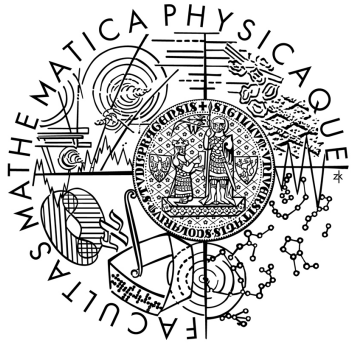
Charles University in Prague
Faculty of Mathematics and Physics

MASTER THESIS



Bc. Peter Berta

Study of Kinematics of Top Quark Pairs at the Atlas Experiment

Institute of Particle and Nuclear Physics

Supervisor of the master thesis: Mgr. Jiří Kvita Ph.D.

Study programme: Physics

Specialization: FJF

Prague 2012

I would like to thank to my supervisor Jiří Kvita for his guidance, support and physics insight. He enabled me to work within the Top Differential Cross Section Group at ATLAS. He thoroughly replied to all of my 391 e-mails. Last but not least, he helped me a lot with my English. I am very grateful for all of this.

I thank to the directorate and employees of IPNP institute in Prague for their management of grants which kindly supported my travels to CERN to work on this diploma thesis.

Many thanks belong to my dear parents, brother Milan and my girlfriend Ruženka for their support.

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In Prague, on 12.4.2012 Bc. Peter Berta

Název práce: Studium kinematiky rozpadů top-antitopových párů v experimentu Atlas

Autor: Bc. Peter Berta

Ústav: Ústav částicové a jaderné fyziky

Vedoucí diplomové práce: Mgr. Jiří Kvita Ph.D., Ústav částicové a jaderné fyziky

Abstrakt: V předložené práci se zabývám měřením spektra rozdělení impulzů top kvarků produkovaných přes top-antitopové páry v proton-protonových srážkách na urychlovači LHC s těžišťovou energií 7 TeV. Analýza je provedena v rámci experimentu ATLAS. V rozpadovém kanále s jedním leptonem jsem provedl studie na simulovaných událostech, které byly potřebné k obdržení výsledného spektra ze skutečných dat. Popisuji základy výběru top-antitopových eventů z důvodu přítomnosti velkého pozadí. Zkoumám efektivitu zrekonstruování top-antitopových párů pomocí detekce a asociace produktů rozpadů s kinematickou hypotézou top-antitopových párů. Studuji dekonvoluci změřeného spektra, což převážně odstraní efekty rozlišení. Na závěr ukazuji změřené spektrum rozdělení impulzů top kvarků obdržené v rámci mé analýzy.

Klíčová slova: LHC, top-antitopové páry, výběr eventů, dekonvoluce.

Title: Study of Kinematics of Top Quark Pairs at the Atlas Experiment

Author: Bc. Peter Berta

Department: Institute of Particle and Nuclear Physics

Supervisor: Mgr. Jiří Kvita Ph.D., Institute of Particle and Nuclear Physics

Abstract: In this thesis, I deal with the measurement of the transverse momentum spectrum of the top quark produced in top-antitop pairs at the LHC at center of mass energy 7 TeV. The analysis is carried out within the ATLAS collaboration. In the single lepton decay channel, I have performed studies on the simulation which were necessary to obtain the final spectrum from real data. I describe basic event selection rules to reduce background events. I study the efficiency of top-antitop pairs reconstruction. I study the unfolding of the measured spectrum which corrects for effects caused by imperfect resolutions. At the end, I show the measured top quark transverse momentum spectrum obtained from my analysis.

Keywords: LHC, top-antitop pairs, event selection, unfolding

Contents

Introduction

The main topic of this thesis is the measurement of the transverse momentum spectrum of the top quark produced in top-antitop pairs at the LHC at center of mass energy 7 TeV.

In the first chapter, basic facts about top quark are summarized. Using laws of kinematics, I perform simple computations regarding the production of top-antitop pairs and their decay. I show the theoretical predictions for differential cross section spectra for top-antitop production.

In next chapters, I briefly describe the ATLAS detector and its performance. Further, I am describing the simulation of events, object selection and event selection regarding any analysis concerning the top-antitop pairs decaying in single lepton channel. I performed minor studies on simulated events which help to understood the basic concepts of the event selection.

The last chapter comprises the measurement aimed at the transverse momentum spectrum of the top quark. The analysis flow is performed by me with the help of the package written by people from Top Differential Cross Section Group at ATLAS. I have also developed this package and I am describing my main contributions to the analysis. I performed studies regarding performance of the top-antitop pair reconstruction and unfolding of the measured spectrum. At the end, I am showing my measured spectrum from data as a proof of concept for this measurement as the analysis would require more detailed studies concerning systematics in order to make a more serious comparison to theory.

In this thesis, the natural system of units is used [1, p. 13] in which the reduced Planck constant and the speed of light are equal to 1. For the analysis, I used the data analysis framework ROOT [2], and all unreferenced figures were produced by me using this framework.

1. Basic Facts on the Top Quark

The top quark is the heaviest elementary fermion within the Standard Model of elementary particle physics. With the bottom quark, it belongs to the third generation of quarks, which was predicted by M. Kobayashi and T. Maskawa in 1973. The top quark is the only quark that does not form bound states, because of its extremely short lifetime. The measurement of top-antitop ($t\bar{t}$) pair production probes our understanding of the strong interactions and predictions from perturbative quantum chromodynamics, while the decay of top quarks and the production of single top quarks examine the electroweak interactions. By comparing measured and predicted properties of the top quark, one can search for new physics beyond the Standard Model.

The top quark was experimentally discovered in 1995 at the Tevatron collider in FNAL, Chicago, which had run in years 1983-2011. Large amount of data has been produced and analyzed, a review of the corresponding measurements can be found in [3], while the data are still being analyzed. Nowadays, the only place where the top quarks are produced and detected is the Large Hadron Collider (LHC) in CERN. It operates since 2010 and until the end of year 2011 much more top quark events were produced than at the Tevatron collider. At the LHC, two multipurpose detectors can efficiently detect the $t\bar{t}$ events: ATLAS and CMS.

1.1 Top Quark in the Standard Model

The Standard Model (SM) is a quantum field theory based on the local gauge symmetries $SU(3)_{QCD} \times SU(2)_L \times U(1)_Y$. The elements of this theory are three generations of spin-1/2 fermions (six quarks and six leptons, see Table 1.1), four spin-1 gauge bosons (Table 1.2) and a spin-0 Higgs boson H (until now unobserved). Each of these elementary particles has its charge conjugate partner (antiparticle), although certain particles are identical with their antiparticles.

The SM consists of two separate quantum field theories: Quantum Chromodynamics (QCD) and Electroweak Theory (EWT). The QCD is based on the gauge group $SU(3)_{QCD}$ and it describes the strong interactions of quarks via gluons. It also incorporates confinement of quarks in hadrons. The EWT is based on the gauge group $SU(2)_L \times U(1)_Y$ and it unifies the electromagnetic and weak interaction into the electroweak interaction. A part of the EWT consists of generation of masses via the Higgs mechanism. An introduction to SM can be found in [1].

In particle physics, there are many unsolved theoretical questions, but until now, there is no experimental result, which is inconsistent with the SM.

Table 1.1: The known fundamental fermions.

fermion type	charge [elem. charge]	generation		
		1.	2.	3.
quark (q)	2/3	up (u)	charm (c)	top (t)
	$-1/3$	down (d)	strange (s)	bottom (b)
lepton (ℓ)	1	electron (e)	muon (μ)	tau (τ)
	0	electron neutrino (ν_e)	muon neutrino (ν_μ)	tau neutrino (ν_τ)

Table 1.2: The gauge bosons.

gauge boson	elec. charge [elem. charge]	mass [GeV]	interaction type
photon (γ)	0	0	electromagnetic
W^- boson (W^-)	-1	80.4	weak
Z^0 boson (Z^0)	0	91.2	weak
gluon (g)	0	0	strong

1.1.1 Top Mass

It is important to define which theoretical framework is used when referring to quark masses, see [4, p. 115,583]. It is normally assumed that what is being measured is the pole mass $m^{\text{top}}_{\text{pole}}$ which is the real part of the pole in the top quark propagator, although this statement is not completely true (see [5]).

The mass of the top quark m^{top} was measured by Tevatron's two particle detectors, CDF and D0. The combination of several most precise measurements gives the value $m^{\text{top}} = (173.3 \pm 1.1)$ GeV, [6]. More precise measurement are expected at the LHC.

The top quark is at least 11 orders of magnitude heavier than the lightest elementary fermion and it is the only fermion which is heavier than the gauge bosons W and Z. This obscure mass hierarchy can provide further insights into the process of mass generation.

1.2 Production of Top Quarks in Proton-Proton Collisions at the LHC

The LHC [7] is the world's largest and highest-energy particle accelerator located $\sim$ 100 m underground at CERN. It is a circular collider with circumference $\sim$ 27 km which collides protons or lead ions. For top quark physics, only the proton-proton (pp) collisions are relevant. The protons in colliding beams have energy 3.5 TeV, therefore the center-of-mass energy of pp interactions is $\sqrt{S} = 7$ TeV.

The theoretical description and prediction of top quark production is complicated, because the colliding protons are composited of partons (quarks, antiquarks and gluons). This can be solved through the QCD factorization theorem [8] which divides the pp collision into two components by introducing a factorization scale μ_F^2 in the calculations. The first component is called universal long-distance phenomena and the second is called perturbatively calculable short-distance phenomena. Using this approach the partons inside proton can be considered quasi-free.

In hadron collisions, top quarks are created in top-antitop ($t\bar{t}$) pair production or in single top production. The $t\bar{t}$ pairs are produced through QCD processes (gluon-gluon fusion $gg \to t\bar{t}$ or quark-antiquark annihilation $q\bar{q} \to t\bar{t}$). In the single top production, top or antitop quarks are produced through electroweak processes ($q_1\bar{q}_2 \to t\bar{b}$ or $q_1 b \to q_2 t$) or through QCD-electroweak processes ($gb \to tW^-$). In my thesis, I will study only the $t\bar{t}$ pair production. The leading order Feynman diagrams for $t\bar{t}$ pair production are shown in the Figure 1.5. It should be stressed that in the diagram (a) there can be also a different propagator (γ, Z,

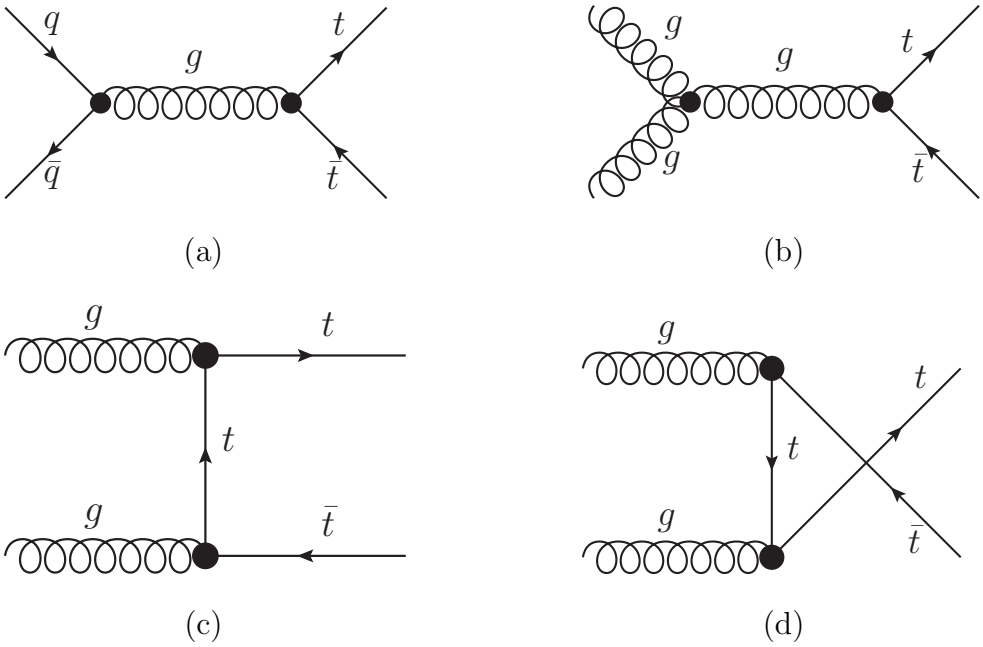


Figure 1.5: Leading order Feynman diagrams for $t\bar{t}$ production.

Higgs boson), but because the strong interaction has much higher contribution, the diagrams with other propagators can be neglected.

The partonic cross sections $\sigma_{gg\to t\bar{t}}$ and $\sigma_{q\bar{q}\to t\bar{t}}$ can be calculated using perturbative QCD by introducing a renormalization scale μ_R^2, along with the corresponding running constant $\alpha_s(\mu_R^2)$.

A certain type of parton A carries fraction x_A of proton's momentum. This fraction is a random variable with a density called the parton distribution function $f_{A/p}(x_A, \mu_F^2)$ (PDF), [4, p. 204]. I compute what should abide the fractions x_A and x_B of four-momenta P_A and P_B of colliding protons. For the square of center-of-mass energy s of the interacting parton system, it holds

$$s = (x_A P_A + x_B P_B)^2 \approx x_A x_B S, \tag{1.1}$$

when the proton mass is neglected. The kinematic condition for an interaction is that variable s cannot be smaller then the square of sum of product masses. Therefore for fractions x_A and x_B, it holds

$$x_A x_B \geq \frac{4\left(m^{\text{top}}\right)^2}{S} \approx 0.0025. \tag{1.2}$$

This condition does not take into account the nonzero top quark width, but is a good approximation. Using the factorization theorem, the $t\bar{t}$ production cross section $\sigma_{pp\to t\bar{t}X}$ is described via formula

$$\begin{aligned}\sigma_{pp\to t\bar{t}X} = &\int \sigma_{gg\to t\bar{t}}(s, \alpha_S(\mu_R^2), \mu_R^2, \mu_F^2, m^{\text{top}}) f_{g/p}(x_{g1}, \mu_F^2) f_{g/p}(x_{g2}, \mu_F^2) \mathrm{d}x_{g1} \mathrm{d}x_{g2} \\ &+ \sum_q \int \sigma_{q\bar{q}\to t\bar{t}}(s, \alpha_S(\mu_R^2), \mu_R^2, \mu_F^2, m^{\text{top}}) f_{q/p}(x_q, \mu_F^2) f_{\bar{q}/p}(x_{\bar{q}}, \mu_F^2) \mathrm{d}x_q \mathrm{d}x_{\bar{q}} \\ &+ \ldots,\end{aligned} \tag{1.3}$$

where the dots incorporate terms from $2 \longrightarrow 3$ processes. In this equation, the Condition 1.1 has to be met. The first term on the right-hand side corresponds to the $t\bar{t}$ production via gluon-gluon fusion, the sum corresponds to the $t\bar{t}$ production via quark-antiquark annihilation. By a glimpse at PDFs, one can observe that large amount of gluon pairs contained in colliding protons abide Condition 1.2 which results in $\sim 85\%$ fraction of the gluon-gluon fusion at the LHC. At the upgraded Tevatron energies, $x_A x_B > 0.0313$ which leads to only 15% fraction of gluon-gluon fusion.

The cross section 1.3 is a physical quantity and does not depend on the scales μ_R^2 and μ_F^2, if it is computed to all orders. It was calculated in the next-to-next-to-leading approximation: $\sigma = 163^{+7+9}_{-5-9}$ pb [9], assuming $m^{\text{top}}_{\text{pole}} = 173$ GeV. The uncertainties come from the uncertainties of PDFs and the dependence on scales μ_R^2 and μ_F^2. The theoretical prediction is in good agreement with the measured total cross section $\sigma_{\text{meas}} = 165.8 \pm 2.2(\text{stat.}) \pm 10.6(\text{syst.}) \pm 7.8(\text{lumi.})$ pb, [10].

1.3 Top Quark Decay

According to the SM, the top quark has three decay modes ($t \to W^+ d$, $t \to W^+ s$ and $t \to W^+ b$) with branching ratios related to the elements of the CKM matrix V_{td}, V_{ts} and V_{tb}, see [4, p. 150]. From the global fit in the SM one gets $V_{tb} > 0.999$ which implies that the channel $t \to W^+ b$ is dominant, with branching ratio almost 100%. The antitop quark $\bar{t}$ decays accordingly to the charge conjugation: $\bar{t} \to W^- \bar{b}$.

The decay width of the top quark Γ_t is predicted in the SM in next-to-leading-order: $\Gamma^{\text{top}} = 1.33$ GeV for $m^{\text{top}}_{\text{pole}} = 172.5$ GeV, [11]. The decay width was also determined at the Tevatron in D0 experiment: $\Gamma^{\text{top}}_{\text{meas}} = 2.00^{+0.47}_{-0.43}$ GeV, [12], which is the most precise measurement of top decay width by now.

The corresponding lifetime to the theoretical decay width is only $\sim 5.10^{-25}$ s, therefore the top quark is expected to decay instead of forming top-flavoured hadrons or $t\bar{t}$-quarkonium bound states, for details see [13].

Due to its large width Γ_t, no detector can directly detect a top quark, it is only possible to detect its decay products. The b-quark will produce a b-jet, more about jets can be found for example in [14]. The W boson has also very large width $\Gamma_W = (2.085 \pm 0.042)$ GeV, and therefore, it almost immediately decays. The $t\bar{t}$ pair signature in a detector depends on the decay mode of two W bosons from the decay $t\bar{t} \to W^- b W^+ \bar{b}$, see Table 1.3. The mean lifetime of μ is $\sim 2.2 \cdot 10^{-6}$ s which is long enough not to decay in a detector. The τ has mean lifetime $\sim 290 \cdot 10^{-15}$ s which means that only the decay products can be detected. What do different mean lifetimes mean in terms of passed range will be described in the next section and compared with the size of the detector.

There are three decay channels in which the basic properties of top quark and $t\bar{t}$ production are studied: all-hadronic, single lepton and dilepton channels, for both W, one W and no W decaying hadronicaly, respectively. Exact definitions of these channels depend on the classification of decay chains containing τ which depends on the concrete analysis method. Until Chapter 5, I use classification which is in Table 1.4. By now, lepton always means electron or muon. The W decaying into τ and consecutive leptonic decay of this τ has a similar signature as W decaying into leptons (not τ), and therefore, these two decay modes cannot be

efficiently separated. This is the reason why single lepton channel contains two different decay modes of a $t\bar{t}$ pair. The difference in these decay modes is that the mode $t\bar{t} \to W^- b W^+ \bar{b} \to 4 \text{ jets} + \ell\nu_\ell$ has in final state one neutrino while the second mode $t\bar{t} \to W^- b W^+ \bar{b} \to 4 \text{ jets} + \ell\nu_\ell \bar{\nu_\tau} \nu_\tau$ has in final state three neutrinos. The important knowledge is that these three neutrinos comes from the same W. Although the three neutrinos do not have zero invariant mass, the reconstruction can be similar as for the first decay mode, but less efficient as is described in Section 6.3.

Each channel has its own unique analysis outline. In my thesis, I will study only the single lepton channel (ℓ+jets channel). This channel can be very effectively distinguished from background and the $t\bar{t}$ pairs can be fully reconstructed.

Table 1.3: Main decay modes of W^+ with branching ratios and the next decays/evolution of W^+ products, [4, p. 420, 527].

	first decay	next decay / evolution
W^+	$q_1\bar{q}_2$ $(67.60 \pm 0.27)\%$	2 jets
	$e^+\nu_e$ $(10.75 \pm 0.13)\%$	$e^+\nu_e$
	$\mu^+\nu_\mu$ $(10.57 \pm 0.15)\%$	$\mu^+\nu_\mu$
	$\tau^+\nu_\tau$ $(11.25 \pm 0.20)\%$	$e^+\nu_e\bar{\nu_\tau}\nu_\tau$ $(17.85 \pm 0.05)\%$
		$\mu^+\nu_\mu\bar{\nu_\tau}\nu_\tau$ $(17.36 \pm 0.05)\%$
		hadrons $\bar{\nu_\tau}\nu_\tau$ 63.0%

Table 1.4: Decay channels of the $t\bar{t}$ pair. Lepton ℓ is e or μ. The charges are not indicated. MET is the missing transverse momentum.

Channel	one W	the other W	br. ratio	signature
all-hadronic	jets	jets	$\sim 45.7\%$	6 jets (at least 2 b-jets)
single lepton	jets	$l\nu_l$	$\sim 28.8\%$	ℓ, MET, 4 jets (at least 2 b-jet)
	jets	$\tau\nu_\tau \to l\nu_l\bar{\nu_\tau}\nu_\tau$	$\sim 5.4\%$	
dilepton	$l\nu_l$	$l\nu_l$	$\sim 4.5\%$	2ℓ, MET, 2 b-jets
	$l\nu_l$	$\tau\nu_\tau \to l\nu_l\bar{\nu_\tau}\nu_\tau$	$\sim 1.7\%$	
	$\tau\nu_\tau \to l\nu_l\bar{\nu_\tau}\nu_\tau$	$\tau\nu_\tau \to l\nu_l\bar{\nu_\tau}\nu_\tau$	$\sim 0.2\%$	

Soon after the b-quark is created through the top quark decay, the b-quark hadronises. The hadronisation of b-quarks is not a simple process and it is described using QCD models. During this process, the b-quark can radiate gluons and eventually forms a B-meson (B^-, B^0 or B_s^0). These mesons have masses around 5 GeV and mean lifetimes around $1.5 \cdot 10^{-12}$ s. Any motion of the b-quark during hadronisation cannot be distinguished by any detector. But the mean lifetime of the B-mesons inside the b-jet is long enough to pass a measurable distance. Therefore, the decay vertex is shifted w.r.t. primary interaction vertex. Detection of such displaced vertex is used to tag a jet as a b-jet which is described in Section 4.5.

1.4 Kinematics of $t\bar{t}$ Pairs at the LHC

I will now introduce the coordinate system and variables which will be used later. The origin of the coordinate system is at the interaction point of the pp collision

producing a $t\bar{t}$ event. The x-axis points towards the center of the LHC ring. The z-axis is parallel to the anti-clockwise beam viewed from above, and consequently the y-axis points in direction to the surface of the Earth. The azimuthal angle $\phi \in [-\pi, \pi)$, the polar angle $\theta \in [0, \pi)$ and the radius $r \in (0, \infty)$ are defined through transformation

$$x = r \sin\theta \cos\phi,$$
$$y = r \sin\theta \sin\phi,$$
$$z = r \cos\theta.$$

The pseudorapidity η is defined as $\eta \equiv -\ln\tan(\theta/2)$. Defined like this, pseudorapidity η has one important property for high-energetic particles. The difference of two pseudorapidities is invariant with respect to a boost along the z-axis. Transverse momentum p_T is the magnitude of the momentum $\vec{p}$ projected to the plane xy ($p_T = |\vec{p}|\sin\theta$).

The angular difference of two four-momenta can be expressed through the variable $\Delta R(\vec{p}_1, \vec{p}_2) \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2}$ where $\Delta\eta$ and $\Delta\phi$ are the difference of the rapidities and the azimuthal angles of two momenta $\vec{p}_1$ and $\vec{p}_2$, respectively. As both $\Delta\eta$ and $\Delta\phi$ are invariant with respect to a boost along the z-axis, also the variable ΔR is invariant. I will assume that momentum $\vec{p}_1$ with polar angle θ_1 is constant. To meet condition $\Delta R(\vec{p}_1, \vec{p}_2) < \Delta R_0$, where ΔR_0 is a constant, the second momentum $\vec{p}_2$ can point to certain solid angle $\Delta\Omega$. The size of this solid angle is a non-constant function of the polar angle θ_1. Also the solid angle $\Delta\Omega$ is not symmetrically distributed around momentum $\vec{p}_1$.

The colliding protons have zero p_{T}. A $t\bar{t}$ pair from interaction $pp \to t\bar{t}X$ can have nonzero p_{T} due to initial and final state radiation (ISR and FSR).

I compute what energies can take on the decay products of a $t\bar{t}$ pair in the laboratory reference frame (LRF), where the LHC ring and the detector are at rest. The upper and lower limits for $t\bar{t}$ products can be helpful in the rejection of certain solutions from the $t\bar{t}$ events reconstruction which is described in 6.3. I assume that the partons producing the $t\bar{t}$ pair have zero p_{T} and I neglect the ISR and FSR. I also neglect the non-zero decay width of particles. I use only energy and momentum conversation laws and the special relativity.

The energy E of a particle P with rest mass m in LRF can be obtained by a Lorentz transformation of energy E_{RF2} in reference frame 2 (RF2)

$$E = \frac{1}{\sqrt{1-\beta^2}} \left(E_{\mathrm{RF2}} + \beta \cos\alpha |\vec{p}_{\mathrm{RF2}}|\right), \tag{1.4}$$

where β is the size of velocity $\vec{\beta}$ of RF2 relatively to LRF, $\vec{p}_{\mathrm{RF2}}$ is the momentum of particle P in RF2 ($|\vec{p}_{\mathrm{RF2}}| = \sqrt{E_{\mathrm{RF2}}^2 - m^2}$) and α is the angle between velocity $\vec{\beta}$ and the momentum $\vec{p}_{\mathrm{RF2}}$. It is obvious that extremal values of E are obtained for parallel vectors $\vec{\beta}$ and $\vec{p}$ ($\cos\alpha = 1$) or anti-parallel vectors $\vec{\beta}$ and $\vec{p}$ ($\cos\alpha = -1$).

I find the energy ranges for top quark from interaction $pp \to t\bar{t}X$ (for antitop quark are the same). I define the parton four-momenta

$$P_A \equiv \left(\frac{x_A\sqrt{S}}{2}, 0, 0, \frac{x_A\sqrt{S}}{2}\right),$$

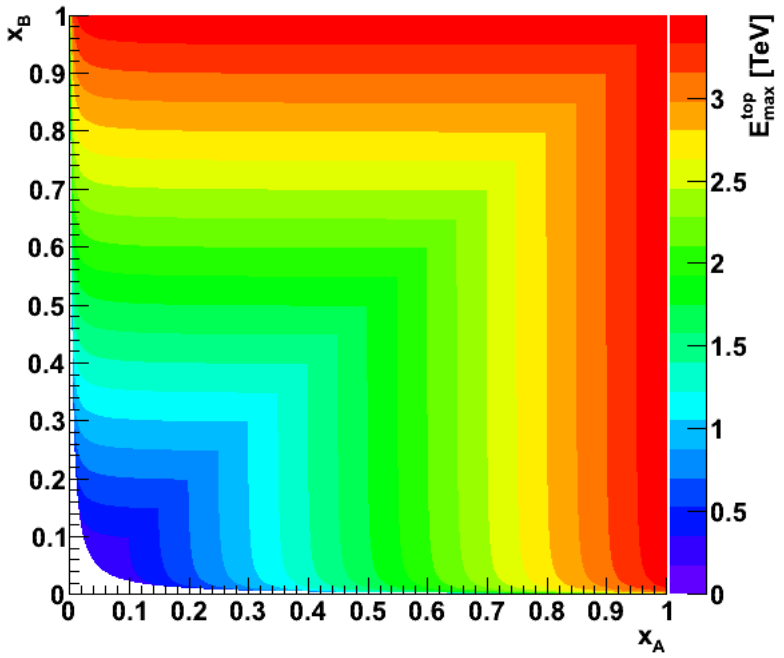


Figure 1.6: Dependence of the maximal energy of the top quark on fractions x_A and x_B.

$$P_B \equiv \left(\frac{x_B\sqrt{S}}{2}, 0, 0, -\frac{x_B\sqrt{S}}{2} \right).$$

In the center of mass system (CMS) of $t\bar{t}$ pair, the top quark has energy $E^{\text{top}}_{\text{CMS}} = \sqrt{x_A x_B S}/2$. Relatively to the LRF, the $t\bar{t}$ system has a velocity $\vec{\beta}_{t\bar{t}}$ with the only non-zero z-component equal to $\frac{x_A - x_B}{x_A + x_B}$. The angle α_t between momentum $\vec{p}^{\,\text{top}}_{\text{CMS}}$ and velocity $\vec{\beta}_{t\bar{t}}$ can be arbitrary, its distribution depends on the dynamics of interaction $pp \to t\bar{t}X$. Using Equation 1.4, the maximal energy $E^{\text{top}}_{\text{max}}$ can be obtained when the angle α_t is 0. This energy depends on the fractions x_A and x_B:

$$E^{\text{top}}_{\text{max}}(x_A, x_B) = \frac{\sqrt{S}}{4} \left(x_A + x_B + |x_A - x_B| \sqrt{1 - \frac{4\left(m^{\text{top}}\right)^2}{x_A x_B S}} \right). \tag{1.5}$$

This dependence is plotted in Figure 1.6. The maximal energy of a top quark over all possible values of x_A and x_B is $E^{\text{top}}_{\text{max}} = \sqrt{S}/2$ which occurs when the fractions $x_A = x_B = 1$. The evaluation of energy $E^{\text{top}}_{\text{max}}$ is straightforward, I did it with the introduction of the fractions x_A and x_B.

The minimal energy of a top quark $E^{\text{top}}_{\text{min}}$ is the rest mass m^{top} which occurs when $x_A = x_B = \frac{2m^{\text{top}}}{\sqrt{S}}$ (production at threshold). Therefore, the top quark energy range is $\left[m^{\text{top}}, \sqrt{S}/2\right]$. The velocity of a top quark $\beta^{\text{top}} \in \left[0, \sqrt{1 - \frac{4(m^{\text{top}})^2}{S}}\right]$.

I want to find the energy limits for the decay products of $t\bar{t}$ pair. I assume that the velocity β in Equation 1.4 can have values from certain interval $[0, \beta_{\text{max}}]$. By mathematical analysis of Equation 1.4 as function of β with two values of

$\cos\alpha$, one can find the maximal and minimal values of the transformed energy $E_{\max}$ and $E_{\min}$:

$$E_{\max} = \frac{1}{\sqrt{1-\beta_{\max}^2}}\left(E_{\mathrm{RF2}} + \beta_{\max}\sqrt{E_{\mathrm{RF2}}^2 - m^2}\right), \tag{1.6}$$

$$E_{\min} = \begin{cases} m & \text{if } \beta_{\max} > \beta_{\mathrm{RF2}}, \\ \frac{1}{\sqrt{1-\beta_{\max}^2}}\left(E_{\mathrm{RF2}} - \beta_{\max}\sqrt{E_{\mathrm{RF2}}^2 - m^2}\right) & \text{otherwise}, \end{cases} \tag{1.7}$$

where the velocity of particle P in RF2 is $\beta_{\mathrm{RF2}} = \frac{\sqrt{E_{\mathrm{RF2}}^2 - m^2}}{E_{\mathrm{RF2}}}$. If this velocity is less than $\beta_{\max}$, then the particle P can be at rest in LRF (in case when the LRF has velocity β_{RF2} relatively to RF2 and $\cos\alpha = -1$).

From basic kinematics, one can derive the energies of a two-body decay products. If a particle with rest mass m_{12} is decaying into particle 1 with rest mass m_1 and particle 2 with rest mass m_2, then the energy E_1 of particle 1 in the CMS of decaying particle is

$$E_1 = \frac{m_{12}^2 + m_1^2 - m_2^2}{2m_{12}}. \tag{1.8}$$

Using Equations 1.6, 1.7 and 1.8, I have computed the minimal and maximal energies of different decay products of a $t\bar{t}$ pair, see Table 1.5. The energies E_{CMS} and velocities β_{CMS} are computed for applicable reference frames (for W and b-quark the CMS of decaying t-quark, for lepton ℓ and neutrino ν_l the CMS of the decaying W). The velocity $\beta_{\max}$ is computed only for t-quark and W in LRF and these velocities represent the relative velocity of two considered reference frames (the $\beta_{\max}$ for t-quark is velocity of CMS of t-quark relatively to LRF and is used in the computation of $E_{\max}$ and $E_{\min}$ for W and b-quark,...). In order to use Equation 1.7, I have to remark that the minimal velocity for t-quark and W is 0. I used rest mass of b-quark $m_b = 4.67$ GeV, [4, p. 30]. I assumed zero neutrino masses.

Table 1.5: Minimal and maximal energies of $t\bar{t}$ pair decay products.

Particle	E_{CMS} [GeV]	$1-\beta_{\mathrm{CMS}}$	$E_{\min}$ [GeV]	$E_{\max}$ [GeV]	$1-\beta_{\max}$
t	–	–	m^{top}	3500	$1.2\cdot10^{-3}$
W	105.2	0.35	m_W	3495	$2.6\cdot10^{-4}$
b	68.1	$2.4\cdot10^{-3}$	m_b	2744	–
e	40.2	$8.1\cdot10^{-11}$	0.46	3495	–
ν_e	40.2	0	0.46	3495	–
μ	40.2	$3.5\cdot10^{-6}$	0.47	3495	–
ν_μ	40.2	0	0.46	3495	–

The observation is that the energy ranges are very wide to have any practical use in $t\bar{t}$ pair reconstruction, because most of the decay products have upper energy limit $\sim 0.5\sqrt{S}$. This is caused by high CMS energy $\sqrt{S}$. If the invariant mass would be only $\sqrt{S_2} = 500$ GeV then the previous table would look as is listed in Table 1.6. One can see that the allowed energy intervals are stricter compared to the previous case and they could be used in the rejection of certain solutions of the $t\bar{t}$ pair reconstruction.

Table 1.6: Minimal and maximal energies of $t\bar{t}$ pair decay products in case $\sqrt{S_2} = 500$ GeV.

Particle	E_{CMS} [GeV]	$1-\beta_{\mathrm{CMS}}$	E_{min} [GeV]	E_{max} [GeV]	$1-\beta_{\mathrm{max}}$
t	–	–	m^{top}	250	0.28
W	105.2	0.35	m_W	222	$6.8 \cdot 10^{-2}$
b	68.1	$2.4 \cdot 10^{-3}$	27.6	169	–
e	40.2	$8.1 \cdot 10^{-11}$	7.52	215	–
ν_e	40.2	0	7.52	215	–
μ	40.2	$3.5 \cdot 10^{-6}$	7.52	215	–
ν_μ	40.2	0	7.52	215	–

1.5 Differential Cross Sections in $t\bar{t}$ Production at the LHC

At the LHC, one of the important measurable spectra is the differential cross section $\mathrm{d}\sigma/\mathrm{d}x$ in the $t\bar{t}$ production, where this differential cross section can be expressed in any variable x. The most interesting differential cross sections are for transverse momentum of the top quark $\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}^{\mathrm{top}}$, transverse momentum of the $t\bar{t}$ pair $\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}^{t\bar{t}}$ and invariant mass of the $t\bar{t}$ pair $\mathrm{d}\sigma/\mathrm{d}M^{t\bar{t}}$. I will call them $p_{\mathrm{T}}^{\mathrm{top}}$, $p_{\mathrm{T}}^{t\bar{t}}$ and $M^{t\bar{t}}$ spectra, respectively. There are different theoretical techniques how to obtain predictions for these spectra in the SM. In QCD perturbation series, the leading order (LO) is insufficient to reasonably describe data and it is necessary to include higher order corrections in QCD. There are predictions available at the next-to-leading order (NLO). The uncertainties of these predictions are caused by uncertainties of parton distribution functions and by introducing unphysical fragmentation scale μ_F and renormalization scale μ_R. Results at next-to-next-to-leading order (NNLO) are not yet available due to the complexity of the calculations. However, an improvement to the NLO results can be achieved by adding approximate NNLO corrections. One way is using threshold resummation methods in which the logarithmic contributions associated to the emission of soft gluons from the initial state can be added at leading logarithmic (LL) order, next-to-leading-logarithmic (NLL) order or next-to-next-to-leading-logarithmic (NNLL) order, [15]. Nowadays, the most precise predictions are calculated using the approximate NNLL order.

The prediction for $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum is shown in Figure 1.7. The NNLO soft-gluon corrections enhance the $p_{\mathrm{T}}^{\mathrm{top}}$ distribution but the shape is approximately unchanged in the $p_{\mathrm{T}}^{\mathrm{top}}$ range shown. The scale dependence for the NNLO soft-gluon corrections is significantly reduced relatively to the NLO prediction. Any deviation of the measured spectrum from the theoretical spectrum can be a sign of the physics beyond SM.

The $M^{t\bar{t}}$ spectrum is more interesting due to the fact that a significant bump in its spectrum can be a signal of a new particle decaying into a $t\bar{t}$ pair. A theoretical candidate of such a particle can be the Z' boson. Also, any other deviation of the measurement from theory can point to physics beyond SM. The prediction of the $M^{t\bar{t}}$ spectrum in the SM is shown in Figure 1.8.

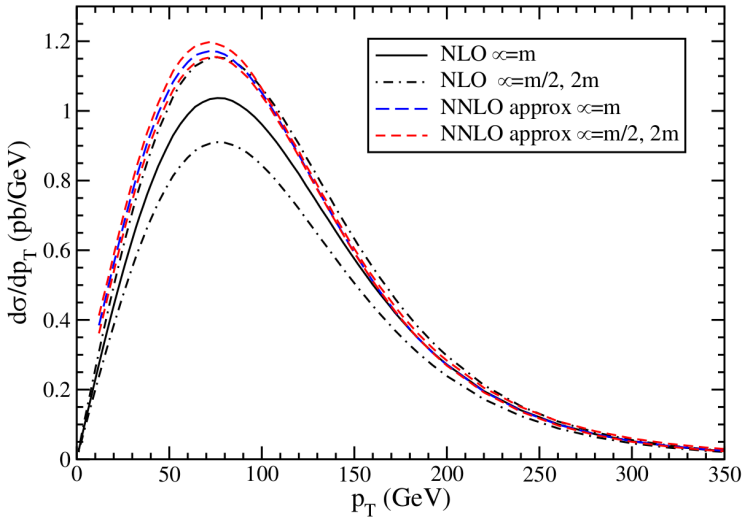


Figure 1.7: The differential cross section $\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}^{\mathrm{top}}$ for $t\bar{t}$ production at the LHC for $\sqrt{S} = 7$ TeV and top quark mass $m = 173$ GeV. The distributions are calculated with NLO (labeled as NLO) and with NNLO corrections derived from NNLL soft-gluon resummation (labeled as NNLO approx) for both fragmentation and renormalization scales equal to the top quark mass (labeled as $\propto= m$) and varied to half and double of the top quark mass (labeled as $\propto= m/2, 2m$), [16].

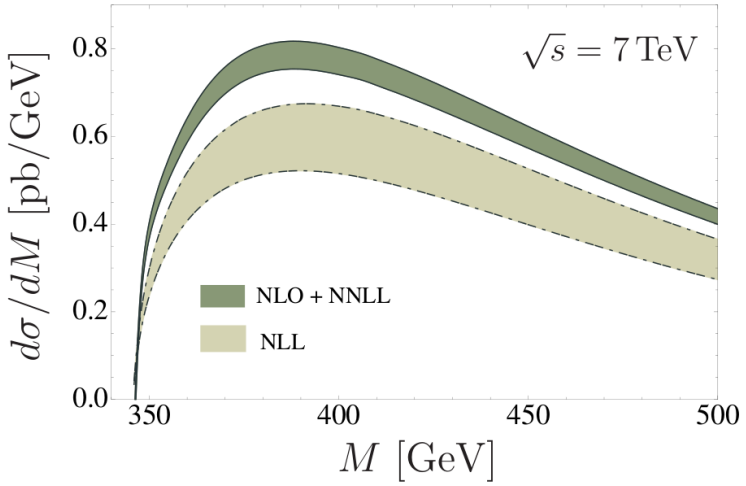


Figure 1.8: The differential cross section $\mathrm{d}\sigma/\mathrm{d}M^{t\bar{t}}$ for $t\bar{t}$ production at the LHC for $\sqrt{S} = 7$ TeV and top quark mass $m^{\mathrm{top}} = 173.1$ GeV. The distributions are calculated using resummed perturbation theory at NLL order (labeled as NLL) and NNLL order matched to the fixed NLO (labeled as NLO + NNLL). The width of the bands reflects the uncertainty of the spectrum under variations of the scale, [17].

2. The ATLAS Experiment

The ATLAS (A Toroidal LHC ApparatuS) experiment [18] is one of the main experiments at the LHC. In this experiment the pp collisions are detected via the world's largest particle detector, see Figure 2.1. The ATLAS detector is about 44 meters long, more than 25 meters high, and weighs about 7000 tons. The detector has subdetectors arranged symmetrically in layers around the interaction point and the beam pipe, therefore, the detector covers almost the entire solid angle around the interaction point.

To the interaction point, the closest subdetector system is the ATLAS Inner Detector, [20]. It can estimate the position of interaction point and momenta and impact parameters of charged particles. It is placed in a solenoidal magnetic field of 2 T to determine charged particles momenta. The acceptance in pseudorapidity is $|\eta| < 2.5$ for any azimuthal angle ϕ. The detector has been designed to provide a transverse momentum resolution of $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}} = 0.0005p_{\mathrm{T}}[\mathrm{GeV}] \oplus 0.01$ and a transverse impact parameter resolution of 10 µm for high momentum particles in the central η region. The Inner Detector consists of three different tracking technologies symmetrically distributed around the beam pipe: the Pixel Detector, the SemiConductor Tracker and the Transition Radiation Tracker.

The next layers of subdetectors are the calorimeters. The calorimeters measure the direction and energy of charged and neutral particles. The principle of calorimetry is to induce from the original particle an electromagnetic or hadronic shower using absorbers and to measure the deposited energy in the active material. The electromagnetic shower is induced by a high-energetic electron or a photon, while the hadronic shower is induced by a high-energetic hadron (assuming the mean lifetimes of hadrons, it can be p, n, π^0, $\pi^{\pm}$, K^0 or $K^{\pm}$). The energy resolution of a calorimeter is generally described by the equation

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}}[\mathrm{GeV}] \oplus \frac{b}{E}[\mathrm{GeV}] \oplus c, \tag{2.1}$$

where a is the stochastic term, b is the noise term and c is the constant term. In the ATLAS detector, there are two types of calorimeters: the ATLAS Liquid Argon (LAr) Calorimeters and the ATLAS Tile Calorimeter (TileCal).

The calorimeters closer to interaction point are the LAr Calorimeters [21] composited of four calorimeters: electromagnetic barrel calorimeter (EMB), electromagnetic endcap calorimeter (EMEC), hadronic endcap calorimeter (HEC) and forward calorimeter (FCal). Since the LAr is used as an active medium and it has to be cold, the LAr calorimeters are located in three separate cryostats. The EMB and EMEC are sampling electromagnetic calorimeters using lead as passive material. They are designated to completely contain the electromagnetic showers. Their pseudorapidity coverage ranges as $|\eta| < 1.475$ (EMB) and $1.375 < |\eta| < 3.2$ (EMEC). Their designed resolution terms are: $a = 10\%$, $b = 30\%$ and $c = 0.7\%$. Their angular resolution is better than $50\ \mathrm{mrad}/\sqrt{E}[\mathrm{GeV}]$. The EMB and EMEC contain only part of the hadronic showers, because hadronic showers have much broader longitudinal profiles than electromagnetic showers. The HEC detects the part of the hadronic showers which remain from EMEC in pseudorapidity region $1.5 < |\eta| < 3.2$ with designed energy resolution terms: $a = 50\%$ and $c = 3\%$.

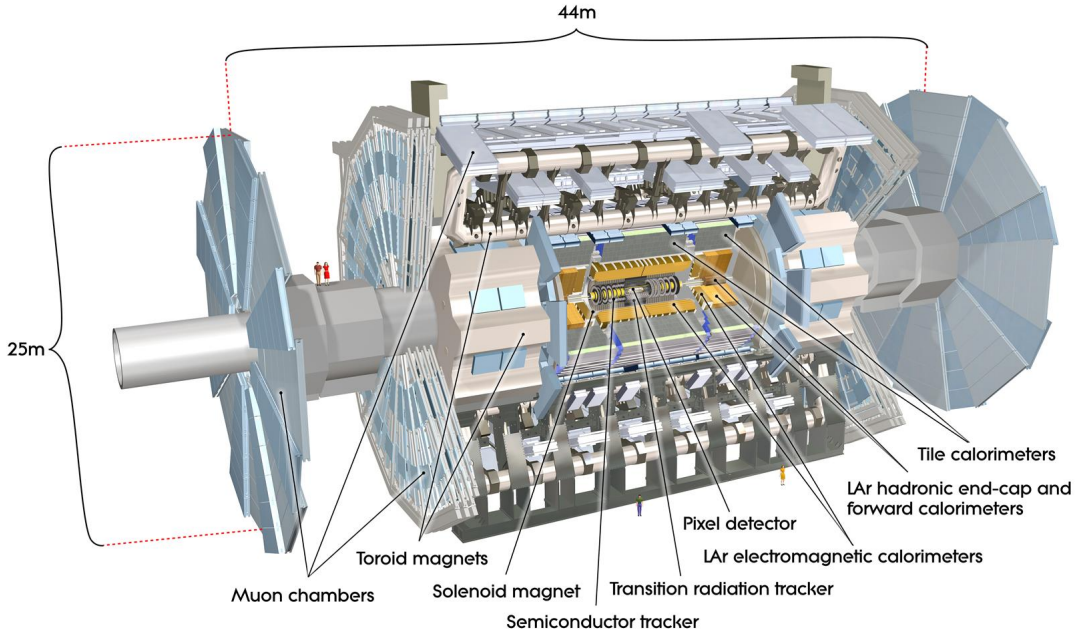


Figure 2.1: The ATLAS detector with indicated subdetectors and magnets, [19]. For scale demonstration, there are two people depicted just behind the first muon chamber on the left.

The FCal detects all kind of particles in the forward region with a pseudorapidity $3.2 < |\eta| < 4.8$. Its designed energy resolution terms are: $a = 100\%$ and $c = 10\%$.

The TileCal [22] is a hadronic sampling calorimeter using plastic scintillator as the active medium and iron as the absorber. It detects the remaining big part of hadronic showers from the EMB. It covers pseudorapidities $|\eta| < 1.7$. Its designed energy resolution terms are: $a = 50\%$ and $c = 3\%$.

The most distant subdetector system from the beam axis is the ATLAS Muon Spectrometer [23] which detects muons only, since the muons are the only particles which can pass through Tilecal (neutrinos also pass through Tilecal, but they cannot be detected). The muon spectrometer measures the muon paths in the magnetic field from which the muon's transverse momenta can be determined. The magnetic field is provided by toroidal magnets and is non-uniform. The designed transverse momentum resolution $\sigma p_{\mathrm{T}}/p_{\mathrm{T}}$ is 4% for muons with $p_{\mathrm{T}} \in (3\ \mathrm{GeV}, 100\ \mathrm{GeV})$ and it is increasing up to 10% for muons with $p_{\mathrm{T}} = 1$ TeV.

I will now point out why are the mean lifetimes of W or t-quark "too short" and mean lifetime of μ "long enough" which I used in the previous chapter. I assume a particle with rest mass m and momentum p. The basic assumption in particle decays is that the probability of a particle decay is same in every moment and do not depend on the history, therefore, the distribution of lifetime of the particle has exponential shape. The passed distance d before decay (range) is equal to lifetime multiplied by velocity. Therefore, the range d is distributed with probability density

$$f(d) = \frac{m}{p\tau} e^{-\frac{md}{p\tau}}, \tag{2.2}$$

where τ is the mean lifetime of the particle in the rest reference frame of the

particle and the distance d is in the LRF. The mean of d is a function of momentum p:

$$d_m(p) = \frac{p\tau}{m}. \tag{2.3}$$

Using the minimal and maximal energies from Table 1.5, I calculated the d_m for unstable particles from a $t\bar{t}$ event, Table 2.1. For b-mesons, I used the mass of 5 GeV and the mean lifetime of $1.5 \cdot 10^{-12}$ s from which an estimate of b-mesons mean range can be obtained.

Table 2.1: Mean lifetimes τ [4] and the computed mean distances $d_m(p_{\min})$ and $d_m(p_{\max})$ of $t\bar{t}$ decay products for their minimal and maximal energies.

Particle	τ[m]	$d_m(p_{\min})$	$d_m(p_{\max})$
t	$1.5 \cdot 10^{-16}$	0	3.1 fm
W	$9.5 \cdot 10^{-17}$	0	4.1 fm
B-mezons	$\sim 4.5 \cdot 10^{-4}$	0	250 mm
μ	$6.6 \cdot 10^{2}$	2.9 km	2200 km
τ	$8.7 \cdot 10^{-5}$	0	170 mm

It is clear that any motion of t-quark and W is negligible in comparison to the detector spacial resolution. The mean ranges of B-mesons and the τ particle are at the order of the size of the inner detector, thereby the secondary vertex of the b-jet event can be recognized. On the other hand, the lower limit for the range of μ is much higher than the detector's size, although the decay is inherently a probabilistic process and there is a probability that the μ will decay in the detector. This probability is certainly less than

$$\int_0^{x_{\max}} \frac{1}{2.9 \text{ km}} e^{-\frac{x}{2.9 \text{ km}}} \mathrm{d}x \approx 8.6 \cdot 10^{-8},$$

assuming that the most distant point of the detector is far $x_{\max} \approx 26$ m from the interaction point. Therefore, it can be safely said that muons from a $t\bar{t}$ event do not decay in the detector.

The total integrated luminosity of the proton-proton run recorded by the ATLAS detector until the end of year 2011 was $L = (5.25 \pm 0.19)$ fb^{-1}, [24]. The expected number of $t\bar{t}$ events is $N = L\sigma \approx 856000$. The number of events is a random variable with Poisson distribution ($L\sigma$ is the mean value). Therefore, even if the luminosity L and the cross section σ would be precisely known, the number of events can be arbitrary (there is a probability of e^{-N}, that no $t\bar{t}$ event was produced). But the number of events N is so high that this effect is negligible against the uncertainties of L and σ (for Poisson distribution the relative mean error is $1/\sqrt{N}$).

3. Simulation of the $t\bar{t}$ events

In most of current experimental particle physics experiments, one cannot perform any measurements without simulating the events under study using computer programs. The simulated events should look as the measured events, therefore, it is necessary to simulate the interaction of colliding protons, the decay of primary particles produced as well as the whole detector to simulate the detection of particles. The simulated events also help to understand the detector performance. The deviation of results obtained with simulated events from results obtained with measured events can point to a problem in the analysis or to new phenomena in the particle physics, but it can also point to wrong detector description (alignment, geometry), shower models, hadronisation models in the simulated events.

The kinematic variables of the decay products of a $t\bar{t}$ event are all random variables, because they depend on following probabilistic distributions:

- the parton distribution functions (PDFs) which are the probability density functions for fractions x_A and x_B of protons four-momenta carried by incidental partons. The PDFs are determined at certain scale in scattering experiments and evolved using theory,
- the angular distribution of the outgoing top and antitop quarks from the interaction point. For certain parton pair, this distribution can be calculated to certain order of the perturbation series. It depends on the previous random variables x_A and x_B,
- the branching ratios of the top quark decay. They are derived from theory: due to magnitude of the elements of the CKM matrix, the top quark decays mostly as $t \to Wb$.
- the angular distribution of the W boson and the b-quark from the decaying top quark. It can be computed to a certain order of perturbation series. It depends on the energy and polarization of the decaying top quark,
- the branching ratios of the W decay. They are precisely measured and also theoretically computed,
- the angular distribution of W decay products. It holds the same as for the top quark decay products.

With the knowledge of all these distributions, events can be simulated by the Monte Carlo method which is a numerical technique for calculating probabilistic distributions using sequences of random numbers. The usual steps of the Monte Carlo method are:

1. generating a sequence of random numbers on interval $[0, 1]$, from the uniform distribution making sure that elements of the sequence has no correlations. This is done by using random number generators,
2. the sequence from the first step is used to generate another sequence distributed according to some probability density function,

3. the sequence from the second step is used for further transformations.

The general software which produces simulated events is called Monte Carlo event generator. Additional software is necessary which models the hadronisation of partons. To produce simulated events which are comparable with the measured one, it is need to further simulate the detection process by simulating different detector responses (electromagnetic and hadronic showers, multiple Coulomb scattering,...). The final output is simulated raw "data" which is processed by reconstruction softwares equally as the measured raw data.

3.1 Monte Carlo event generators and detector simulation

For my analysis, I used simulated $t\bar{t}$ events prepared by the ATLAS group which are designed for the analysis focusing on the top quark. I studied two types of samples containing no hadronic $t\bar{t}$ events which I will denote as:

- MC@NLO - the event generator used was MC@NLO [25] version 3.41. It combines a Monte Carlo event generator with NLO calculations of rates for QCD processes. For simulating soft parton emission with interfering gluons the HERWIG [26] version 6.510 package was used. The JIMMY Generator [27] version 4.31 generated multiple parton scattering events in proton-proton collisions. An event can have weight -1 or +1 originating from the MC@NLO approach. This weight has to be correctly taken into account when the histograms are filled,

- POWHEG - events were generated by POWHEG BOX [28] interfaced with PYTHIA [29]. The POWHEG BOX is a general computer framework for implementing NLO calculations in shower Monte Carlo programs. The PYTHIA program simulates the initial-state and final-state parton showers and hadronisation.

The following information is valid for both samples. The ATLAS detector simulation was performed with GEANT4, [30]. The pole mass of the top quark is 172.5 GeV. The samples do not contain interactions of other proton-proton pairs in the colliding bunches (pile-up) which simplifies the analysis at MC level.

In the MC@NLO sample, the distribution of the top quark mass is a δ-function at the pole mass unlike in the POWHEG sample where the top mass has a Breit-Wigner distribution (BW) with a width (full width at half of the maximum) of 1.9 GeV.

The samples contain ROOT objects of TTree class with information sufficient for any analysis concerning the top quark. These samples are called D3PDs at ATLAS. I will call the set of simulated events prepared by Monte Carlo simulation as **MC** and the real events detected by the ATLAS detector as **data**. In the MC, the information on parton level is available which I will call the **truth** information. Of course, the data samples do not have any truth information.

I inspected the basic differences of these two samples using the truth information. In Figure 3.1, one can see that the fraction of production types $gq \to t\bar{t}q$ is much higher for POWHEG sample than for MC@NLO sample which is caused

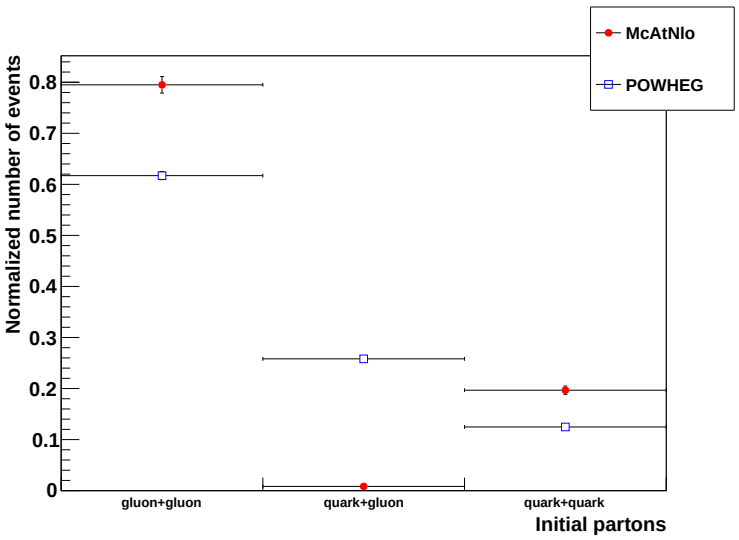


Figure 3.1: Distribution of initial interacting parton types for $t\bar{t}$ pair production for two MC generators.

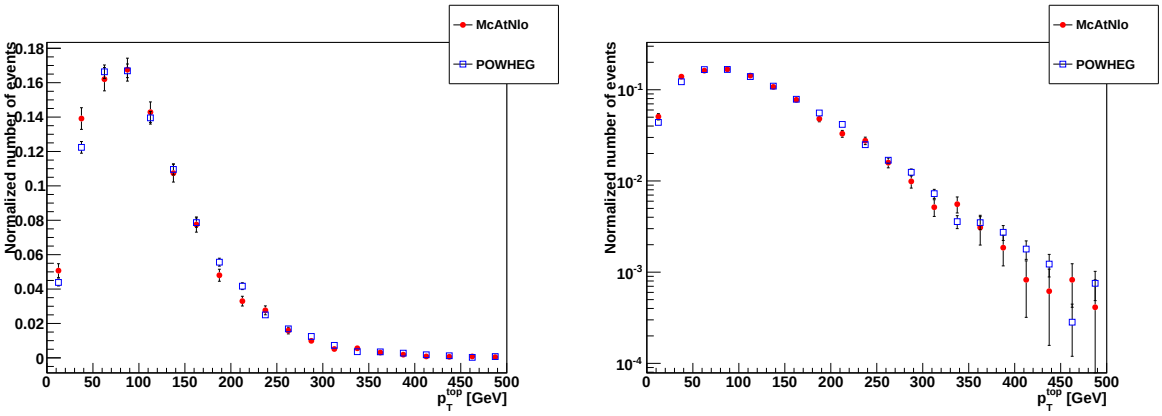


Figure 3.2: Top quark transverse momentum $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum for two MC generators in linear and logarithmic scales on the left and right, respectively.

by using different set of PDFs and different physics model involved in the description of hard scattering.

In Figures 3.2 and 3.3, two differential cross section spectra are shown. There is no large difference in the $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum, just the $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum for MC@NLO sample is slightly shifted to lower values.

From the comparison of the $M^{t\bar{t}}$ spectrum in the two samples, the effect of the non-zero width of the top quark BW is clearly seen. The POWHEG sample, which has non-zero width, has a tail to lower $M^{t\bar{t}}$ values while the MC@NLO sample with zero width has a sharp threshold at $M^{t\bar{t}} = 2m^{\mathrm{top}}$. The higher $M^{t\bar{t}}$ tail for both samples has a similar shape which is very important as this tail can contain a bump signaling a new particle decaying into $t\bar{t}$ pair. The possible candidate can be Z'.

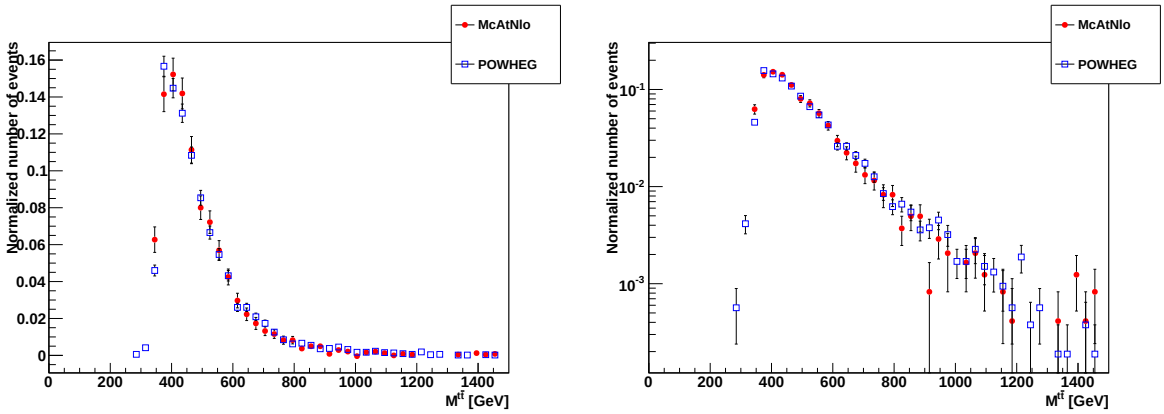


Figure 3.3: The invariant mass of the $t\bar{t}$ pair $M^{t\bar{t}}$ for two MC generators in linear and logarithmic scales on the left and right, respectively.

4. Object Reconstruction

To study the $t\bar{t}$ events, the corresponding decay products must be detected and reconstructed. Therefore, electrons, muons, jets and missing transverse energy have to be reconstructed. This cannot be accomplish without thorough understanding of performance and calibration of all detector components. I summarize the basic ideas which are important for object reconstruction.

4.1 Electrons

The reconstruction of electrons is done by the information measured in the inner detector and electromagnetic calorimeters. The signature of electrons in the ATLAS detector can be misidentified as coming from other particles or a jet. This is quantified in the variables electron identification efficiency and fake-rate. The electron identification efficiency is the probability that an electron is identified as electron. The reciprocal of fake-rate is the probability that a non-electron object is identified as electron. The reconstruction of an electron can fail due to insufficient response of subdetectors which is quantified by electron reconstruction efficiency. Using decays of the Z, W and J/Ψ particles, the electron identification and reconstruction efficiencies can be measured as is described in [31]. Also the energy resolution of electrons can be determined by that method. All these efficiencies and fake-rate depends on the p_{T} and η of the electron.

Within the information on the $t\bar{t}$ event, there is not only one electron or zero electron in case of failed reconstruction. There can be other electrons, like background electrons (primarily from photon conversions and Dalitz decays) and jets faking electrons, see the distribution of total number of electrons in Figure 4.1. In Section 5, I used preselected set of electrons which passed certain criteria on shape of electromagnetic shower and other criteria to reject background electrons.

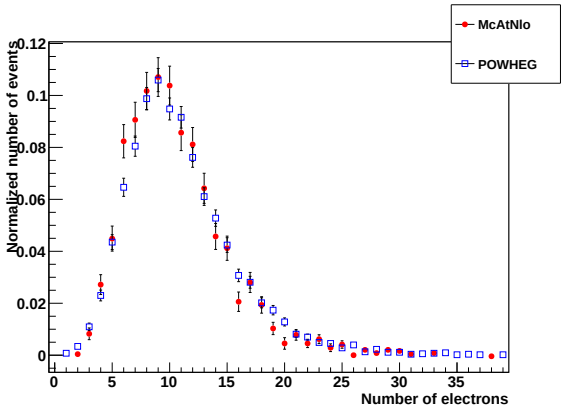


Figure 4.1: Distribution of number of electrons per event for two MC generators.

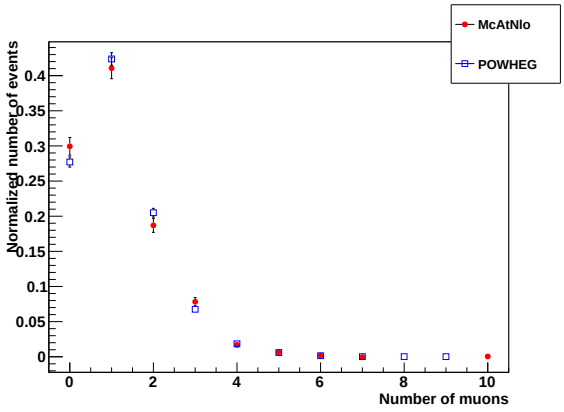


Figure 4.2: Distribution of number of muons per event for two MC generators.

4.2 Muons

The main subdetectors used to reconstruct muons are the muon spectrometer and the inner detector, but the information from calorimeters can also be used.

Similarly as for electrons, the muons can be also misidentified as other particles or they are not reconstructed. These effects are quantified by muon identification and reconstruction efficiencies.

In Figure 4.2, the number of muon candidates is shown (for both MC generators to show no significant difference between two MC generators). The background muons are also present as for electrons, but in lower rate which is caused by the fact that photon converts mainly to electron-positron pairs and much less often to muon-antimuon pairs. Also from other processes the presence of electrons is more frequent. Although, cosmic muons can be detected and if they go closely through the interaction point, they contribute to the count of background muons.

4.3 Jets

The inner detector and calorimeters are used to detect and identify jets. Jets are reconstructed with the anti-k_t algorithm, [32]. This is a cone algorithm with $\Delta R = 0.4$.

Similarly as for leptons, jet identification and reconstruction efficiencies are imposed. A jet can be misidentified as lepton or the reconstruction of jet can fail.

The distribution of number of jet candidates detected per event is shown in Figure 4.3. Although the simulation of $t\bar{t}$ events and jet production is quite different for the two samples used, the distribution of number of jets has a very similar shape. From the $t\bar{t}$ pair decay in single lepton channel, four jets arise, but there are many other sources of jets like soft and hard emissions, initial or final state radiation, other parton-parton interactions of colliding protons or pile-up. And these jets can also have relatively high p_{T}. In Figure 4.4, one can see that the first four highest p_{T} jets match the four truth partons only in $\sim 28\%$. It can happen that more truth $t\bar{t}$ partons are matched to the same jet. It is not sure that jet matched to truth parton really originate from this parton which mainly depends on the detector angular resolution. But one can say that in high number of $t\bar{t}$ events, not the first four highest p_{T} jets come from the original four partons. It is an important fact regarding the $t\bar{t}$ pair reconstruction.

4.4 Missing Transverse Energy

The missing transverse energy MET is determined using information from calorimeters and muon spectrometer as it is in detail described in [33]. Roughly, the MET is calculated using equations

$$\begin{aligned} \mathrm{MET} &= \sqrt{\mathrm{MET}_x^2 + \mathrm{MET}_y^2}, \\ \mathrm{MET}_x &= -p_x, \\ \mathrm{MET}_y &= -p_y, \end{aligned} \tag{4.1}$$

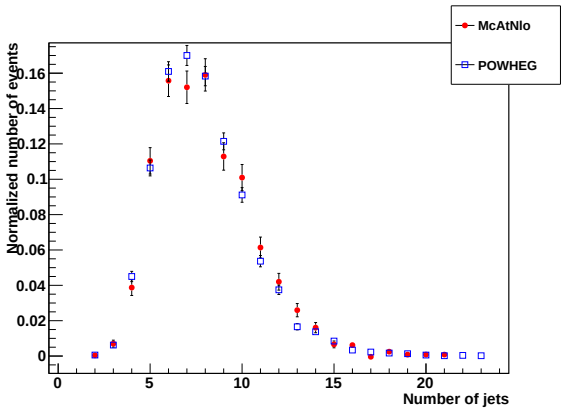


Figure 4.3: Distribution of number of jets per event for two MC generators.

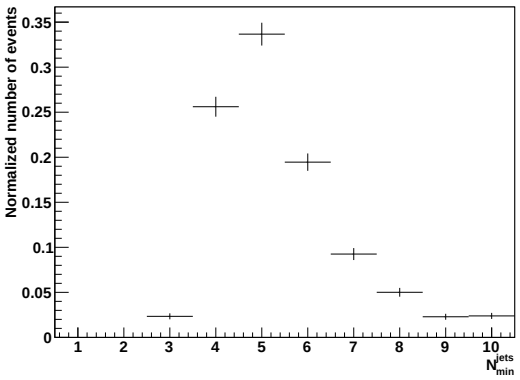


Figure 4.4: The minimal number of the highest p_{T} jets $N_{\mathrm{min}}^{\mathrm{jets}}$ which match the four truth $t\bar{t}$ partons using matching condition $\Delta R = 0.4$. The last bin is the overflow bin. The used sample was MC@NLO.

where p_x and p_y is the x-component and y-component of momentum of all reconstructed objects, respectively.

The MET has two important roles in an analysis of $t\bar{t}$ events in single lepton channel. High MET is an important signature of these events which can be used in the event selection process. The second role is in the $t\bar{t}$ pair reconstruction where the transverse momenta of neutrino are assumed to be close to the measured MET_x and MET_y, see Section 6.3.

4.5 Tagging of *b*-jets

The $t\bar{t}$ events contain at least two b-jets which is their significant property as b-jets arise in event types with relatively low cross sections. Therefore, it is very important to have the ability of distinguishing b-jets from other jets. Identification of b-jets (b-tagging) is based on specific properties of B-hadrons: long lifetime, large mass and large branching ratio into leptons. The discrimination of b from light quark jets originates mainly in the relatively long lifetime of b-flavoured hadrons, resulting in a significant flight path length. This leads to measurable secondary vertices and impact parameters of the decay products.

There are several b-tagging algorithms available: SV0, JetProb,... All of them have certain efficiency ϵ_q which is defined as the fraction of jets originating from a q-quark that are tagged by the tagging algorithm. The efficiency ϵ_b is called b-tagging efficiency and efficiency ϵ_q for $q = u, d, s, c$ is called as mistagging efficiency. For $q = u, d, s, c$, one can define the rejection rate $R_q \equiv 1/\epsilon_q$. The performance of any b-tagging algorithm is heavily dependent on p_T and η. The basic algorithms and comparison of different b-taggers can be found in [34].

The b-tagging has two main functions in an analysis of the $t\bar{t}$ events. In the event selection process, one can select events in which is at least one b-tagged jet to reduce background. The second exploitation is in the $t\bar{t}$ pair reconstruction, see Section 6.3.

5. Event Selection

At the CMS energy of 7 TeV, the measured total pp cross section is $\sim 10^{11}$ pb [35] which is $\sim 2 \cdot 10^9$ times higher than the cross section of $t\bar{t}$ production in the ℓ+jets channel. Therefore, a thorough event selection is necessary in order to distinguish between signal and background events.

At the maximal luminosity of 10^{34} cm^{-2}s^{-1} in the interaction region of the ATLAS detector, roughly 10^9 inelastic collisions occur every second. Technically, it is possible to record only a small fraction of all events due to this high collision rate. Most of the events are background events, because the interesting physical processes have smaller cross sections (like production of $t\bar{t}$ pairs). Most of the background events originate from QCD processes like dijet production, and this type of background is called QCD. In order to skip the QCD background events and record candidates for interesting events (like $t\bar{t}$), a fast online filter is necessary that selects the events of interest for offline analysis which is called trigger. For the analysis of $t\bar{t}$ events, the single lepton trigger was used. The detailed trigger requirements vary through the data-taking period due to the rapidly increasing LHC luminosity and the commissioning of the trigger system. They vary also for the type of lepton (electron or muon). The main selection requirement of the single lepton trigger is that only events with a lepton with $p_{\mathrm{T}} > 20$ GeV are recorded. The probability of accepting a certain type of events by the trigger is called trigger efficiency ϵ_{trig}. It depends on the p_{T}, η and other properties of the signature of the detected lepton. Events with signature of two leptons have generally higher single lepton trigger efficiency as events with signature of one lepton, roughly $\epsilon_{trig}^2 + 2\epsilon_{trig}(1 - \epsilon_{trig})$ when the dependence of ϵ_{trig} on lepton p_{T} and η is neglected. Throughout this chapter, I am assuming the same trigger efficiency for all types of events which is certainly not true, but it simplifies this part of my analysis.

5.1 Backgrounds to the $t\bar{t}$ Signal Process after the Trigger Selection

After the trigger selection, event sample enriched in high lepton p_{T} is recorded. But there still are background processes of higher or comparable cross section, see Table 5.1 listing the main background types. They all can have very similar signature as the $t\bar{t}$ events. One of them is presence of at least four high p_{T} jets. This signature can be reproduced in background events by jets originating from soft and hard emissions, initial or final state radiation, other parton-parton interactions of colliding protons or pile-up. The main difference in signature of signal and background events is that in general the b-jets are produced less often in background events, until in $t\bar{t}$ events, at least two b-jets are present.

The W+jets background is the most significant background which has very similar signature as the $t\bar{t}$ events: missing transverse energy, one high p_{T} lepton and high p_{T} jets. But in general, the number of high p_{T} jets are smaller.

The background events containing two leptons (Z+jets, dilepton) can have similar signature as $t\bar{t}$ events, because one lepton can be outside of the detector

Table 5.1: Signal and background processes with their theoretical cross-sections σ (CMS energy = 7 TeV) obtained from [36] and [37]. The $t\bar{t}$ cross section was referenced in Chapter 1. The cross section for the QCD background cannot be easily computed and is estimated from data. The X means any other particles.

process	denomination	σ [pb]
$pp \to t\bar{t}X$, $t\bar{t}$ decaying in single lepton channel	signal	57
$pp \to WX$, $W \to \ell\nu_\ell$	W+jets	10500
$pp \to ZX$, $Z \to \ell^+\ell^-$	Z+jets	960
$pp \to t\bar{t}X$, $t\bar{t}$ decaying in dilepton channel	dilepton	11
$pp \to tX$, W from t decaying leptonically	single top	69
$pp \to tWX$, one W decaying leptonically	single top	16
$pp \to$ jets, one jet is identified as lepton	QCD	-

acceptance and then only one lepton is detected and moreover missing transverse energy is generated. Another possibility is that one lepton is not reconstructed or not identified or one lepton overlaps with a jet.

The single top background has the most similar signature: the top quark produces a b-jet, lepton and missing transverse energy and other other jets can be produced along. Fortunately, the single top background has a small cross section.

After the trigger selection, there is still significant background from QCD processes as a jet can be misidentified as a lepton. The QCD background is called instrumental background.

From Table 5.1, it can be observed that the background is still $\sim$ 200-times higher then the signal. Therefore, one has to apply some additional selection criteria (cuts) on the data.

To quantify the performance of cuts, I introduce these terms: signal efficiency, background efficiency, purity and significance. The signal efficiency ϵ_s is the probability to accept a signal event as signal. The probability to accept a background event as signal is called the background efficiency ϵ_b. To determine these probabilities, one has to have large number of signal or background events $N_{\text{total},s/b}$ and count the number of events $N_{\text{pass},s/b}$ which passed the applied cut. With $N_{\text{total},s/b} \to \infty$, the signal or background efficiency will be

$$\epsilon_{s/b} = \frac{N_{\text{pass},s/b}}{N_{\text{total},s/b}}. \tag{5.1}$$

The purity P is the probability that the event is signal if it was accepted by cuts. It is intuitive that the purity can be computed as

$$P = \frac{\epsilon_s \pi_s}{\epsilon_s \pi_s + \sum_{i=1}^{n} \epsilon_{bi} \pi_{bi}}, \tag{5.2}$$

where n is the number of background processes, π_s and π_{bi} is the probability of finding a signal event and i-th background event in the data before cuts, respectively. This formula is also called the Bayes theorem in which the conditional probability P is expressed with the help of the conditional probabilities ϵ_s and ϵ_{bi} and prior probabilities π_s and π_{bi}. The prior probabilities can be obtained from

the magnitudes of cross sections σ_s and σ_b for signal and background events, respectively. Using simulated signal and background samples with large statistics, the purity can be obtained as

$$P = \frac{N_{\text{pass},s}}{N_{\text{pass},s} + \sum_{i=1}^{n} N_{\text{pass},bi} \frac{\sigma_{bi}}{\sigma_s}}. \tag{5.3}$$

The higher the purity, the higher the fraction of signal events in the resulting data sample after cuts. The maximization of purity can be done by maximizing the ratio $N_{\text{pass},s}/N_{\text{pass},bi}$. In the process of developing cuts where this ratio is higher and higher, the number of signal events can be very low and then the quality of the data sample after cuts suffers from low statistics. Therefore, cuts have to be developed in the way where both purity and $N_{\text{pass},s}$ are sufficiently high. For optimization, one can maximize the product of purity and number of passed signal events or the root square of this product. For this, I define variable significance S

$$S = \frac{N_{\text{pass},s}}{\sqrt{N_{\text{pass},s} + \sum_{i=1}^{n} N_{\text{pass},bi} \frac{\sigma_{bi}}{\sigma_s}}}. \tag{5.4}$$

But the choice of maximization of P or S, depends on the available statistics and background magnitude.

After applying cuts, certain quantities like $p_{\text{T}}^{\text{top}}$ spectrum can be biased (its shape changes). This is caused by the dependence of signal efficiency on the studied quantity. This dependence has to be taken into account when one wants to compare theoretical and measured spectra. The signal efficiency dependence on certain quantity can be quantified only in MC.

Other quantities like the mass of the top quark or total cross section of the $t\bar{t}$ production should not be biased by cuts.

5.2 Selection Cuts

The cut optimization is performed using MC samples where the exact effect of cuts can be quantified. In order to enhance the data sample, one has to find differences in signatures of signal and background events.

I inspected the distribution of variable missing transverse energy MET defined in 4.4 for signal and two main backgrounds W+jets and Z+jets which is shown in Figure 5.1. For my simple study, I neglected the other backgrounds. My definition of applied cut is: event passes, if MET $>$ t$_{\text{MET}}$. My goal was to find the optimal cut value of t_{MET}.

In Figure 5.2, the dependence of the cut efficiency on t_{MET} is shown. The Z+jets background is largely suppressed already at low t_{MET} which is the result of the fact that Z+jets events do not contain high p_{T} neutrinos. There is no difference of efficiency between the two simulated signal samples MC@NLOand POWHEG. As the events from W+jets background contain high energy neutrino, the cut has small suppression effect to this background with respect to the signal. For rejection of events from W+jets background, another cut has to be developed.

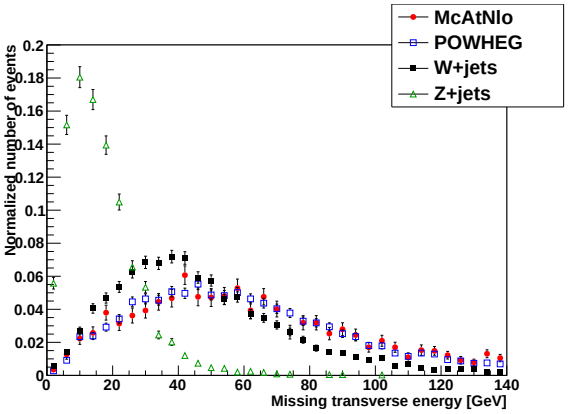


Figure 5.1: Missing transverse energy for MC signal, W+jets and Z+jets samples.

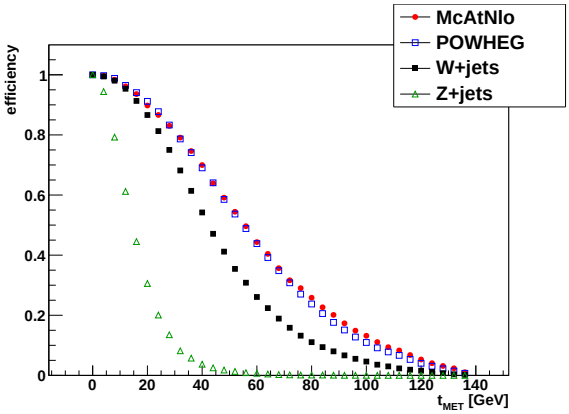


Figure 5.2: Dependence of cut efficiencies on t_{MET}.

I inspected the purity and significance, where only the Z+jets or only the W+jets or both Z+jets and W+jets backgrounds are taken into account in formulas 5.3 and 5.4, see Figures 5.3 and 5.4. The purity P_Z corresponding to the case of assuming only Z+jets as background gets much better than P_W with the higher cut value t_{MET}. Due to much higher cross section of W+jets background, there is no difference between P_W and P_{WZ} or between S_W and S_{WZ}. It is better to survey only the purity P_Z and significance S_Z, because to suppress W+jets background, one needs a cut which better distinguishes between signature of $t\bar{t}$ events and W+jets events than such a cut on MET.

If there would be only Z+jets background, the optimal value would be $t_{\mathrm{MET}} = 48$ GeV which has the maximal significance, but the signal efficiency is only $\epsilon_s = 55\%$.

This approach does not have to be necessarily the best. Because there are other backgrounds, one has to optimize other cuts to suppress them and the overall ϵ_s of cuts can be very small and the statistics of the final sample low. This is why maybe lower t_{MET} should be chosen. I chose value $t_{\mathrm{MET}} = 32$ GeV which has $\epsilon_s = 81\%$.

Other significant differences between signal and background spectra are the p_{T} of lepton and number of high p_{T} jets. I did not optimize further cuts, just inspected the efficiencies, purity and significance of these cuts:

- Cut 1: missing transverse energy larger than 32 GeV. This cut reduces the background types without neutrinos (Z+jets, QCD),
- Cut 2: exactly one electron or muon passing the lepton selection criteria from Top Group (main are p_{T} larger than 20 GeV and with $\eta \in [-2.5, 2.5]$). This cut reduces background types with no or more than one lepton (Z+jets, dilepton, QCD). But a fraction of these events can pass because the lepton reconstruction is not 100% efficient (the lepton can be identified as a jet),
- Cut 3: at least four jets which passed the jet selection criteria from Top Group (main are p_{T} larger than 25 GeV and with $\eta \in [-2.5, 2.5]$). This cut reduces all types of background and it is the main cut for W+jets background, because events from this background have usually lower number of high p_{T} jets as $t\bar{t}$ events.

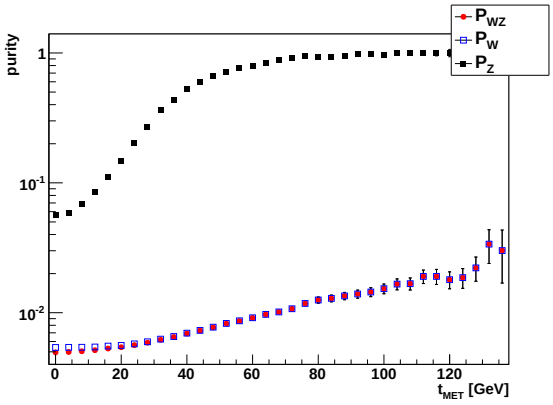


Figure 5.3: Dependence of purity on t_{MET}. The used MC signal sample was MC@NLO. The purity P_W and P_Z was computed assuming only background W+jets and Z+jets in Equation 5.3, respectively. The purity P_{WZ} were computed assuming both backgrounds W+jets and Z+jets in Equation 5.3.

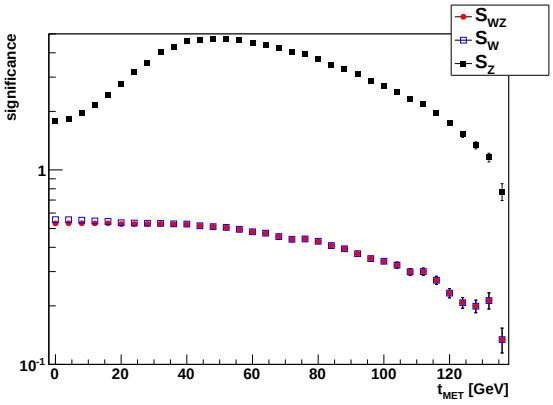


Figure 5.4: Dependence of significance on t_{MET}. The used MC signal sample was MC@NLO. The significance S_W and S_Z was computed assuming only background W+jets and Z+jets in Equation 5.4, respectively. The significance S_{WZ} were computed assuming both backgrounds W+jets and Z+jets in Equation 5.4.

For the one lepton and four jets, the condition $\eta \in [-2.5, 2.5]$ is applied not because of background reduction, but because the resolution of detector is lower outside of the region (limited detector acceptance). I applied these cuts on signal, W+jets and Z+jets MC samples. In Figure 5.5, one can see the cumulative efficiency after each cut (as the total number of events N_{total} was used the original number of events). In Figure 5.6, the relative efficiency is plotted which is the ratio of number of events after and right before applying of cut. One can see that to reject the W+jets background, the Cut 3 is the most useful and to reject Z+jets background, the Cut 1 is the most useful. In Figure 5.5, one can see that the cut efficiency for signal events also depends on the type of MC simulation. This can harm an analysis on real data as it is written in Section 6.6.

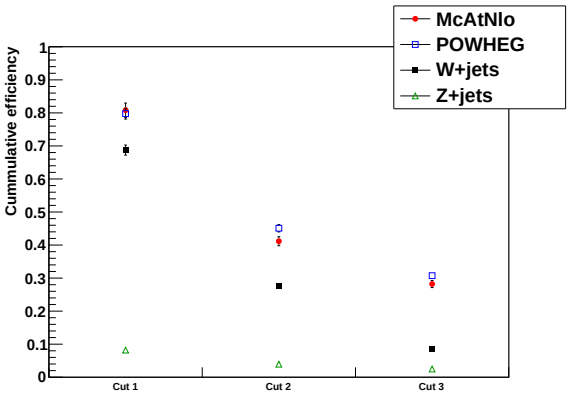


Figure 5.5: Cumulative efficiency of applied cuts for signal and background.

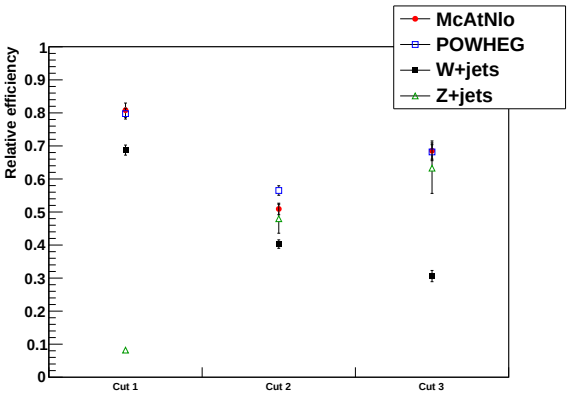


Figure 5.6: Relative efficiency of applied cuts for signal and background.

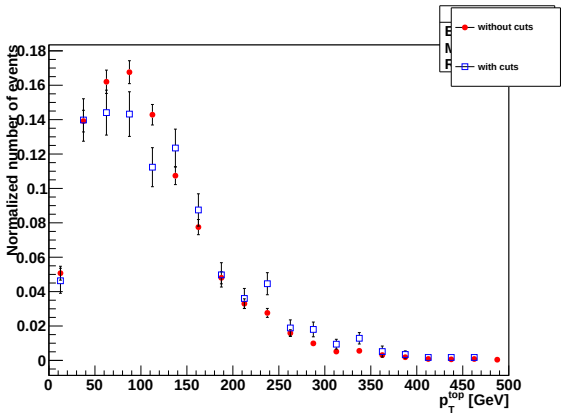


Figure 5.7: Normalized distributions of $p_{\mathrm{T}}^{\mathrm{top}}$ before and after applied cuts obtained with MC@NLO sample.

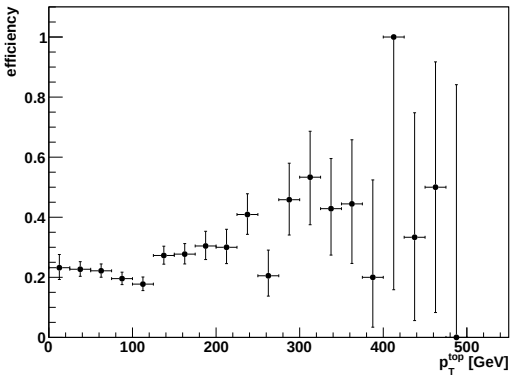


Figure 5.8: Dependence of the cut efficiency on the $p_{\mathrm{T}}^{\mathrm{top}}$ obtained with MC@NLO sample.

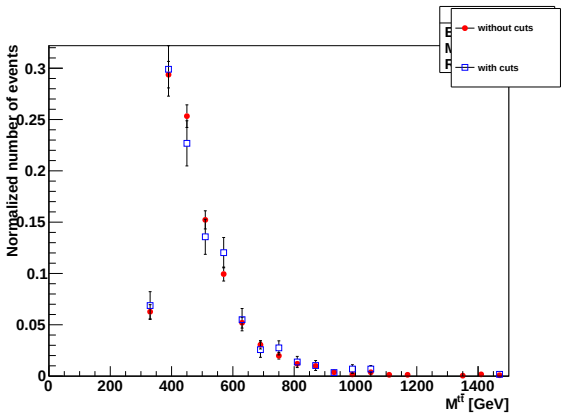


Figure 5.9: Normalized distributions of $M^{t\bar{t}}$ before and after applied cuts obtained with MC@NLO sample.

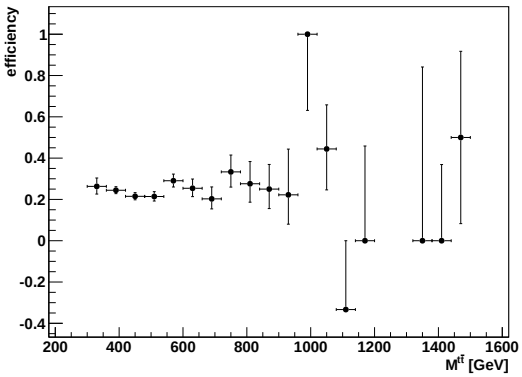


Figure 5.10: Dependence of the cut efficiency on the $M^{t\bar{t}}$ obtained with MC@NLO sample.

The signal efficiency dependence on $p_{\mathrm{T}}^{\mathrm{top}}$ and $M^{t\bar{t}}$ spectra is shown in Figures 5.7-5.10. This dependence has to be taken into account when one wants to compare theoretical and measured spectra.

In this chapter, I used only one dimensional cuts. I introduced the basics for optimization of cut values. I tried to optimize the cut value for variable MET for two backgrounds W+jets and Z+jets. The cut on MET can effectively reduce only Z+jets background which is caused by the large difference of the shape of MET spectrum between signal and Z+jets events. I showed that the purity rise with the cut value, but the significance has its peak which points to the optimal cut value. For high number of backgrounds, the cut optimization has to take into account the largeness of resulting statistics.

One can also devise more general cuts with the help of multivariate statistical methods: neural networks, boosted decision trees,... The development of such methods is complicated due to high number of background types.

6. Measurement of Differential Cross Sections in $t\bar{t}$ Production at ATLAS

In the ATLAS experiment, there are several groups with the aim of measuring the differential cross section spectra which are important distributions regarding the physics beyond SM as was discussed in Section 1.5. In this chapter, I will describe my contributions to the analysis focused on the measurement of the transverse momentum spectrum of the top quark, $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum. However, I will also study distributions of $p_{\mathrm{T}}^{t\bar{t}}$ and $M^{t\bar{t}}$ spectra as it was easy to obtain them during the analysis. But the main steps of the analysis are described for $p_{\mathrm{T}}^{\mathrm{top}}$spectrum. This analysis has no direct connection with my studies in previous chapter where I was showing the basic principles of event selection for background reduction.

To contribute to the analysis, I used the TopNtupleAnalysis package available for ATLAS members. This package was written and developed by people from the ATLAS Top Differential Cross Section Group (Group). I took part in the developing of this package and performed several studies which I am describing in this chapter.

This analysis studies the single lepton channel (ℓ+jets) with a slightly different definition as I had in the previous chapters: the leptonicaly decaying W can decay into e, μ or τ not regarding how the τ decays. This is the choice of Top Differential Cross Section Group. There are two orthogonal selection paths focusing on finding a high p_{T} electron (and vetoing high p_{T} muon) or a high p_{T} muon (and vetoing high p_{T} electron). Both are applied on the same samples and thereby the analysis is divided into two parts: called the e+jets channel and μ+jets channel. But it should be clear that for example the e+jets channel can contain events where in reality the leptonic W decays to μ, just the muon reconstruction failed, and background electron was reconstructed and the event was selected by selection criteria focusing on high p_{T} electron. Therefore the e+jets and μ+jets channels are defined by selection criteria, not by the truth lepton type in the event. At the end of the analysis, both spectra obtained in different channels are combined into one to obtain the desired spectrum in ℓ+jets channel. In my studies, I am showing only figures for e+jets channel as the figures for μ+jets channel are very similar, and therefore, the same conclusions apply.

The analysis uses simulated $t\bar{t}$ events (MC signal), simulated background events (MC background) and detected events (data). Data sample corresponding to integrated luminosity of $L = 4.713$ fb^{-1} collected by the ATLAS detector at $\sqrt{S} = 7$ TeV in 2011 were used in the analysis. The expected number of signal events in ℓ+jets channel before the trigger requirement or other selection is $\sim$ 340 000. The statistics of simulated events is much higher than of data (number of events for MC signal in ℓ+jets channel is $\sim$ 9 200 000). The process of event generation was described in Section 3. MC signal is simulated with the MC@NLO generator, parton showering and the underlying event are modeled using HERWIG and JIMMY. In the simulation, these constant were used: pole mass of the top quark $m_{\mathrm{MC}}^{\mathrm{top}} = 172.5$ GeV, pole mass of the W boson

$m^W_{\mathrm{MC}} = 80.399$ GeV, decay width of the top quark $\Gamma^{\mathrm{top}}_{\mathrm{MC}} = 1.32$ GeV and decay width of the W boson $\Gamma^W_{\mathrm{MC}} = 2.085$ GeV.

The main background types were listed in Table 5.1. The W+jets and Z+jets backgrounds were simulated with the ALPGEN [38] generator interfaced with HERWIG and JIMMY. The single top background was simulated with AcerMC [39]. The dilepton background was simulated with the same generator as MC signal sample. There is one type of background which is not obtained from MC but from data (data-driven background). This is the multijet (QCD) background which incorporates events from ordinary QCD processes where a jet is reconstructed as lepton and the event comply the selection criteria.

I introduce terms **truth** and **reco** levels as follows. The truth level means all information about the $t\bar{t}$ pair event on parton level which is theoretically motivated as a quark does not develop a parton shower. The reco level means all information which can be obtained from the detector. In data, there is obviously only the reco level. The aim of the measurement is to correct the measured reco spectrum in data to the truth level in terms of resolutions and reconstruction effects.

The final goal of the analysis is to obtain corrected $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum in some optimal binning. The number of bins in histogram for $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum is $N_{\mathrm{bins}} = 8$ and their widths are based on certain optimization algorithm.

6.1 Analysis Flow

The analysis is performed along the following outline:

1. corrections applied to the simulation to match the data – it is performed on all MC-based samples, because the detector performance cannot be precisely simulated in MC. The quantification of imperfect simulation is obtained from comparison of other measurements with MC. Two main types of MC corrections are implemented: smearing of energies of objects due to imperfect detector resolution in simulation, and weighting of events due to imperfect simulation of object reconstruction efficiencies. The first correction changes the four-vector of reco objects and the truth objects are unchanged, the second correction gives to each event a weight affecting both truth and reco levels which is important for reco level to improve simulation of detector, but it has no meaning on truth level and it just shows a part of detector effects on truth spectra. It has meaning after step 3.

2. object definition and selection – the prescription from the Top Group is used to define appropriate objects at reco level. The aim of the object selection is to remove objects which are outside well-understood detector regions or not fulfilling tighter object definition requirements. The same object selection is performed for MC and data.

3. event selection – the prescription from the Top Group is used to select events at the reco level. The aim of the event selection is to essentially reduce the background and to remove events which do not contain sufficient number of good objects which is important for the $t\bar{t}$ pair reconstruction. The same event selection is performed for MC and data. Moreover, in MC, cut on

trigger simulation is applied, because in data the trigger requirement is applied automatically at online recording due to high collision rate. The event selection tightly relates with object selection as the selection criteria depend on the number and properties of selected objects. Also, the event selection in MC depends on corrections in step 1. Assuming perfect MC corrections and correspondence between simulated spectra and real spectra, the truth spectra from data would correspond to truth spectra in MC after event selection. The main selection criteria are listed in Section 6.2.

4. $t\bar{t}$ pair reconstruction – a kinematic likelihood fit is used to reconstruct the $t\bar{t}$ pairs with the help of the KLFitter tool. The same reconstruction method is used for MC and data. After reconstruction one can get the measured (reco) $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum for MC and data and the corresponding migration matrix for MC only (see the definition below). The KLFitter tool is described in Section 6.3. I studied the performance of this tool.

5. control plots – the purpose of data/prediction control plots is to show agreement or disagreement between distributions of data and MC. Various spectra from data and stacked MC signal and background are compared and validated.

6. background subtraction – the number of events in j-th bin of MC background BG_j is subtracted from number of events in j-th bin of data D_j. By seeing good control plots, it is assumed that the background description is credible and it can be used for subtraction from data.

7. unfolding – corrects for all effects responsible for non-equivalence of truth and reco $p_{\mathrm{T}}^{\mathrm{top}}$ in a MC event. These are mainly the detector resolution effects and effects with imperfect $t\bar{t}$ pair reconstruction. The pile-up effects are also responsible, because they make the $t\bar{t}$ pair reconstruction more difficult. The unfolding is done on data after background subtraction but its effects are studied in MC. I performed unfolding with matrix inversion method which is described in Section 6.5.

8. event selection correction – corrects for effects connected with rejecting events because of event selection and consequently biasing the final spectrum (caused by non-100% trigger efficiency, all cuts and failed $t\bar{t}$ reconstruction). To perform this correction, each bin j of the unfolded $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum has to be divided by efficiency ϵ_j. This efficiency can be obtained from MC when one divides the truth $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum after reco level cuts and reconstruction from step 4 by truth $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum from the beginning of the analysis (without any correction or cuts). This correction is separately done for both channels: e+jets and μ+jets. Because the selection cuts were applied on the whole single lepton channel, after this correction, the $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum in both channels will correspond to the inclusive single lepton channel.

9. systematic uncertainties (systematics) – quantify the effects of systematics errors of jet energy resolution, jet energy scale, luminosity,... Most of them can be obtained only from MC. To quantify the systematics, the whole analysis is re-run with variation of all MC corrections one by one and the final

spectra are compared to spectra without any variation of MC corrections and the systematic error is derived.

10. combine the two channels e+jets and μ+jets into one ℓ+jets channel – this is not implemented in this analysis yet.

Finally, the evaluation of the resulting measured differential cross section $\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}^{\mathrm{top}}$ can be expressed by the formula

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^{\mathrm{top}}}\right)_j \equiv \frac{\sum_{i=1}^{N_{\mathrm{bins}}} (M^{-1})_{ji}(\mathrm{D}_i - \mathrm{BG}_i)}{\epsilon_j L \Delta x_j}, \tag{6.1}$$

where $\left(\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}^{\mathrm{top}}\right)_j$ is the averaged measured value of $\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}^{\mathrm{top}}$ in interval pertaining to the j-th bin with width of Δx_j. The multiplication with matrix M^{-1} represents any unfolding method. The integral of $\left(\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}^{\mathrm{top}}\right)_j$ over all bins gives the measured total cross section of $t\bar{t}$ production in the studied channel.

Because a $t\bar{t}$ event contains two top quarks, there are two $p_{\mathrm{T}}^{\mathrm{top}}$ values from each event called leptonic $p_{\mathrm{T}}^{\mathrm{top}}$ and hadronic $p_{\mathrm{T}}^{\mathrm{top}}$ (depending on the leptonic or hadronic decay of corresponding W). In my analysis, I always fill histograms with both $p_{\mathrm{T}}^{\mathrm{top}}$. For variables $M^{t\bar{t}}$ and $p_{\mathrm{T}}^{t\bar{t}}$, only one entry is obtained from each $t\bar{t}$ event. One can get the wrong impression that the statistics of $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum is twice as high as the statistics of $M^{t\bar{t}}$ or $p_{\mathrm{T}}^{t\bar{t}}$ spectra. But there is a correlation between the leptonic $p_{\mathrm{T}}^{\mathrm{top}}$ and hadronic $p_{\mathrm{T}}^{\mathrm{top}}$. One should study this correlation and also the individual leptonic $p_{\mathrm{T}}^{\mathrm{top}}$ and hadronic $p_{\mathrm{T}}^{\mathrm{top}}$ spectra. In this analysis, it is not done yet. One has to remember that the variables D_i and BG_i for $p_{\mathrm{T}}^{\mathrm{top}}$ in equation 6.1 are twice as high as they should be and to obtain the correct final spectrum, one has to divide the resulted $\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}^{\mathrm{top}}$ histogram by two.

During the analysis, it is useful to define and study the **migration matrix** which is a two-dimensional histogram with the x-axis corresponding to the truth $p_{\mathrm{T}}^{\mathrm{top}}$ and y-axis corresponding to the reco $p_{\mathrm{T}}^{\mathrm{top}}$ obtained after the $t\bar{t}$ pairs are reconstructed. It can be plotted only for MC signal. For the measure of goodness of reconstructed $p_{\mathrm{T}}^{\mathrm{top}}$, one can take the correlation r defined as

$$r \equiv \frac{\mathrm{E}\left[(\text{truth } p_{\mathrm{T}}^{\mathrm{top}} - \mathrm{E}\left[\text{truth } p_{\mathrm{T}}^{\mathrm{top}}\right])(\text{reco } p_{\mathrm{T}}^{\mathrm{top}} - \mathrm{E}\left[\text{reco } p_{\mathrm{T}}^{\mathrm{top}}\right])\right]}{\mathrm{RMS}\left(\text{truth } p_{\mathrm{T}}^{\mathrm{top}}\right)\mathrm{RMS}(\text{reco } p_{\mathrm{T}}^{\mathrm{top}})}, \tag{6.2}$$

where $\mathrm{E}(X)$ and $\mathrm{RMS}(X)$ are the mean value and the root mean square of variable X, respectively. For evaluation of $\mathrm{E}(X)$ and $\mathrm{RMS}(X)$, the center of bins are used. The higher the correlation between the reco and truth $p_{\mathrm{T}}^{\mathrm{top}}$, the more is the matrix diagonal, and the reconstructed $p_{\mathrm{T}}^{\mathrm{top}}$ is closer to the truth $p_{\mathrm{T}}^{\mathrm{top}}$. But the correlation is only one number and also other measures of goodness of migration matrix have to be taken into account. By the help of migration matrix, one can construct the **transition matrix** by normalizing the columns to one (each element of the migration matrix is divided by the sum of the elements in the corresponding column). An array for certain truth $p_{\mathrm{T}}^{\mathrm{top}}$ shows the distributions of fractions for reco $p_{\mathrm{T}}^{\mathrm{top}}$. Concretely, the elements on diagonal shows the fraction of events reconstructed correctly for certain truth $p_{\mathrm{T}}^{\mathrm{top}}$. Therefore, the transition matrix shows the effectiveness of $p_{\mathrm{T}}^{\mathrm{top}}$ reconstruction and it is important during the unfolding.

In all following one-dimensional histograms, the last bin is filled also with entries larger than the upper range value of the histogram's x-axis and this bin is called the overflow bin. Similarly, it applies also for two-dimensional histograms. Entries outside the range of x or y axis are filled to the last bin of the corresponding array or row.

I define few terms useful for classification of the decay products of a $t\bar{t}$ pair decaying in single lepton channel. The hadronic and leptonic top is the top quark which decays to hadronicaly and leptonicaly decaying W, respectively. The hadronic and leptonic b-quark is the b-quark from the decay of hadronic and leptonic top, respectively. The hadronisation of hadronic and leptonic b-quark gives hadronic and leptonic b-jet, respectively. In SM, there is no correlation between the type of top quark (leptonic or hadronic) and the charge of the top quark (top or antitop).

6.2 Event Selection

In this section, I summarize the event selection cuts. In both e+jets and μ+jets channels, these selection criteria were used:

- at least four jets with $p_{\mathrm{T}} > 25$ GeV, $|\eta| < 2.5$,
- at least one b-tagged jet.

In the e+jets channel these selection criteria were used:

- exactly one electron with $p_{\mathrm{T}} > 25$ GeV, $|\eta| < 2.5$,
- no muon with $p_{\mathrm{T}} > 20$ GeV, $|\eta| < 2.5$,
- MET > 30 GeV,
- $m_{\mathrm{T}}^{W} > 30$ GeV.

In the μ+jets channel these selection criteria were used:

- exactly one muon with $p_{\mathrm{T}} > 25$ GeV, $|\eta| < 2.5$,
- no electron with $p_{\mathrm{T}} > 20$ GeV, $|\eta| < 2.5$,
- MET > 20 GeV,
- $m_{\mathrm{T}}^{W} + \mathrm{MET} > 60$ GeV.

The W boson transverse mass m_{T}^{W} is defined as

$$m_{\mathrm{T}}^{W} = \sqrt{2p_{\mathrm{T}}^{\ell}p_{\mathrm{T}}^{\nu}(1 - \cos(\phi^{\ell} - \phi^{\nu}))}, \tag{6.3}$$

where p_{T}^{x} and ϕ^{x} is the transverse momentum and the azimuthal angle of particle x, respectively. The particle x can be lepton ℓ or neutrino ν. On reco level, the MET_x and MET_y are used for the transverse momentum components of neutrino.

One can observe that the selection criteria for e+jets are orthogonal to μ+jets channel in terms of the lepton type requirements.

It can happen that in some events, the reconstruction of $t\bar{t}$ pair does not succeed and these events are rejected. Therefore, the reconstruction of $t\bar{t}$ pair is also accounted as an event selection criterion.

6.3 Reconstruction of $t\bar{t}$ Pairs and KLFitter Studies

After the object and event selection, only events with exactly one selected lepton and $N^{\text{jets}} \geq 4$ selected jets remained. The four highest p_{T} jets need not to originate from the four partons originating from $t\bar{t}$ pair as it was shown in my study on smaller MC sample in Figure 4.4. Therefore, one has to decide the number of jets $N^{\text{jets}}_{\text{max}}$ which will be used for $t\bar{t}$ pair reconstruction. If $N^{\text{jets}} < N^{\text{jets}}_{\text{max}}$, then only N^{jets} will be used for reconstruction. From $N^{\text{jets}}_{\text{max}}$ jets, one has to assign one jet to leptonic b-jet, one jet to hadronic b-jet and two jets to jets from the hadronicaly decaying W. There are

$$N_{\text{comb}} = N^{\text{jets}}_{\text{max}}(N^{\text{jets}}_{\text{max}} - 1)(N^{\text{jets}}_{\text{max}} - 2)(N^{\text{jets}}_{\text{max}} - 3)/2$$

combinations how to do it neglecting the swap of light jets from the W, see Table 6.1 for concrete numbers.

Table 6.1: The number of combinations N_{comb} for $N^{\text{jets}}_{\text{max}}$ jets.

$N^{\text{jets}}_{\text{max}}$	4	5	6	7
N_{comb}	12	60	180	420

For each combination, a kinematic fit based on the likelihood approach can be used to find the momentum of neutrino and adjust the energy and angular variables of jets and the lepton belonging to the $t\bar{t}$ pair. One can define the likelihood function L as a product of kinematic constraints on reconstructed candidates from the decaying top quarks:

$$\begin{aligned} L(P^{\ell}, P^{\nu}, P^{j1}, P^{j2}, P^{hb}, P^{lb} | m^{\text{top}}_{\text{MC}}, m^{W}_{\text{MC}}, \Gamma^{\text{top}}_{\text{MC}}, \Gamma^{W}_{\text{MC}}, P^{\ell}_{\text{meas}}, P^{lb}_{\text{meas}}, P^{hb}_{\text{meas}}, P^{j1}_{\text{meas}}, P^{j2}_{\text{meas}}, \\ \text{MET}_x, \text{MET}_y) \equiv \text{BW}(M_{\text{inv}}(P^{\ell}+P^{\nu}) | m^{W}_{\text{MC}}, \Gamma^{W}_{\text{MC}}) \cdot \text{BW}(M_{\text{inv}}(P^{j1}+P^{j2}) | m^{W}_{\text{MC}}, \Gamma^{W}_{\text{MC}}) \cdot \\ \text{BW}(M_{\text{inv}}(P^{j1}+P^{j2}+P^{hb}) | m^{\text{top}}_{\text{MC}}, \Gamma^{\text{top}}_{\text{MC}}) \cdot \text{BW}(M_{\text{inv}}(P^{\ell}+P^{\nu}+P^{lb}) | m^{\text{top}}_{\text{MC}}, \Gamma^{\text{top}}_{\text{MC}}) \cdot \\ \text{TF}(E^{hb}|E^{hb}_{\text{meas}}) \cdot \text{TF}(E^{lb}|E^{lb}_{\text{meas}}) \cdot \text{TF}(E^{j1}|E^{j1}_{\text{meas}}) \cdot \text{TF}(E^{j2}|E^{j2}_{\text{meas}}) \cdot \text{TF}(E^{\ell}|E^{\ell}_{\text{meas}}) \cdot \\ \text{TF}(\Omega^{hb}|\Omega^{hb}_{\text{meas}}) \cdot \text{TF}(\Omega^{lb}|\Omega^{lb}_{\text{meas}}) \cdot \text{TF}(\Omega^{j1}|\Omega^{j1}_{\text{meas}}) \cdot \text{TF}(\Omega^{j2}|\Omega^{j2}_{\text{meas}}) \cdot \text{TF}(p^{\nu}_x|\text{MET}_x) \cdot \\ \text{TF}(p^{\nu}_y|\text{MET}_y) \cdot P(b\text{−tagging}), \end{aligned} \quad (6.4)$$

where P^x is the four-momentum of object x which can be lepton ℓ, neutrino ν, leptonic b-jet lb, hadronic b-jet hb or one jet from W hadronic decay $j1$ or $j2$. Similarly, E^x is the energy of object x and Ω^x stands for angular variables of object x. The term $M_{\text{inv}}(P)$ is the invariant mass of the four-momentum P. The transfer function $\text{TF}(A|A_{\text{meas}})$ is the conditional probability density function for variable A in assumption that the measured value of this variable is A_{meas}. For transverse momenta of neutrino p^{ν}_x and p^{ν}_y, the measured values are assumed to be the missing transverse energies MET_x and MET_y. It is assumed that the mass m of the top quark or W boson has a distribution of relativistic Breit-Wigner function centered around the pole mass m_{pole} with width Γ:

$$\text{BW}(m|m_{\text{pole}}, \Gamma) = \frac{1}{(m^2 - m^2_{\text{pole}})^2 + m^2_{\text{pole}}\Gamma^2}. \quad (6.5)$$

The term $P(b-\text{tagging})$ represents the probability that jets assigned to truth b-jets and truth light jets have the same flavor in reality. This depends on whether the individual jets are b-tagged or not, and the probability is computed with corresponding b-tagging efficiencies and rejection rates.

The likelihood approach says that the optimal variables of all objects will be obtained by maximizing the likelihood function w.r.t. free parameters. There are 16 free parameters for this function: three components of momentum of neutrino, five energies of jets and lepton, eight angular variables of jets. The mass of the top quark can be also used as a free parameter, if one wants to study the top quark mass.

The purpose of transfer functions in the likelihood function is to adjust the measured energies or angular variables due to imperfect detector resolution. This adjusting is done with the use of kinematic constraints which are implemented by Breit-Wigner functions. For the lepton there is no need for tuning of angular variables as the resolution of lepton's angular variables are very good. Also, the longitudinal momentum of neutrino p_z^ν does not have any transfer function because there is no constraint for longitudinal momentum of the initial parton-parton system.

With such high number of free parameters, the success of maximization highly depends on initial values of parameters. For all parameters except p_z^ν the measured values are used as initial values. The initial value for momentum p_z^ν can be obtained from the kinematic equation

$$M_{\text{inv}}(P^\ell_{\text{meas}} + P^\nu_{\text{meas}}) = m^W, \tag{6.6}$$

where in the four-vector P^ν_{meas} the measured values are assumed to be MET_x and MET_y. This is a quadratic equation for p_z^ν with solutions

$$\begin{aligned}
p^\nu_{z\,1,2} &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \\
a &= \left(E^\ell_{\text{meas}}\right)^2 - \left(p^\ell_{z,\text{meas}}\right)^2, \\
b &= -p^\ell_{z,\text{meas}} \left(\left(m^W\right)^2 + 2p^\ell_{x,\text{meas}}\text{MET}_x + 2p^\ell_{y,\text{meas}}\text{MET}_y \right), \\
c &= \left(E^\ell_{\text{meas}}\right)^2 \left(\text{MET}_x^2 + \text{MET}_y^2\right) - \left(\frac{\left(m^W\right)^2}{2} + p^\ell_{x,\text{meas}}\text{MET}_x + p^\ell_{y,\text{meas}}\text{MET}_y \right)^2,
\end{aligned} \tag{6.7}$$

where the lepton mass was neglected. One has to choose one of initial values $p^\nu_{z\,1,2}$ or perform the maximization for both initial values and choose the solution with higher likelihood. Due to the detector resolution or due to the Breit-Wigner distribution of m_W, it can happen that the discriminant $b^2 - 4ac < 0$ and then the initial value for maximization can be chosen as $p_z^\nu = \frac{-b}{2a}$.

After the maximization of likelihood function is done for all combinations of four jets from $N^{\text{jets}}_{\text{max}}$ jets, the combination with the highest maximized value of L is called the best combination and it is assumed as the correct combination. From best combination, the 16 parameters, for which the L has maximal value, are used to construct four-vectors of $t\bar{t}$ decay products on parton level using the corresponding masses of objects. The mass of the neutrino is assumed to be zero. Then any variable concerning the $t\bar{t}$ pair can be obtained ($p_{\text{T}}^{\text{top}}$, $M^{t\bar{t}}$, $p_{\text{T}}^{t\bar{t}}$,...) and it is called as reco variable.

On principle described above, the tool KLFitter (Kinematic Likelihood Fitter) [40] works and it was used in this analysis. The transfer functions are derived in MC signal samples and they are parametrized with double Gaussian functions with ten parameters except for neutrino transfer functions which are parametrized only with simple Gaussian function. For simplicity and numerical stability, it is not the function L maximized, but rather the $[-\log L]$ function is minimized.

I studied the performance of the KLFitter and optimized several settings of this tool. In this section, all further studies were done on MC signal.

6.3.1 Number of Jets Provided to the KLFitter

In this section, I am studying what choice of maximal number of jets provided to the KLFitter is better: $N_{\text{max}}^{\text{jets}} = 4$ or $N_{\text{max}}^{\text{jets}} = 5$. From Figure 4.4, it is obvious that using $N_{\text{max}}^{\text{jets}} = 5$ instead of $N_{\text{max}}^{\text{jets}} = 4$ more than doubles the number of events where the KLFitter inspects also the combination of four jets most probably originating from the four partons from the $t\bar{t}$ pair. The main disadvantage of using $N_{\text{max}}^{\text{jets}} = 5$ instead of $N_{\text{max}}^{\text{jets}} = 4$ is that there are five times more possibilities how to assign the four jets coming from $t\bar{t}$ event. In events where the four partons from $t\bar{t}$ pair corresponds to four jets with the highest p_{T}, the reconstruction efficiency can be only worse, because the KLFitter can accidentally choose possibility containing the new fifth jet and this possibility will be certainly wrong (although, the surveyed $p_{\text{T}}^{\text{top}}$ can be the same in terms of used binning). But in the other events, the reconstruction efficiency can be only better for $N_{\text{max}}^{\text{jets}} = 5$ than $N_{\text{max}}^{\text{jets}} = 4$.

One can also study option $N_{\text{max}}^{\text{jets}} = 6$ where the KLFitter chooses the best combination from 180 possibilities, but the choice of the correct combination is then quite improbable.

I studied the impact of different $N_{\text{max}}^{\text{jets}}$ on the reconstruction of $p_{\text{T}}^{\text{top}}$, $M^{t\bar{t}}$ and $p_{\text{T}}^{t\bar{t}}$. In Figure 6.1, the comparison of migration matrices for $p_{\text{T}}^{\text{top}}$ is shown. By inspecting the diagonal elements, one can see that for option $N_{\text{max}}^{\text{jets}} = 5$ the reconstructed $p_{\text{T}}^{\text{top}}$ is more often the same as the truth $p_{\text{T}}^{\text{top}}$. Also, the correlation between truth and reco $p_{\text{T}}^{\text{top}}$ is much higher. the same holds for $M^{t\bar{t}}$ migration matrices in Figure 6.2. Similarly for $p_{\text{T}}^{t\bar{t}}$ in Figure 6.3, the correlation factor is much higher for $N_{\text{max}}^{\text{jets}} = 5$ than for $N_{\text{max}}^{\text{jets}} = 4$. Although the lower reco values of $p_{\text{T}}^{t\bar{t}}$ are more often reconstructed correctly for $N_{\text{max}}^{\text{jets}} = 4$. This can be explained as follows. In events with low $p_{\text{T}}^{t\bar{t}}$, the contribution of high p_{T} jets from initial state radiation is smaller, and therefore, there is higher probability that the four highest p_{T} jets are really the four jets coming from the $t\bar{t}$ pair decay products. The reconstruction for $N_{\text{max}}^{\text{jets}} = 4$ can be more efficient as there is five times less combinations. On the other hand, events with higher $p_{\text{T}}^{t\bar{t}}$ have frequently more high p_{T} jets not coming from the $t\bar{t}$ pair decay products. And therefore, the choice $N_{\text{max}}^{\text{jets}} = 5$ is more efficient for higher $p_{\text{T}}^{t\bar{t}}$.

Further, I studied how the different option for $N_{\text{max}}^{\text{jets}}$ shifts the variables $p_{\text{T}}^{\text{top}}$, $M^{t\bar{t}}$ and $p_{\text{T}}^{t\bar{t}}$. I filled two-dimensional histograms, where the x-axis displays the desired variable evaluated with option $N_{\text{max}}^{\text{jets}} = 4$ and the y-axis displays the same variable evaluated with option $N_{\text{max}}^{\text{jets}} = 5$. In Figure 6.4, these histograms are shown.

It is obvious that events on the diagonal give the same result for both settings in terms of used binning. The fraction of same results for $p_{\text{T}}^{\text{top}}$ is $\sim 78\%$, for $M^{t\bar{t}}$

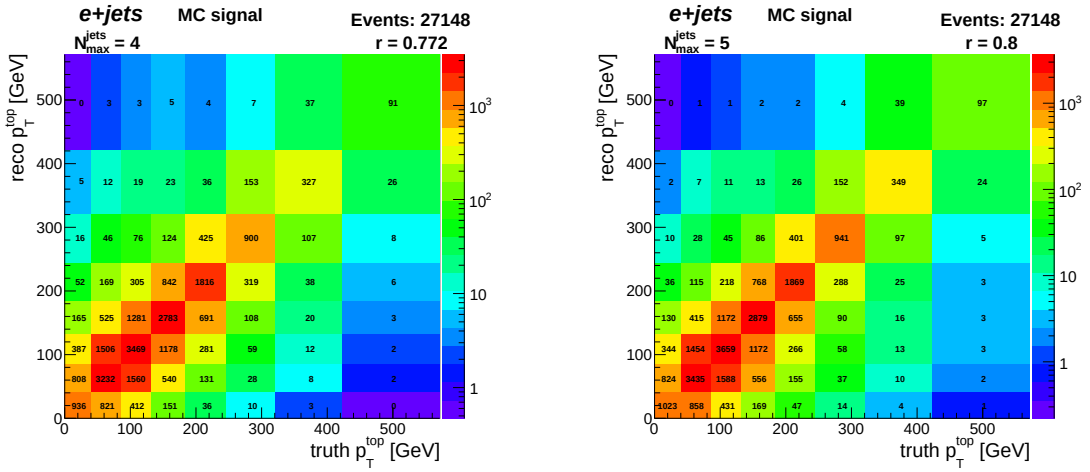


Figure 6.1: Migration matrices for $p_{\mathrm{T}}^{\mathrm{top}}$ obtained with $N_{\mathrm{max}}^{\mathrm{jets}} = 4$ and $N_{\mathrm{max}}^{\mathrm{jets}} = 5$ on the left and right, respectively. The migration matrix for $N_{\mathrm{max}}^{\mathrm{jets}} = 5$ has higher correlation factor r and higher diagonal elements.

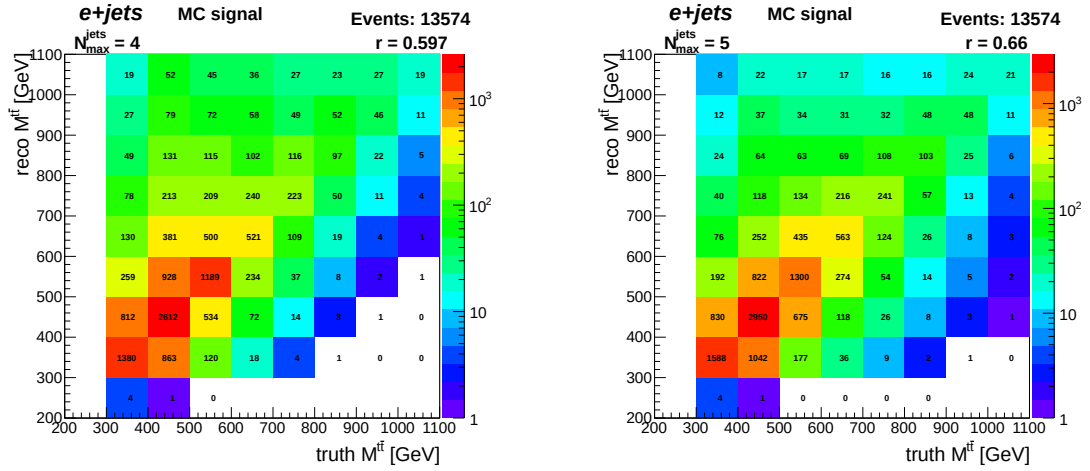


Figure 6.2: Migration matrices for $M^{t\bar{t}}$ obtained with $N_{\mathrm{max}}^{\mathrm{jets}} = 4$ and $N_{\mathrm{max}}^{\mathrm{jets}} = 5$ on the left and right, respectively. The migration matrix for $N_{\mathrm{max}}^{\mathrm{jets}} = 5$ has higher correlation factor r and higher diagonal elements.

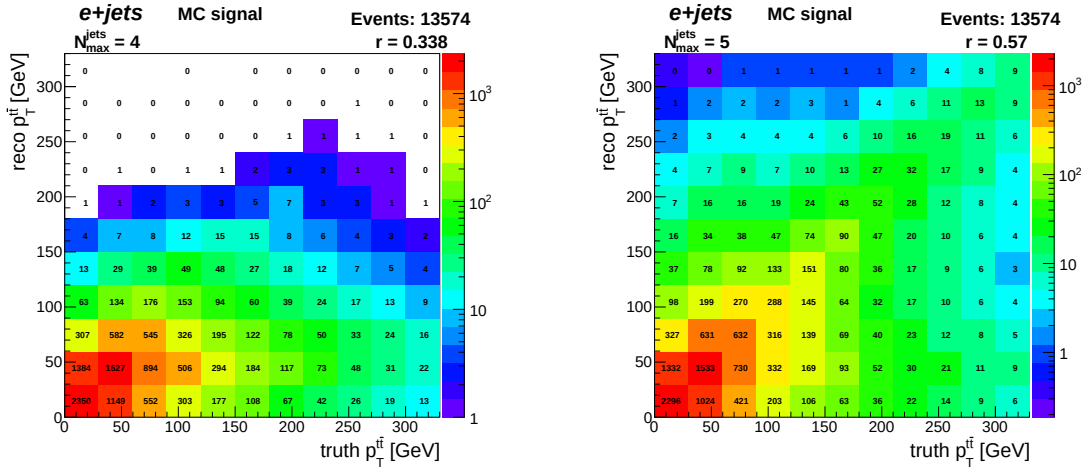


Figure 6.3: Migration matrices for $p_{\mathrm{T}}^{t\bar{t}}$ obtained with $N_{\mathrm{max}}^{\mathrm{jets}} = 4$ and $N_{\mathrm{max}}^{\mathrm{jets}} = 5$ on the left and right, respectively. The migration matrix for $N_{\mathrm{max}}^{\mathrm{jets}} = 5$ has higher correlation factor r and higher diagonal elements except the first two diagonal elements.

it is $\sim 73\%$ and for $p_{\mathrm{T}}^{t\bar{t}}$ it is $\sim 69\%$. From figures, one can see that the different settings do not have a sizable impact on $p_{\mathrm{T}}^{\mathrm{top}}$. On the other hand, the different settings do have large impact on the $M^{t\bar{t}}$ or $p_{\mathrm{T}}^{t\bar{t}}$ spectrum. For $N_{\mathrm{max}}^{\mathrm{jets}} = 4$, the reconstructed $M^{t\bar{t}}$ is shifted to higher values and the reconstructed $p_{\mathrm{T}}^{t\bar{t}}$ is shifted to lower values. It is hard to explain what is the reason of such shifts.

The conclusion from this Section is that for $p_{\mathrm{T}}^{\mathrm{top}}$ and $M^{t\bar{t}}$ it is certainly better to use $N_{\mathrm{max}}^{\mathrm{jets}} = 5$ due to higher reconstruction efficiency. For the variable $p_{\mathrm{T}}^{t\bar{t}}$, it cannot be clearly concluded until more optimal binning is chosen and also one should perform unfolding to see the full effect of the choice $N_{\mathrm{max}}^{\mathrm{jets}} = 4$ or $N_{\mathrm{max}}^{\mathrm{jets}} = 5$.

In the following part of my analysis, I used $N_{\mathrm{max}}^{\mathrm{jets}} = 5$.

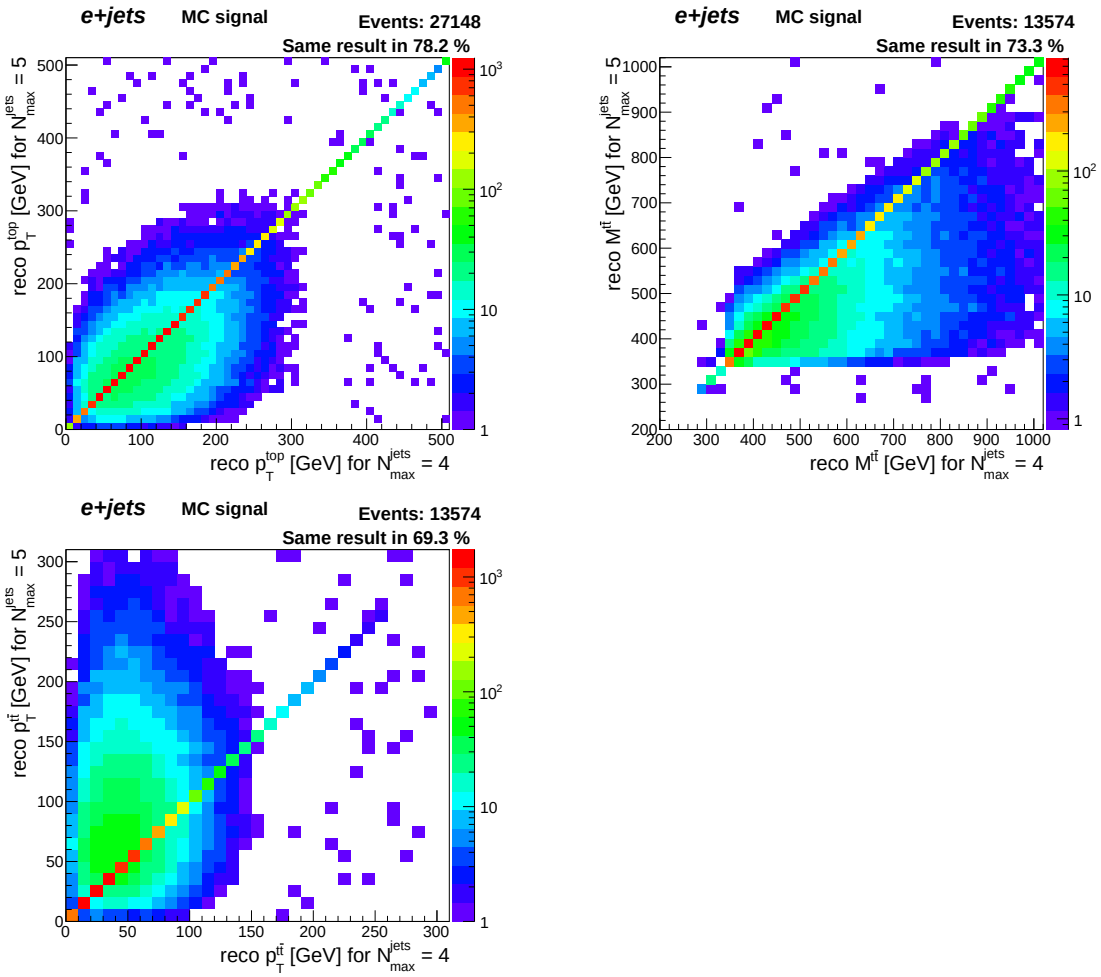


Figure 6.4: Two-dimensional histograms of reco $p_{\mathrm{T}}^{\mathrm{top}}$, $M^{t\bar{t}}$ and $p_{\mathrm{T}}^{t\bar{t}}$ for the two options of the KLFitter: $N_{\mathrm{max}}^{\mathrm{jets}} = 4$ on the x-axis and $N_{\mathrm{max}}^{\mathrm{jets}} = 5$ on the y-axis obtained from the MC signal sample. The bin size for $p_{\mathrm{T}}^{\mathrm{top}}$ and $p_{\mathrm{T}}^{t\bar{t}}$ is 10 GeV $\times$ 10 GeV and for $M^{t\bar{t}}$ it is 20 GeV $\times$ 20 GeV.

6.3.2 Performance of the KLFitter for Tauonic Events

It was already discussed that the signature of an $t\bar{t}$ event is determined by the decay of two W bosons from decay $t\bar{t} \to W^{-}\bar{b}W^{+}b$ and also by the decay of τ in case that one or two Ws decayed tauonicaly. The definition of single lepton

channel in this analysis is that the hadronic W decays to jets and the leptonic W decays to e, μ or τ. I will call the events where the leptonic W decays to τ as tauonic events. I study the performance of the KLFitter on tauonic events. Large part of tauonic events are removed by selection criteria as in 63 % the τ decays to hadrons and then in the $t\bar{t}$ pair, one does not have a real lepton as a decay product (although small fraction of these events can pass the selection criteria as a high p_{T} background lepton can always occur in an event as was discussed in Section 4.1). These events will certainly have smaller reconstruction efficiency, but their fraction from all events which passed the selection criteria is negligible. The question is how the reconstruction of $t\bar{t}$ pair performs for tauonic events where the τ decays to lepton.

There is a problem to use the KLFitter for the tauonic events with the decay chain $t\bar{t} \to W^- b W^+ \bar{b} \to 4$ jets $+ \ell\nu_\ell \bar{\nu_\tau} \nu_\tau$. The three neutrinos come from the same W, which is important for the KLFitter as the sum of the three neutrino's four-momenta can be used as the four-momentum of one neutrino which is input for the KLFitter. But this sum of three neutrino's four-momenta has one large dissimilarity from the one neutrino coming from the decay $t\bar{t} \to W^- b W^+ \bar{b} \to$ 4 jets $+ \ell\nu_\ell$. This sum of four-momenta does not have zero invariant mass which holds for the neutrino alone and is assumed in the KLFitter. I investigated what are the kinematic limits for this invariant mass. It holds:

$$P^W = P^\tau + P^{\nu_\tau} = P^{\bar{\nu_\tau}} + P^{\nu_\ell} + P^\ell + P^{\nu_\tau}, \tag{6.8}$$

where P^x is the four-momentum for particle x. This is the kinematics for tauonicaly decaying W. The square of the invariant mass of the three neutrinos M can be computed as

$$M^2 = (P^{\bar{\nu_\tau}} + P^{\nu_\ell} + P^{\nu_\tau})^2 = (P^W - P^\ell)^2. \tag{6.9}$$

In the rest reference frame of W (RRFW) this will be

$$M^2 = \left(m^W\right)^2 + \left(m^\ell\right)^2 - 2m^W E^\ell, \tag{6.10}$$

where E^ℓ is the energy of the lepton in the RRFW. It can be shown that $E^\ell \in \left[m^\ell, \frac{\left(m^W\right)^2 + \left(m^\ell\right)^2}{2m^W}\right]$, and consequently $M \in \left[0, m^W - m^\ell\right]$.

Therefore, the tree neutrinos in a tauonic event can have any invariant mass with the upper limit of $\sim$ 80 GeV which certainly harms the reconstruction of $t\bar{t}$ pair with the KLFitter. The issue of my study was whether the tauonic events should be or should not be classified as a part of the single lepton channel. Although the reconstruction can be worse on these events, by excluding them from the single lepton channel, the statistics is smaller which can harm the analysis. If the tauonic events would be excluded, they should be treated as a background to the newly defined single lepton channel.

I prepared migration matrices for $p_{\mathrm{T}}^{\mathrm{top}}$ and reconstructed $p_{\mathrm{T}}^{\mathrm{top}}$ spectra for single lepton channel with tauonic events (SL with Tauon), without tauonic events (SL without Tauon) and only for tauonic events (Tauon). I am also showing results for reconstructed $p_{\mathrm{T}}^{\mathrm{top}}$ in dilepton channel to show that the reconstruction of $t\bar{t}$ pairs in this channel is not very good. This is understandable as the two neutrinos from different W are combined into one, and used in the KLFitter.

The migration matrix for $p_{\mathrm{T}}^{\mathrm{top}}$ in dilepton channel cannot be obtained as it is not clear the matching between reconstructed two tops and truth tops (the truth tops are not assigned as leptonic and hadronic because both tops are leptonic).

In Figures 6.5-6.7, one can see that the migration matrices look very similarly for the three channels SL with Tauon, SL without Tauon and Tauon. The correlation for Tauon channel is only slightly worse then the correlation for SL without Tauon channel. Interesting is the fraction of tauonic events in SL with Tauon channel after e+jets selection cuts which is only $\sim 4.4\%$. Using branching ratios for W and τ decay, it can be computed that the fraction of tauonic events where the lepton in SL with Tauon channel is electron on truth level before any cuts is $\sim 15.7\%$. This means that the selection cuts have much smaller efficiency on tauonic events which can be caused by different MET and lepton p_{T} spectrum. One can also inspect the transition matrices in Figures 6.5-6.7. The SL without Tauon transition matrix has higher diagonal elements than the Tauon transition matrix. The first diagonal element is $\sim 14\%$ higher, the second is $\sim 7\%$ higher and with increasing $p_{\mathrm{T}}^{\mathrm{top}}$, this relative size descends. This fact implies that the reconstruction of tauonic events is less efficient. Although by comparing transition matrices for SL with Tauon and SL without Tauon channel, one can see maximally 0.5% higher diagonal elements due to low fraction of tauonic events in SL with tauon channel.

The truth $p_{\mathrm{T}}^{\mathrm{top}}$ spectra look quite differently for Tauon and SL without Tauon channels, see left Figure 6.8. This is caused by different $p_{\mathrm{T}}^{\mathrm{top}}$ dependence of selection efficiency for these two channels. The spectra for SL with Tauon and SL without Tauon channel are very similar, because the tauonic events have small weight as the fraction of tauonic events is small.

Similarly, the reco $p_{\mathrm{T}}^{\mathrm{top}}$ spectra look quite differently for Tauon and SL without Tauon channels, see right Figure 6.8. This is caused by different $p_{\mathrm{T}}^{\mathrm{top}}$ dependence of selection efficiency for these two channels, but in addition the effect of different $t\bar{t}$ reconstruction efficiencies adds.

The crucial plots are in Figure 6.9 where the ratios of reco and truth spectra are shown. This ratio shows how the reconstruction performed. If the ratio of reco and truth spectra are close to one, it means that the reco spectrum is close to the truth spectrum. One can see that the reco spectrum is worse for Tauon channel than for SL without Tauon. Overall, there is no large difference between SL with Tauon and SL without Tauon channels due to low statistics of tauonic events. Just a slight improvement for SL without Tauon is visible.

From the figures, it is also clear that the reconstruction of $p_{\mathrm{T}}^{\mathrm{top}}$ in the dilepton channel is very poor and this is the reason why this channel is assumed as background to the $t\bar{t}$ signal.

The decision of the Group after this study was to use the tauonic events in the definition of the single lepton channel. This decision is based on the small differences between migration matrices and ratios of truth and reco $p_{\mathrm{T}}^{\mathrm{top}}$ spectra for SL with tauon and SL without Tauon channels. Although, one should also try to use the unfolding separately for SL with Tauon and SL without Tauon channels. Only after that, a final conclusion can be derived. In general, the more diagonal is the transition matrix, the more effective is the unfolding procedure. The only advantage of using the tauonic events in the single lepton channel is that the statistical error of data is smaller. Therefore in the following part of my

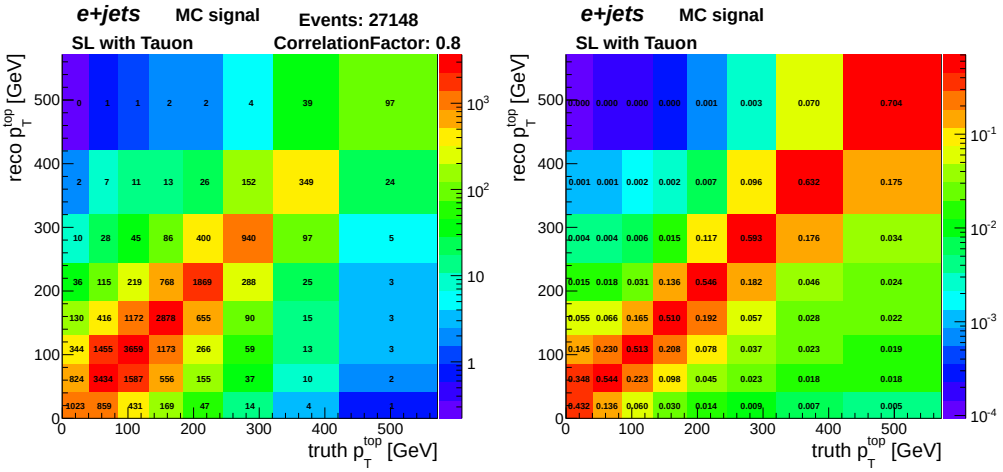


Figure 6.5: Migration and transition matrix for the SL with Tauon channel on the left and on the right, respectively.

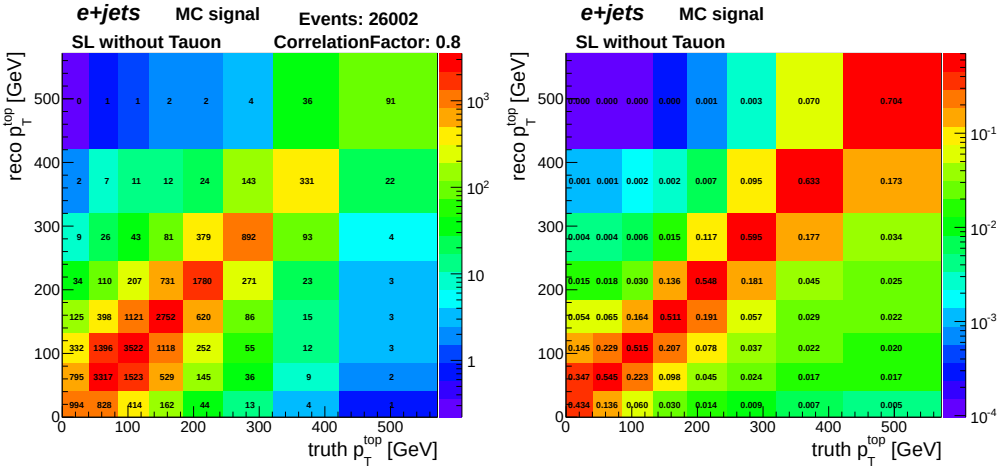


Figure 6.6: Migration and transition matrix for the SL without Tauon channel on the left and on the right, respectively.

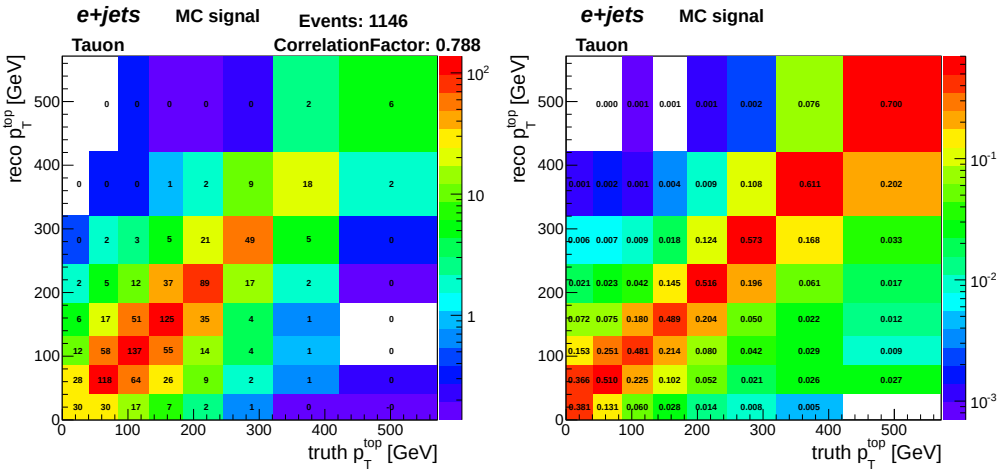


Figure 6.7: Migration and transition matrix for the Tauon channel on the left and on the right, respectively.

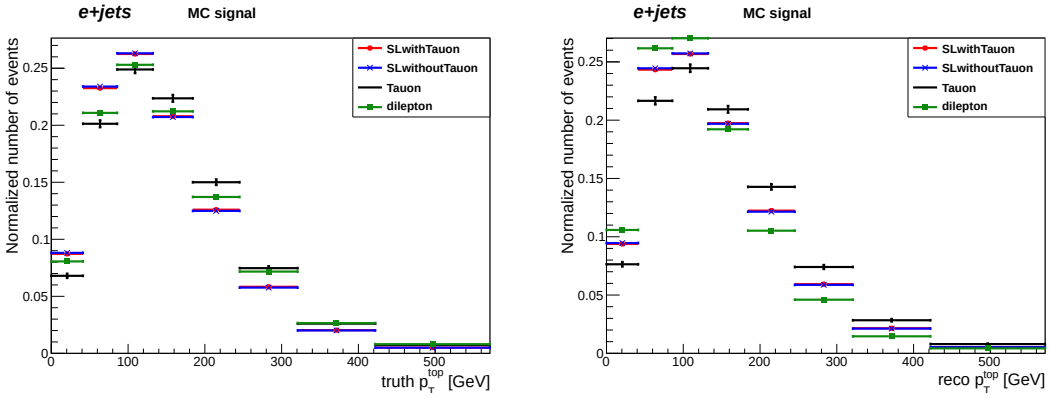


Figure 6.8: Truth and reco $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum for individual channels after $t\bar{t}$ reconstruction on the left and on the right, respectively.

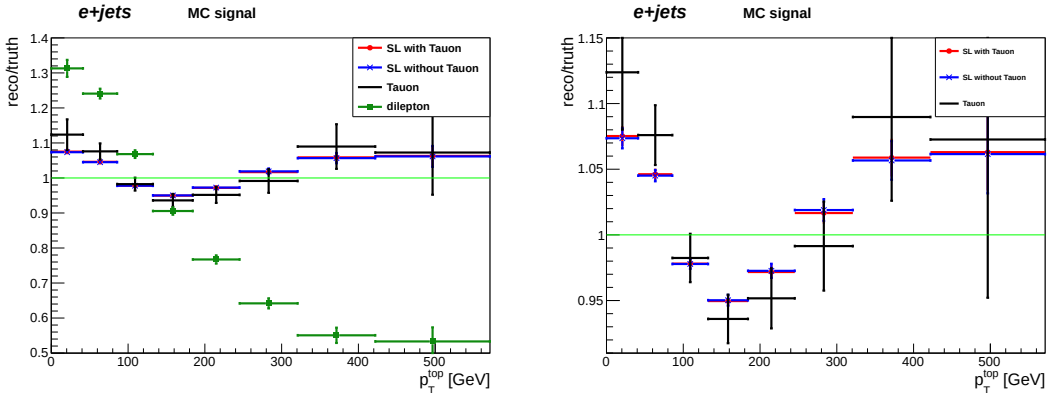


Figure 6.9: Ratio of reco to truth $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum for individual channels after $t\bar{t}$ reconstruction on the left. On the right, the detail is shown without dilepton channel.

analysis, the single lepton channel contains tauonic events.

6.3.3 Rejecting Poorly Reconstructed Events

There are two main signals that the $t\bar{t}$ pair was not reconstructed properly. The first is that the reconstructed top quark masses (one or both) are far away from the pole mass $m_{\mathrm{MC}}^{\mathrm{top}}$. The second is that the maximized likelihood L for the best combination is lower than for other reconstructed $t\bar{t}$ pairs. I reject these types of events and study if the migration matrix and the reco spectrum improve.

The most of the reconstructed top masses have the value of $m_{\mathrm{MC}}^{\mathrm{top}}$ in terms of the used binning, see Figure 6.10. There is no sign of Breit-Wigner shape of the reco m^{top} in contrast with the Breit-Wignericaly distributed truth m^{top} spectrum. I integrated that in range $[132, 214]$ GeV, all truth top masses are present, but only $\sim 93\%$ of all reco top masses are present in this range. Therefore, events where at least one reco top mass is outside this range were certainly not reconstructed properly (although the reco $p_{\mathrm{T}}^{\mathrm{top}}$ can be correct in terms of the used binning). I apply cut to select out events with reco m^{top} tightly around $m_{\mathrm{MC}}^{\mathrm{top}}$.

There is an explanation for the weird shape of the reco m^{top}. One can see that tightly around highly populated bin containing $m_{\mathrm{MC}}^{\mathrm{top}}$, there are bins with very low number of events and at 150 GeV and 190 GeV this number significantly rise. This can be caused by terms $\mathrm{BW}(M_{\mathrm{inv}}(P^{j1}+P^{j2}+P^{hb})|m^{\mathrm{top}}, \Gamma^{\mathrm{top}})$ and $\mathrm{BW}(M_{\mathrm{inv}}(P^{\ell}+P^{\nu}+P^{lb})|m^{\mathrm{top}}, \Gamma^{\mathrm{top}})$ in the likelihood function 6.4. Around the $m_{\mathrm{MC}}^{\mathrm{top}}$, the Breit-Wigner rises or declines rapidly and therefore by small variation of $M_{\mathrm{inv}}(P^{j1}+P^{j2}+P^{hb})$ or $M_{\mathrm{inv}}(P^{\ell}+P^{\nu}+P^{lb})$ during the change of parameters the likelihood function changes a lot. For masses more distant from $m_{\mathrm{MC}}^{\mathrm{top}}$, the small change in invariant mass during the change of parameters does not change the value of these Breit-Wigner terms a lot. This means that once during the maximization, the invariant mass is in a close region around $m_{\mathrm{MC}}^{\mathrm{top}}$ the maximization process quickly pushes this invariant mass to $m_{\mathrm{MC}}^{\mathrm{top}}$. If this invariant mass is far away from $m_{\mathrm{MC}}^{\mathrm{top}}$, the maximization process need not change this value a lot as it has no large effect on the value of likelihood.

In Figure 6.11, there is the distribution of obtained likelihood values. It has a high peak around $\log L = -45$ and a tail for $\log L < -50$. It can be assumed that the reconstruction performed less efficiently for events from this tail, but it is hard to say what is the exact origin of this tail. I apply cut on this events.

I define this terminology:

- top mass cut – event passes if both reconstructed top masses are in the interval $[m_{\mathrm{MC}}^{\mathrm{top}} - 3\ \mathrm{GeV}, m_{\mathrm{MC}}^{\mathrm{top}} + 3\ \mathrm{GeV}]$,
- likelihood cut – event passes if for the maximized likelihood L for the best combination holds $\log L > t_{\mathrm{lhood}}$, I used value $t_{\mathrm{lhood}} = -50$.
- good mass events – events which passed the top mass cut,
- bad mass events – events which not passed the top mass cut,
- good likelihood events – events which did not pass the likelihood cut,
- bad likelihood events – events which did not pass the likelihood cut.

I inspect what is the correlation between the likelihood and top mass cut performance. In Figure 6.12, one can see the connection between reconstructed top mass and likelihood. Events with $\log L > -50$ have reconstructed top masses at $m_{\mathrm{MC}}^{\mathrm{top}}$. From Figure 6.13, one can see a large correlation between likelihood and top mass cut. 99.99% of all events which passed the likelihood cut also passed the top mass cut. Therefore applying both cuts simultaneously has no meaning as it has the same effect as applying only the likelihood cut. Also the cut efficiencies can be computed. Cut efficiency for top mass cut is 85% and for likelihood cut is 75%.

The comparison of migration matrices for good mass events and bad mass events is in Figure 6.14 (the migration matrix without this cut was already shown in Figure 6.5). The correlation of migration matrix for good mass events is much higher than for bad mass events. The comparison of transition matrices is in Figure 6.15. The diagonal elements are much higher for good mass events which means that for good mass events, the reconstructed $p_{\mathrm{T}}^{\mathrm{top}}$ is more often correct in terms of the used binning. The same conclusions hold for good and bad likelihood events where the correlation and diagonal elements in transition matrix are even higher for good likelihood events, see Figure 6.16 and 6.17. It looks that both cuts improve the migration matrices of $p_{\mathrm{T}}^{\mathrm{top}}$ which can help in unfolding, but one should also compare the reconstructed and truth $p_{\mathrm{T}}^{\mathrm{top}}$ spectra with these cuts.

In Figure 6.18, the ratio of reco and truth spectra are shown. Both spectra were obtained after the same event selection. One can see that the reco $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum which best corresponds to the truth $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum is the one without any cut. This fact can impeach the goodness of top mass or likelihood cut for $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum. The final decision of using one of these cuts can be done after observing the performance of the unfolding.

The usage of top mass or likelihood cuts has much visible effect on $M^{t\bar{t}}$ spectrum, see Figure 6.19. Both cuts strongly improve the difference between reco and truth $M^{t\bar{t}}$ spectrum.

Except the enhancement of the resolution of measured spectra, the top mass or likelihood cuts can also enhance the purity of the data sample, because the cut efficiency on background events should be lower (KLFitter can perform worse as in background are no real $t\bar{t}$ events). However, the negative effect of using these cuts is the lower statistics of $t\bar{t}$ events in data sample. Ideally, one could also optimize the value of likelihood cut t_{lhood}.

There is also a suggestion for improving the $t\bar{t}$ reconstruction method. From Figure 6.10, it is clear that one cannot really expect a Breit-Wigner shape of reco m^{top}. Although, the $t\bar{t}$ reconstruction method assumes that the top mass has Breit-Wigner shape. For this reason, it would be interesting to change the values Γ^{top} in likelihood function 6.4 to very small non-physical values. With this change, the reconstructed top mass will be always forced to be very close to $m_{\mathrm{MC}}^{\mathrm{top}}$ which can help in the reconstruction of $p_{\mathrm{T}}^{\mathrm{top}}$. This suggestion should be tested in the future.

In the following part of my analysis, no top mass or likelihood cut was applied.

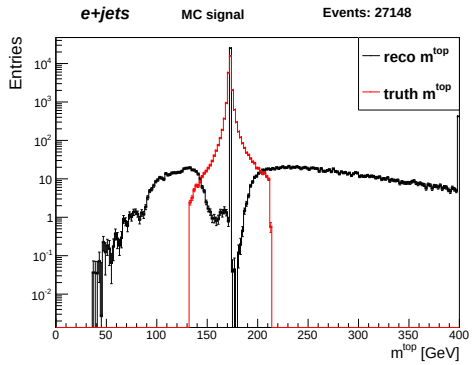


Figure 6.10: Reco and truth top quark mass m^{top} from MC. Both leptonic and hadronic top masses are filled in histogram. The width of the bins is 2 GeV. The most populated bin is $[172, 174]$ GeV. The used pole mass in MC was 172.5 GeV.

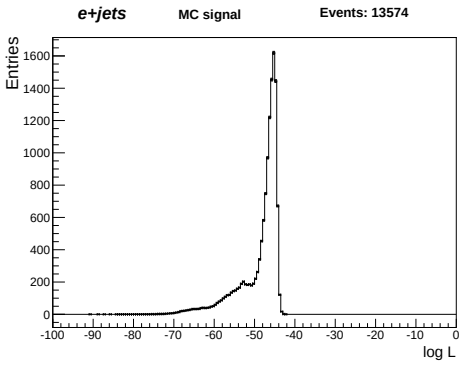


Figure 6.11: The logarithm of maximal likelihood L from best combination of jets.

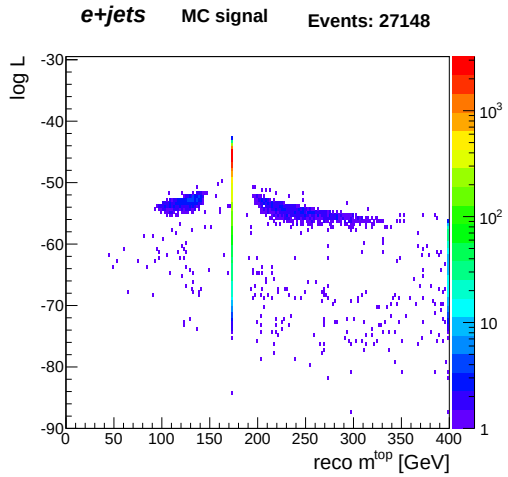


Figure 6.12: Logarithm of likelihood versus the reconstructed top mass. For each event, the histogram is filled by both leptonic and hadronic top masses.

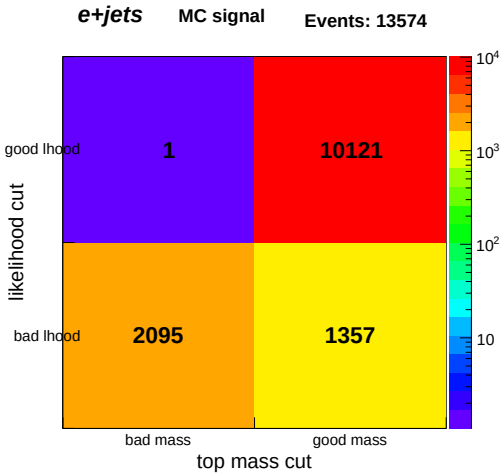


Figure 6.13: Correlation between likelihood cut and top mass cut. In the bins, there are the number of events with good or bad mass versus good or bad likelihood (lhood).

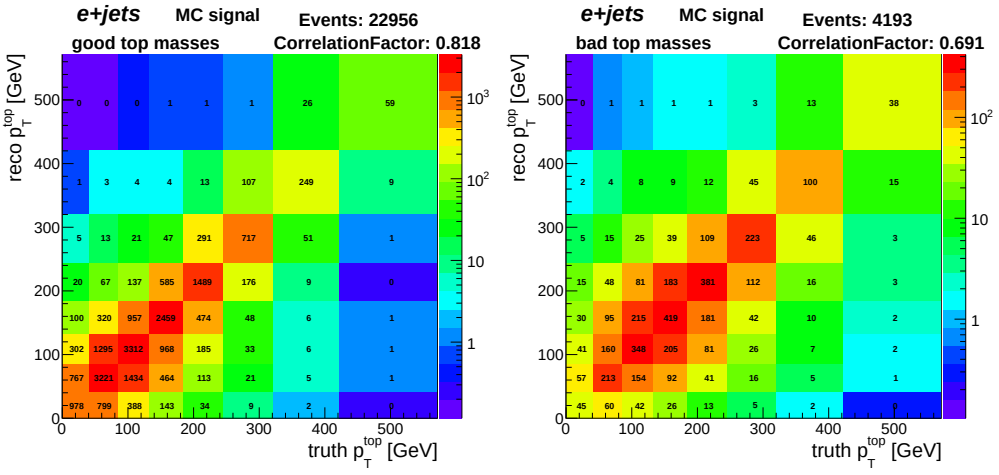


Figure 6.14: Migration matrices for $p_{\mathrm{T}}^{\mathrm{top}}$ for good and bad masses events on the left and right, respectively. The correlation between reco and truth $p_{\mathrm{T}}^{\mathrm{top}}$ for good mass events is 0.82 and for bad mass events it is 0.69.

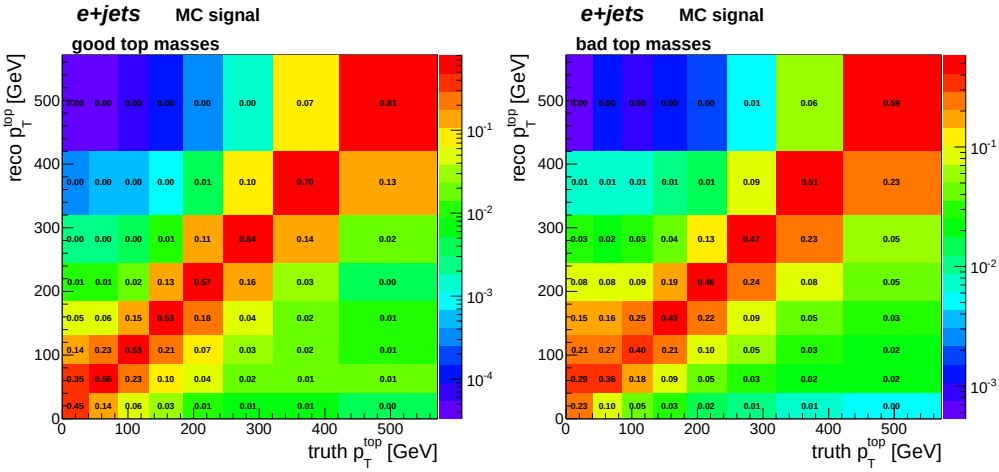


Figure 6.15: Transition matrices for $p_{\mathrm{T}}^{\mathrm{top}}$ for good and bad masses events on the left and right, respectively.

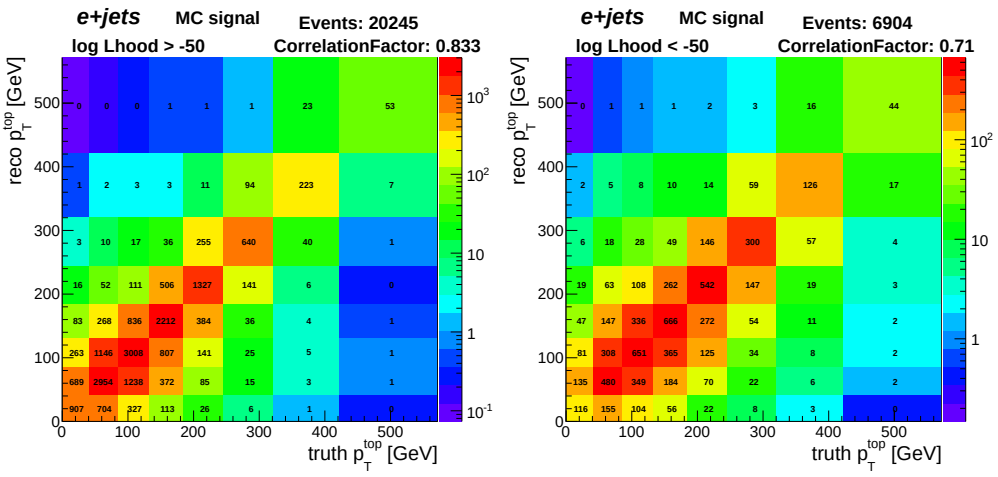


Figure 6.16: Migration matrices for $p_{\mathrm{T}}^{\mathrm{top}}$ for good and bad likelihood events on the left and right, respectively. The correlation between reco and truth $p_{\mathrm{T}}^{\mathrm{top}}$ for good likelihood events is 0.83 and for bad likelihood events 0.71.

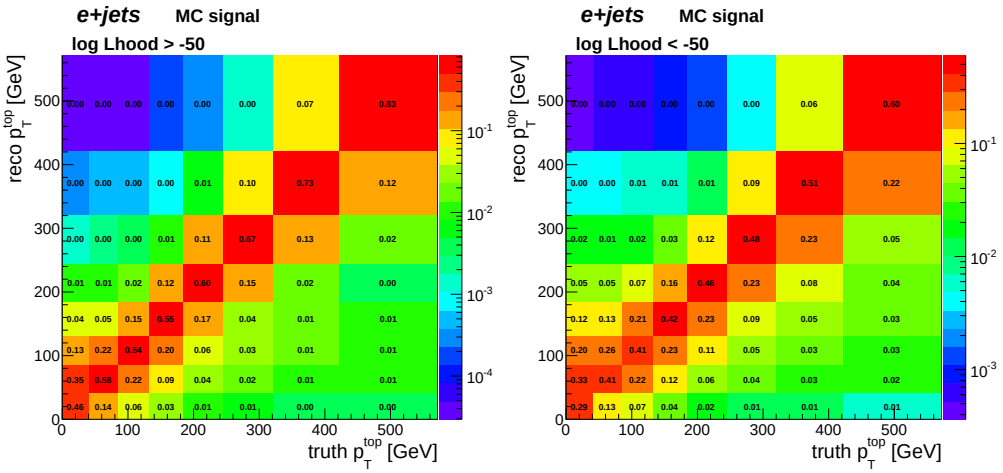


Figure 6.17: Transition matrices for $p_{\mathrm{T}}^{\mathrm{top}}$ for good and bad likelihood events on the left and right, respectively.

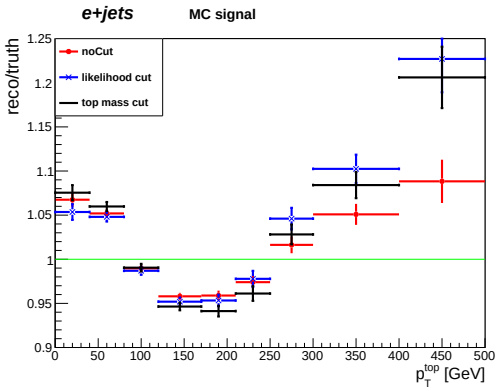


Figure 6.18: Ratio of reco over truth $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum for good mass, good likelihood and all events.

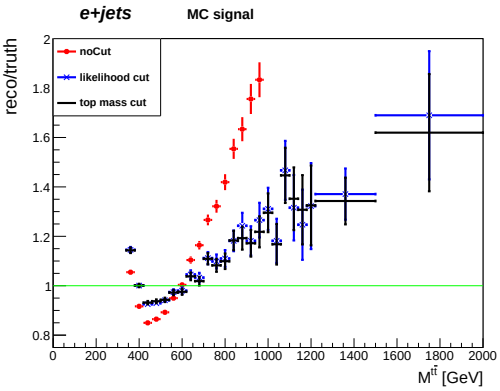


Figure 6.19: Ratio of reco over truth $M^{t\bar{t}}$ spectrum for good mass, good likelihood and all events. For $M^{t\bar{t}} >$ 1000 GeV, this ratio is higher than two.

6.4 Control Plots

After event selection, one can plot various distributions e.g. MET, lepton p_{T}, p_{T} of highest p_{T} jets, azimuthal angle,... This can be done for data and also for the sum of all backgrounds and MC signal, which is a prediction for data. In case of perfect simulation and treatment of MC and no new physics in data, the control plots should show agreement between data and prediction.

I am showing only control plots for lepton p_{T} and MET in Figure 6.20. The total number of events for prediction is overestimated by 5.6% which is not in range of statistical errors. There are also bins with 10% differences between data and prediction (neglecting the higher values of spectra where the statistical errors are high). To resolve the significance of this small disagreement, the systematic errors have to be determined.

After the $t\bar{t}$ pairs reconstruction, the control plots can be also displayed for distributions closely connected with $t\bar{t}$ pair like azimuthal angle of top quarks, η of top quarks, η of $t\bar{t}$ pairs and also the measured spectra $p_{\mathrm{T}}^{\mathrm{top}}$, $M^{t\bar{t}}$ or $p_{\mathrm{T}}^{t\bar{t}}$. I am showing the $p_{\mathrm{T}}^{\mathrm{top}}$ and $M^{t\bar{t}}$ spectra in Figures 6.21.

6.5 Unfolding

The unfolding corrects the change of $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum caused by the resolution of detector and imperfect $t\bar{t}$ pair reconstruction. I define the following N_{bins}-dimensional vectors:

- r_{MC} – element $r_{\mathrm{MC}i}$ is the number of entries in in i-th bin of reco $p_{\mathrm{T}}^{\mathrm{top}}$ histogram from MC signal.
- t_{MC} – element $t_{\mathrm{MC}i}$ is the number of entries in i-th bin of truth $p_{\mathrm{T}}^{\mathrm{top}}$ histogram from MC signal. This truth $p_{\mathrm{T}}^{\mathrm{top}}$ histogram is obtained after same event selection as reco $p_{\mathrm{T}}^{\mathrm{top}}$ histogram.
- r_{data} – element $r_{\mathrm{data}i}$ is the number of entries in i-th bin of reco $p_{\mathrm{T}}^{\mathrm{top}}$ histogram in data. For real data, $r_{\mathrm{data}i} = D_i - BG_i$.
- t_{data} – element $t_{\mathrm{data}i}$ is the number of entries in i-th bin of truth $p_{\mathrm{T}}^{\mathrm{top}}$ histogram in data. For real data, this vector is not known.
- u - element u_i is the number of entries in i-th bin of the unfolded $p_{\mathrm{T}}^{\mathrm{top}}$ histogram. The goal of the unfolding is to obtain this vector from vector r_{data} as close as possible to the vector t_{data}.

There are several unfolding methods, of which the simplest method is called bin-by-bin correction. It is based on the knowledge of relative change $r_{\mathrm{MC}i}/t_{\mathrm{MC}i}$ of detector effects in i-th bin which is applied on data:

$$t_{\mathrm{data}i} = r_{\mathrm{data}i}\frac{t_{\mathrm{MC}i}}{r_{\mathrm{MC}i}}. \tag{6.11}$$

I have showed the ratio $r_{\mathrm{MC}i}/t_{\mathrm{MC}i}$ in Figure 6.18. This method is dependent on the truth $p_{\mathrm{T}}^{\mathrm{top}}$spectrum in MC, and therefore, it is hard to see effects of new

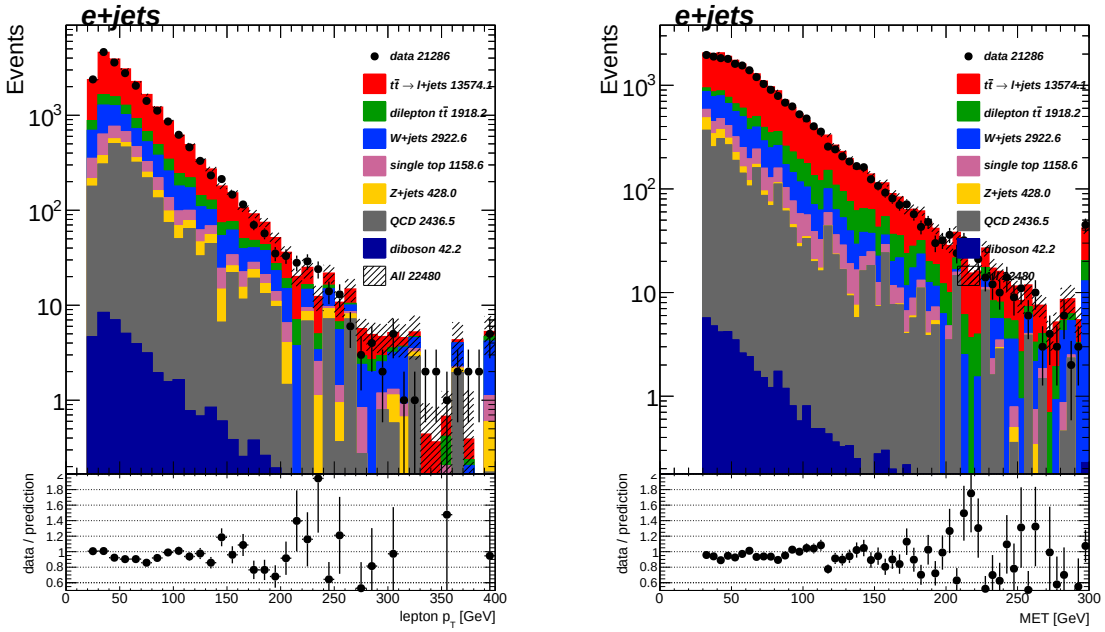


Figure 6.20: Distributions of lepton p_{T} and MET of data with prediction on the left and right, respectively. The prediction is a stack of all backgrounds and MC signal. All simulated samples are normalized to the luminosity of data. Only the statistical errors are displayed.

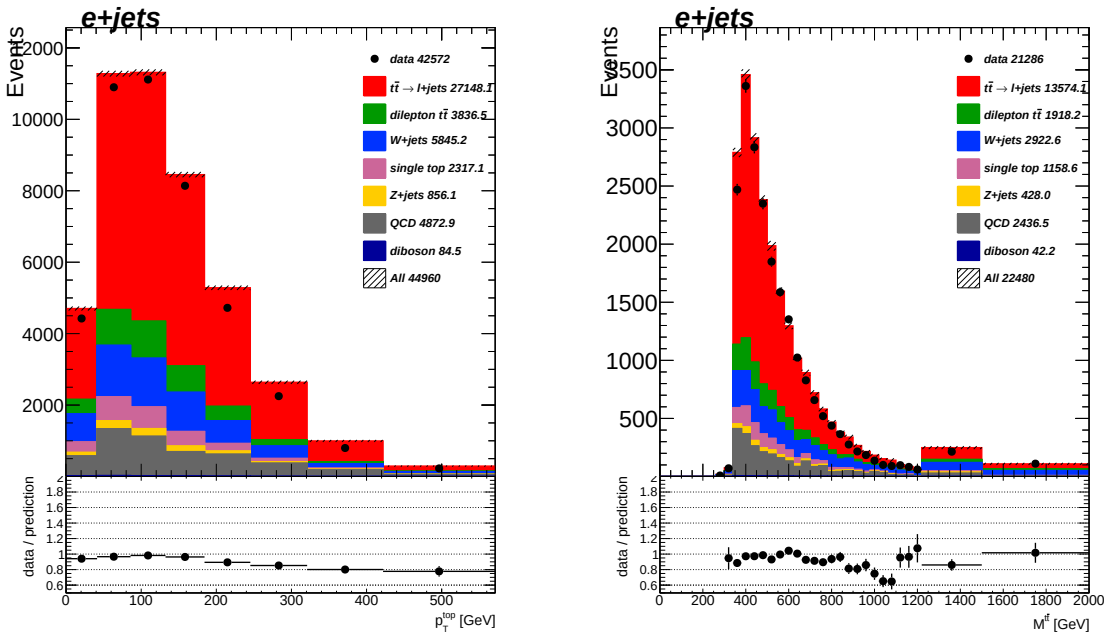


Figure 6.21: Reconstructed $p_{\mathrm{T}}^{\mathrm{top}}$ and $M^{t\bar{t}}$ of data with prediction on the left and right, respectively. The prediction for the reconstructed spectrum is a stack of all backgrounds and MC signal. All simulated samples are normalized to the luminosity of data. Only the statistical errors are displayed.

physics in data. Moreover, it neglects the correlations between adjacent bins. Therefore, the bin-by-bin method is insufficient for realistic unfolding.

In next, I introduce the matrix inversion method. I denote the transition matrix as M. The element in i-th array and j-th column M_{ij} shows the fraction of events which were originally in i-th bin of truth spectrum and ended in j-th bin of reco spectrum. I am reminding that the evaluation of M_{ij} does not depend on the truth spectrum in MC signal. From the definition of transition matrix M, it certainly holds:

$$r_{\mathrm{MC}i} = \sum_{j=1}^{N_{\mathrm{bins}}} M_{ji} t_{\mathrm{MC}j}. \tag{6.12}$$

I define $P(\text{reco } i|\text{truth } j)$ as the conditional probability that if the truth $p_{\mathrm{T}}^{\mathrm{top}}$ of a top quark is in j-th bin then the reco $p_{\mathrm{T}}^{\mathrm{top}}$ of this top quark is in i-th bin. For the expected reconstructed $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum obtained in data, it holds:

$$r_{\mathrm{data}i}^{\mathrm{exp}} = \sum_{k=1}^{N_{\mathrm{bins}}} P(\text{reco } i|\text{truth } k) t_{\mathrm{data}k}. \tag{6.13}$$

I will assume that this expected number is what was really obtained from the measurement:

$$r_{\mathrm{data}}^{\mathrm{exp}} \approx r_{\mathrm{data}}, \tag{6.14}$$

which should hold in case of high statistics of data.

One of the most important assumptions for most unfolding methods and also for matrix inversion method is

$$M_{ji} \approx P(\text{reco } i|\text{truth } j), \tag{6.15}$$

which is sufficiently fulfilled in case of high statistics of MC sample. Also, there should be no difference in detector resolution and efficiency of $t\bar{t}$ pairs reconstruction in MC and data events.

Using the last two assumptions, one gets

$$r_{\mathrm{data}i} = \sum_{k=1}^{N_{\mathrm{bins}}} M_{ji} t_{\mathrm{data}j}. \tag{6.16}$$

If the inverted matrix M^{-1} exists, then this equation can be multiplied by this matrix M^{-1} and the truth spectrum is obtained. Therefore, the unfolding consists of multiplication by inverted matrix:

$$u_i = \left(M^{-1}\right)_{ji} r_{\mathrm{data}j}. \tag{6.17}$$

If the matrix M is singular, the inversion does not exist and this method cannot be used. Also from a probabilistic point of view, the matrix M^{-1} is not a transition matrix and it can contain negative elements. These negative elements can yield negative elements of the unfolded vector u in case of large statistical fluctuations of vector r_{data}. The advantage of this method is the independence on the truth $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum in MC. The matrix inversion method has its justification based on two assumptions in Equations 6.14 and 6.15. The first depends only on the statistics of the data. The second depends on the statistics of the simulation, but

it can depend also on the simulation model (e.g. two different parton showering models can lead to different transition matrices). The simulation model should be as realistic as possible. The assumption 6.15 can be verified by comparing the transition matrices from MC signal samples with different simulation models.

I studied this method in MC with the goal to check the numerical stability and statistical fluctuations. At the end of this section, I apply this method to measured data.

I divide the MC signal sample into two halves with a similar statistics. One half is used to obtain the transition matrix M. The second half is for testing the unfolding method and I call it an ensemble. It is divided into 13 subsamples which are used to test this method. One subsample has statistics similar to data (corresponding to luminosity 4.7 fb^{-1}). Each subsample or the whole ensemble can be handled as a pseudo-experiment. In contrast with real data sample, I have both r_{data} and t_{data} in each subsample which allows me to control the performance of the unfolding. The truth spectrum t_{data} in each subsample can have different shape due to statistical fluctuations.

The migration and transition matrices are shown in Figure 6.22. One can see that the fraction of events with correctly reconstructed $p_{\text{T}}^{\text{top}}$ are between 44%-70%, see the diagonal of the transition matrix M. The conspicuous rise of the reconstruction efficiency with $p_{\text{T}}^{\text{top}}$ can be caused by the rise of the bin size. The inversion of the transition matrix M^{-1} is shown in Figure 6.23, together with the matrix $A = M \cdot M^{-1}$ to crosscheck the correctness of the inversion. The inversion succeeded because the rounded values of diagonal elements are 1 and the absolute values of off-diagonal elements of matrix A are $< 10^{-15}$. It is interesting to examine the inverted matrix. Roughly, sum of elements in each row gives 1. On the diagonal, there are the highest values of ~ 2 and the other elements with larger absolute values are right next to diagonal elements and they are negative. Therefore, the unfolding of reco spectrum can be roughly described as follows: each bin is multiplied by ~ 2 and a fraction of neighboring bins are subtracted.

After multiplication of matrix M^{-1} with the vector containing the reco spectrum, I get the unfolded spectrum. The question is how to determine the uncertainties of this unfolded spectrum. There is no easy way how to compute the uncertainties for the elements of matrix M^{-1}. But I can take into account the statistical uncertainties $\sigma r_{\text{data}i}$ of reco spectrum. I computed the uncertainties σu_i by the use of error propagation formula, because each bin is a linear combination of all bins from the reco spectrum:

$$\sigma u_i = \sqrt{\sum_{i=1}^{N_{\text{bins}}} \left(\left(M^{-1}\right)_{ij} \sigma r_{\text{data}i} \right)^2}. \tag{6.18}$$

The use of this formula is not entirely correct, because the bins of reconstructed spectrum are correlated between each other, but I will use this formula for evaluating the uncertainties of unfolded spectrum.

The obtained spectrum after unfolding is assumed as the measured spectrum before event selection correction. The authenticity of this measured spectrum depends on the detector, efficiency of $t\bar{t}$ reconstruction and success of the unfolding. To quantify the authenticity of the measured spectra in MC subsamples, I define

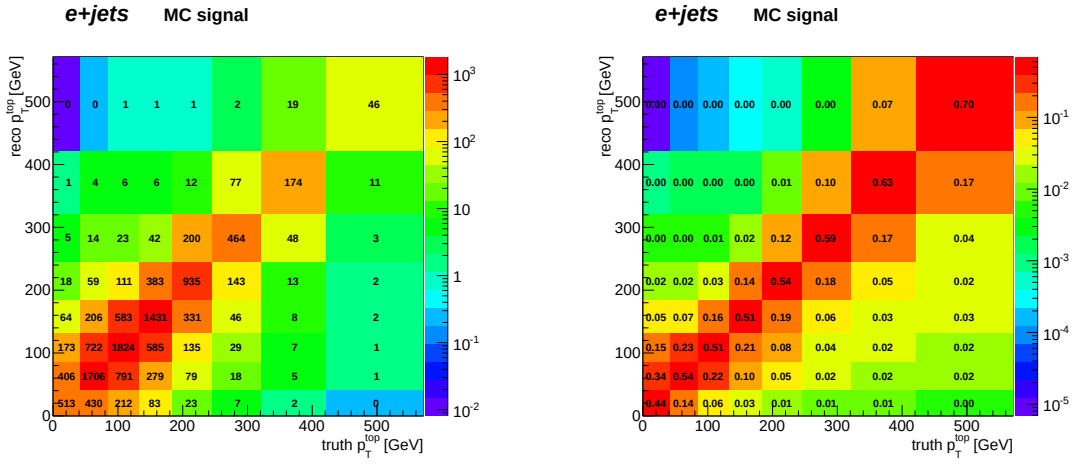


Figure 6.22: Migration matrix obtained from half of MC signal and its corresponding transition matrix M on the left and right, respectively.

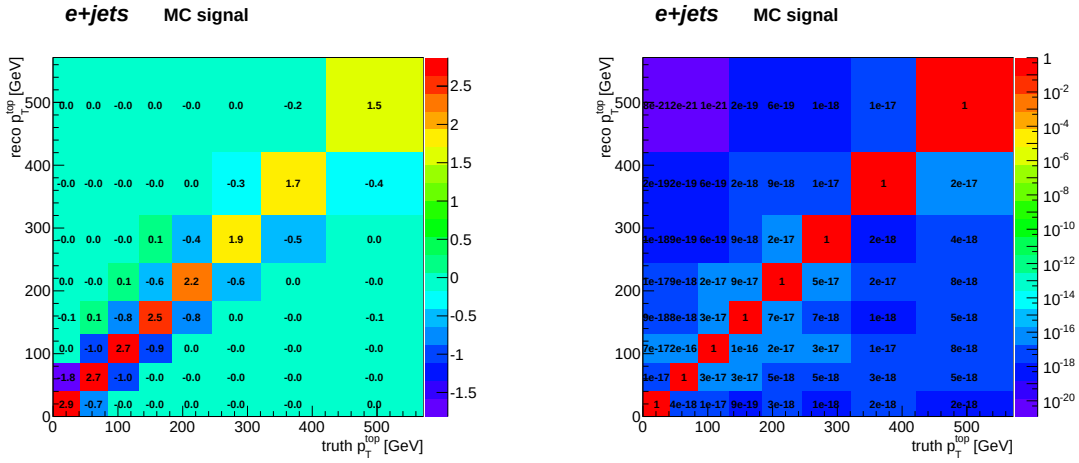


Figure 6.23: Inversion of transition matrix M^{-1} and matrix A which is multiplication if two matrices M and M^{-1} on the left and right, respectively. In matrix A, the off-diagonal elements have absolute values to show the plot in logarithmic scale.

value χ^2 as

$$\chi^2 = \sum_{i=1}^{N_{\text{bins}}} \frac{(u_i - t_{\text{data}i})^2}{(\sigma u_i)^2}. \tag{6.19}$$

Assuming that u_i are N_{bins} independent random variables Gaussialy distributed with mean values $t_{\text{data}i}$ and variances $(\sigma u_i)^2$, the χ^2 is a random variable with Chi-squared distribution with parameter number of degrees of freedom (NDF) equal to N_{bins}. This random variable has mean value equal of NDF. Therefore, the closer is the value χ^2/NDF to 1, the probable is the correctness of the measurement (reconstruction+unfolding+determination of uncertainties). Also, the value χ^2 is the measure of goodness of the measured spectrum. The lower the χ^2, the closer is the measured spectrum to the truth spectrum. But if the $\chi^2/\text{NDF} \ll 1$, the measurement is not probable, being a lucky fluctuations.

Firstly, I use the matrix M^{-1} to unfold the reconstructed spectrum from the sum of ensemble. In Figure 6.24, one can see the comparison of truth, reco and unfolded spectra. The reco spectrum has significantly different shape as the truth spectrum and the unfolded spectrum is much closer to the truth spectrum which emphasizes the importance of the unfolding. The improvement is also visible in right Figure 6.24 where the unfolded spectrum is much closer to truth spectrum for low $p_{\text{T}}^{\text{top}}$. Although, for higher $p_{\text{T}}^{\text{top}}$ the unfolding has no such distinct effect.The value $\chi^2/\text{NDF} \approx 1.4$ which implies sizable probability of the measurement. All in all, it seems that the matrix inversion method succeeded. But this test sample has 13-times higher statistics as the statistics of data.

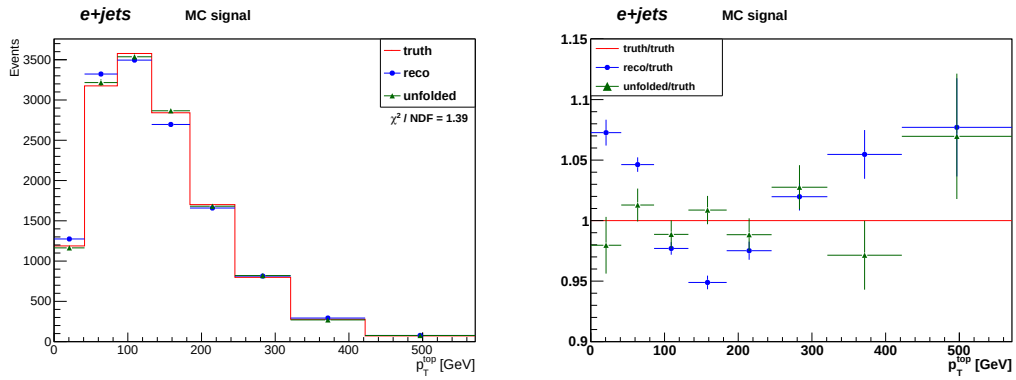


Figure 6.24: Truth, reco and unfolded $p_{\text{T}}^{\text{top}}$ spectra obtained from the half of MC signal sample. The used transition matrix was obtained from second half of MC signal sample. On the right are these spectra divided by the truth spectrum.

Secondly, I use the matrix M^{-1} to unfold the reconstructed spectrum for each subsample to illustrate the possible outcome of a real experiment. I am showing few examples, see Figure 6.25. These examples are the extreme cases of what can happen. The first two subsamples are examples where the unfolded spectrum is very close to the truth spectrum. The next two subsamples are examples where the unfolded spectrum is far away from the truth spectrum. This fact is also expressed by the relative size of the χ^2/NDF: for the first two unfolded spectra the χ^2/NDF is much lower then for the second two unfolded spectra. One can conclude that the outcome can vary a lot and sometimes the unfolding can give biased results. This is a negative feature of the matrix inversion method.

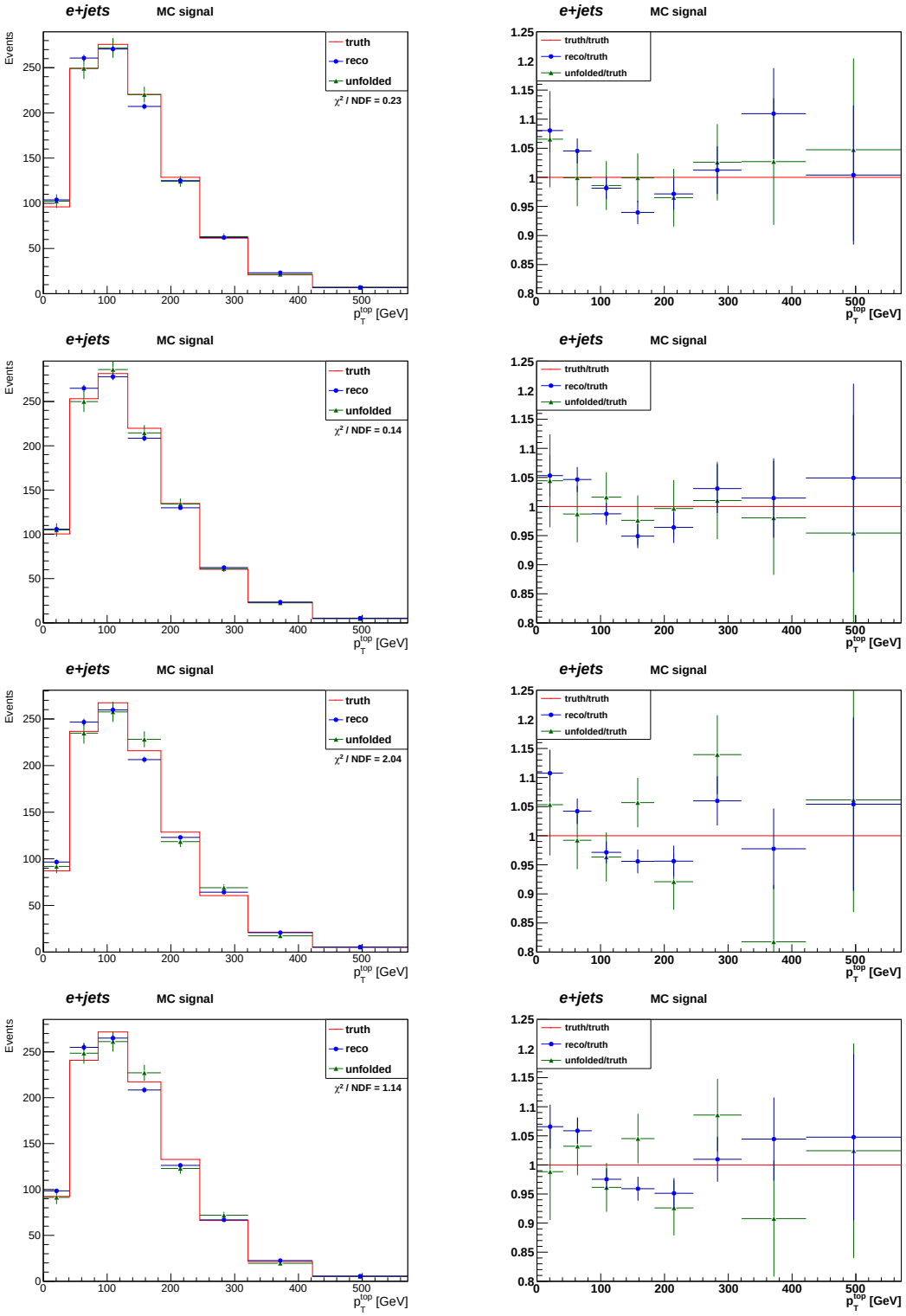


Figure 6.25: Truth, reconstructed and unfolded $p_{\mathrm{T}}^{\mathrm{top}}$ spectra for chosen subsamples and their ratio to truth spectrum on the left and right, respectively. The upper two rows are examples of best outcome and the bottom two rows are examples of worst outcome.

The unfolded spectra of all subsamples are shown in Figure 6.26. One can see a large spread in the unfolded spectra. Although each subsample has its own truth spectrum (due to statistical fluctuations), it is useful to compare the unfolded results to the mean of the truth spectra which is the truth spectrum obtained from all ensemble normalized to the statistics of one subsample. This shows the joint effect of statistical fluctuations and unfolding on the measured spectrum.

Next, I wanted to quantify the spread of the unfolded spectra obtained from all subsamples. For each bin, I computed the mean of the values in the bin and the root mean square of these values. The uncertainties obtained with Equation 6.18 were not taken into account. I define the relative spread of the subsamples as the ratio of computed root mean square and mean. This quantity is plotted in Figure 6.27. The relative spread is smallest ($\sim 3\%$) in the region of $p_{\mathrm{T}}^{\mathrm{top}} \in [40, 190]$ GeV where the unfolding performs the best. One can see that in the last bin the relative spread is much higher than in other bins. It seems that the relative spread rise with the lowering statistics of the bins. Th plot 6.27 should be compared to results from other unfolding methods.

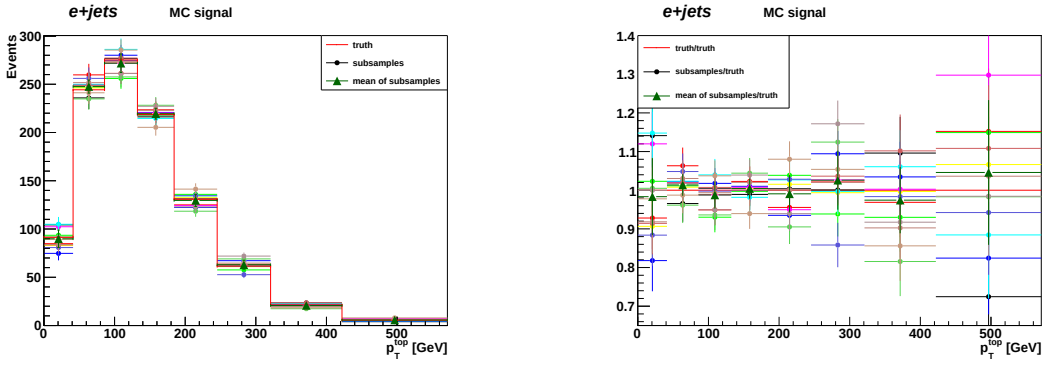


Figure 6.26: Unfolded $p_{\mathrm{T}}^{\mathrm{top}}$ spectra obtained from all simulation subsamples and their mean in comparison with truth $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum from all ensemble normalized to statistics of one subsample (divided by 13). On the right, these spectra are divided by the same truth spectrum.

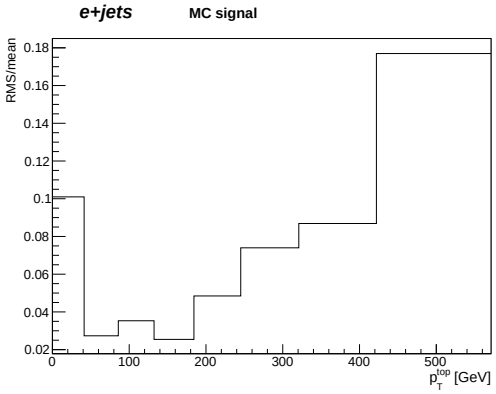


Figure 6.27: Relative spread of the unfolded spectra obtained from all subsamples.

I summarize my findings about the unfolding studies. The performance of the unfolding on the sample with 13-times higher statistics was visibly good. But it is

not known, if it is only a statistical fluctuation. The unfolding on subsamples with statistics of data can lead to very badly measured spectra, but also to very good results. There is quite high relative spread of unfolded spectra for subsamples mainly for higher $p_{\mathrm{T}}^{\mathrm{top}}$ where the statistics of bins is lower. All in all, the matrix inversion method seems unstable due to the low statistics of data. It is necessary to compare the unfolded subsample's spectra with some other unfolding methods which would be less sensitive to these statistical fluctuations.

There are more advanced methods for unfolding e.g. method based on Bayes theorem which iteratively gets to the truth spectrum. I briefly describe this method. For the expected truth spectrum, it holds

$$t_{\mathrm{data}j}^{\mathrm{exp}} = \sum_{i=1}^{N_{\mathrm{bins}}} P(\text{truth } j|\text{reco } i) r_{\mathrm{data}i}, \tag{6.20}$$

where $P(\text{truth } j|\text{reco } i)$ is the conditional probability that if the reco $p_{\mathrm{T}}^{\mathrm{top}}$ of a top quark is in j-th bin then the truth $p_{\mathrm{T}}^{\mathrm{top}}$ of this top quark is in i-th bin. This probability can be obtained also from migration matrix by normalizing the rows, but this probability depends on the initial truth spectrum in MC. This dependence is expressed by the Bayes theorem:

$$P(\text{truth } j|\text{reco } i) = \frac{P(\text{reco } i|\text{truth } j)P(\text{truth } j)}{\sum_{k=1}^{N_{\mathrm{bins}}} P(\text{reco } i|\text{truth } k)P(\text{truth } k)}, \tag{6.21}$$

where $P(\text{reco } i|\text{truth } j)$ can be assumed as the elements of the transition matrix (Figure 6.22) which does not depend on the truth spectrum in MC. $P(\text{truth } j)$ is the probability that the $p_{\mathrm{T}}^{\mathrm{top}}$ of a random top quark belongs to j-th bin of the truth spectrum and in case of high statistics, it can be approximated as

$$P(\text{truth } j) = \frac{t_{\mathrm{data}j}}{\sum_{k=1}^{N_{\mathrm{bins}}} t_{\mathrm{data}k}}. \tag{6.22}$$

After combining the last three equations into one formula, one gets

$$t_{\mathrm{data}j}^{\mathrm{exp}} = \sum_{i=1}^{N_{\mathrm{bins}}} \frac{P(\text{reco } i|\text{truth } j)t_{\mathrm{data}j}}{\sum_{k=1}^{N_{\mathrm{bins}}} P(\text{reco } i|\text{truth } k)t_{\mathrm{data}k}} r_{\mathrm{data}i}. \tag{6.23}$$

In the first step of the unfolding, certain initial t_{data} is proposed, for example the measured r_{data}. This is used in Equation 6.23 to obtain the expected spectrum $t_{\mathrm{data}}^{\mathrm{exp}}$. In the next step, this spectrum is used as t_{data} and a new expected spectrum $t_{\mathrm{data}}^{\mathrm{exp}}$ is obtained. These steps are repeated iteratively until no change is apparent on the expected $t_{\mathrm{data}}^{\mathrm{exp}}$. This result is then assumed as the unfolded spectrum. It can happen that after too much iterations the unfolded spectrum is getting worse. Another important part of the unfolding is the estimation of the uncertainties, for more details about the Bayesian method see [41].

The main difference between the matrix inversion method and the Bayesian method is that the Bayesian method is based on a probabilistic approach while

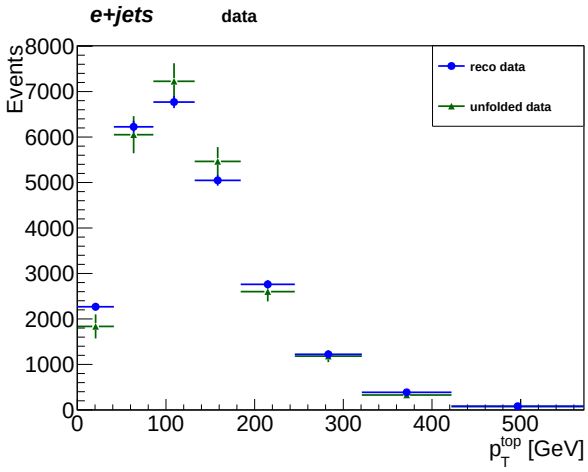


Figure 6.28: The reconstructed $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum obtained after background subtraction from data in Figure 6.21 and the unfolded $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum with matrix inversion method. The errors of unfolded spectrum were computed with Equation 6.18.

the inverted transition matrix has no meaning in probability terms. This is why the Bayesian method can perform better, but the number of iterations has to be tested.

Other unfolding method uses singular value decomposition, see [42].

In spite of the drawbacks of the matrix inversion method, I use this method to unfold the measured data after background subtraction to exercise the machinery of obtaining the differential cross section. In Figure 6.28, the result of unfolding is shown on data.

6.6 The Resulting Differential Cross Section Spectrum

The spectrum $p_{\mathrm{T}}^{\mathrm{top}}$ in Figure 6.28 is the numerator of Equation 6.1, i.e. the unfolded data previously subtracted by the background. Next step is the event selection efficiency correction. In Figure 6.29, the event selection efficiency ϵ is shown. I investigated that the $t\bar{t}$ pair reconstruction efficiency is nearly 100 %, and therefore the selection efficiency is determined mainly by the selection criteria mentioned in Section 6.2. The overall event selection efficiency for e+jets channel is $\sim 4\%$.

The disadvantage of the event selection correction is the dependence on MC as it was shown in my studies in Figure 5.5. In case of other simulation, the distributions like i.e. lepton p_{T} or MET can have slightly different shape and therefore the cut efficiencies would be different. But there is no other way to perform the event selection correction o the full phase space level.

The final $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum is shown in Figure 6.30. I integrated this spectrum to obtain the total cross section just to crosscheck the validity of the measurement. The result (67.4 ± 2.5) pb is in the order of the theoretical MC value of $t\bar{t}$ production in ℓ+jets channel which validates the concept of the measurement. But it has no meaning to compare this number to theoretical value and deduce some

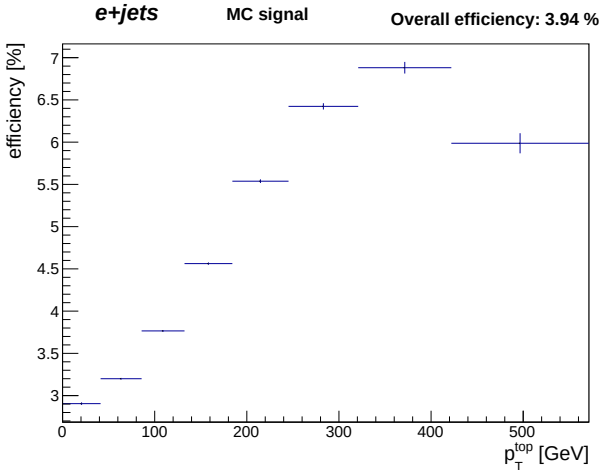


Figure 6.29: The $p_{\mathrm{T}}^{\mathrm{top}}$ dependence of binned efficiency ϵ_i for i-th bin (ratio of truth $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum after $t\bar{t}$ pair reconstruction to the initial truth $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum in ℓ+jets channel) obtained from MC signal sample.

conclusions as, many important steps of the analysis would have to be improved. First of all, the corrections for MC simulation have to be improved to obtain better agreement in control plots 6.20 or 6.21. There is no need for perfect agreement but the overestimation of MC in these plots points to imperfect handling of MC. The unfolding has to be also improved. And the important step of analysis is to quantify the systematic errors. In the same plot with the measured $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum, there is the MC truth $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum. My final measured $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum is in reasonable agreement with this theoretical prediction. Due to reasons mentioned above, it would be premature to deduce any physics conclusions.

In Figure 6.31, there is shown the $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum normalized to one. The systematic error from luminosity measurement has no effect on this spectrum. Therefore, this spectrum has smaller relative errors.

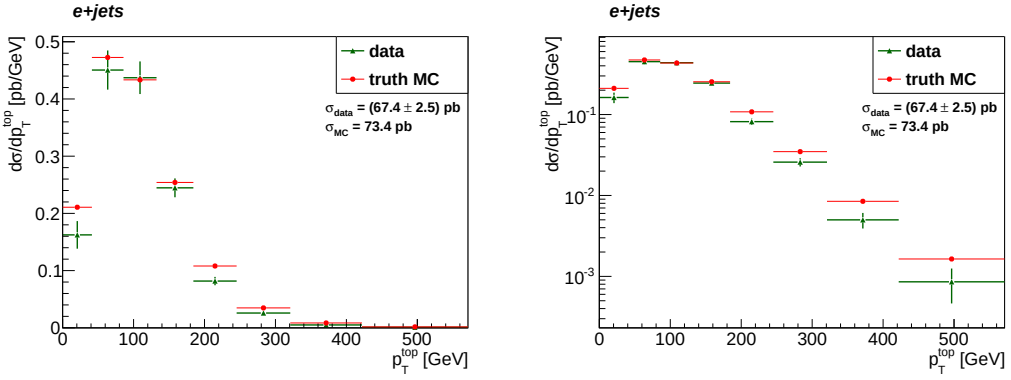


Figure 6.30: The final measured $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum in ℓ+jets channel obtained from data using Equation 6.1 and divided by two because the histogram 6.28 contains leptonic and hadronic tops. The luminosity error of 3.5 % was taken into account. It is compared to MC signal $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum. The total cross section was computed, the error is only statistical and the luminosity error. On the left (right), the linear (logarithmic) scales are shown.

6.7 Final Remarks

This chapter was a proof of concept of the correctness of steps in the analysis. My studies helped to resolve important issues in the analysis.

It would be interesting to perform the analysis in different way regarding the definitions of signal. Now, the definition of signal is the whole single lepton channel not regarding the type of lepton. Two selection criteria (e+jets and μ+jets) are applied and the analysis bifurcate into two parts. The e+jets events can contain also events where the truth lepton is muon and the reconstruction of these events can be worse. Moreover there can be events where the leptonic W decay to τ which decays to hadrons. For these events the reconstruction cannot be successful.

The alternative way would be to divide the analysis into two part from the beginning: single electron channel and single muon channel. The definition of signal events for single electron channel would be: events where the leptonicaly decaying W decays to electron (eventually also to τ which decays to electron). The other decays of leptonic W would be accounted as background. This background as other background types would be also subtracted from measured data. In this channel, only e+jets selection criteria would be applied. Similarly for single muon channel. The $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum from these two channels would be combined into one as the shape of this spectrum is the same in both channels (assuming SM physics). The advantage of this approach would be the rejection of badly reconstructed events. On the other hand, the statistics of data would be lower, but hopefully by poorly reconstructed events. This proposal of improvements should be studied in the future.

There are several systematics uncertainties which have to be taken into account, notably the jet energy scale, jet energy resolution, muon and electron identification efficiency,... Evaluation of the systematics effects is beyond the scope of this thesis.

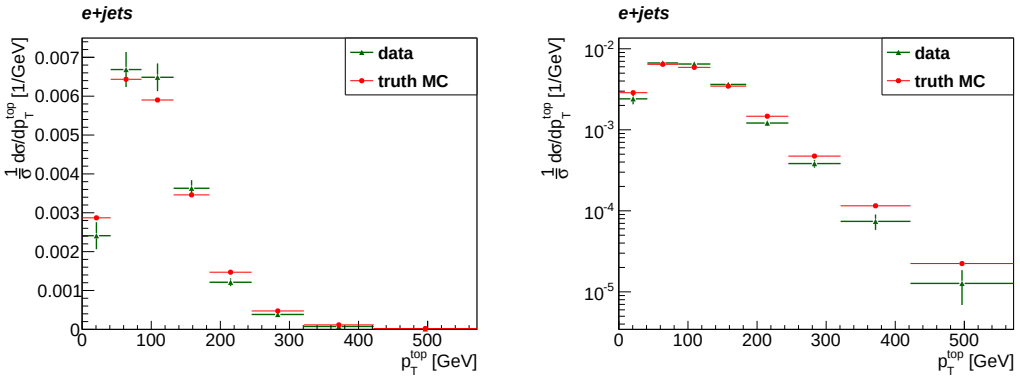


Figure 6.31: The measured final $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum in ℓ+jets channel normalized to one compared to MC signal $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum. On the left the linear scale, on the right the logarithmic scale is shown.

Conclusions

In this thesis, I studied the basic concepts of measurement of kinematic spectra from $t\bar{t}$ production in the ATLAS experiment. The measurement of these kinematic spectra is an important crosscheck of the Standard Model and also it can lead to discoveries of new forces and particles beyond the Standard Model. The main emphasis was put on the measurement of the transverse momentum spectrum of top quark.

I summarized the main facts about design and performance of the ATLAS detector. I described the basics of data simulation, object reconstruction and event selection.

I contributed to the analysis focused on the measurement of the transverse momentum spectrum of top quark. I performed several studies regarding the reconstruction of $t\bar{t}$ events. I concluded that the optimal maximum number of highest p_{T} jets used for the reconstruction of $t\bar{t}$ pairs is five. I showed that the $t\bar{t}$ events where the leptonicaly decaying W decays to tauon should be also used as signal events in spite of the fact that the reconstruction algorithm is not entirely designed for them. I also showed that the reconstructed $t\bar{t}$ pairs with a reconstructed top quark mass not agreeing with the generated MC should be rejected. The signal efficiency of this cut is $\sim 85\%$. Also cut on lower maximized likelihood values obtained after $t\bar{t}$ pair reconstruction is useful with lower signal efficiency of $\sim 75\%$.

As the unfolding procedure is an essential step toward evaluating the real spectrum, I studied one unfolding method: the matrix inversion method. This method appears to be unstable due to low statistics of data. However, outcomes from other unfolding methods should be compared to this method. More statistics of data will not be available for this analysis as the CMS energy at LHC changes from 7 TeV to 8 TeV in 2012. But there is still space for optimization of the analysis and also, the whole analysis can be performed for data with CMS energy of 8 TeV for which the cross section for $t\bar{t}$ production is higher. Also in 2012, it is expected higher recorded luminosity.

Using the studied unfolding method, I performed the whole analysis to obtain the $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum without systematic uncertainties. My final measured $p_{\mathrm{T}}^{\mathrm{top}}$ spectrum is in reasonable agreement with the theoretical prediction obtained from simulated events. This agreement cannot be quantified due to missing systematics studies. Also the analysis needs optimization.

Bibliography

[1] Ho-Kim Q. and Pham X., *Elementary Particles and Their Interactions: Concepts and Phenomena* (Springer, 2004), 1 edition, ISBN 3540636676

[2] The ROOT Team, *ROOT* (web page, last revision 26.2.2012, cited 11.4.2012)
URL http://root.cern.ch/drupal/content/about

[3] Pleier M., *Review of Properties of the Top Quark from Measurements at the Tevatron*, Int. J. Mod. Phys., **A24**, pp. 2899 (October 2008)
URL http://arxiv.org/abs/0810.5226

[4] Nakamura K. et al., *Review of Particle Physics*, Journal of Physics G, **37**, p. 075021 (July 2010)
URL http://iopscience.iop.org/0954-3899/37/7A/075021

[5] Hoang A.H. and Stewart I.W., *Top Mass Measurements from Jets and the Tevatron Top-Quark Mass*, Nucl. Phys. Proc. Suppl., **185**, pp. 220 (August 2008)
URL http://arxiv.org/abs/0808.0222

[6] Galtieri A.B. et al., *Precision measurements of the top quark mass from the Tevatron in the pre-LHC era*, arXiv:1109.2163 (September 2011)
URL http://arxiv.org/abs/1109.2163

[7] Brüning O.S. et al., *LHC Design Report* (CERN (Geneva), 2004), ISBN 9789290832249
URL http://cdsweb.cern.ch/record/782076

[8] Collins J.C. et al., *Heavy particle production in high-energy hadron collisions*, Nuclear Physics B, **263**, 1, pp. 37 (January 1986)
URL http://www.sciencedirect.com/science/article/pii/055032138690026X

[9] Kidonakis N., *Top quark cross sections and differential distributions*, arXiv:1105.3481 (May 2011)
URL http://arxiv.org/abs/1105.3481

[10] The CMS Collaboration, *Combination of top pair production cross section measurements* (2011)
URL http://cdsweb.cern.ch/record/1401250

[11] Jeżabek M. and Kühn J., *QCD corrections to semileptonic decays of heavy quarks*, Nuclear Physics B, **314**, 1, pp. 1 (February 1989)
URL http://www.sciencedirect.com/science/article/pii/0550321389901089

[12] D0 Collaboration, *An improved determination of the width of the top quark*, arXiv:1201.4156 (January 2012)
URL http://arxiv.org/abs/1201.4156

[13] Bigi I. et al., *Production and decay properties of ultra-heavy quarks*, Physics Letters B, **181**, 1-2, pp. 157 (November 1986)
URL http://www.sciencedirect.com/science/article/pii/037026938691275X

[14] Sterman G., *QCD and Jets*, arXiv:hep-ph/0412013 (December 2004)
URL http://arxiv.org/abs/hep-ph/0412013

[15] Catani S. et al., *The Top Cross Section in Hadronic Collisions*, arXiv:hep-ph/9602208 (February 1996), Phys.Lett. B378 (1996) 329-336
URL http://arxiv.org/abs/hep-ph/9602208

[16] Kidonakis N., *Top Quark Theoretical Cross Sections and pT and Rapidity Distributions*, arXiv:1109.3231 (September 2011)
URL http://arxiv.org/abs/1109.3231

[17] Ahrens V. et al., *Renormalization-Group Improved Predictions for Top-Quark Pair Production at Hadron Colliders*, arXiv:1003.5827 (March 2010)
URL http://arxiv.org/abs/1003.5827

[18] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST, **3** (August 2008)
URL http://iopscience.iop.org/1748-0221/3/08/S08003

[19] Pequenao J., *Computer generated image of the whole ATLAS detector* (2008)
URL http://cdsweb.cern.ch/record/1095924

[20] The ATLAS Collaboration, *The ATLAS Inner Detector commissioning and calibration*, arXiv:1004.5293 (April 2010), Eur.Phys.J.C70:787-821,2010
URL http://arxiv.org/abs/1004.5293

[21] The ATLAS Liquid Argon Calorimeter Group, *Commissioning and Performance of the ATLAS Liquid Argon Calorimeters*, arXiv:0809.2672 (September 2008)
URL http://arxiv.org/abs/0809.2672

[22] The ATLAS Collaboration, *Readiness of the ATLAS Tile Calorimeter for LHC collisions*, arXiv:1007.5423 (July 2010), Eur.Phys.J.C70:1193-1236,2010
URL http://arxiv.org/abs/1007.5423

[23] The ATLAS Collaboration, *Calibration and Performance of the ATLAS Muon Spectrometer*, arXiv:1109.6933 (September 2011)
URL http://arxiv.org/abs/1109.6933

[24] Heinemann B., *Luminosity* (web page, last revision 8.11.2011, cited 11.4.2012)
URL https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults

[25] Frixione S. et al., *Matching NLO QCD and parton showers in heavy flavour production*, arXiv:hep-ph/0305252 (May 2003), JHEP 0308 (2003) 007
URL http://arxiv.org/abs/hep-ph/0305252

[26] Corcella G. et al., *HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons*, arXiv:hep-ph/0011363 (November 2000), JHEP 0101:010,2001
URL http://arxiv.org/abs/hep-ph/0011363

[27] Butterworth J. et al., *Jimmy Generator* (web page, last revision 11.4.2012, cited 11.4.2012)
URL http://jimmy.hepforge.org/news

[28] Frixione S. et al., *A Positive-Weight Next-to-Leading-Order Monte Carlo for Heavy Flavour Hadroproduction*, arXiv:0707.3088 (July 2007), JHEP 0709:126,2007
URL http://arxiv.org/abs/0707.3088

[29] Sjöstrand T. et al., *PYTHIA* (web page, last revision 11.4.2012, cited 11.4.2012)
URL http://home.thep.lu.se/~torbjorn/Pythia.html

[30] Agostinelli S. et al., *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **506**, 3, pp. 250 (July 2003)
URL http://www.sciencedirect.com/science/article/pii/S0168900203013688

[31] The ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, arXiv:1110.3174 (October 2011)
URL http://arxiv.org/abs/1110.3174

[32] Cacciari M. et al., *The anti-k_t jet clustering algorithm*, arXiv:0802.1189 (February 2008), JHEP 0804:063,2008
URL http://arxiv.org/abs/0802.1189

[33] The ATLAS Collaboration, *Performance of missing transverse momentum reconstruction in proton-proton collisions at 7 TeV with ATLAS*, arXiv:1108.5602 (August 2011), Eur.Phys.J.C 72 (2012) 1844
URL http://arxiv.org/abs/1108.5602

[34] The ATLAS Collaboration, *Vertexing for b-Tagging* (2009)
URL http://cdsweb.cern.ch/record/1174260

[35] Antchev G. et al., *First measurement of the total proton-proton cross section at the LHC energy of \surd s =7 TeV*, arXiv:1110.1395 (October 2011)
URL http://arxiv.org/abs/1110.1395

[36] The ATLAS Collaboration, *Measurement of the W -> lnu and Z/gamma* -> ll production cross sections in proton-proton collisions at sqrt(s) = 7 TeV with the ATLAS detector*, arXiv:1010.2130 (October 2010), JHEP 1012:060,2010
URL http://arxiv.org/abs/1010.2130

[37] Kidonakis N., *Top Quark Theoretical Cross Sections and pT and Rapidity Distributions*, arXiv:1109.3231 (September 2011)
URL http://arxiv.org/abs/1109.3231

[38] Mangano M.L. et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, arXiv:hep-ph/0206293 (June 2002), JHEP 0307 (2003) 001
URL http://arxiv.org/abs/hep-ph/0206293

[39] Kersevan B.P. and Richter-Was E., *The Monte Carlo Event Generator AcerMC 2.0 with Interfaces to PYTHIA 6.2 and HERWIG 6.5*, arXiv:hep-ph/0405247 (May 2004)
URL http://arxiv.org/abs/hep-ph/0405247

[40] Nackenhorst O., *Top Quark Pair Reconstruction in the Electron + Jets Channel with a Kinematic Likelihood Fitter in the ATLAS Experiment*, diploma thesis, GEORG-AUGUST-UNIVERSITAT GOTTINGEN (July 2010)
URL http://physik2.uni-goettingen.de/research/high-energy/publications/DiplomNackenhorst.pdf

[41] D'Agostini G., *A multidimensional unfolding method based on Bayes' theorem*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **362**, 2–3, pp. 487 (August 1995)
URL http://www.sciencedirect.com/science/article/pii/016890029500274X

[42] Hoecker A. and Kartvelishvili V., *SVD Approach to Data Unfolding*, arXiv:hep-ph/9509307 (September 1995), Nucl.Instrum.Meth.A372:469-481,1996
URL http://arxiv.org/abs/hep-ph/9509307