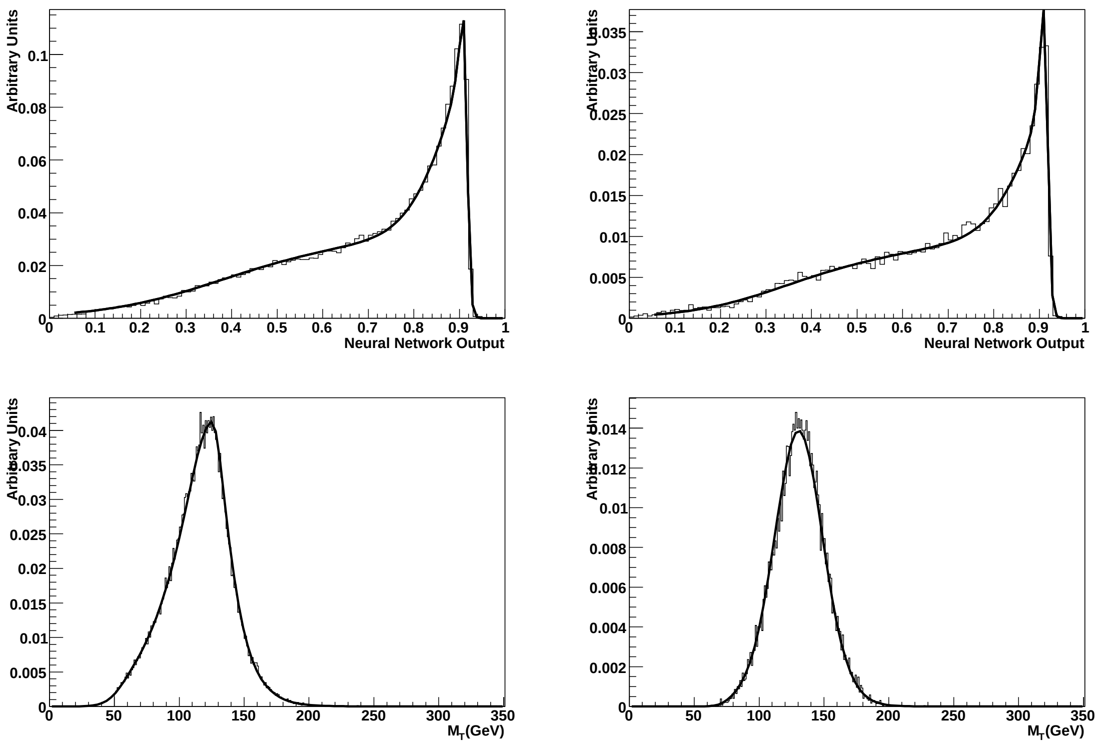

Figure B.1: Upper left: the distribution of the Neural Network output in the signal box, for signal with a Higgs mass of 130 GeV. Upper right: the Neural Network output distribution in the control region. Lower left and lower right: the corresponding M_T distributions.

The Neural Network output distribution is modelled as a rising exponential plus a double Gaussian centered at 1, cut off at either end of the range by Fermi-Dirac functions:

$$\frac{e^{p_1 X} + |p_5|\mathrm{Gaus}(X, 1, p_2) + |p_6|\mathrm{Gaus}(X, 1, p_7)}{(1 + e^{p_3(X-p_4)})(1 + e^{-|p_8|(X-p_9)})} \tag{B.2}$$

where X is the Neural Network output.

Because there are so many parameters in these descriptions of the signal shape, they are not all allowed to float in the fit. Instead, they are parameterized as functions of the Higgs mass. The parameterizations used for signal in the signal box are:

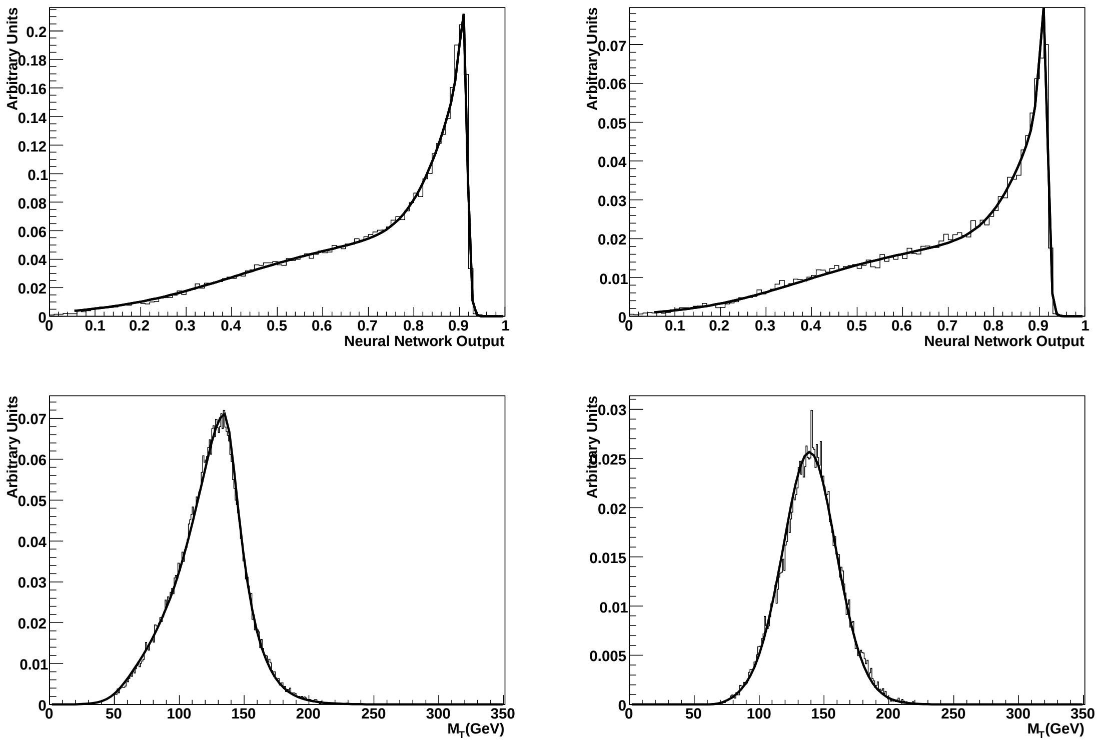

Figure B.2: Upper left: the distribution of the Neural Network output in the signal box, for signal with a Higgs mass of 140 GeV. Upper right: the Neural Network output distribution in the control region. Lower left and lower right: the corresponding M_T distributions.

- $m_1 = 0.0159078\ \text{Gaus}(M_H, 166.764, 10.7134) + 0.229062 - (1.69815 \times 10^{-3})M_H + (4.15806 \times 10^{-6})M_H^2$

- $m_2 = 0.1463783 - (2.583270 \times 10^{-3})M_H + (8.966613 \times 10^{-6})M_H^2$

- $m_3 = -89.38117 + 2.269337M_H - (4.759316 \times 10^{-3})M_H^2$

- $m_4 = -17.37645 + 0.6750543M_H - (6.872558 \times 10^{-3})M_H^2 + (2.108216 \times 10^{-5})M_H^3$

- $m_5 = -1.099018 + 0.01178982M_H - (3.433610 \times 10^{-5})M_H^2$

- $m_6 = -36.32650 + 1.066344M_H - (3.066954 \times 10^{-3})M_H^2$

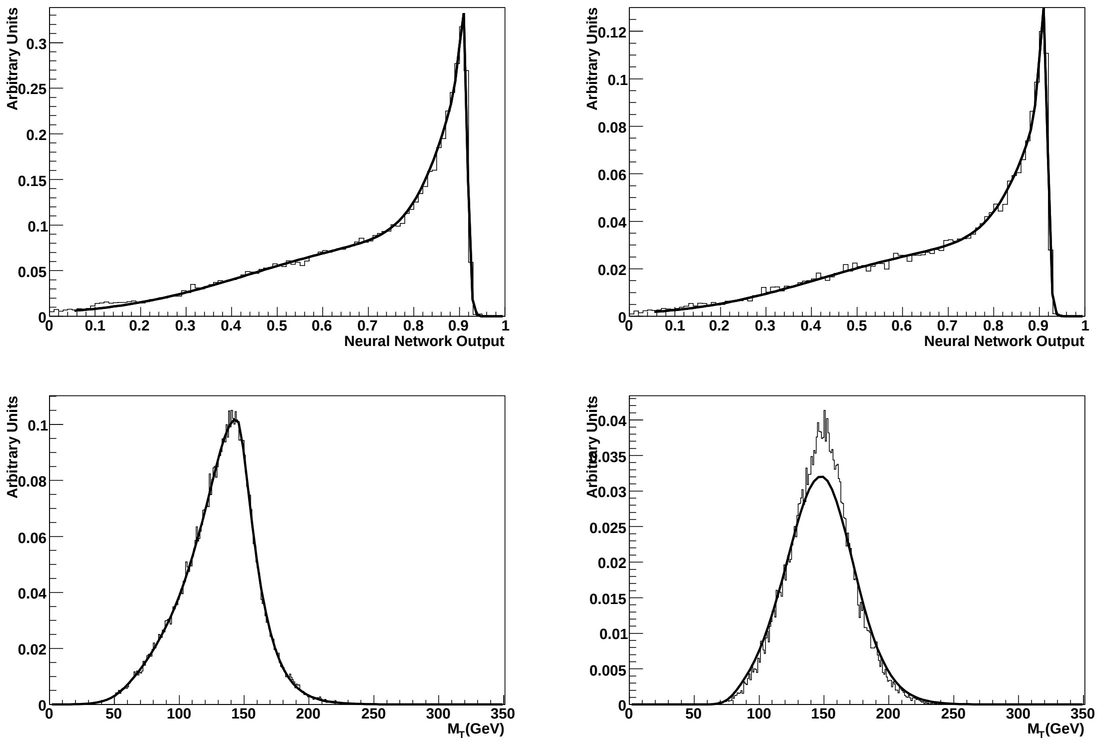

Figure B.3: Upper left: the distribution of the Neural Network output in the signal box, for signal with a Higgs mass of 150 GeV. Upper right: the Neural Network output distribution in the control region. Lower left and lower right: the corresponding M_T distributions.

- $p_1 = 0.07908943 + 3.772234 \times 10^{-3} M_H$
- $p_2 = 0.2900676 - 4.599199 \times 10^{-3} M_H + 2.690864 \times 10^{-5} M_H^2 - 5.219334 \times 10^{-8} M_H^3$
- $p_3 = 339.8866 - 0.2519229 M_H$
- $p_4 = 0.9176503 - 1.197831 \times 10^{-5} M_H$
- $p_5 = -1687.262 + 20.12875 M_H - 0.04996809 M_H^2$
- $p_6 = 1.717439 + 0.02079167 M_H$
- $p_7 = 0.1250334 - 1.670046 \times 10^{-4} M_H$

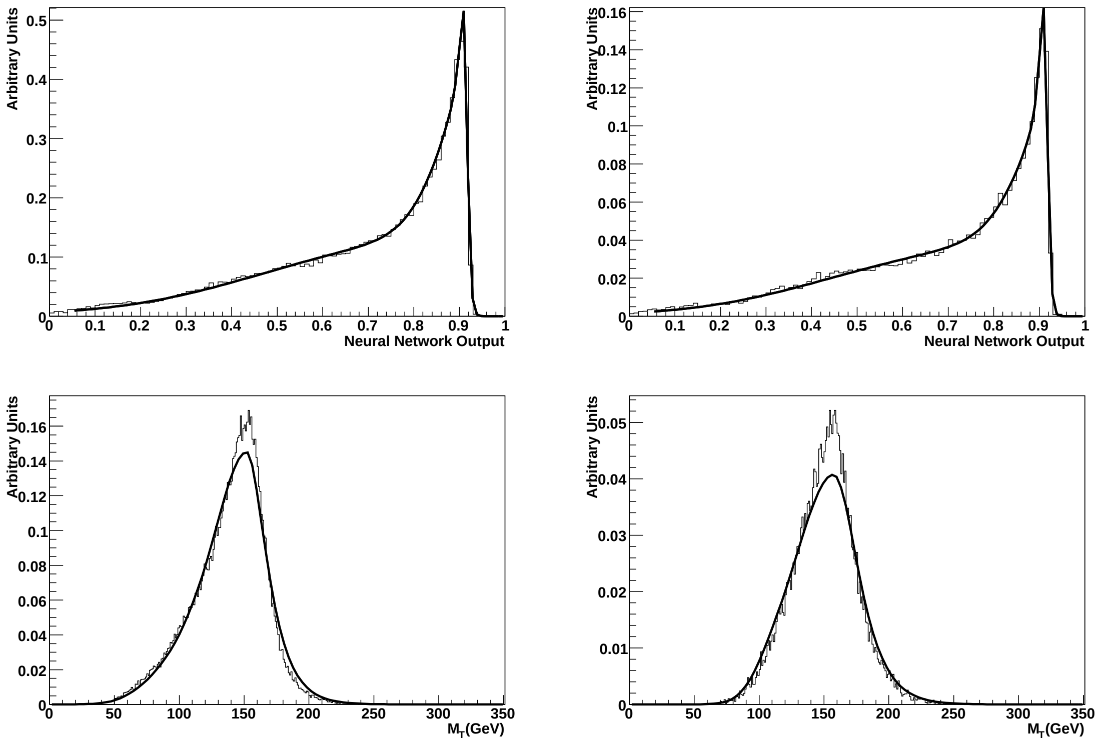

Figure B.4: Upper left: the distribution of the Neural Network output in the signal box, for signal with a Higgs mass of 160 GeV. Upper right: the Neural Network output distribution in the control region. Lower left and lower right: the corresponding M_T distributions.

- $p_8 = 13.98680 - 0.04827293 M_H$
- $p_9 = 0.1066077 + 1.783891 \times 10^{-3} M_H$

The parameterizations for the signal model in the control region are:

- $m_1 = 0.239885 - (5.14617 \times 10^{-4})M_H - (1.37976 \times 10^{-6})M_H^2$

  $- 0.0571278 \text{ Gaus}(M_H, 156.834, 5.95814)$

- $m_2 = -0.363567 - (6.16791 \times 10^{-4})M_H + (8.69095 \times 10^{-6})M_H^2$

  $+ 0.145338 \text{ Gaus}(M_H, 157.001, 5.49402)$

- $m_3 = 81.98915 - 0.05030877M_H + (3.259596 \times 10^{-3})M_H^2$
- $m_4 = -631.845 + 8.73049M_H - 0.0264748M_H^2 - 85.0000\ \mathrm{Gaus}(M_H, 160, 8.45817)$
- $m_5 = -0.427266 + (5.79953 \times 10^{-4})M_H + (5.28324 \times 10^{-6})M_H^2$
  $+ 0.0931400\ \mathrm{Gaus}(M_H, 162.030, 6.02601)$
- $m_6 = 109.823 - 0.796375M_H + (3.67017 \times 10^{-3})M_H^2$
  $+ 15.5418\ \mathrm{Gaus}(M_H, 159.375, 5.88306)$
- $p_1 = 0.6769414 + (4.586457 \times 10^{-3})M_H$
- $p_2 = 0.04143522$
- $p_3 = 339.8866 - 0.2519229M_H$
- $p_4 = 0.9676503 - (1.197831 \times 10^{-5})M_H$
- $p_5 = -3.142953 + 0.1315060M_H$
- $p_6 = 1.717439 + 0.02079167M_H$
- $p_7 = 0.1397914 - (9.837363 \times 10^{-5})M_H$
- $p_8 = 25.34811 - 0.02524389M_H$
- $p_9 = 0.1962854 - (8.634453 \times 10^{-4})M_H + (2.217899 \times 10^{-6})M_H^2$

Figures B.1 through B.7 show comparisons of the signal model and the Monte Carlo in the signal box and control region for various Higgs masses. Generally, the parameterization works quite well, with the exception of the transverse mass distributions for a few points near the WW threshold.

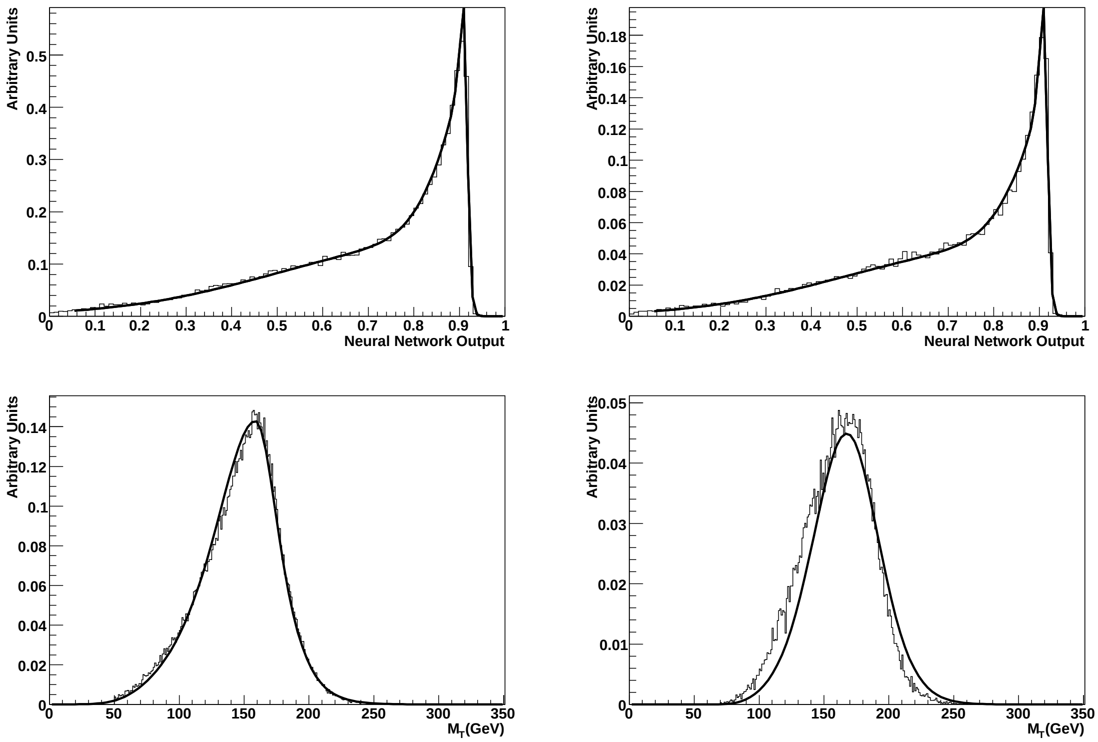

Figure B.5: Upper left: the distribution of the Neural Network output in the signal box, for signal with a Higgs mass of 170 GeV. Upper right: the Neural Network output distribution in the control region. Lower left and lower right: the corresponding M_T distributions.

The ratio of signal cross-sections in the signal box and the contol region is taken from Monte Carlo and parameterized as a function of M_H. $R = \sigma(control)/\sigma(sigbox)$ is parameterized as:

$$R = \frac{(a_1 - a_2 M_H + a_3 M_H^2 - a_4 M_H^3)}{(1 + e^{0.4\ (M_H - 160)})} + \frac{(a_6 - a_7 M_H + a_8 M_H^2 - a_9 M_H^3)}{(1 + e^{0.4(160 - M_H)})} \tag{B.3}$$

where $a_1 = 6.234552$, $a_2 = 0.1594717$, $a_3 = 1.358589 \times 10^{-3}$, $a_4 = 3.711165 \times 10^{-6}$, $a_5 = 36.79881$, $a_7 = 0.5916528$, $a_8 = 3.127904 \times 10^{-3}$, and $a_9 = 5.349667 \times 10^{-6}$. Figure B.8 shows a comparison between this model and the actual values of the ratio obtained from fast simulation; the agreement is quite good.

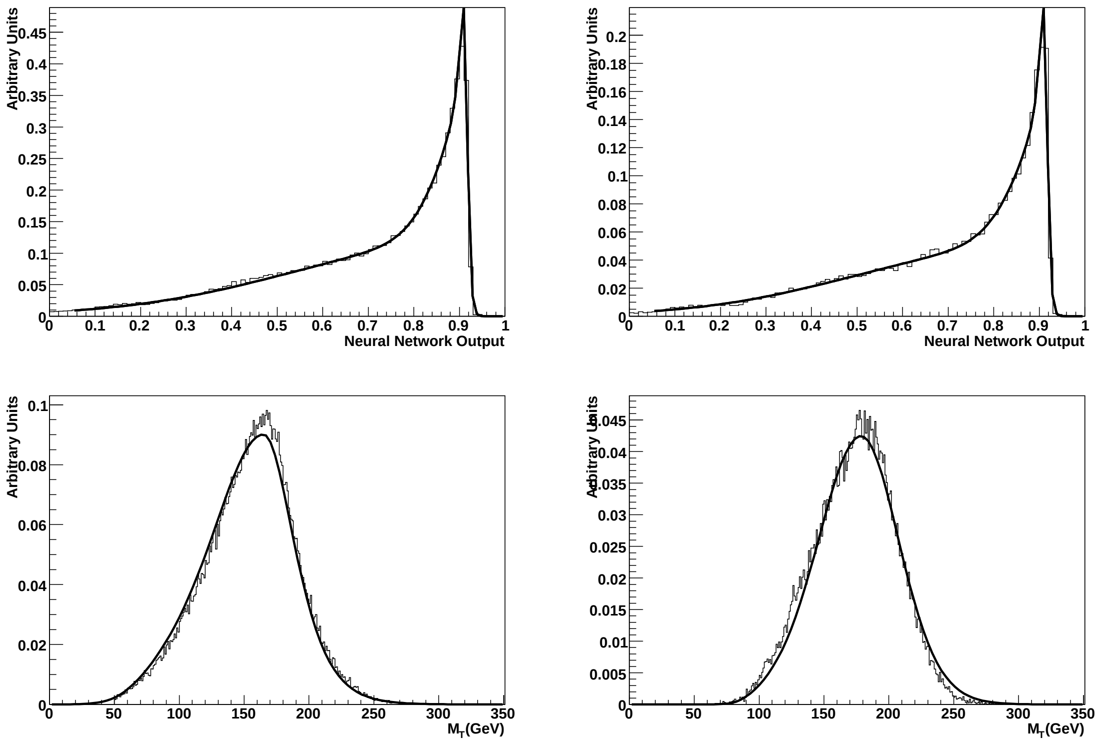

Figure B.6: Upper left: the distribution of the Neural Network output in the signal box, for signal with a Higgs mass of 180 GeV. Upper right: the Neural Network output distribution in the control region. Lower left and lower right: the corresponding M_T distributions.

Alternative signal models have been derived from fully simulated samples generated with Sherpa and Pythia. These models are used to generate toy Monte Carlo outcomes only, but not to fit pseudoexperimental outcomes. Table B.1 shows the best-fit values of the parameters for the two different samples in the signal box and the control region. Figures B.9 and B.9 show that the models reproduce the distributions observed in Monte Carlo.

Parameter	Pythia(Sigbox)	Pythia(Control)	Sherpa(Sigbox)	Sherpa(Control)
m_1	0.0765564	0.106990	0.0765564	0.159510
m_2	-0.037907	-8.88820$\times10^{-3}$	-0.037907	-0.0564399
m_3	161.892	162.282	161.892	194.445
m_4	2.69383	59.3338	2.69383	69.5343
m_5	-4.01285	-0.0870619	-4.01285	0.0511718
m_6	52.8153	56.3170	52.8153	146.089
p_1	0.313847	-9.11396	1.60061	0.342794
p_2	0.313847	0.482649	0.0769954	-0.0650703
p_3	187.871	187.566	409.798	409.798
p_4	0.913285	0.912947	0.922237	0.922878
p_5	2.01023	1.51477	12.6131	-20.6553
p_6	12.2649	17.5214	4.40098	5.28657
p_7	0.0676128	0.0581133	0.216724	0.216724
p_8	7.88538$\times10^{-3}$	8.301$\times10^{-5}$	4.79856$\times10^{-4}$	14.9136
p_9	-0.166285	19.632	1.36669	0.0672049
p_{10}	1.01627	0.276417	0	0
p_{11}	61.7928	11.2424	0	0

Table B.1: The true values of the fit parameters for the fully simulated Pythia and Sherpa signal samples in the signal box ("Sigbox") and the control region ("Control"). These samples have only been studied for $M_H = 170$ GeV.

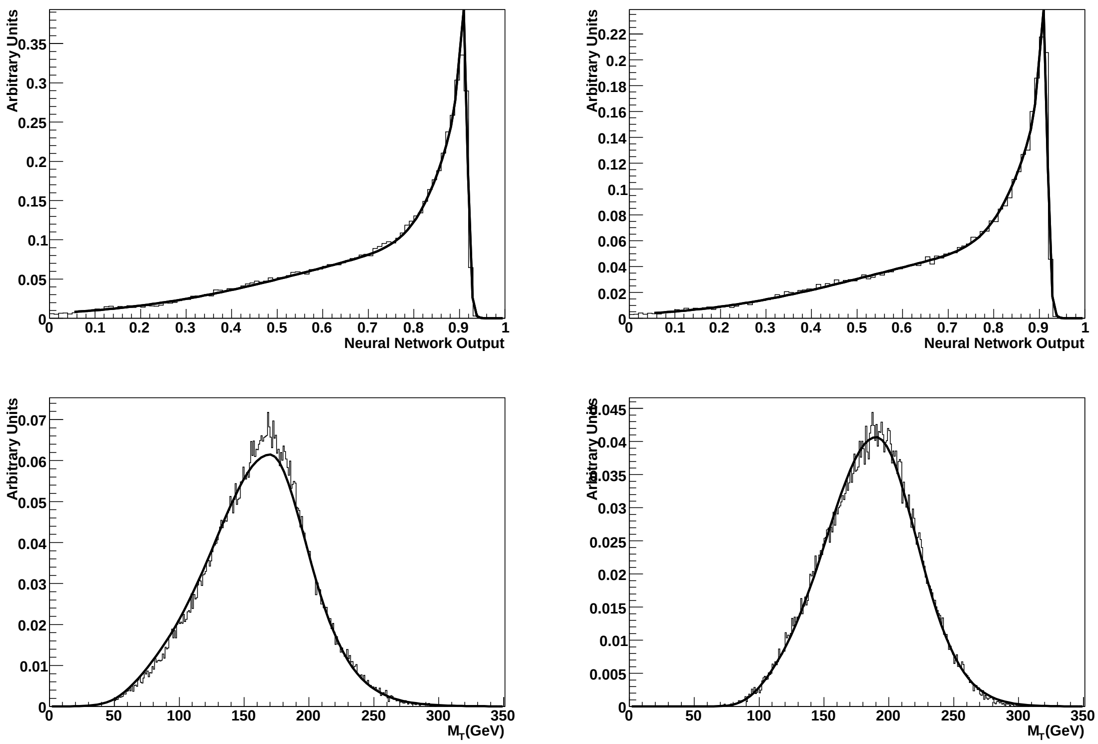

Figure B.7: Upper left: the distribution of the Neural Network output in the signal box, for signal with a Higgs mass of 190 GeV. Upper right: the Neural Network output distribution in the control region. Lower left and lower right: the corresponding M_T distributions.

B.2 Model for Irreducible Backgrounds

Like the signal model, the model for the irreducible backgrounds (namely $t\bar{t}$ and WW+jets) is an uncorrelated product PDF. The transverse mass distribution of the background is modeled as a sum of two exponentials modulated by a product of two Fermi-Dirac cutoffs at low M_T:

$$\frac{e^{-|m_1|M_T} + |m_5|e^{-(|m_2|+|m_1|)M_T}}{(1+e^{-m_3(M_T-m_4)})(1+e^{-(|m_6|+|m_3|)(M_T-m_4-|m_7|)})} \tag{B.4}$$

The functional form used to model the Neural Network output distribution in the control region consists of a $1+xe^x$ component cut off by a Fermi-Dirac function at

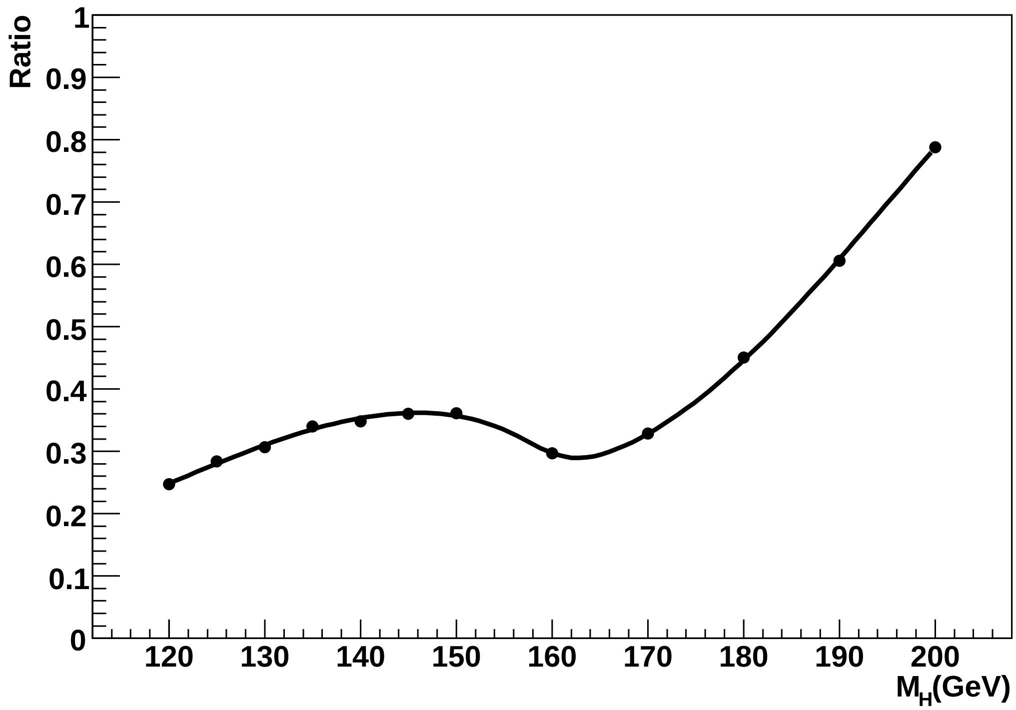

Figure B.8: The ratio of the signal cross-sections in the signal box and the control region as a function of mass. The points show the values obtained from fast simulation, and the curve shows the parameterization used in the fit.

large Neural Network output plus a Gaussian to model the bump in the distribution for Neural Network outputs around 0.3, plus a $1/x$ term to model the spike at zero, plus a Gaussian centered at zero:

$$\frac{(p_1 X + p_9)e^{p_8 X}}{1 + e^{|p_2|(X - p_6)}} + p_3 \mathrm{Gaus}(X, p_4, p_5) + \frac{p_7}{X} + p_{10} \mathrm{Gaus}(X, 0, p_{11}) \tag{B.5}$$

In the signal box, the model used for the truth comes from an independent fit of the same functional form to the Monte Carlo, but in fits to real data or pseudodata, the values of the parameters are constrained to be the same as in the control region, but

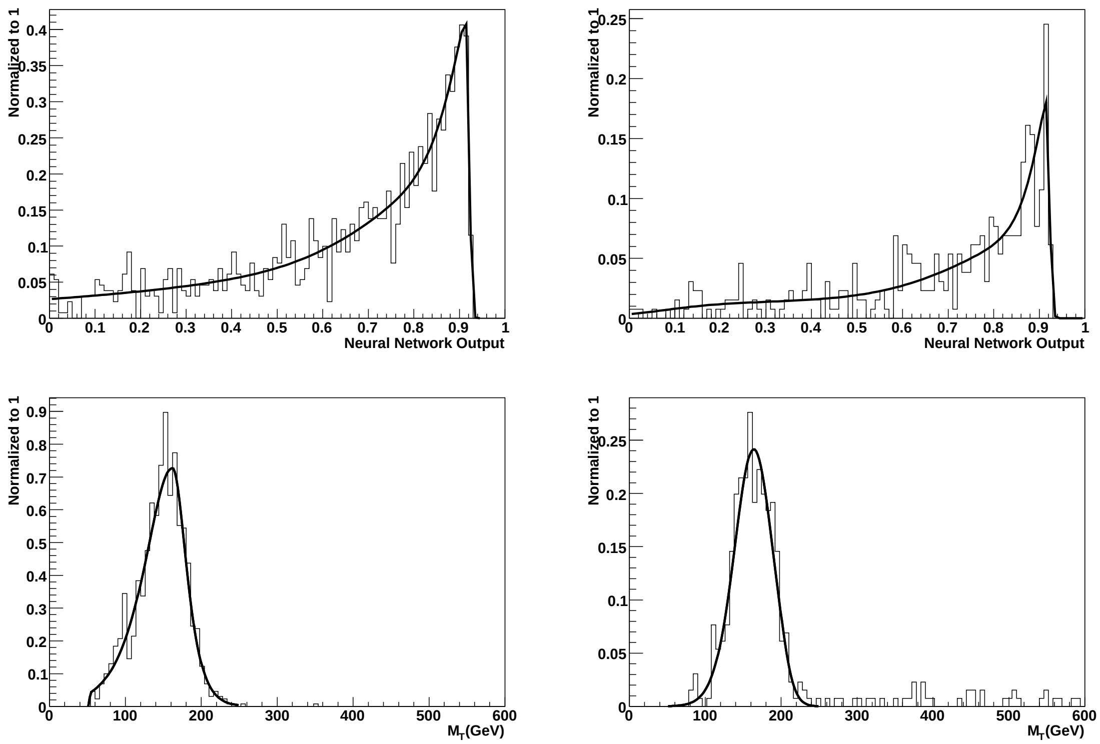

Figure B.9: Upper left: the distribution of the Neural Network output in the signal box, for the fully simulated Sherpa signal sample. Upper right: the Neural Network output distribution in the control region. Lower left and lower right: the corresponding M_T distributions.

multiplied by a linear extrapolation factor:

$$\left(\frac{(p_1X+p_9)e^{p_8X}}{1+e^{|p_2|(X-p_6)}}+p_3\mathrm{Gaus}(X,p_4,p_5)+\frac{p_7}{X}+p_{10}\mathrm{Gaus}(X,0,p_{11})\right)(1+p_{12}X) \quad \text{(B.6)}$$

Table B.2 shows the parameter values used as truth. Figure B.11 shows a comparison between this background model and the sum of all backgrounds using fast simulation Monte Carlo; the agreement is good.

Parameter	Signal Box	Control Region
p_1	2.51420	0.678445
p_2	2188.91	2267.87
p_3	0.172513	0.629091
p_4	0.371385	0.369532
p_5	-0.0938936	0.137029
p_6	0.915134	0.923950
p_7	0.0283267	0.0325936
p_8	-2.49538	-2.77589
p_9	2.67859×10^{-3}	0.772266
p_{10}	0.753330	2.13944
p_{11}	0.118378	0.0868435
m_1	0.0112476	7.36309×10^{-3}
m_2	0.0172139	0
m_3	0.0492454	0.0370160
m_4	160.914	165.985
m_5	317.563	0
m_6	0.0157362	0.0699518
m_7	93.9404	72.1067

Table B.2: The true values of the fit parameters for the irreducible background in the $H + 2j$ (dilepton) analysis.

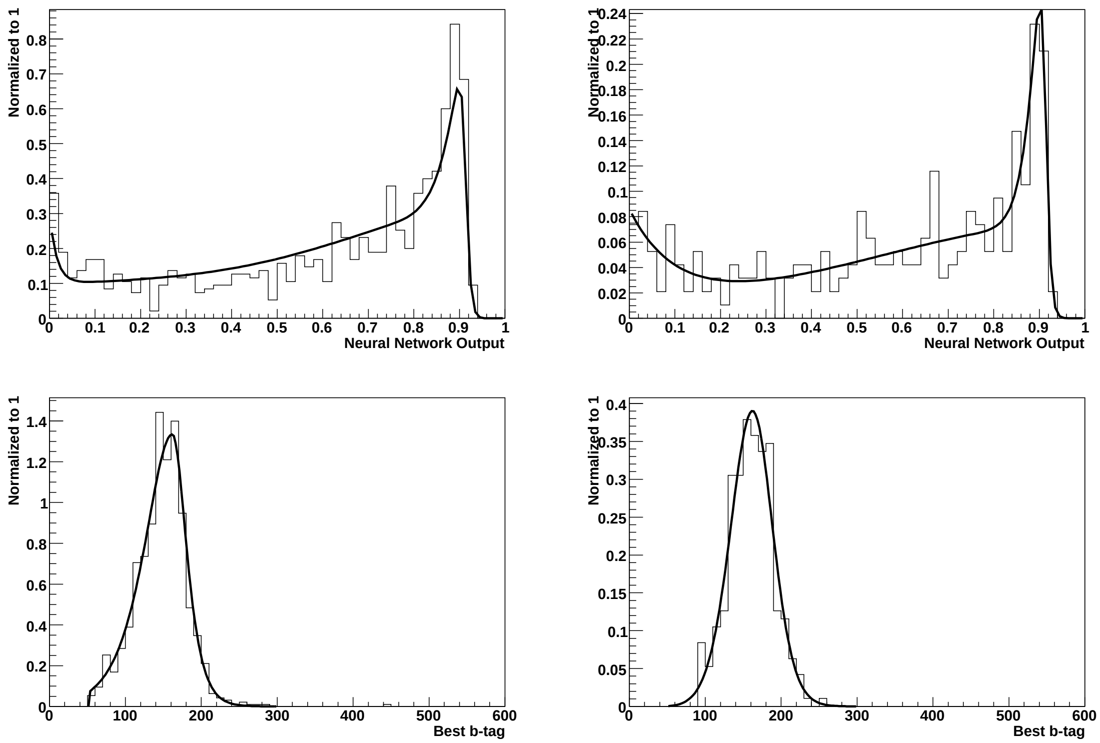

Figure B.10: Upper left: the distribution of the Neural Network output in the signal box, for the fully simulated Pythia signal sample. Upper right: the Neural Network output distribution in the control region. Lower left and lower right: the corresponding M_T distributions.

B.3 Model for Fake Backgrounds

The model used for the fake background from W+jets is the same functional form as the model used for the irreducible background, except that the xe^x term and the Gaussian centered at zero are switched off. The model is fit to a sample of W+jets Monte Carlo events which pass an event selection that is loosened with respect to the main analysis cuts. In particular, the requirement on electron ID is loosened from "medium" to "loose" electrons, and no cuts are applied on calorimeter isolation or tracking isolation variables. However, the rest of the kinematic cuts of the analysis

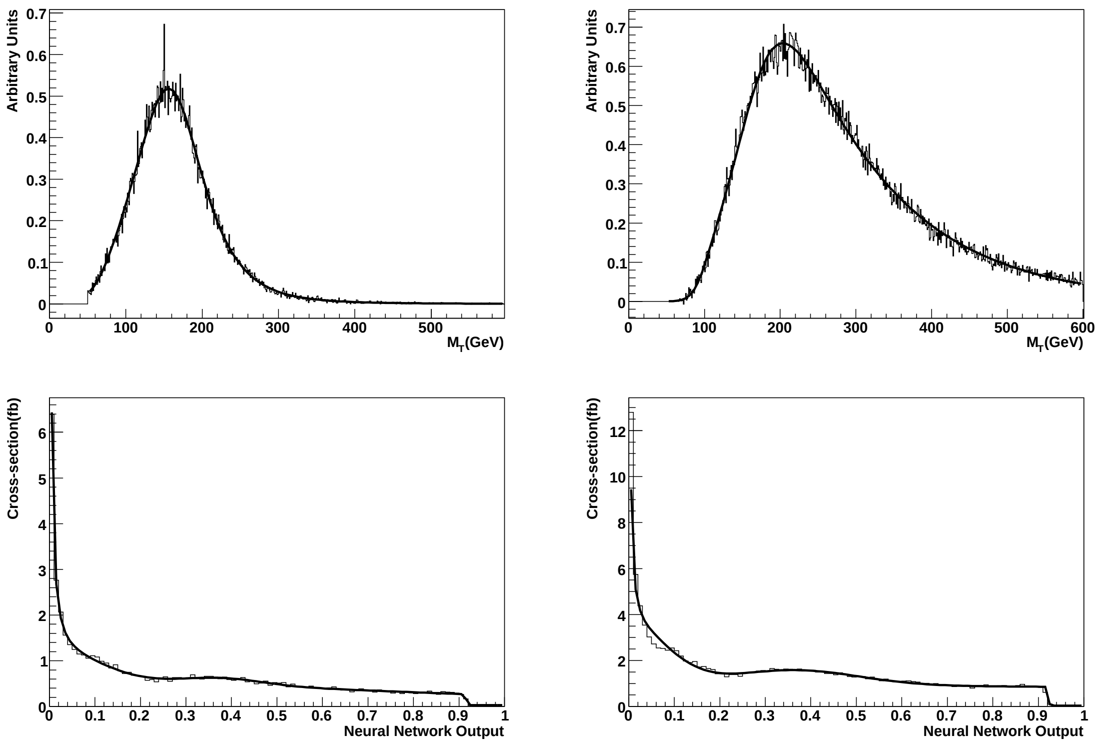

Figure B.11: Upper left: the distribution of the transverse mass for background in the signal box. Upper right: the distribution of transverse mass in the control region. Lower left: the distribution of the Neural Network output for background in the signal box. Lower right: the distribution of the Neural Network output in the control region.

are applied as usual. It is assumed that the Neural Network output shape is the same in the signal box and the control region. This assumption is necessary because of the poor statistics of the W+jets Monte Carlo sample.

Table B.3 shows the values of the shape parameters used to model the W+jets background. Figure B.12 shows a comparison of the model with the available fully simulated Monte Carlo. The agreement is good to within statistics. These parameters are used for generating toy Monte Carlo outcomes, and they are also used in the fit. Both the shape and normalization of the W+jets background is assumed to be

Parameter	Signal Box	Control Region
p_1	0	0
p_2	1985.15	1985.15
p_3	74.754	74.754
p_4	0.444006	0.444006
p_5	0.0761708	0.0761708
p_6	0.94	0.94
p_7	0	0
p_8	0	0
p_9	14.4676	14.4676
p_{10}	0	0
p_{11}	0	0
m_1	8.71303×10^{-3}	-7.34639×10^{-11}
m_2	0.0383162	8.35824×10^{-3}
m_3	-0.184227	0.0766223
m_4	83.2653	115.497
m_5	5340.58	75.0042
m_6	0.202217	4.20732×10^{-7}
m_7	128.338	0.171923

Table B.3: The true values of the fit parameters for the fake background in the $H+2j$ (dilepton) analysis.

well-predicted in the present study; however, it is intended that a future study will investigate ways to normalize this background and understand its shape using control samples.

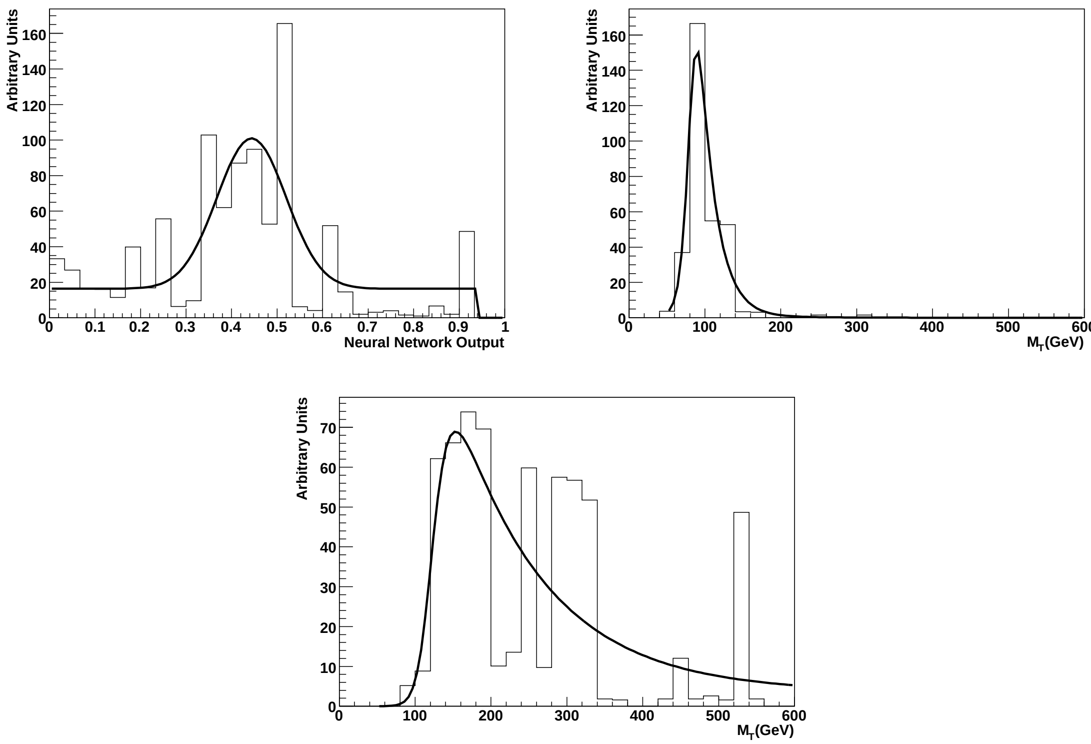

Figure B.12: Upper left: the distribution of the Neural Network output for the W+jets background, summed over the signal box and the control region. Upper right: the transverse mass distribution for the W+jets background in the signal box. Bottom: the transverse mass distribution for W+jets in the control region.