UNIVERSITÀ DEGLI STUDI DI ROMA
"LA SAPIENZA"

DIPARTIMENTO DI FISICA

Search for new physics in the $t\bar{t} + \not{E}_T$ channel using fully hadronic top pair decays with the CDF detector at Tevatron

Tesi di Laurea Specialistica in Fisica

Relatore
Dott. Marco Rescigno

Laureando
Marco Bentivegna

Correlatore
Dott. Fabrizio Margaroli

ANNO ACCADEMICO 2010-2011

ai miei genitori
a cui devo ogni cosa

Acknowledgements

I want to express my gratitude to my supervisor, Marco Rescigno, for the invaluable help in all these years, and for having given me the first occasion to work in America. I would like also to thank my co-supervisor, Fabrizio Margaroli, for having pushed me to the limit, and having been a guide in my moments of discouragement. I want to thank my workmate and friend Qiuguang Liu, for his infinite patience in the difficult art of understanding someone else's programming errors. My gratitude goes also to Daniela Bortoletto, for having placed trust in me.
Thanks to Margherita, because she was close to me, even when thousand miles away. Looking back at the last three years at *la Sapienza*, i want also to thank Giovanni C., Sergio, Linda, Giovanni Z., Marianna and Chiara for having made this place my second home.
But most of all, thanks to my parents and my brothers, because all good things i've done in my life are due to them.

Contents

Introduction

The existence of dark matter is an outstanding theme for both astronomy and particle physics. It was first introduced in the 30's to explain anomalies in the orbital velocities of galaxies and clusters of galaxies, and this first hint was supplemented in the following decades by observations of gravitational lensing, hot gas in galaxy clusters, distant supernovae and cosmic microwave background. Togheter, this data provide now clear evidence for the fact that not only is there non-luminous matter in the Universe, but most of it is not composed of baryons or any of the other known particles. Howewer, all of the evidence for dark matter noted above is based on its gravitational effects. Given the universality of gravity, such evidence don't allow to pinpoint what exactly dark matter is. In recent times, some direct detection experiments based on ultra-low noise detectors have shown signals compatible with expectations of some theoretical models. Nevertheless, some of these results are in contrast with each other, and in general are not completely understood. Indeed, the greatest advance in the path to understand dark matter would be to create it and study it at particle colliders. There are several theoretical models that provide a dark matter candidate which could be produced at colliders. Many of them predict an associated production with top quarks, the third generation up-type quarks discovered at Tevatron in 1995. The present search is focused on a model in which the stability of dark matter is provided by a new symmetry, and its interaction with standard model particles happens through fourth generations T' quarks, acting as "connector particles". The process we are searching for is the pair production of T' quarks, wich decay subsequently to top quarks plus dark matter. Because of the very low rate of interactions with standard matter, the dark matter particles are expected to leave no signal in the detectors, giving rise to missing transverse

energy. Therefore, the signal we want to investigate is that of a top quark pair plus missing transverse energy. In particular, our search is the first one focused on the fully hadronic channel, in which both the W bosons produced by top quarks decay hadronically. This channel is favoured by the larger branching ratio and low physics background. However, QCD multijet events with large missing transverse energy due to jet energy mismeasurement constitute an overwhelming background, which must be reduced and correctly modeled. This search use a novel technique, based on the angular correlation between missing transverse energy and missing transverse momentum, to model the QCD multijet processes through a data-driven method. Our results are consistent with standard model expectations, and we put new exclusion limits in the $(m_{T'}, m_X)$ phase space, where X is the dark matter particle. Since the top pair plus missing energy signature is typical of many other theories beyond Standard Model, the same techniques could be applied soon at the Large Hadron Collider to explore many other new physics scenarios. The discussion is organized as follows. Chapter 1 provides an overview of the experimental evidences for dark matter and of the theoretical models that provide a dark matter candidates which is expected to be produced in association with top quarks. Chapter 2 and 3 describe respectively the experimental apparatus and the standard techniques used at CDF to reconstruct the physical particles from detector informations. Chapter 4 provides an experimental overview of the present analysis, compared with previous searches in the semileptonic channel. In Chapter 5 we report the techniques to simulate the signal and the standard model non-QCD backgrounds, togheter with a detailed descriptions of the data-driven QCD model. Chapter 6 contains a description of the methods we use to suppress backgrounds and to check the validity of the overall background modeling. Finally, Chapter 7 describes the analysis techniques used to extract expected and observed limits from the data, which we report in the end of the document.

Chapter 1

The Dark Matter Puzzle

1.1 Motivations for dark matter hypothesis

The belief that the visible and known matter constitute only a fraction of the matter of our universe is now widely accepted. The effects of the presence of some kind of long-lived, non strong-interacting neutral particles are visible in a wide range of observations, and also supported by many of the most credited cosmological theories.

1.1.1 Evidences from astronomical observations

The "classical" evidences that much more than the visible matter should fill the Universe are all based on anomalies of the motion and distribution of velocities of stars and galaxies. This strategy has a long tradition in astronomy, as it lead, for instance, to discovery of Neptune in 1846 from unexplained anomalies in the orbit of Uranus. The first evidence was obtained from F. Zwicky in 1933 [1]. From observation of eight galaxies in the Coma cluster, Zwicky found an unexpectedly large dispersion of the galaxies' velocities. He noticed that, for a velocity dispersion of $\simeq 1000\ km/s^{-1}$, the mean density of the Coma cluster would have to be about 400 times larger than the one inferred from the luminous matter. He than concluded that the Coma cluster must contain a huge amount of some kind of "dark matter". Three years later, the same anomaly was found in the Virgo cluster by S. Smith [2]. After

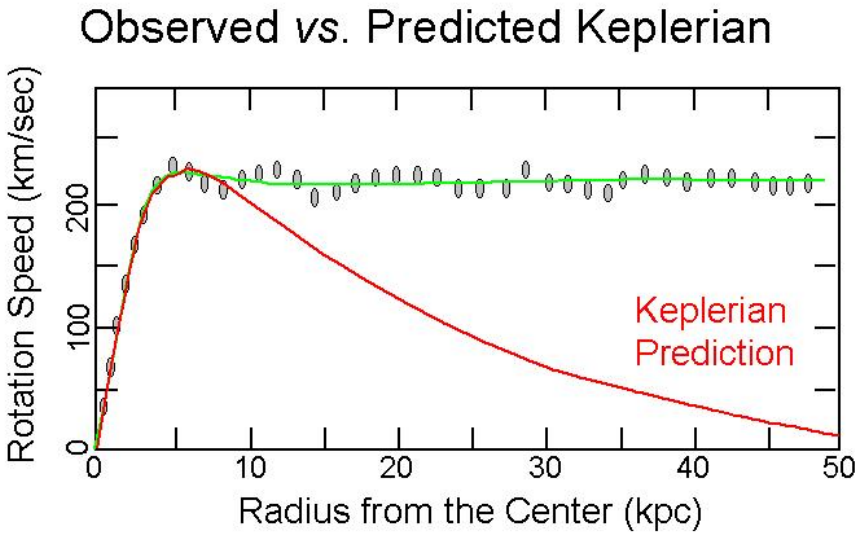


Figure 1.1: Relationship between the velocity of rotation and the distance from the galactic nucleus in a typical spiral galaxy. The red curve shows the keplerian prediction, assuming that only visible matter is present.

that, in 1939, H. Babcock started a similar study on different scales, analyzing the spectra of the Andromeda galaxy [3]. He found that, in its outer regions, stars were rotating with a far higher velocity than what expected from Kepler's laws. Howewer, only about 40 years later a systematic study was performed independently by V. Rubin and A. Bosma, which measured the rotation curves of many spiral galaxies. They found that the velocity curves in the galaxy plane as a function of the distance from the galaxy center stay flat even outside of the luminous disk, crediting the presence of a dark halo (see Fig. 1.1).

At the beginning of the 1980's the presence of dark matter on the basis of its gravitational effect on luminous matter was directly confirmed by many other kinds of sources, on very different lenght scale, such as:

- the motion of the Large Magellanic Cloud
- the distribution of X-ray emitting gas clouds surrounding elliptical galaxies
- the velocity distribution of hot intergalactic plasma
- the gravitational lensing effect on the light coming from distant objects

1.1.2 The role of dark matter in cosmology

In the 80's it was also already known that the structures in the Universe are formed by gravitational collapse starting from small density fluctuations in the early universe [4]. In order to give rise to the structures we now observe, the amplitudes of density fluctuations at the epoch of recombination, when the universe became transparent to light, is required to be at least one thousandth of the density itself. The first detection of the relic radiation coming from this epoch was measured by Penzias and Wilson in 1965 [5]. This radiation, known as the Cosmic Microwave Background, evidenced an almost isotropic behaviour. A first precise measurement of the CMB anisotropy, performed in 2000 by de Bernardis et al. [6] revealed fluctuations of about two order of magnitude lower than those required for the density evolution of the luminous (baryonic) mass. In the same years, many theorists independently suggested dark matter as a possible "booster" of structures formation. In fact, if the dominating component of the universe is some kind of matter non interacting with light, than the density fluctuations could have started to grow much earlier than the epoch of recombination, because the radiation pressure would not have slowed their early growth. In this scenario, at the time of recombination dark matter could have already formed "dark structures", whose gravitational field would eventually have captured ordinary matter, giving rise to galaxies, clusters of galaxies and so on. This explanation imply that the bulk of dark matter can only be either Cold (CDM) or Warm (WDM), i.e. non relativistic or becoming non relativistic when the temperature of the Universe was T$\simeq$ keV, otherwise gravitational collapse would not have been possible at the required time. For this reason, Hot DM candidates as neutrinos, which could be the only particle of the Standard Model (SM) that could do for the purpose, are generally ruled out.

There are also other cosmological problems requiring the existence of dark matter. The standard inflationary cosmology assumes that the Universe arose from an initial singularity and went on expanding. The average density of matter and energy in the Universe determine both its global curvature, and its evolution: a density ρ equal to the critical one ρ_c correspond to a flat universe, while ρ higher or lower than ρ_c correspond respectively to an open, ever-expanding Universe and to a closed, finally recollapsing Universe. Current measurements from CMB give a value

of $\Omega \equiv \frac{\rho}{\rho_c}$ equal to 1.02 ± 0.02 [7, 8]. However, the average density of the Universe as measured by photometric methods is $\Omega_{ph} \simeq 0.007$. For the sake of completeness, we have to mention the fact that this riddle cannot be solved only introducing dark matter in our model: the most recent results [7, 8] privileges a scenario in which the measured value $\Omega \simeq 1$ is due only at $\simeq 30\%$ to matter, while all the rest would be due to some kind of vacuum energy responsible for the acceleration of the Universe's expansion, called dark energy. Anyway, dark matter is still necessary to account for more than 80% of matter density.

Dark matter plays also an important role in the theory of Big Bang Nucleosynthesis, which predicts that roughly 25% the mass of the Universe consists of Helium. It also predicts about 0.01 % deuterium, and even smaller quantities of lithium. The important point is that the prediction depends critically on the density of baryons (*i.e.* neutrons and protons) at the time of nucleosynthesis. Furthermore, one value of this baryon density can explain all the abundances at once. In terms of the present day critical density of matter, the required density of baryons is a few percent (the exact value depends on the assumed value of the Hubble constant). This relatively low value means that a huge amount of non-baryonic matter is required to fill the gap.

1.1.3 Search for DM-SM non-gravitational interactions

All the above evidences are based on gravitational effect, and so could be explained by the existence of some new particle interacting with SM particles only through gravity. If this is the case, DM will never be seen in other ways. Nevertheless, many theories predict some other kind of interaction between DM and SM particles. In Weakly Interacting Massive Particles (WIMP) models, the dark matter can interact through weak force. Other models predict DM-SM interaction through more exotic mechanism (see next section). In any case, DM-SM non-gravitational interactions open the way to three other different types of search, besides of gravitational effects:

- **Indirect searches**

  This means searching for products of the annihilations of DM particles in photons or fermions. Depending on theoretical model, signatures can range

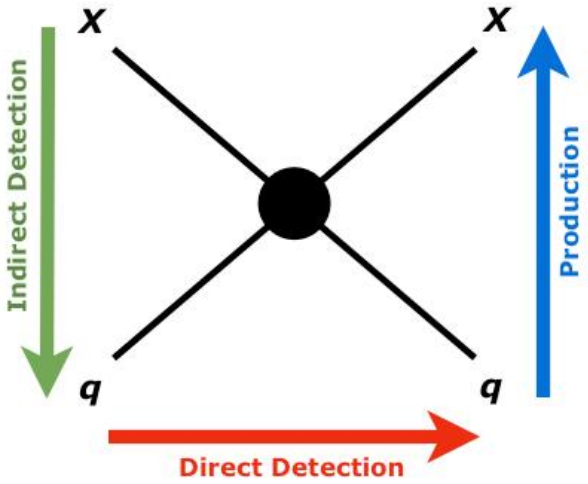


Figure 1.2: Schematic view of the three different types of search for DM - SM non-gravitational interactions.

over a variety of products: mainly high-energy cosmic rays (electrons or protons), neutrinos, or photons. This many different signatures are reflected in a large variety of underground, terrestrial and extraterrestrial experiments. Another problem consist in where to focus the research. For example, searches for gamma rays coming from the galactic center, where DM should be more dense, are difficult because of the large background, while searches from outer halo can suffer for low statistics.

Results from this branch of search are still controversial.

- **Direct searches**

  This kind of analysis look for the effects of DM-SM elastic scattering, tipically measuring the recoil energy of a scattered nucleus in a target-detector. The additional possibility to investigate the DM-nucleus inelastic scattering producing low-lying excited nuclear states is disfavoured by the very low expected counting rate. Since these experiments require a very low background rate, they are performed in underground laboratories. It was shown [9, 10] that the motion of the earth about the sun introduces an annual modulation in the flux of dark matter particles reaching the earth (see Fig. 1.3). This is the type of signal searched for by the DAMA experiment [11], a set of scintillating

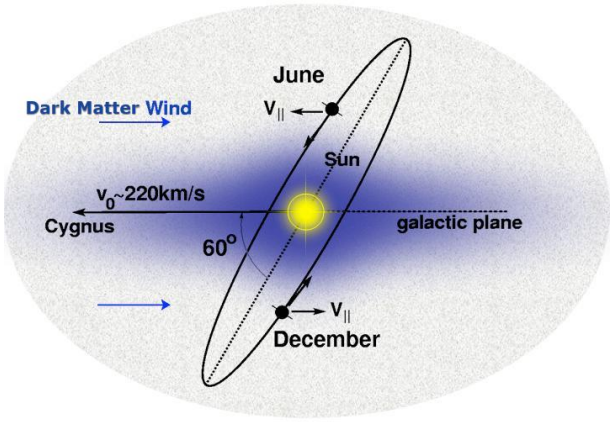


Figure 1.3: Schematic explanation of the annual modulation signature reported by DAMA and CoGeNT.

high purity NaI(Tl) crystals located at the Gran Sasso National Laboratory, which have now reached a result compatible with the one searched for at 8.9 σ. Also the CoGeNT experiment [12], using an high-purity germanium crystal cooled to nitrogen temperature located at Soudan Underground Laboratory, seems to be able to detect an eccess of events, and a modulation. There are still controversies on the compatibility of DAMA and CoGeNT results with the measures of other experiments (such as XENON100 [21] and EDELWEISS [22]), and between each other. Uncertainties from both astrophysics and detector response open a non-conflicting window in the (mass - cross section) phase space [13], favouring a DM candidates of mass of about 5-10 GeV/c^2 and an interaction cross section with ordinary matter of about $100 \div 500$ fb (see Fig. 1.4). Another interesting result has been recently reported by the CRESST collaboration, which found an excess of events ascribable to dark matter elastic scattering off nuclei in the detector [23].

- **Production in colliders** If DM can have non-gravity interactions with SM particles, like DAMA and CoGeNT seems to suggest, than it could be produced at colliders. The signature of DM production would be analogous to the one of SM neutrinos. However, the production cross section and, generically, the characteristics of the event to search for will (partially) depend on the

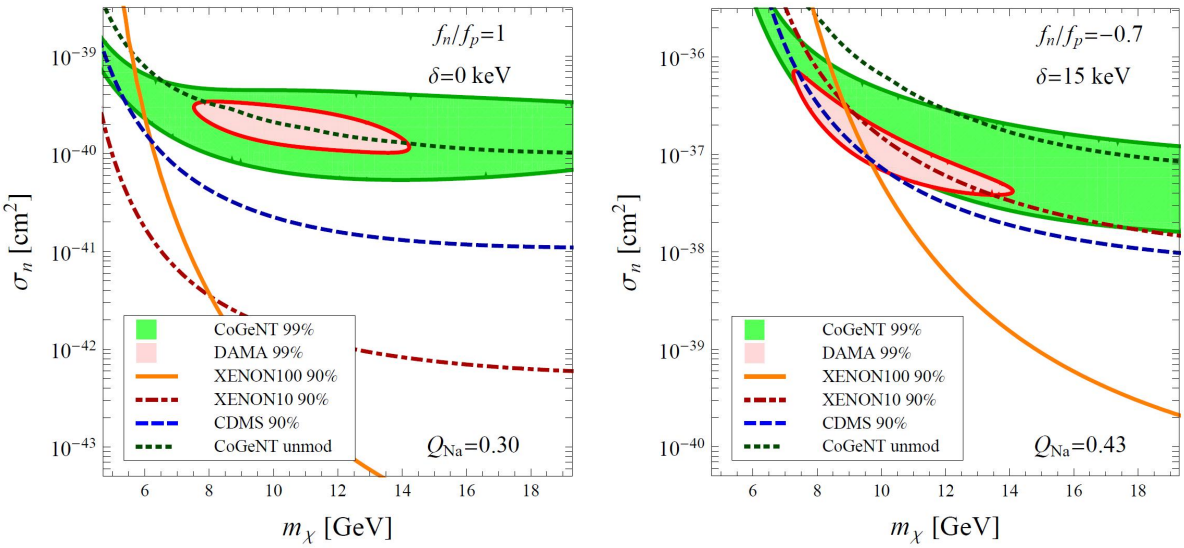


Figure 1.4: Best-fit parameter regions for DAMA and CoGeNT (coloured regions) as well as exclusion limits from XENON10, XENON100, CDMS II and the unmodulated CoGeNT signal. The left panel show the common choices for theoretical assumption :equal DM-proton and DM-neutron scattering probabilities, no changes in the DM mass after scattering and sodium quenching factor (0.3). The right panel shows how the tension is alleviated if we assume an isospin-violating scattering ($f_n \neq f_p$), an inelastic DM-SM scattering ($\delta \neq 0$ KeV) and we vary the sodium quenching factor inside the one σ experimental uncertainty. From [13]

particular theoretical model. The analysis reported in this thesis is focused on a signature which is predicted by many theoretical frameworks, as we will see in the next section.

1.2 The top quark

The search for the top quark started right after the discovery of the b quark and lasted almost 20 years when it was finally discovered at the Tevatron collider in 1995 by the CDF [14] and D0 [15] collaborations. Its mass proved to be surprisingly large, having been measured at the end of the 1992-1995 data taking period to be equal to 178 ± 4.8 GeV/c^2, i.e. about the mass of a gold atom. The last combination of the measurements of the top quark mass at the Tevatron gives a value of 173.2 ± 0.9 GeV/c^2 [16]. The large mass of the top quark gives rise to large radiative corrections, for example to the W propagator, which causes a strong correlation between m_t, m_W and the Higgs boson mass m_H. The current predicted value of m_H are influenced by precise measurements of m_t and m_W. Calculations in perturbative QCD predict that the dominating subprocess of the production of $t\bar{t}$ at the Tevatron is $q\bar{q}$ annihilation (85%), while gg fusion contribute is 15%.

Quarks are not observed as free particles but are confined to form hadronic bound states. The top quark is special in that regard with respect to other quarks: being its lifetime $\tau \simeq 10^{-24}$ s shorter than the typical hadronization time, which is estimated to be $\Lambda_{QCD}^{-1} \simeq \mathrm{O}(100\,\mathrm{MeV})^{-1} \simeq \mathrm{O}(10^{-23}\,\mathrm{s})$, the top quark decays before hadronizing. As a consequence we do not expect a $t\bar{t}$ resonant state as with c and b quarks, but we can detect it through its decay products. In fact the daughter particles will retain information on the mass and quantum numbers of the parent top quark, including polarization effects in the angular distributions.

According to the Standard Model, the top quark decays to a W boson and a b quark almost 100% of the times. The decays $t \to Ws$ and $t \to Wd$ are allowed too; but the former is suppressed with respect to the $t \to Wb$ by a factor $V_{ts}^2/V_{tb}^2 \simeq 10^{-3}$, and the latter by a factor $V_{td}^2/V_{tb}^2 \simeq 10^{-4}$. The W boson decays 1/3 of the times into a lepton and a neutrino and 2/3 of the times into quarks. The final states are thus determined by the decays of the two W bosons; using the calculated branching

Signature	***Branching ratio***
$q_1\bar{q}'_2 q_3\bar{q}'_4$	36/81
$e + q\bar{q}'$	12/81
$\mu + q\bar{q}'$	12/81
$\tau + q\bar{q}'$	12/81
ee	2/81
$e\mu$	2/81
$e\tau$	2/81
$\mu\tau$	1/81
$\mu\mu$	1/81
$\tau\tau$	1/81

Table 1.1: Branching ratios for the various top pair decay modes.

ratios, (see Table 1.1), we expect that the fully hadronic and the semileptonic decays of top quark pairs make up the majority of the events.
Here we describe briefly the typical signatures which pair produced top quarks will leave in the detector:

- **The dilepton channel**

  The nominal signature for this channel has two high-P_T leptons, missing transverse energy from the two neutrinos and two jets from the b quarks. The yield is pretty small, mostly due to the low branching ratio in this channel, 5%. Moreover, the presence of two high energy neutrinos complicates the event reconstruction. On the other hand, the background is very small, mainly coming from Drell-Yan events, so we expect a very clean sample.

- **The lepton+jets channel**

  The nominal signature has a high-P_T lepton, missing transverse energy from the neutrino, and four jets out of which two are expected to contain B hadrons from the hadronization of the b quarks. Without considering the events with τ's, we expect $BR \sim 30\%$. The main background contribution comes from W+multijet production.

- **The all-hadronic channel**

  In this decay mode we expect six final state jets, four of which come from the

hadronic decays of the two W's and two from the b quarks. Approximately 44% of the $t\bar{t}$ events have this decay signature. The major challenge here is to overcome the huge background coming from QCD multijet production. This channel is the one used for the measurement in this work.; its characteristics will be explained in more details in Chapter 6.

Since m_t is so close to the energy scale where the electroweak scale breaks down (vacuum expectation value of the Higgs field), it has been hypotized that the top quark could be part of some mechanism which causes the electroweak symmetry breaking. In any case, giving access to the highest energy scale, the top quark offers the chance to find hints of new physics, for example in possible resonant production. A number of measurements have been performed setting limits to new particle production, in particular on production of $t\bar{t}$ resonances, of W' in association with single top. Also, CDF and D0 have searched for fourth generation T' quarks decaying into Wq where q is a generic quarks, setting limits up to 358 GeV/c^2 [18]. Up to now, at the end of 16 years of studies, top quark seems to show all the expected SM features [19]. The only one discrepancy, and possible hint for new physics, has been found measuring the forward backward asymmetry.

Forward-backward asymmetry

Due to interference effects at next-to-leading order (NLO) QCD predict a forward-backward asymmetry effect in the $t\bar{t}$ production. Top quarks tend to be produced in the proton direction more than antitop quarks. The point is that the value measured at Tevatron for this asymmetry is substantially larger than the SM prediction. Using events in the lepton+jets channel, CDF have investigated the charge asymmetry [17], finding a value for the total laboratory-frame asymmetry of

$$A_{FB}^{lab} = \frac{N(-Q_l \cdot y_{had} > 0) - N(-Q_l \cdot y_{had} < 0)}{N(-Q_l \cdot y_{had} > 0) + N(-Q_l \cdot y_{had} < 0)} = 0.150 \pm 0.055(stat + syst) \quad (1.1)$$

in the $t\bar{t}$ rest frame. There is less than 1% probability that this value is a fluctuation of the SM value. It has been also shown that all the discrepancy is due to event with $t\bar{t}$ invariant masses greater than 450 GeV/c^2. The latest measurements

of the $t\bar{t}$ FB asymmetry are summarized in Tab. 1.5. It has been argued [20] that this asymmetry may be due to an alternative $t\bar{t}$ production mechanism, involving scalar top and neutralino (see section 1.3.1). The fact that the discrepancy seems related to events with high $t\bar{t}$ invariant masses is in fact an hint for the possible underlying production of new particles. As we will see in the next section, the production of the new particles involved in the mechanism described in [20] could provide the signature we are searching for in this analysis.

1.3 Production of dark matter and top quark pairs: the "connector particles" approach

Though the precise nature of dark matter is still elusive, it is clear that it must be stable or with a lifetime comparable with the age of universe. This is typically achieved giving DM a charge under an unbroken, discrete or continuous symmetry. Since none of the known symmetries could serve the purpose, is necessary to introduce a new, unbroken symmetry in the model. The lightest particle carrying a charge under an unbroken symmetry will be completely stable, and a neutral, non strong-interacting massive particle of this kind will be a perfect DM candidate.

At this point, DM could directly interact with SM particles through weak interaction, or only through gravitational interaction. In the last case is referred to as "hidden dark matter". This kind of particles X can interact with SM particles f only through the exchange of some "connector particles" Y carrying both SM and "dark" charges, making XYf coupling possible. In this hypothesis, connector particles carry SM charges and so they can be produced, in even numbers, at colliders. Their subsequent decay will necessarily include the lightest dark-charged particle, *i.e.* dark matter.

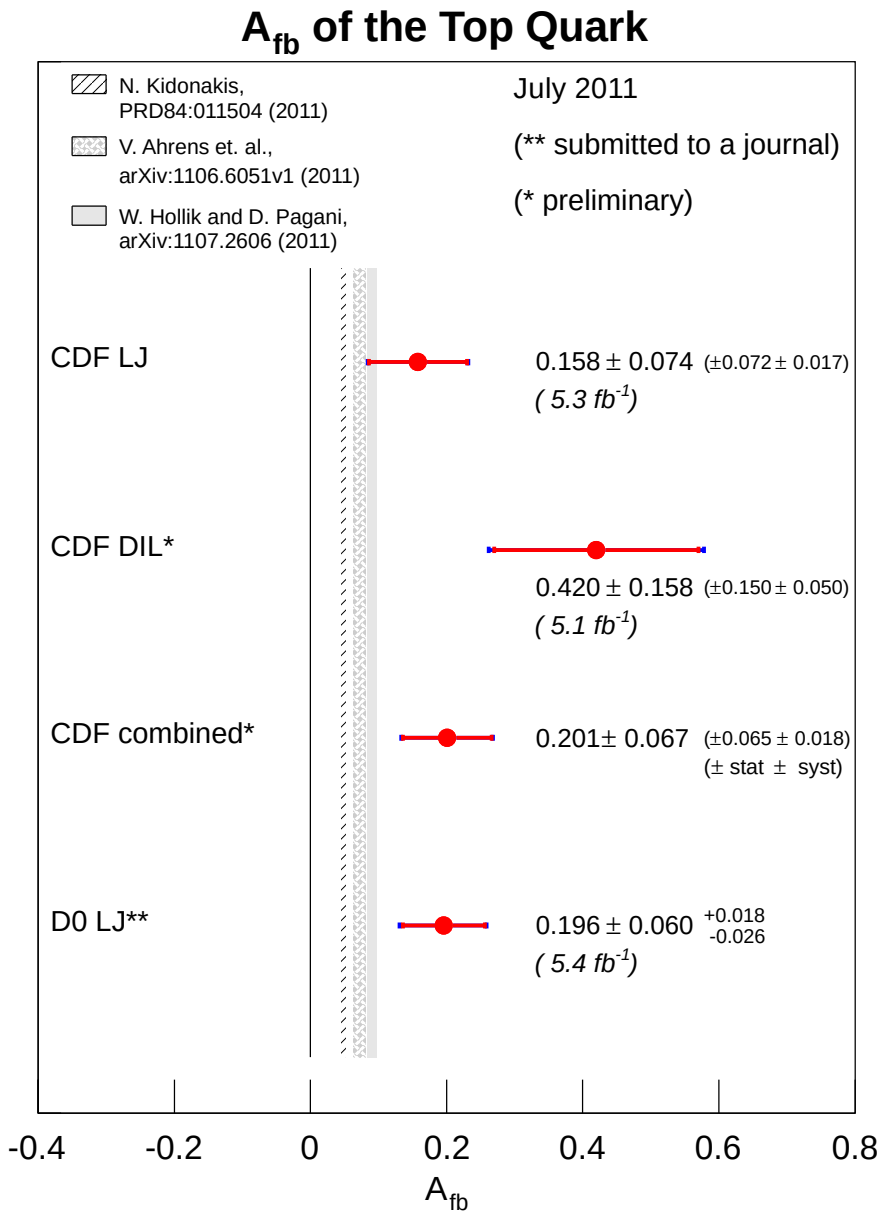


Figure 1.5: Summary of recent measurements of $t\bar{t}$ FB asymmetry.

1.3.1 Models predicting top pairs plus dark matter signature

Among the dark matter models ascribable to the connector particles approach, there are many that favour an XYt coupling, where t is a top quark, so that the process following a YY production would be $YY \rightarrow tt + XX$, leading to the top pairs plus missing energy signature searched for in this analysis.

Models involving fourth generation quarks

The existence of a fourth generation of fermions is neither ruled out by theory nor excluded by experiments. The only theoretical restriction to the number of generations comes from asymptotic freedom constraints in QCD, which requires it to be less than nine. Also the bounds from electroweak precision measurement don't forbid a fourth generation [24, 25]. On the contrary, such an expansion of the SM may play an important role in understanding unanswered questions [26]. One of these is the observed unbalance of matter-antimatter in the universe, which requires some CP violation source of about ten order of magnitudes larger then what is observed in SM.

Fourth generation quarks are expected to receive their mass through electroweak symmetry breaking, just like SM quarks, so that $m_{Q'} = y_{Q'}v/\sqrt{2}$, where $v \simeq 246$ GeV/c^2. Perturbativity places an upper bound on $m_{Q'}$, since requiring $\alpha_{Q'} < 1$ implies $m_{Q'} < 600$ GeV/c^2.

A recent model [27] hypotize a fourth generation T' quark as a connector particle to a hidden DM candidate X, i.e. a real or complex scalar charge under a new continuous or discrete symmetry. An example of how this could be realized is given in [28, 29]. Here, the DM particle couples to the SM through Yukawa couplings:

$$V = \lambda[X\bar{Q}'_l q_l + X\bar{B}'_r b_r + X\bar{T}'_r t_r] \tag{1.2}$$

where X is the DM particle, a complex scalar charged under a discrete symmetry (hidden parity); q_l, b_r and t_r are the third generation quarks and Q'_l, B'_r and T'_r are exotic fourth generation quarks, which works as connectors. The Q' have hidden

parity, and in the $SU(3) \times SU(2) \times U(1)_Y$ SM representation they are:

$$Q'_l : (3, 2, \frac{1}{6});\ B'_r : (3, 1, -\frac{1}{3});\ T'_r : (3, 1, \frac{2}{3}) \tag{1.3}$$

Here the l and r subscripts mean $SU(2)$ "doublets" and "singlets" respectively, not chirality, since chirality of fourth generation quarks in this model is opposite to those of their SM counterparts (they are "mirror" quarks). The DM coupling to SM described by previous equation make possible the elastic scattering $Xq \to Q' \to Xq$ where q is a SM third generation quark. Interactions with nucleons can thus happen because of one-loop coupling to the gluons of the nucleon. These interactions have high spin-independent cross sections if, as in this case, X is a scalar and connectors are chiral fermions. In fact, since Q' are heavy, they can easily provide the chirality flip necessary to a spin-independent scattering. As a result of this, as shown in [29], this model can easily provide explanation for the results of DAMA and CoGeNT. The range of the model parameters allowed by the two experiments are $m_X \sim 1 - 10 GeV/c^2$, $m_{Q'} \sim 300 - 500 GeV/c2$ and $\lambda \sim 0.3 - 1$. The best fit point to the CoGeNT data can be obtained with $m_X \sim 9 GeV/c^2$, $m_{Q'} \sim 400 GeV/c2$ and $\lambda \sim 0.7$. Since this T' quark carries dark charge, once produced in colliders it cannot decays only to SM particles. If $|m_{T'} - m_{B'}| \leq m_W$, as required by constraints from precision electroweak measurements (which are not modified by the exotic and mirror features of the quarks we consider here) the decay $T' \to W^{+*} B'$ is strongly suppressed by kinematics. The process searched for in this analysis is the pair production of T' fourth generation quarks followed by decay to the lightest dark-charged particle X (DM) and a top quark ($T' \to tX$).

Another recent model [39] involve fourth generation quarks in the attempt to treat Baryon and Lepton numbers as local gauge symmetries. Decays of the new quarks, necessary to avoid stable coloured particles, are induced by adding a new scalar field X, which is stable when is the lightest particle with baryon number. Again, this model predict the process $p\bar{p} \to t'\bar{t}' \to t\bar{t}XX$, giving rise to the top pairs plus missing energy signature. Those scenarios clearly require a different analysis from the "classical" fourth generation searches at colliders, in which the T' is supposed to decay similarly to a t quark ($T' \to Wb$). Since T' cross sections are

large enough for small enough $m_{T'}$ (see Fig. 1.6), and since from unitarity we know that the fourth generation quarks' masses must be less than about 600 GeV/c^2, the Tevatron has the potential to exclude large part of the parameter space of those models [27].

Supersymmetric models

Supersymmetry (SUSY) is one of the most promising theories for physics beyond the SM, because of the solution it provides to many unresolved questions and also because of its intrinsic elegance. It basically consist of introducing an underlying symmetry between every fermion (boson) and a bosonic (fermionic) "superpartner" of the same mass, coupling constants and gauge quantum numbers. Since none of these correnspondence is seen in nature, this symmetry has to be broken, so that the masses are no longer degenerate and the SM superpartners are all at energies beyond the reach of present experiments.

To forbid the possibility of events like proton decays, one has to introduce a symmetry that constrains the interactions of particles with their supersymmetric partners. The most common way is to introduce the R-parity, which is defined from the baryon number B, lepton number L and spin s as $R = (-1)^{3(B-L)+2s}$. This multiplicative quantum number is +1 for all SM particles and -1 for all SUSY partners. Its conservation have important consequences:

- SUSY particles can only be produced in pairs
- The lightest partcle with $R = -1$ is stable. If neutral, is a good DM candidate
- Two $R = -1$ particles can annihilate and produce only $R = +1$ particles ($\rightarrow$ indirect detection)

It is clear that R-parity in supersymmetry plays the role of dark charge we have discussed before. The exact nature of the lightest supersymmetric particle (LSP) depends on the SUSY breaking mechanism, which determine the particles spectrum. In many models, the LSP is the lightest neutralino (χ_1, or simply χ), i.e. a linear combination of the fermionic partner of the photon ("photino"), of the hypercharge gauge boson ("bino") and of the two Higgs bosons ("higgsinos"). The

next-to-lightest supersymmetric particle (NLSP) will work as the connector particle. In the SUSY lagrangian squarks and sleptons receives negative mass contributions proportional to their Yukawa coupling . Also, the scalar partner of the left-handed and right-handed quarks are mixed to form two mass eigenstates. The mixing term for up-type squarks is proportional to their Yukawa coupling and to the ratio between the two Higgses vacuum expectation values, $< H_{up} >$ and $< H_{down} >$, defined as $\tan\beta$. So, in particular if $\tan\beta$ is large, the mass splitting between the mass eigenstates will be large for the stop quarks. For these reasons, the lighter stop mass eigenstate ($\tilde{t}_1$) is a good candidate to NLSP.

If stop and neutralino are the NLSP and LSP, the main deay mode of the stop will be in top (which have the same quantum numbers) plus neutralino. If $m_{stop} > m_{top} + m_{\chi}$, then the top quark will be real. After producing a stop pair in a hadron collider through quark-antiquark annihilation or gluon-gluon fusion, the signal will be two top quarks plus missing transverse energy from the undetected neutralinos. This process has too small a theretical cross section at the Tevatron (see Fig. 1.6) to be probed even with the full luminosity that is scheduled to be acquired by the end of the run. On the other hand, the supersymmetric cross section for this scenario will be soon accessible with LHC data. Up to now, the stop is excluded for $m_{\tilde{t}} \leq$ 150-180 GeV depending on theoretical assumptions [30].

Models with extra dimensions

As noted long ago by Kaluza and Klein, if particles propagate in extra space-time dimensions (ED), they will have an infinite set of partner with identical quantum numbers and increasing masses, called Kaluza-Klein (KK) towers. This is because, if the new dimension is a circle of radius R, the quantization of the particle momentum on a circle give rise to an infinite set of excited states n with increasing masses. As a result of this, for example, a single massless scalar in five dimensions is equivalent to a massless scalar (0 level of the tower) and a collection of massive scalars in four dimensions. There are many ways to explain the hidden nature of additionals dimensions; most of the theorists argue that they are compactified and sufficiently small, or that the matter we know is "stuck" in the four-dimensional world.

ED theories provide good candidates to dark matter [33]. The KK states can

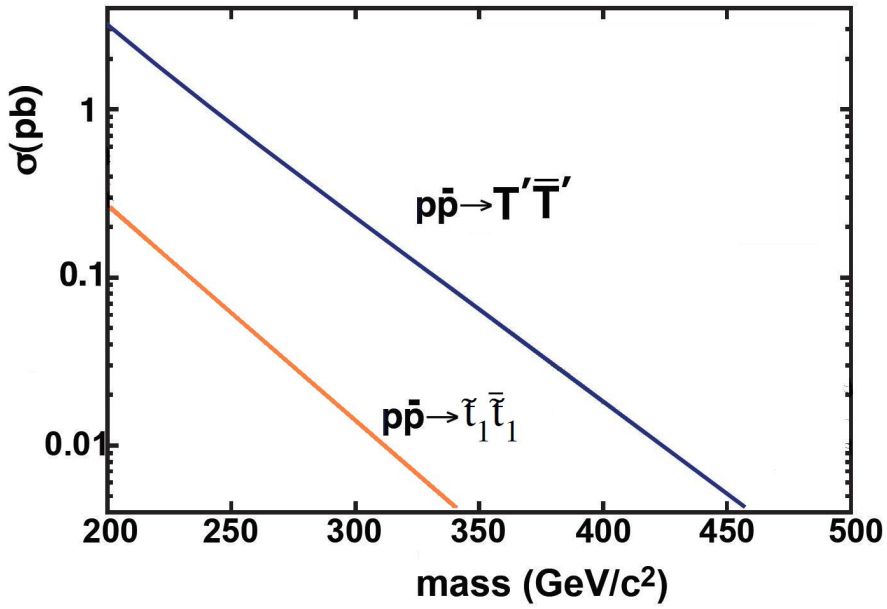


Figure 1.6: Theoretical cross sections for pair production of color-triplet fermionic top partners (like fourth generation T' quarks) [31] and of supersimmetric scalar top $\tilde{t}$ (adapted from [32]) in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, depending on their mass.

be distinguished in even and odd levels. It is possible to introduce a KK-parity, under which KK-modes with odd level numbers are charged. In the simplest cases, this KK-parity correspond to the symmetry of reflection about the midpoint in the extra dimension, and its conservation implies that the lightest level-1 KK particle is stable. The most viable candidates for KK dark matter are the KK excited state of electroweak boson and the KK-graviton. It is clear now that this scenario can be considered a "connector particle" one, with the KK-parity as the new symmetry, and that it can predict the pair production of KK-excited top quarks with subsequent decay into top quarks and dark matter. In particular, in Randall-Sundrum models with warped extra-dimensions [34], (*i.e.* models where the space-time metric depends from the position in the extra dimension) the right-handed top quark is the only quark expected to have the first KK-state below 1 TeV [35], thus is a good candidate to be the next-to-lightest Kaluza-Klein particle.

It is worth to point out that KK-quarks are color triplets, so that their pair production via quark-antiquar annihilation and/or gluon-gluon fusion in hadron

colliders follow the same mass-cross section dependence as in the case of the fourth generation quarks.

Chapter 2

The Experimental Apparatus

2.1 The accelerator chain

Fermilab is a high energy physics research center; it is located in Batavia, in the suburbs of Chicago, Illinois. Inside the complex is located the accelerator which produces high energy $p\bar{p}$ collisions. The accelerator is actually an accelerating chain composed mainly by four subsystems: the proton source, the main injector, the antiproton source and the Tevatron. A pictorial representation of the accelerating complex can be seen in Fig. 2.1.

2.1.1 Proton Source

The first device in the chain is the Cockroft-Walton electrostatic accelerator. Here a hydrogen gas is ionized by the addition of an electron. The resulting ions are accelerated up to an energy of 750 keV. The ions enter then into a 130 m long linear accelerator (Linac) in which a series of radiofrequency cavities accelerate the ions up to 400 MeV. At the outer end of the Linac the ions smash through carbon foils which strip the atoms of their electrons. The resulting protons are then inserted into the Booster, a synchroton accelerator with a circumference of about 0.5 km. There, the protons reach the energy of 8 GeV, and subsequentely enter the Main Injector.

2.1.2 Main Injector

The main injector is a synchroton accelerator with a circumference of about 3 km. Its purpose is to:

- accept protons coming from the Booster or antiprotons from the Accumulator;
- accelerate protons up to 120 GeV and send them either to the Target Station, or to the fixed target aerea for other experiment's use, or in alternative to a neutrino facility;
- accelerate protons and antiprotons from 8 GeV to 150 GeV and send them to the Tevatron collider;
- accept antiprotons from the antiproton source or the Recycler, accelerate them to 150 GeV and send them to the Tevatron

2.1.3 Antiproton Source

Protons coming out of the main injector at 120 GeV are smashed on a target in the Target Station. The number of antiprotons collected is very low: about 20 $\bar{p}$ per 10^6 protons. The resulting particles are focused into a beam using magnet quadrupoles and a 8 GeV beam of antiprotons is obtained; thereafter the beam is sent to the Accumulator, a storage ring used to collect antiprotons until about $1.3 \times 10^{12}\,\bar{p}$ are produced. The accumulation rate is about $7 \times 10^{10}\,\bar{p}$ per hour so about 15 hours are needed to collect an adequate amount. Prior to collider Run II, which began in 2001, the Accumulator core was the final destination for antiprotons before transfer to the Tevatron via the Main Ring. When the Main Injector ring was designed as a replacement for the Main Ring, another antiproton storage ring, known as the Recycler, was proposed. A separate storage ring dedicated to cooling the antiprotons prior to transfer to the Tevatron could allow the Antiproton Source to increase the pbar accumulation rate. The Recycler ring installation was not completed until after the beginning of Run II. A lengthy period of commissioning and upgrades followed, which included the installation of electron cooling. In 2005, the Recycler

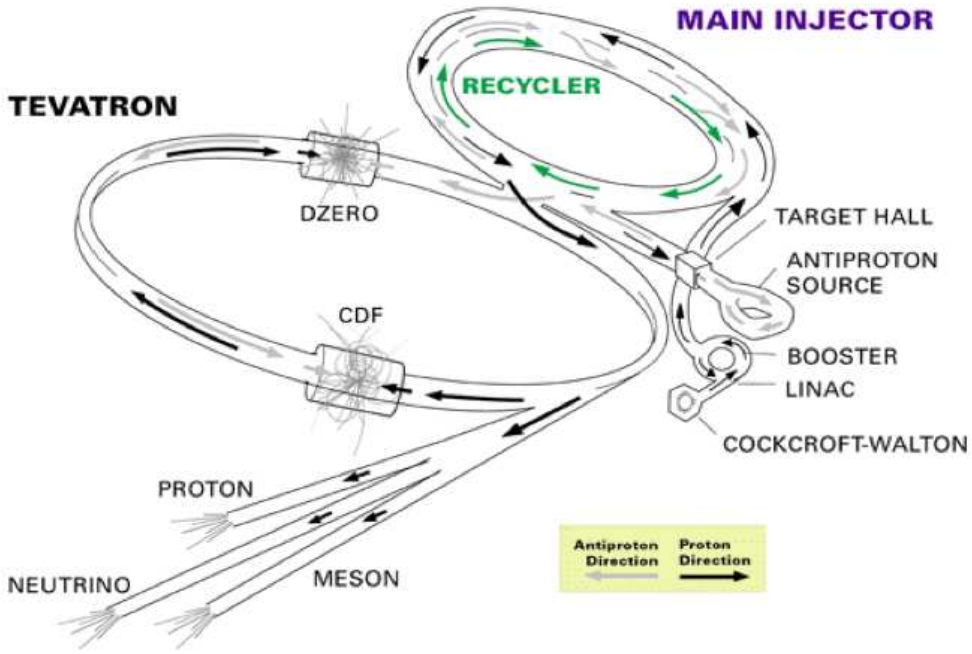


Figure 2.1: Schematic view of the Tevatron accelerator complex.

ring and electron cooling entered operation, boosting the instantaneous luminosity well beyond the Run II design of $2 \times 10^3 2$ cm^{-2} s^{-1}

2.1.4 The Tevatron

The Tevatron is a synchroton with a 1 km radius. Particles are bent thanks to superconducting magnets with $B \simeq 5.7$ T, and energies reach 980 GeV per beam. Beams are subdivided in 36 bunches each, which meet at 72 interaction points along the ring. Most of them are parasitical, while in the two regions where the CDF II and D0 detectors are located the beams are further focused to increase the luminosity. The latter is defined (for a machine in which beams collide head-on) as

$$\mathcal{L} = \frac{N_p \, N_{\bar{p}} \, f_R \, N_B}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} \tag{2.1}$$

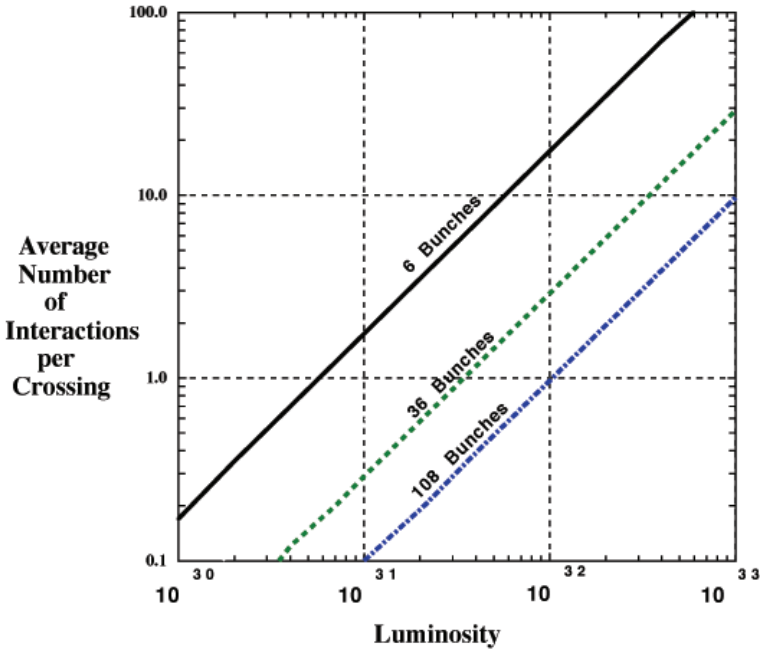


Figure 2.2: Average number of $p\bar{p}$ interactions per bunch crossing as a function of the instantaneous luminosity and of the number of circulating bunches. The current configuration is set to 32×32 bunches.

where $N_p(N_{\bar{p}})$ is the number of $p(\bar{p})$ circulating, f_R is the revolution frequency, N_B is the number of circulating bunches and $\sigma_{p(\bar{p})}$ is the width of the spatial distribution of $p(\bar{p})$ on the transverse plane at the interaction point. The interaction region has a spatial distribution of about 30 cm along the beam direction ($\sigma_z \simeq 30\,\text{cm}$) while on the transverse plane the beam is approximately circular with $\sigma_T^{beam} \simeq 25\,\mu\text{m}$. The number of multiple interactions is a Poisson variable whose mean is shown in Fig. 2.2 as a function of the number of circulating bunches. To date, the maximum luminosity achieved with the Fermilab accelerator complex reached $4 \cdot 10^{32}\,\text{cm}^{-2}\text{s}^{-1}$. This search is based on the data collected in the Run II, up to the end of 2009. The Tevatron most important parameters are summarized in Table 2.1.

2.2 The CDF II detector

The Collider Detector at Fermilab (CDF II) (Fig. 2.4) is a multi-purpose detector; it is designed to study a wide range of physics processes produced at proton-antiproton

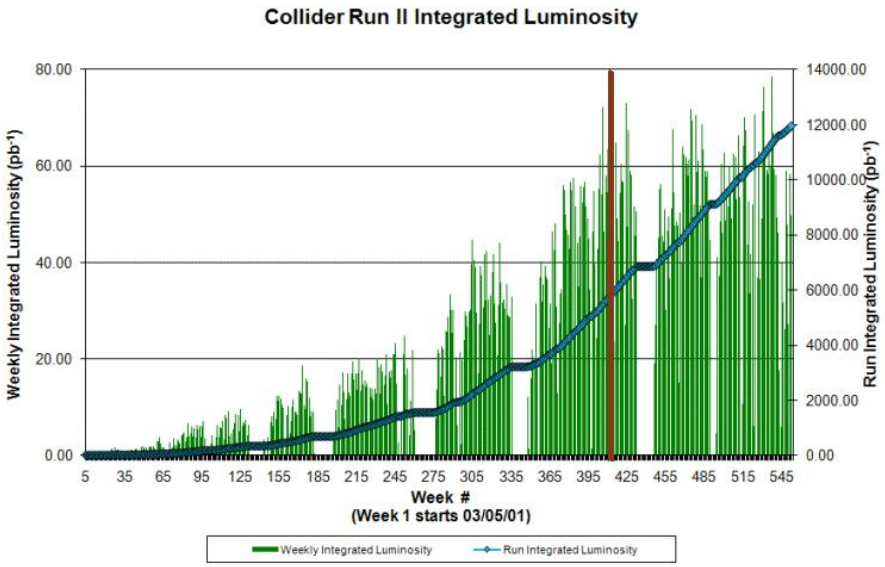


Figure 2.3: Tevatron integrated luminosity as a function of time (Run II data taking period, 2001-2011). The brown line shows the amount of data used in this analysis.

Parameters	***Value***
$p \times \bar{p}$ bunches	36×36
Number of p per bunch	$3.3 \cdot 10^{11}$
Number of $\bar{p}$ per bunch	$3.6 \cdot 10^{10}$
Total number of $\bar{p}$	$1.1 \cdot 10^{12}$
p emittance (mm mrad)	30
$\bar{p}$ emittance (mm mrad)	20
Energy $(p + \bar{p})$ (GeV)	980+980
Bunch spacing (ns)	396
$\mathcal{L}$ ($\mathrm{cm}^{-2}\mathrm{s}^{-1}$) (peak)	$4.0 \cdot 10^{32}$
Number of interactions/collisions (peak)	10

Table 2.1: Summary of the most important Tevatron parameters.

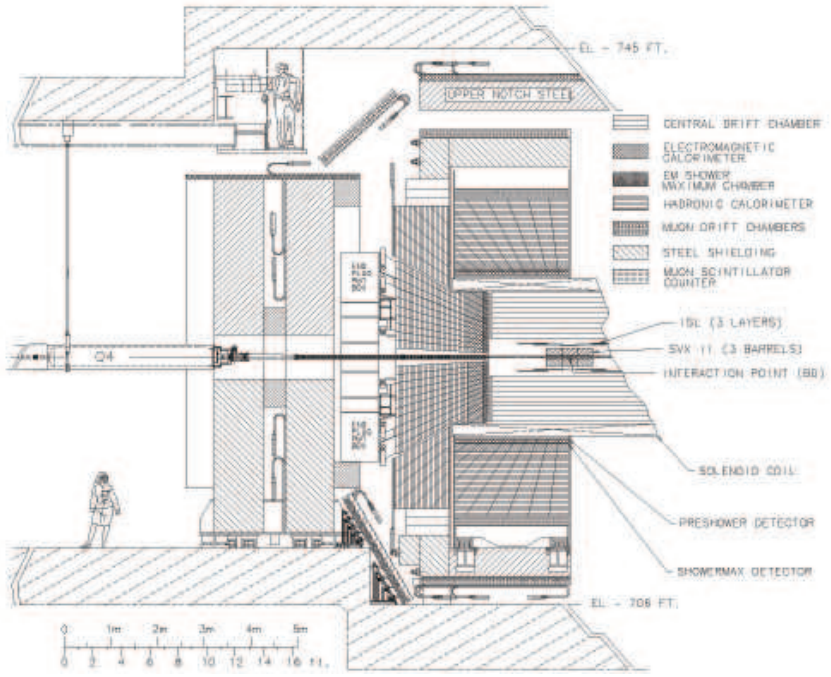


Figure 2.4: Side view of the CDF II detector.

interactions and characterized by final states with high transverse momenta particles. Since the two beams collide head-on, the detector exhibits a forward-backward symmetry, and a cylindrical symmetry around the beam-pipe. Starting from the beam-pipe and proceeding radially outwards, we first encounter the silicon vertex detector, surrounded by a drift chamber, both contained in a magnetic field of 1.4 T produced by a superconducting solenoid. The energy of photons, electrons and hadrons is measured by electromagnetic and hadronic sampling calorimeters. Finally muons are identified and their four-momenta measured in the proportional chambers located outside the calorimeters. In the forward region we also have some additional detectors, out of which a very important one is a Cherenkov luminosity counter.

There are too many collisions to be recorded, but luckily most of them are of little interest. A trigger system made of three levels decide whether or not to record the outcome of the collisions. We will now describe the CDF II reference system and define some quantities which will be used in the following; a detailed description of the subdetectors listed above will follow.

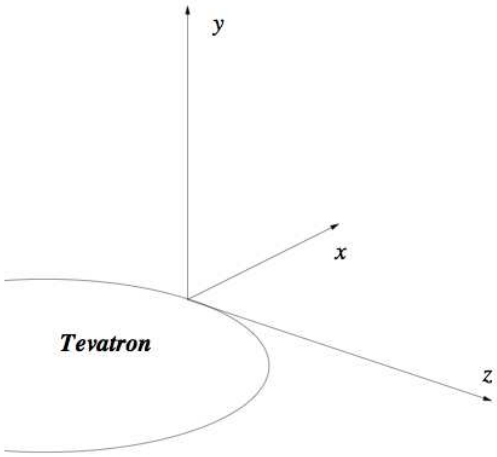


Figure 2.5: CDF II (x, y, z) reference system.

Reference systems

Since the detector has a cylindrical symmetry around the beam axis, a convenient choice for the coordinate system is the cylindrical one, where the z axis coincides with the proton beam direction, and ρ and ϕ are measured on the orthogonal plane intersecting the nominal interaction vertex. We define also a cartesian reference system using x and y where the former coordinate points outside the accelerator plane and the latter perpendicularly to it, as shown in Fig. 2.5. The ϕ angle is measured with respect to the positive direction of the x axis. It is also useful to define θ as the angle with respect to the positive z direction. Since we do not know the Lorentz boost of the proton-antiproton center of mass with respect to the laboratory reference system, we introduce a quantity which transforms peculiarly under Lorentz transformations to describe the forward direction, while the transverse component is of course invariant. Let's consider a particle with energy E and momentum P. The quantity called rapidity y is defined as:

$$y = \frac{1}{2} \ln \left(\frac{E + P_z}{E - P_z} \right).$$

The rapidity transforms as:

$$y \to y + \text{const} = y + \tanh^{-1} \beta_z$$

(being $v_z = \beta_z c$ the velocity in the reference frame of the partons). The difference in rapidity between two particles is thus unaltered by a Lorentz boost along the beam axis. The rapidity is particularly useful in the limit $P \gg m$ (which is generally correct at the Tevatron energies):

$$y(P \gg m) = \frac{1}{2} \ln \left(\frac{P + P_z}{P - P_z} \right) = - \ln \left(\tan \frac{\theta}{2} \right) = \eta. \tag{2.2}$$

The quantity η, called pseudorapidity, is a function of the θ angle only; from now on we will use η to describe the direction of particles along the (y, z) plane.
A related quantity, $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. is typically used to estabilish criteria of closeness, labeling as close those particles contained inside a circumference of arbitrary radius ΔR on a (η, ϕ) plane.

2.2.1 Tracking and vertexing systems

The tracking system occupies the inner volume of the solenoid (see Fig. 2.6). It is made up of the following detectors (proceeding from the beam pipe radially outward): the "Layer 00", the "SVX II" (*Silicon VerteX detector*), the "ISL" (*Intermediate Silicon Layer*) and the "COT" (*Central OuTer chamber*). All these detectors have cylindrical simmetry. In CDF II the silicon detectors are comprised between the beam pipe and the COT and constitute a stand-alone tracking system that has a pseudorapidity coverage which extends up to $|\eta| \leq 2$.

Layer 00

It is constituted from a single layer of silicon sensors with microstrips aligned with the beam axis. The spatial resolution on the crossing point of charged particles is approximately $6\,\mu$m. Layer 00 lies on the external surface of the beam pipe, therefore at an average distance $r \simeq 1.6\,$cm from the nominal beam axis and covering longitudinally the region $|z| < 40\,$cm.

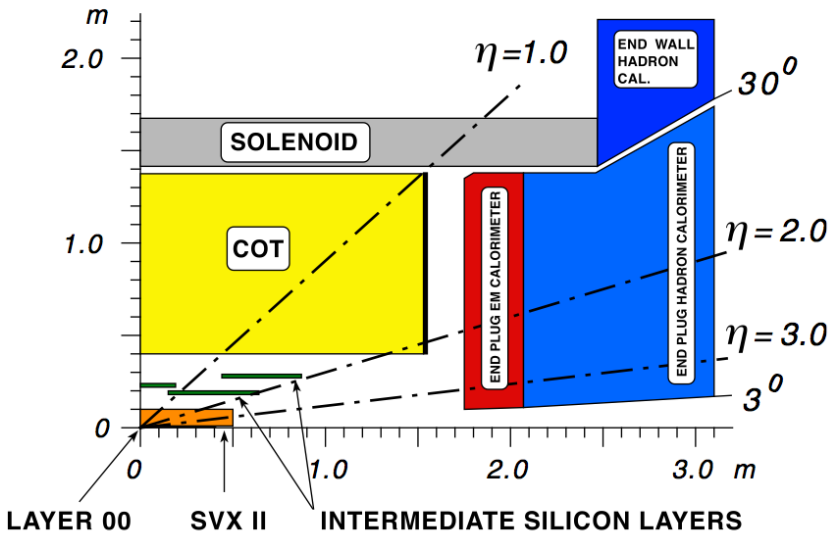


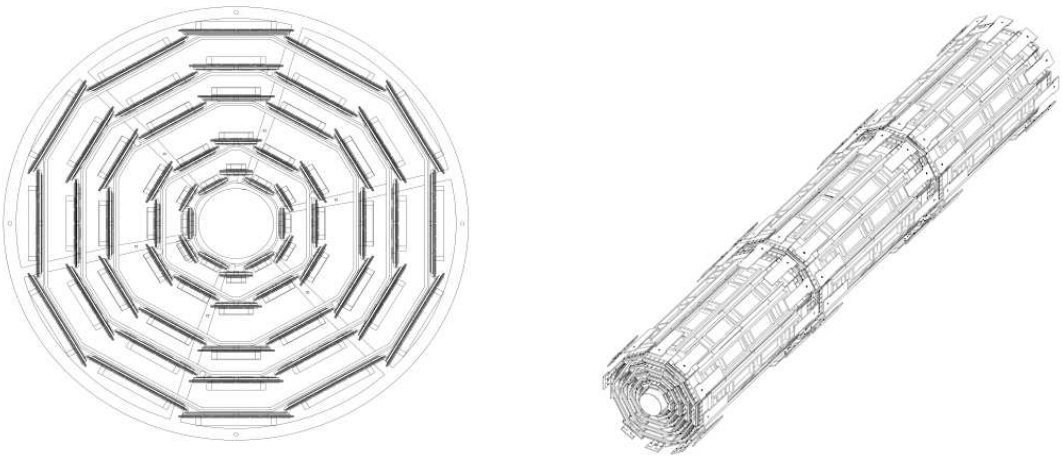
Figure 2.6: Cross section view of the CDF II tracking and calorimetry subsystems.

SVX II

At a distance from the axis comprised between 2.4 cm and 10.7 cm, 5 silicon layers are arranged radially on 3 barrels, each subdivided in 12 wedges (Fig. 2.7). Each layer has microstrips on both sides: on one side they are aligned along the z-axis, on the other orthogonally (3 layers) or to a stereo angle $\pm 1.2^o$ (2 layers). Such a geometry of the strips allows an optimal reconstruction of the tracks in the transverse plane and in the $r-z$ plane. The detector extends longitudinally for approximately 96 cm covering a fraction of the luminous region corresponding to $2.5\,\sigma$.

The microstrips' size is approximately 60 μm on the $r-\phi$ plane. The resolution on the position of the single hit is approximately 16 μm for axial strip and for the small stereo angle layer (38 μm for the stereo orthogonal layer). The main informations are summarized in Table 2.2. Approximately 406000 strips (or channels) of SVX II are read using the fast chips SVX3D in less than 10 μs. The information on the $r-\phi$ coordinates is made available first to the trigger system. The electronic devices and the sensors are designed in order to resist to large radiation doses ($\sim 0.5\,\mathrm{MRad/fb^{-1}}$).

The Layer 00 guarantees five layers of sensors in the case of damage of the first silicon layer. The thickness of the detector in term of the radiation length (X_0) is



(a) (b)

Figure 2.7: (a) SVX II cross sectional view. (b) The three SVX II barrels.

Parameter	***Value***
Distance from beam pipe (cm)	$2.4 \leftrightarrow 10.7$
Pseudorapidity range	$\|\eta\| < 2.0$
Barrels/layers/wedges	3 / 5 / 12
Length/active length (cm)	$96 / 29 \times 3$
$r - \phi$ pitch (μm)	60, 62, 60, 60, 65
$r - z$ pitch (μm)	141, 125.5, 60, 141, 65
Stereo angle (degrees)	90, 90, 1.2, 90, 1.2
Channels	$211968\,(r-\phi) + 193536\,(stereo)$

Table 2.2: Characteristic SVX II parameters.

$0.1\,X_0$ ($\theta = 90^o$) that becomes $0.2\,X_0$ for the regions where the readout electronics is installed.

ISL

The Intermediate Silicon Layer is located between the vertex detector and the drift chamber. It consists of three layers of silicon with axial microstrips on one side, separated by $110\,\mu$m, and stereo, with an angle of $\pm 1.2^o$, on the other ($146\,\mu$m). The resolution on the position of the single hit is about $16\,\mu$m for axial strips and $23\,\mu$m for the stereo ones. Of the three layers, the central has an average distance of 22 cm from the beam axis and covers the pseudorapidity range $|\eta| < 1$ while the remaining two are respectively at 20 cm and 29 cm, both covering $1 < |\eta| < 2$. The total length is approximately 174 cm. The layers are partially overlapped on the $r - z$ plane (Fig. 2.6) and on the $r - \phi$ plane. The thickness of a detector layer is, in average, $0.5\,X_0$. The tracking system composed from SVX II and ISL covers the entire luminous region and has 6 layers available overall in the central region ($|\eta| < 1$) and 7 in the plug ($1 < |\eta| < 2$). SVX II, because of its insufficient radial extension, does not allow a good resolution on the transverse momentum P_T (and therefore on the impact parameter (d_0) and on the variables relative to the $r-z$ plane ($\cot\theta$ and z_0). The information supplied from ISL is useful for three-dimensional reconstruction in the central region (where the COT measurement is also available), and in the plug region where the presence of an additional silicon layer compensates the reduced coverage of the drift chamber.

COT

The Central Outer Tracker is an open-cell drift chamber located at radii between 40 cm and 132 cm. The wires are subdivided in 8 superlayers (SLs), divided in 4 axial superlayers for the measurement on the transverse plane, and 4 stereo superlayers (stereo angle $\pm 2^o$), for the measurement of the z coordinate. Each superlayer contains 12 wires for the collection of signal for a total of 96 measurement points of each charge particle trajectory. The maximum drift time is approximately 100 ns. This allows a correct operation of the chamber with a bunch spacing of 396 ns and to use the information of the COT at the first level of the trigger. The resolution

Parameter	***Value***
Radius (cm)	$40 \leftrightarrow 132$
Length (cm)	310
Gas	Ar-Et-CF_4 (50:35:15)
Max drift distance (cm)	0.88
Max drift time (ns)	100
Drift field (kV·cm^{-1})	2.6
Lorentz angle	35^o
Superlayer × axial wires	4×12
Stereo angle	$\pm 3^o$
Superlayers × stereo wires	4×12
Total wire number	63000
Thickness (X_0)	1.7%

Table 2.3: Main COT parameters.

on the position measurement of a single hit is approximately 180 μm. The tracking efficiency is estimated, for pions, to be approximately 95% in a wide range of occupancy, falling steeply for $|\eta|$ larger than 1 (see Fig. 2.8). The material of the COT is equivalent to approximately 1.7% of a radiation length ($\theta = 90^o$). The main informations are given in Table 2.3. The COT resolution on the charged particles transverse momentum is $\sigma_{P_T}/P_T^2 \sim 1.7 \cdot 10^{-3}$ (GeV/c)$^{-1}$. If we are to consider altogether the three tracking systems (SVX II + ISL + COT), then the resolution becomes $\sigma_{P_T}/P_T^2 \sim 1 \cdot 10^{-3}$ (GeV/c)$^{-1}$.

2.2.2 Calorimetry

Calorimetry is the main component used for the measurement of the energy of hadrons, jets, electrons and photons. The system employed in CDF II consists of a sampling calorimeter shaped in a projective tower geometry which provides a full azimuthal coverage (2π) and pseudorapidity coverage up to $\eta = 3.6$. The central tower segmentation in η is $\Delta\eta = 0.1$ and in ϕ is $\Delta\phi = 7.5^o - 15^o$. Each tower is actually made of two separate devices, the electromagnetic calorimeter and the hadronic calorimeter. Both are sampling devices, the first being made of alternating layers of lead and scintillator and the second of iron and scintillator. Overall, the CDF II calorimeter is composed of three main parts, which corresponds to two

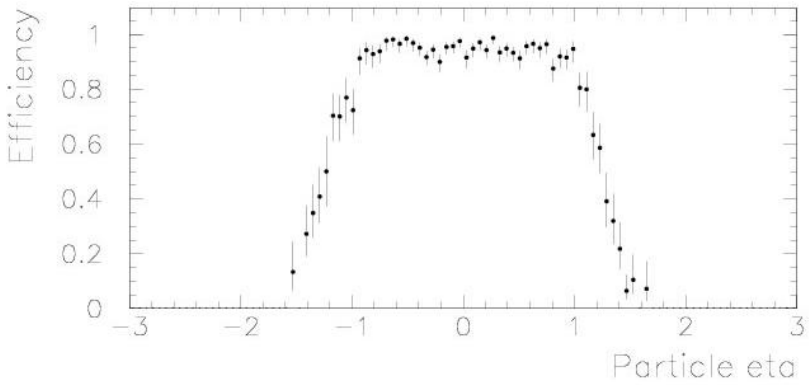


Figure 2.8: Relationship between particles pseudorapidity and COT tracking efficiency.

different regions in η:

- the central and wall calorimeter ($|\eta| < 1$);
- the plug calorimeter ($1 < |\eta| < 3.6$).

We deal with the first two together in the next section (since the wall calorimeter is actually a subset of the central hadronic calorimeter). The description of the plug calorimeter will follow.

Central calorimeter

The central calorimeter is a collection of different devices. Proceeding radially outwards we meet first a preshower detector, then the electromagnetic calorimeter which is equipped with an electromagnetic shower maximum detector, and finally the hadronic calorimeter. Let's now describe the subdetectors in some detail.

Right outside of the solenoid we find the proportional chamber CPR (*Central PRereadiator detector*), which supplies the information on the position of electromagnetic showers that are produced in the solenoid.

Then the electromagnetic calorimeter (or CEM – Central ElectroMagnetic) [40] follows. It is divided in two halves symmetrical with respect to the plane $z = 0$. The segmentation $\Delta\eta \times \Delta\phi = 0.10 \times 15^o$ corresponds to having 24 wedges along the azimuthal direction, where each of the two halves is subdivided in 10 wedges.

The CEM consists of 31 layers of plastic scintillator 5 mm thick alternated with 30 layers of lead 3 mm thick each, for a total of 18 X_0. Light guides collect from two sides of the tower the photons coming from the scintillators and carry them to two photomultipliers located in the external region of the calorimeter. The space for the light guides constitutes a dead zone in the regions of separation in ϕ between the towers, called ϕ-cracks, that amounts to the 4.5% of the entire coverage in ϕ. Not-instrumented zones are also present along the η direction (η-crack) because of the aluminum sheets of ~ 0.4 mm of thickness that cover absorber and scintillators. Another η crack is constituted from the region $\eta = 0$ corresponding to the mechanical separation between the two halves. The electromagnetic calorimeter is calibrated with electrons from test beam and monitored with LED, xenon, and with radiation from ^{60}Co and ^{137}Cs sources. The energy resolution amounts to $\frac{13.5\,\%}{\sqrt{E_T}} \oplus (1.5\,\%)$.

Another proportional chamber, the CES (*Central Electromagnetic Strip detector*), is located within the calorimeter at a distance of approximately 6 X_0) from the lower face of the CEM, that corresponds to the distance in which on average the development of the electromagnetic shower is maximum. Its purpose is to facilitate the identification of $e^{\pm}$ and γ through the observation of the electromagnetic shower's shaper that characterizes their interactions in the scintillator, and to allow the separation of π^0 from γ through the different shape of the shower[1].

The central hadronic calorimeter is divided in two separated mechanical parts: Central HAdronic and Wall HAdronic, respectively CHA and WHA [41]. The first one contains 8 towers for side, each made of 32 layers of alternating scintillator and absorber, where the scintillator layers are 1.0 cm thick and the absorber is constituted from sheets of 2.5 cm of iron; the second is made of 6 towers for each side with 15 layers, of which half of scintillators (1.0 cm thick) and the rest of iron absorber of 5.0 cm. Both cover in total 5 interaction lengths (λ_i). Every tower is read by two photomultipliers. The calibration is done with π's, while the monitoring with mixed techniques that include light laser, γ emissions from ^{137}Cs, sources of β

[1]The CES is made of anodic wires that run along the z axis of the CDF II detector and measure the x coordinate, and orthogonal cathodic strips that measure the z coordinate. The CES fiducial coverage corresponds to $|x| < 22$ cm and $14\,\text{cm} < z < 217$ cm. A clustering algorithm identifies the strip or wire clusters, that will be associates to photons or π^0, or to electrons in the case there are the tracks that point to the cluster.

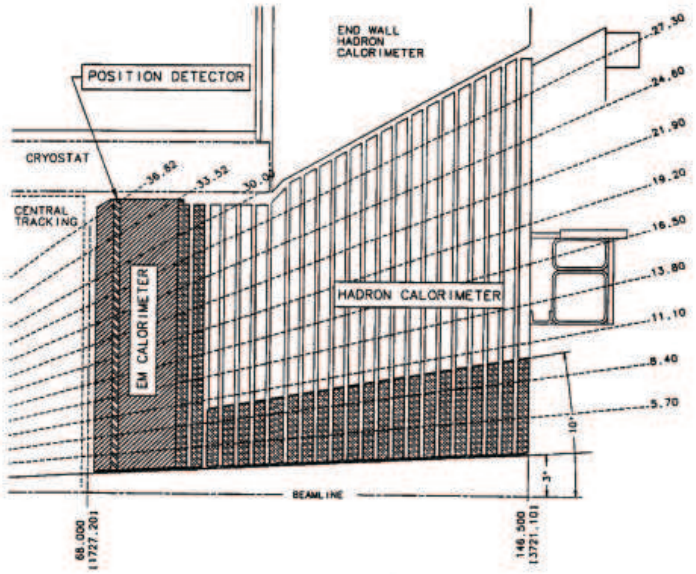


Figure 2.9: Quadrant view of the plug calorimeter CDF II.

radiation. The energy resolution is $\frac{75.0\%}{\sqrt{E_T}} \oplus 3\%$ for the CHA and $\frac{80.0\%}{\sqrt{E_t}} \oplus 4\%$ for the WHA.

Plug calorimeter

The plug calorimeter covers the pseudorapidity range $1.1 < |\eta| < 3.6$. The scintillator is composed of a mixture of 50% argon e 50% ethane. Segmentation in $\eta \times \phi$ varies from 0.11×15^o to 0.11×5^o depending on the region. The system is subdivided, analogously to the other calorimeters, into an electromagnetic device and an hadronic device (Fig. 2.9). The electromagnetic calorimeter [42] is composed of 23 alternated layers of 4.5 mm of lead absorber and 4 mm of scintillator for a total of $\sim 21\,X_0$. The first layer is a scintillator 10 mm thick. The hadronic calorimeter is also a sampling device with 23 layers where sheets of 5 mm of iron are interleaved with sheets of 6 mm of scintillator. The segmentation $\Delta\eta \times \Delta\phi$ varies from 0.1×7.5^o to 0.6×15^o. The resolution for the measurement of energy is approximately $\frac{14\,\%}{\sqrt{E}} \oplus 1\,\%$ and $\frac{80\,\%}{\sqrt{E}} \oplus 5\%$ for the electromagnetic and hadronic calorimeters respectively. As was the case for the central calorimeter, the plug calorimeter is equipped with a shower maximum detector [43]. It is made of scintillator strips at approximately

Electromagnetic calorimeter					
Subsystem	*η Region*	*Type*	$\sigma_{E_t}/E_t\,(\%)$	*Thickness*	$\Delta\eta \times \Delta\phi$
CEM	$\lvert\eta\rvert < 1.0$	*Pb-scint.*	$13.5/\sqrt{E_t} \oplus 1.5$	$18\,X_0$	0.1×15^o
PEM	$1.1 < \lvert\eta\rvert < 3.6$	*Pb-scint.*	$14.0/\sqrt{E_t} \oplus 1$	$18 - 21\,X_0$	0.1×5^o
Hadronic calorimeter					
Subsystem	*Region*	*Type*	$\sigma_{E_t}/E_t\,(\%)$	*Thickness*	$\Delta\eta \times \Delta\phi$
CHA	$\lvert\eta\rvert < 0.9$	*Fe-scint.*	$75.0/\sqrt{E_t} \oplus 3$	$5.5\,\lambda_i$	0.1×15^o
WHA	$0.6 < \lvert\eta\rvert < 1.3$	*Fe-scint.*	$80.0/\sqrt{E_t} \oplus 4$	$5.5\,\lambda_i$	0.1×15^o
PHA	$1.1 < \lvert\eta\rvert < 3.6$	*Fe-scint.*	$80.0/\sqrt{E_t} \oplus 5$	$8.0\,\lambda_i$	0.1×15^o

Table 2.4: Main characteristics of the CDF II calorimeter. The resolutions for the electromagnetic calorimeters (hadronic) are relative to isolated photons and electrons (pions). $E_t = E \cdot \sin\theta$ (GeV). $A \oplus B \equiv \sqrt{A^2 + B^2}$. The thickness, for particles incidents normally, are indicate in radiation lengths (X_0) and interaction lengths (λ_i) ($21\,X_0 \simeq 1\,\lambda_i$). These represent the average distance traveled such that, respectively, an electron loses $1/e$ of its initial energy for emission of radiation (*bremsstrahlung*) and a pion gives rise to an inelastic interaction

180 cm from the origin of the reference system, corresponding to approximately $6\,X_0$ from the base of the calorimeter, at the depth where in average the extension of the electromagnetic shower is maximum. The longitudinal coverage of hadronic showers amounts to 8 interaction lengths λ_i. A summary of the main informations on the CDF II calorimetry is available in Table 2.4.

2.2.3 Cherenkov Luminosity Counter

The Cherenkov Luminosity Counter (CLC) measures the average number of interactions per bunch crossing, μ. The instantaneous luminosity $\mathcal{L}$ is extracted using the equation

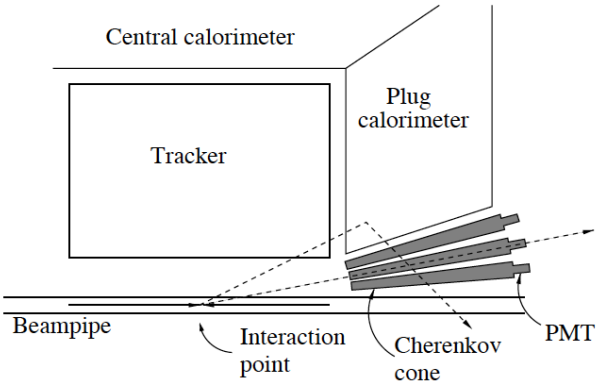
$$\mu \cdot f_{bc} = \sigma_{p\bar{p}} \cdot \mathcal{L} \tag{2.3}$$

where $\sigma_{p\bar{p}}$ is the relatively well known total $p\bar{p}$ cross section at $\sqrt{s} = 1.96$ TeV, and f_{bc} is the rate of bunch crossings at the Tevatron. The CLC uses the effect known as the Cherenkov radiation, where particles traversing a medium at a speed higher than the speed of light in that medium radiate light into a cone around the particle direction; the cone's opening angle depends on the ratio of the two speeds and on the refraction index of the medium. The idea is to use an assembly of long gas

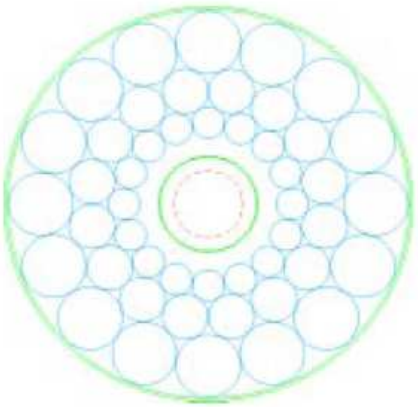
Cherenkov counters positioned in the plug calorimeter 3^o gap so that they point toward the interaction point as schematically shown in Fig. 2.10. This arrangement allows to make the detector much more sensitive to the particles coming directly from the interaction point because their flight path in the gas of the counter is the longest and therefore the amount of light produced the largest. Excellent time resolution ($\sim 50\,\text{ps}$) and clever design allow the CLC to discern multiple interactions within the same bunch crossing and achieve an overall accuracy of the luminosity measurement better than 6%. In depth information on the CLC design and performance is given in reference [80].

2.2.4 Muon detection

Muons are characterized by their high penetrating power through matter. They interact only electromagnetically and weakly, and loose a small amount of energy through brehmsstrahlung as opposed to electrons, thanks to their higher mass. As a result they leave a minimum amount of energy in the calorimeter and reach the dedicated drift chambers located at the outermost part of the detector. In fact, outside the hadronic calorimeter CHA, almost 3.5 m away from the beam line ($\sim 5.4\,\lambda_i$), is located the CMU (*Central MUon*) made of 4 layers of drift chambers (4 hits). A similar device, CMP (*Central Muon uPgrade*), is located behind an iron layer 60 cm thick ($\sim 3\,\lambda_i$). A muon has to have at least 1.4 GeV/c of transverse momentum to reach the CMU and higher than 2.8 GeV/c to reach the CMP. The probability for an hadronic particle to reach the CMU is 1%, while for the CMP this is negligible. CMU and CMP cover the region $|\eta| < 0.6$. Coverage of the region $0.6 < |\eta| < 1$ is assured from the CMX chambers (*Central Muon eXtension*). These are made of 4 layers, and are located between two scintillating layers to be used for triggering (CSX). The muons' "stubs" are reconstructed as segments formed by hits in the CMU, CMP and CMX detectors The single hit resolution is about $250\,\mu$m on the $r-\phi$ plane and 1.2 mm along the z axis. Since outside the solenoid the magnetic field is absent, we can reconstruct only stubs of tracks with the muon chamber. This information has to be integrated with the COT tracks to define a muon. Another constraint is the presence of energy in the hadronic calorimeter compatible with a



(a)



(b)

Figure 2.10: a) Quadrant view and b) cross sectional view of the Cherenkov Luminosity Counters at CDF II.

release from a μ ($\sim 0.5\,\mathrm{GeV}$).

2.2.5 Trigger and data acquisition systems

In collisions between two bunches, a proton and an antiproton interact through an inelastic collisions. The outcome is the production of tens of particles that fly away from the interaction zone. A fraction of these go through the detector in one or more active regions (the silicon of the vertex detector, the gas in the drift chamber, some towers of the calorimeter etc.) where they will interact with the detector material and produce some kind of release.

The data acquisition system (DAQ - Data AcQuisition) is a collection of devices that have the task to manage the informations coming from the detector from the phase of reading until the phase of writing on a support of permanent memory. The acquisition begins with digital conversion of the analogic signal supplied from the single detectors. Then they are collected, elaborated in real time and, finally, recorded on a magnetic tape where they remain available for the offline analyis. The acquisition of the data relative to the collision between two bunches is labeled "event acquisition".

In the current accelerator configuration bunches collide every 396 ns, that is $2.5 \cdot 10^6$ collisions $\times$ second. To acquire the event for every bunch crossing would demand electronic devices of prohibitive speed and storage volume. However, to collect all the events indiscriminately is not necessary: the interesting processes (therefore less known) constitute only a small fraction of the total, in the proportion between their production cross sections and the total $p\bar{p}$ cross section. For example, in our case, theoretical calculations for the $T'T'$ production cross section range from 5 pb to 2 fb in the mass range 200-400 GeV/c^2 [31], while inelastic cross section for $p\bar{p}$ scattering amounts to 61 mb [45], that is to say about 1 $T'T'$ event produced every $10^{10} \div 10^{13}$ inelastic collisions. The task of selecting the events is performed by the trigger system; the DAQ works, therefore, in synchronism with the trigger and is based on the instructions received from the latter.

The basic requirement is is that an inelastic interaction between a proton and an antiproton has undergone. This kind of process is nearly always accompanied from

a large particle multiplicity, produced in the partonic interactions at low transferred moment, emitted with a small angle θ.
The trigger system is structured in three levels. Every level selects a fraction of the events that have survived the previous one, allowing a more detailed analysis at the subsequent level. The electronic devices who make part of the trigger produce a measurement of the quantities of the event that are useful to isolate, *i.e.* the signal. When the measured quantity exceeds the threshold value, it triggers the execution of the following level, otherwise the information is lost and the DAQ goes on to examine the following event.

The first level takes into account all the events after every bunch crossing. The events that survive the third level are written on tape. Generally a trigger does not succeed to decide whether to keep or not an event before the following collision. The dead-time is defined as the fraction of time in which the trigger is busy examining an event and can not accept others. A trigger organized on more levels allows to apply sufficiently effective criteria of selection without introducing a significant dead-time.

The effectiveness of a certain trigger is quantified through the rejection factor and its efficiency. The rejection factor is the ratio between the total number of events analyzed and the number that survive the selection criteria. The fraction of events passing the trigger requirement is called trigger rate. The rejection factor must be set to limit the trigger rate to a value that does not to introduce a dead time higher than a fixed threshold (approximately 10%). The efficiency, instead, refers to a specific signal and is defined as the fraction of the events of signal that survives the action of the trigger. The choice of the selection criteria has the scope of increasing the efficiency maintaining the rejection factor within the limits. For this purpose two processors have been built: XFT (*eXtremely Fast Tracker*) that allows to reconstruct charged tracks in the transverse plane, in a time short enough for the first level, and SVT (*Silicon Vertex Tracker*) to be able to measure the impact parameter of the reconstructed tracks and sending the information to the second level of the trigger system. The fundamental importance of these devices resides in the possibility to introduce selection criteria based, at Level 1, exclusively on the charged tracks and, at Level 2, on the presence of secondary vertices requiring tracks with high impact parameter. Such criterion is suited in order to select the

events in which a b quark is produced, being these characteristic of the presence of a hadron with long lifettime. The time that the first level employs in order to take one decision is approximately $5.5\,\mu$s. In order to avoid introducing dead-time every detector must have the possibility to store up to 42 data of successive collisions. For this reason all the systems of reading of CDF II are equipped with sliding registers with 42 cells (pipeline structure) where to keep the information while waiting for the trigger decision. The diagram of Fig. 2.11 describes the structure of the system of DAQ and trigger of CDF II. We now describe in some detail the three trigger level (Fig. 2.12).

Level 1 (*L1*) Level 1 trigger uses informations from the COT, the calorimeters and the muon chambers. The XFT processor reconstructs the charged tracks using the hits from the axial layers of the COT with reduced resolution ("XFT tracks"). The measured parameters are p_t and ϕ. The average time needed is $2.7\,\mu$s. The trajectories are then extrapolated until intersecting the lower face of the calorimeters and of the muon chambers (XTRP of Fig. 2.12). The possibility to formulate the selection criteria is based on the presence of tracks in the COT that satisfy specific kinematic requirements. Electrons, photons and jet candidates are identified imposing the presence of energy in the single towers of the calorimeters above the threshold values. The value of the sum of the energy released on all the towers is used for the selections based on the total transverse energy and the missing transverse energy. The selection of events containing muons is based on the location of stubs in the muon chamber in coincidence with signal from the scintillators. The availability of the parameters of the charged tracks allows to improve the identification of electromagnetic particles, hadronic and muons by checking the contiguity between the XFT tracks and the calorimetric towers and the stubs in the muon chamber. On charged tracks it is possible, moreover, to impose kinematic cuts (as an example the presence of one pair of tracks with $p_t > 3\,\mathrm{GeV/c}$). At this point the trigger system is able to examine simultaneously 64 sets of requests (each specific to various physical process) in order to decide to whether to accept an event or not. The maximum trigger rate expected for Level 1 is approximately 50 kHz (rejection factor: $\sim$ 150).

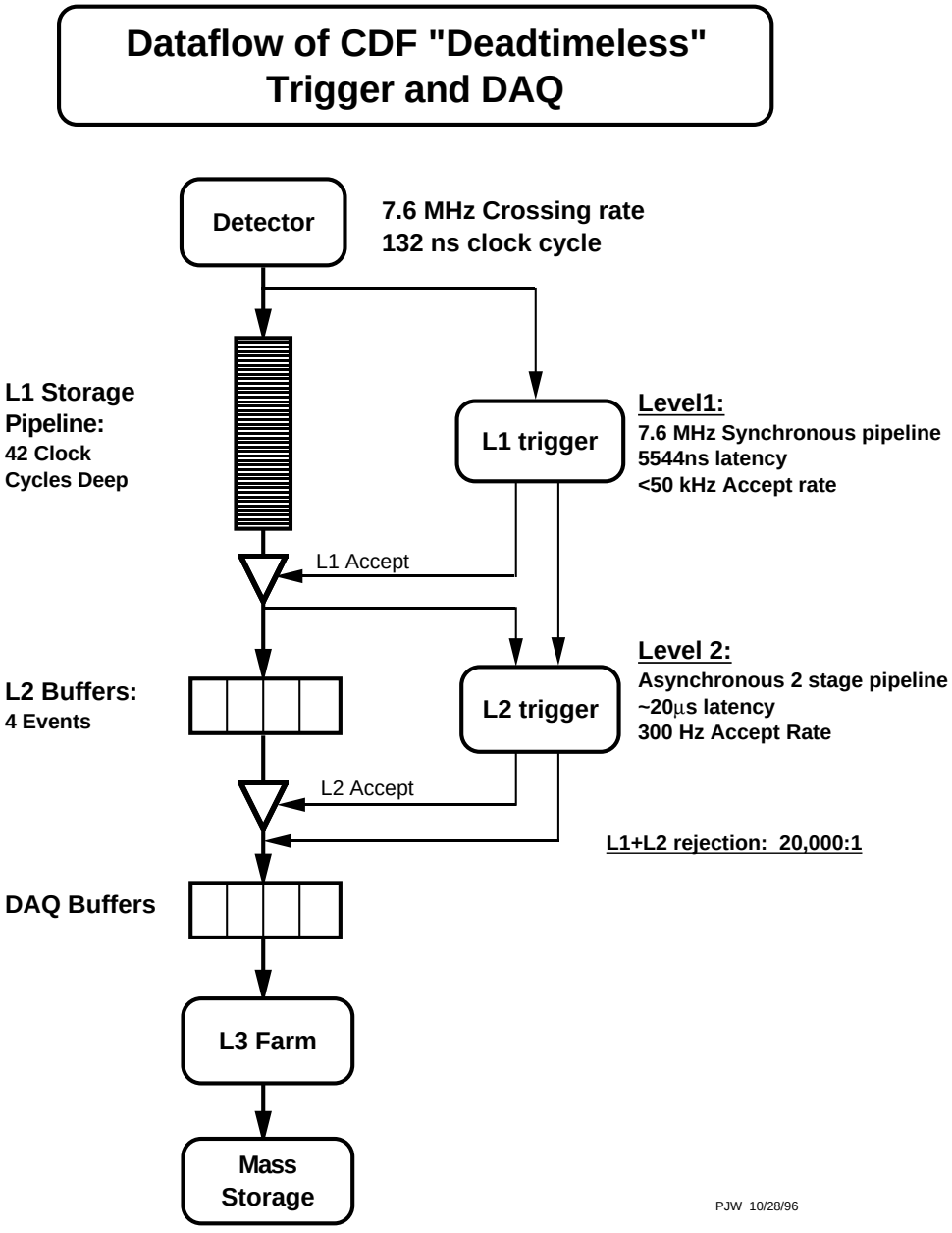


Figure 2.11: Data acquisition and trigger system at CDF II.

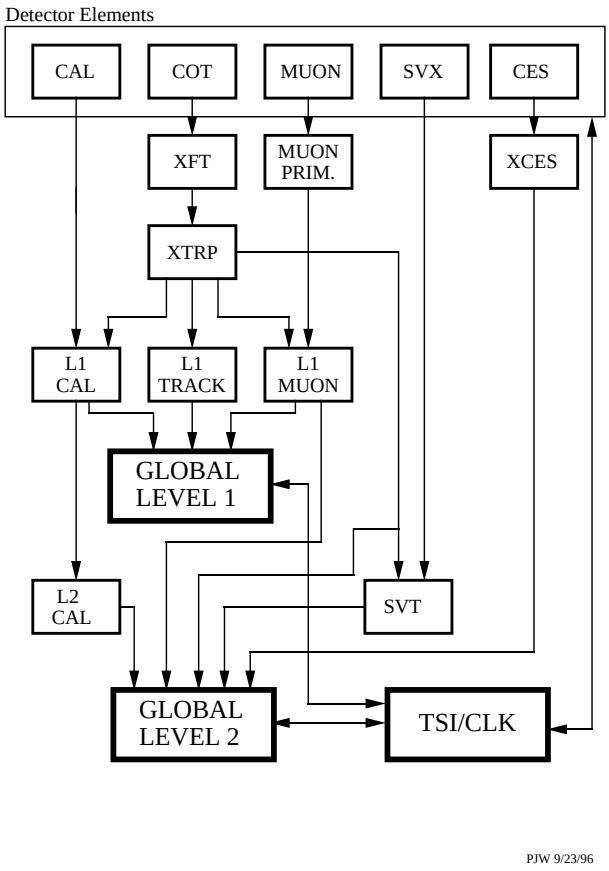


Figure 2.12: Trigger system at CDF II.

Level 2 (*L2*) Accepts events from L1 and records them on a buffer waiting to be examined. There are 4 available buffers. The XFT tracks with 4 hits in SVX II and $p_t > 2\,\mathrm{GeV/c}$ are reconstructed from SVT taking advantage of the additional information of the silicon detector: the measured parameters are p_t, ϕ and d (where d is the impact parameter, *i. e.* the closest distance between the extrapolated track and the primary vertex) with resolution comparable to that achievable with offline procedures. Algorithms are applied to trigger electrons photons and jets: the energy for clusters of adjacent towers, and the information from the detectors of maximum expansion of the shower in the electromagnetic calorimeters (module XCES of the Figure 2.12) is now available. The time of execution is approximately $20\,\mu$s. The frequency of

L2 trigger is limited to 300 Hz (rejection factor of $\sim$ 150). In 2007, because of the increase iof instantaneous luminosity, two major trigger upgrades were implemented in the Level 2: XFT stereo upgrade and the Level 2 calorimer upgrade. The first was thought to reduce the rate of fake tracks. The second makes the full calorimeter trigger tower information directly available to the Level 2 decision CPU, significantly improving the purity as well as the efficiency of the jet and missing energy related triggers.

Level 3 (*L3*) At this level complete informations supplied from the various detectors are available. The events selected from Level 2 are therefore elaborated with algorithms analogue to those of the offline analysis. In particular, the tracking is completed executing the three-dimensional reconstruction of the trajectories in the volume $|\eta| < 2$, and more detailed algorithms reconstruct the energy in the calorimeters.

Chapter 3

Physical Objects at CDF

The topology under study here is characterized by a very striking signature: at least 6 quarks in the final state and large missing energy. Quarks and gluons are not free particles in the Standard Model; they hadronize almost immediately giving rise to a spray of hadronic particles around the parent parton motion. The experimental signature of a jet is thus a cluster of approximately collinear hadrons which leave hits in the tracking system and energy deposit in the calorimeter. After a jet four-momentum is defined, it has to be corrected for both calorimetry and physics effects in order to best reproduce the parton energy; this is done through many steps, each one accounting for different effects. Finally, a systematic uncertainty on the jet energy scale is derived. If a hard interaction occurred, we expect a large number of tracks to point to a certain region along the beam axis z. It is thus useful to exploit the z coordinate of all tracks at the point of closest approach to the beam pipe to look for a primary vertex as a signature of the occurrence of the hard interaction itself. Finally, the production of undetected particles results in an energy imbalance in the transverse plane. The measurement of this missing transverse energy is of fundamental importance in this analysis, and will be described in detail.

3.1 Run requirements

Each run stored by the CDF experiment is associated to a database entry which specifies the amount of integrated luminosity written to tape and the status of all the sub-detectors. The runs used in this analysis are selected from those taken during the data taking period from March 2002 until December 2009. Collider runs which are known to contain problems are removed. The runs are required to have the calorimetry and the silicon tracker on and in good condition, and to contain at least $10\,\mathrm{nb}^{-1}$ of data written to tape.

3.2 Luminosity measurement

The total integrated luminosity is calculated from the rate $R_{p\bar{p}}$ of inelastic $p\bar{p}$ events measured by the luminosity monitor, Cerenkov Luminosity Counters (CLC). The CLC acceptance is estimated to be $(60.2 \pm 2.4)\%$ [80]. The inelastic cross section is derived by scaling the CDF measurement of $\sigma_{in} = 60.4 \pm 2.3$ mb at $\sqrt{s} = 1.8$ TeV to 1.96 TeV, resulting in $\sigma_{in} = 61.7$ nb.

Using these numbers and requiring the goodness of the run as described in Section 3.1, the data used in this analysis correspond to an estimated integrated luminosity of $(5.70 \pm 0.34)\,\mathrm{fb}^{-1}$. The 6% quoted uncertainty is dominated by the measurements of the absolute normalization of the CLC acceptance for a single $p\bar{p}$ inelastic collision.

3.3 Track reconstruction

Charged particle trajectories ("tracks") are reconstructed by combining the measuremenst ("hits") in the tracking system, where a uniform magnetic field provided by the solenoid magnet make them following a helical trajectory. The curvature of the helix depends on the momentum and charge of the particle.

Track reconstruction [46] starts in the COT, which surrounds the silicon tracker. Because of the larger radius this leads to a lower track density and better separated tracks. The track reconstruction is performed in two steps. First, four axial superlayers are used to reconstruct $r-\phi$ tracks. Second, for each $r-\phi$ track the algorithm

performs a 3D fit using all combinations of segments in the stereo super-layers that are consistent with the reference track. The newly estimated trajectory is then used to determine which of the segments in the remaining stereo layers are consistent with belonging to the track. The "Outside-In" tracking algorithm extends the COT tracks into the silicon, by extrapolating them into the silicon detectors and attaching silicon hits to the COT tracks. Hits in the silicon tracker that are not associated with COT tracks are used to construct the "Silicon Stand-Alone" tracks. In order to reduce combinatorics, hits that are attached to any other track are not used in this algorithm. The advantage of the Silicon Stand-Alone algorithm is that it allows to extend the tracking coverage up to $\eta < 2$, while the COT covers only $\eta < 1$. A third way to reconstruct tracks is by using the "Inside-Out" algorithm [47], which extrapolates the Silicon Stand-Alone tracks into COT. This algorithm recovers tracks that did not traverse the entire COT volume due to being in the intermediate rapidity region.

3.4 Primary vertex reconstruction

The location of the hard interaction in the event is referred to as the primary interaction vertex. A precise knowledge of the location of the primary vertex is required to correctly calculate the transverse components of physical observables in the detector, such as transverse energies of jets. Primary vertices are found at the intersection of the prompt tracks that satisfy a certain set of quality requirements. These tracks are iteratively added to the fit to a common origin. If a track causes the χ^2 of the fit to exceed a certain threshold, it is removed from the fit ("pruning"). This procedure continues until either the list of tracks is exhausted and a stable χ^2 of the fit is reached, or no vertex can be formed. In cases when several vertices are found, the one with the highest scalar sum of outgoing track p_T's is considered as the primary vertex of the event. The resulting resolution on the primary vertex position in the transverse plane ranges between 6 and 26 μm, depending on the event track multiplicity.

3.5 Secondary vertex reconstruction

Particles with long lifetime can produce in their decays vertices that are displaced with respect to the primary vertex. High energy b quark, while hadronizing, form B mesons and a shower of other hadronic particles. These jets are thus characterized by the presence of long-lifetime B mesons, $\tau_B \sim \mathcal{O}(1\,\mathrm{ps})$. The secondary vertexes reconstruction allows therefore to identify jets originating from b-quarks ("b-tagging"). In order to reconstruct these vertices, we need the high precision of the silicon detector in the spatial determination of the hits. The algorithm which looks for a secondary vertex, called SecVTX, takes as input objects the tracks from the COT, and the SVX II and ISL. It needs at least two tracks which pass a minimum threshold on transverse momentum (typically 0.5 GeV/c) and whose closest approach to the z axis is sufficiently displaced. Now, in order to tag a jet, the algorithm needs to look if one of these vertices lies inside a jet. We define L_{xy} the projection of the vector connecting the primary to the secondary vertex, $\vec{x}_{sec} - \vec{x}_{prim}$, on the transverse jet direction as:

$$L_{xy} = (x_{sec} - x_{prim}) \cdot cos\phi_{jet} + (y_{sec} - y_{prim}) \cdot sin\phi_{jet} \tag{3.1}$$

SecVTX considers as generated by heavy quarks the jets with L_{xy} positive and whose significance ($L_{xy}/\sigma_{L_{xy}}$) is greater than 7.5. In terms of tagging performance, some difference is found between data and Monte Carlo events. A correction factor (scale factor) is then introduced to compensate for such difference.

In the present analysis we compared the expected sensitivity with and without b-tagging requirements. As we will see in chpt. 6 the higher sensitivity is reached without this restriction.

3.6 Lepton reconstruction

In this work we analyze events in which we expect to find hadrons only. Nevertheless, lepton identification is necessary because we reject semileptonic $t\bar{t}$ and W+jets background with a veto on events containing at least one lepton (see chpt. 6).

Clearly, both electrons and muons need tracks to be reconstructed by the tracking system. Here we briefly describe how leptons are identified at CDF; in particular, the categories we use in this analysis are the Central Electrons, the Plug electrons and the Central Muons. The quantities used to identify lepton candidates are:

- the total transverse energy of the electron cluster in the electromagnetic calorimeter E_T
- the total transverse momentum of the electron track p_T
- the ratio of the total hadronic cluster energy to the total EM energy E^{had}/E^{em}
- the position of the track vertex along the z-axis, Z
- the number of axial superlayers containing a signal, AS, and the number of hits per superlayer, $hpAS$
- the number of stereo superlayers containing a signal, SS, and the number of hits per superlayer, $hpSS$
- the distance between the PES centroid and the PEM centroid, ΔR_{plug}
- the track impact parameter respect to the primary vertex $d0$
- the ratio of the energy collected in 5 layers over the energy collected in 9 layers of the PES, in the u and v orthogonal directions, $PES\_5\times9\_u$ and $PES\_5\times9\_v$
- the "track isolation" $Iso4$, defined as the scalar sum of transverse energies of the tracks in a cone radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$ around the electron candidate track
- The χ^2 resulting from the comparison of the PEM shower profile of the electron candidate with the profile of test beam electrons, using a (3×3) cluster size, $\chi^2(3 \times 3)$
- the matching between the candidate muon track and the stub in the CMU/CMP/CMX, $\Delta X(CMU/CMP/CMX)$

Central Electrons	
E_T	$\geq$ 20 GeV
p_T	$\geq$ 10 GeV
E^{had}/E^{em}	$< 0.055 + (0.00045 \times \mathrm{E})$
$\lvert Z \rvert$	$\leq$ 60 cm
$AS, hpAS$	$>$ 3, 5
$SS, hpSS$	$>$ 2, 5
$Iso4/E_T$	< 0.1

Plug Electrons	
E_T	$\geq$ 20 GeV
E^{had}/E^{em}	< 0.05
$PES\_5 \times 9\_u$	≥ 0.65
$PES\_5 \times 9\_v$	≥ 0.65
$Iso4/E_T$	< 0.1
$\chi^2(3 \times 3)$	< 10
ΔR_{plug}	$\leq$ 3 cm

Table 3.1: Central and Plug electrons identification criteria.

Electrons

Electrons are identified by the electromagnetic calorimeters CEM and PEM. The identification selections are different for CEM and PEM, and are listed in Tab. 3.1. Electron candidates identified in the central calorimeter must also match a track in the COT. The efficiencies of electron identification cuts are 0.923±0.001 and 0.837±0.003 respectively for central and plug electrons [48]. Fake rates are less than 1%.

Muons

We form muon candidates by matching a track to stubs in the muon chambers. All events are required to not pass the cosmic tag [49]. Furthermore, the energy deposited in the calorimeter by the candidate track is required to be consistent with that of minimum ionizing particles to remove fake hadronic particles. Finally, the same isolation requirement as for the electron candidates is applied, with the exception that the energy in the cone is compared to the track p_T instead. All requirements are listed in Tab. 3.2. The efficiencies of these cuts are 90.52±0.37

Central Muons	
p_T	$\geq$ 20 GeV
E^{em}	$<$ 2 GeV + max(0, 0.0115$\times$(p-100))
E^{had}	$<$ 6 GeV + max(0, 0.028$\times$(p-100))
E^{had}/E^{em}	$<$ 0.055 + (0.00045$\times$E)
$\|Z\|$	$\leq$ 60 cm
$AS, hpAS$	$>$ 3, 5
$SS, hpSS$	$>$ 2, 5
$\|d0\|$ (with silicon hits)	$<$ 0.02 cm
$\|d0\|$ (with no silicon hits)	$<$ 0.2 cm
$Iso4/p_T$	$<$ 0.1
$\Delta X(CMU)$ (if CMUP)	$<$ 7 cm
$\Delta X(CMP)$ (if CMUP)	$<$ 5 cm
$\Delta X(CMX)$ (if CMX)	$<$ 6 cm

Table 3.2: Central muons identification criteria.

and 92.75±0.47 respectively for CMUP and CMX [49]. Also for muons, fake rates are less than 1%.

3.7 Jets reconstruction

Quarks and gluons do not exist freely in nature due to the color confinement. They undergo a process called fragmentation where they create partons via a cascade of gluon emissions and decays. Partons then form colorless hadrons in a process called hadronization. The non-stable hadrons decay to stable particles which reach the detector material. The particles reaching the calorimeters produce showers that appear as clusters of energy deposited in localized areas of the calorimeter, called jets. There are several algorithms developed for calorimeter jets. Some algorithms may also incorporate tracking information in searching for charged jets or in measuring their transverse momenta. The jet identification algorithm used in these searches is called JETCLU [50] and is based only on calorimeter informations. The standard CDF jet clustering algorithm [51] is an iterative cone algorithm based on a fixed cone-radius (R) in $\eta - \phi$ space.

The jet-finding algorithm begins by searching seed towers among all the calorimeter towers with a transverse energy above 1 GeV. In the plug calorimeter region,

towers are grouped in ϕ into sets of three to reproduce the central segmentation. The list of seed towers, ordered in decreasing transverse energy, is used to create preclusters. A precluster is formed by merging together an unbroken chain of contiguous seed towers. If a tower is outside a window of 7×7 towers surrounding the starting seed, it is used to form a new precluster. The precluster centroid is calculated taking the E_T-weighted average of the tower positions. The clustering starts using the most energetic precluster, grouping all the towers with an E_T above 100 MeV which lie inside a cone of a radius R around the precluster seed centroid. Then the E_T-weighted centroid of the new cluster is calculated from the set of towers within the cluster cone, and a new cone is defined around this direction. Towers might be added or deleted from the new cone cluster during this procedure. The iterative process stops when the tower assignment to clusters remains unchanged. The algorithm also provides a prescription for treating overlapping clusters. If the towers of one cluster are completely contained within another, the less energetic cluster is dropped. If two clusters have some common towers, the shared energy is computed by summing the E_T of the overlapping towers. If this shared energy is above the 75% of the total E_T of the less energetic cluster, then the two clusters are combined. If the shared energy is smaller than this threshold, the clusters remain separate, and each tower in the overlap region is assigned to the closest cluster in $\eta - \phi$ space. After all towers are uniquely assigned to clusters, the centroids are recalculated. This process of centroid computation and tower reshuffling is iterated until the tower lists remain fixed. A cluster four-vector ($p_x,\ p_y,\ p_z,\ E$) is defined once a stable configuration is reached. A massless four-vector is assigned to each electromagnetic and hadronic tower having a magnitude equal to the energy deposit in the tower, and a direction defined by the unit vector pointing from the nominal detector origin to the center of the tower (calculated at the depth that corresponds to the shower maximum). Then all the tower four-vectors belonging to the same cluster are summed:

$$p_x = \Sigma_i p_x^i \qquad p_y = \Sigma_i p_y^i \qquad p_z = \Sigma_i p_z^i \qquad E = \Sigma_i E^i \tag{3.2}$$

Several jet variables are calculated subsequently from these quantities:

$$E_T = E \cdot \sin\theta \qquad \eta = -\ln\tan\left(\frac{\theta}{2}\right) \qquad \phi = \arctan\left(\frac{p_y}{p_x}\right) \tag{3.3}$$

where

$$\theta = \arcsin\left[\frac{\sqrt{p_x^2 + p_y^2}}{\sqrt{p_x^2 + p_y^2 + p_z^2}}\right]. \tag{3.4}$$

3.8 Jet corrections

The transverse energy and momentum in the above definition are based on the raw calorimeter energy. These uncorrected quantities differ from the true partonic four-momentum for a variety of reasons. Some effects originate from detector performance limitations:

- The calorimeter response to low-energy charged particles shows a non-linearity for momenta below 10 GeV.
- Charged particles with a transverse momentum below 400 MeV/c are confined inside the COT volume by the magnetic field and thus never reach the calorimeter. At slightly higher transverse momenta, the magnetic field can spread particles outside the jet cone.
- Particles showering in a poorly instrumented region of the calorimeter, like boundary regions between calorimeter modules or junctions between central, plug and forward subsystems, have on average a smaller energy response.

Others effects stem from algorithm limitations and physics processes:

- Energy coming from soft-scattering processes or extra $p\bar{p}$ interactions in the same bunch crossing (underlying event), might be collected into the jet cone.
- Soft energy radiation and fragmentation effects might cause energy leakage due to particle loss outside the jet cone.

- Neutrinos and muons contained in jets escape the calorimeter without depositing a substantial energy, resulting in a mismeasured jet energy.

The standard CDF jet correction procedure [88] takes care of all these effects. The raw detector energy is studied as a function of:

- jet cone radius (R)
- the raw jet transverse momentum (p_T)
- the jet detector pseudorapidity (η_d)
- the electromagnetic fraction (em)
- the number of reconstructed vertices (N_V)

The jet correction procedure consists of many steps. a flat response in η, while others account for absolute energy mismeasurements, the underlying event and out-of-cone effects. If one want to get the full "parent parton" transverse energy, it is necessary to correct for all these effects. For completeness, we describe the entire procedure, described by the following equation:

$$P_T(R) = (P_T^{raw}(R) \times f_{rel} - UEM(R)) \times f_{abs}(R) - UE(R) + OC(R) \qquad (3.5)$$

where R is the cone radius. Each correction term appearing in the above equation is detailed in the following.

f_{rel} : Relative jet energy corrections

The first step in jet correction is to adjust the non uniform energy response of the calorimeter making it independent on η (see Fig. 3.1). The reason for this nonuniformity is the presence of uninstrumented regions (cracks) and differing tower segmentation.

The central calorimeters CEM/CHA are the best understood calorimeters in CDF II and they cover the regions far away from cracks. Their response, measured

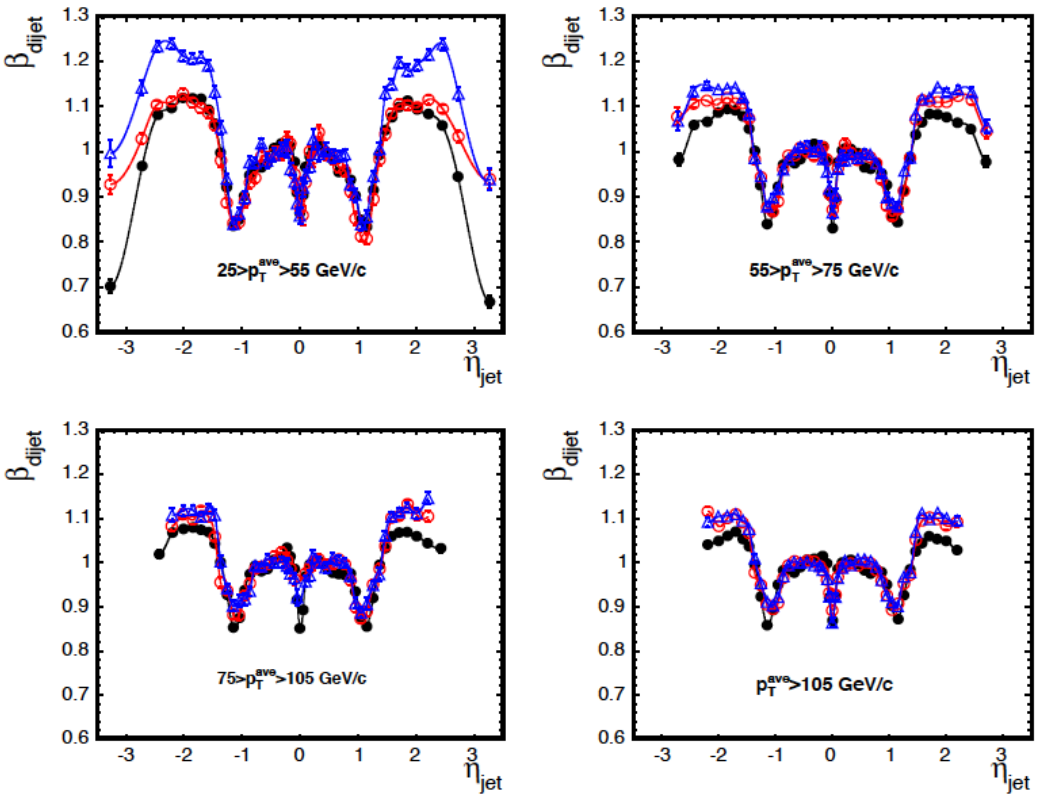


Figure 3.1: Relative calorimeter response of simulated data compared to di-jet data. Jets are reconstructed with a cone radius of 0.4.

with test beam data and checked with the COT information during the data taking, is found to be linear.

The correction procedure is based on di-jet p_T balancing. Assuming the di-jet production to be a $2 \rightarrow 2$ process, the transverse energy of the two jets has to be equal. Confining one jet to lie in the central region, $0.2 < |\eta| < 0.6$ (*trigger jet*), and the other spanning the whole calorimeter (*probe jet*), a scale factor can be derived as a function of η and p_T. This correction is parametrized for three different cone sizes: 0.4, 0.7 and 1.0.

The energy response also changes with time, mostly due to the ageing of the phototubes. Data are calibrated to take into account this variation.

f_{abs} : Absolute jet energy corrections

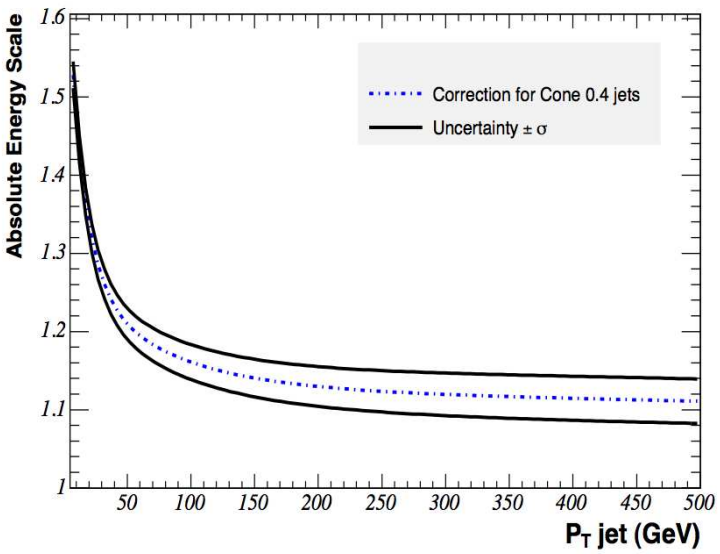


Figure 3.2: Absolute jet energy scale correction factor as a function of the jet P_T for a cone size $R = 0.4$., together with its uncertainty.

The conversion factor from calorimeter energy to particle-level energy is called absolute energy correction. This factor depends on the jet fragmentation properties and on the non-linearity of the calorimeter response due to the difference in the response to π^0 and $\pi^{\pm}$.

The absolute energy corrections is determined using jets generated with tuned Monte Carlo. In this context the reconstructed jet energy is compared to the sum of the four momenta of all the MC particles lying inside the jet cone. In Figure 3.2 the correction function is plotted.

UEM : Multiple interaction corrections

With the current instantaneous luminosity and with 36 proton and anti-proton bunches, several interactions per bunch crossing are expected. Moreover, since the number of interactions per bunch crossing follows a Poisson distribution, some events have more overlapping minimum bias interactions. The energy released in these collisions could fall inside the clustering cone, and should be subtracted from the jet energy. This correction is a linear function of the number of primary vertices in the event, and its value is derived from minimum bias events measuring the transverse energy in a cone of the appropriate radius opened in a random direction.

UE : Underlying event corrections

The underlying event is the result of soft collisions between the spectator partons inside the proton and antiproton. These interactions produce background energy in the whole calorimeter which contributes to the jet energy although it is not associated to the hard scattering. This energy is estimated using single vertex minimum bias events, measuring the transverse energy density (ΣE_T divided by the whole calorimeter surface) in the central calorimeter. The jets are then corrected according to their cone radius by subtracting a definite amount of energy proportional to the energy density mentioned before.

OC : Out-of-Cone corrections

This factor is determined using Monte Carlo events. This correction is independent of the calorimeter performance while it is a function of the jet transverse momentum and the cone radius. The additional energy is parametrized as:

$$p_T^{ooc} = A[conesize] \cdot (1.0 - B[conesize] \cdot e^{-C[conesize] \cdot p_T}) \tag{3.6}$$

	A	B	C
Cone radius 0.4	22.999	0.915	0.00740

Table 3.3: Out-of-cone energy parameters.

The values of the parameters describing the out-of-cone correction are listed in Table 3.3.

In this analysis we don't need to scale the jet energy to the parent parton value, but only to hadron level, so that the out-of-cone correction is not necessary.

3.8.1 Jet corrections systematics

Each correction described above has associated systematic uncertainties [88] which are shown in Fig. 3.3 as a function of the jet E_T. l

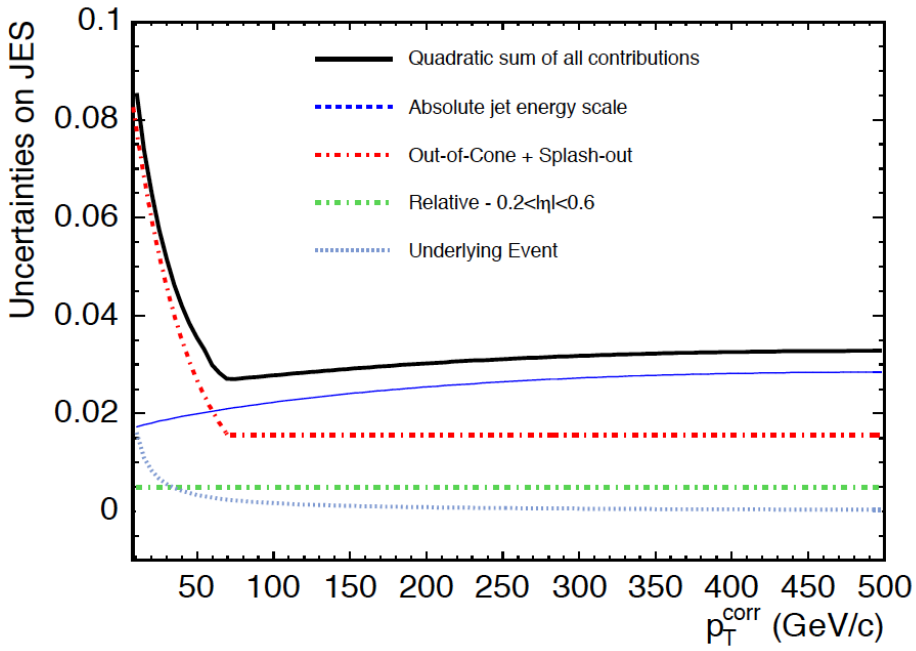


Figure 3.3: Systematic uncertainty on the jet energy scale as a function of the corrected jet E_T.

It has been observed that the Monte Carlo simulation of the calorimeter does not exactly reproduce the CDF II calorimeter energy response (see Fig. 3.1) as a

function of η. Therefore an extra contribution to the relative correction factor has been introduced to account for the difference between data and Monte Carlo.

3.9 Missing transverse energy

Protons and anti-protons that collide at the Tevatron have equal and opposite momenta. Therefore, the total vector momentum sum in an event should be zero. The hard collision happens between the partons of the proton and antiproton, and they can carry any fraction of the parent proton or antiproton. However, since the partons usually have very little momentum in the plane transverse to the beam, the transverse energy can be considered as a conserved quantity to a good approximation. Any transverse energy imbalance in the detector may indicate that a particle left the detector without interacting with its material.

In this analysis we look for a signature containing pairs of massive particles that are supposed not to leave any signature in the detector, thus giving rise to large missing transverse energy ($\vec{\not\!\!E}_T$, or $\not\!\!E_T$ if we refer to its magnitude). Noteworthy, the final fitting procedure will be carried out on a $\not\!\!E_T$-based distribution. Therefore, this quantity is one of the most important for this analysis.

The x and y components of the raw missing transverse energy of the event are obtained from:

$$\not\!\!E_x = -\sum_{i=1}^{N_{towers}} E_T^i \cos\phi_i \tag{3.7}$$

$$\not\!\!E_y = -\sum_{i=1}^{N_{towers}} E_T^i \sin\phi_i \tag{3.8}$$

where the sum is taken over all towers that are above a threshold of 0.1 GeV, and the total electromagnetic and hadronic energy in i^{th} tower is E_T^i. The magnitude of the missing energy is then calculated by:

$$\not\!\!E_T = \sqrt{\not\!\!E_x{}^2 + \not\!\!E_y{}^2} \tag{3.9}$$

The azimuthal direction of the $\not\!\!E_T$ is then given by:

$$\phi_{\not\!\!E_T} = \tan^{-1}\left(\frac{\not\!\!E_y}{\not\!\!E_x}\right) \tag{3.10}$$

While a large $\not\!\!E_T$ is recorded in events that contain undetected particles, other types of processes may also lead to the experimental signature of $\not\!\!E_T$:

- problems with some of the calorimeter tower electronics/calibrations may cause the tower to report a wrong value of the energy of the incident particles.
- when the direction of an energetic jet is near an uninstrumented calorimeter region, the energy of the jet will be underestimated. A configuration of two jets that are produced back-to-back with the same momentum will appear to have a momentum imbalance.
- when protons or antiprotons of the Tevatron beam collide with nuclei of gas atoms or beam collimators, they produce a "halo" of muons, travelling roughly parallel to the beam. Some of these muons cross a row of calorimeter towers along the z-axis depositing energy to the calorimeters asymmetrically in ϕ.
- muons carrying a large momentum can be created in the hard collision. Being minimum ionizing particles, these muons can pass through the calorimeter, without substantial energy loss.
- cosmic muons traveling through the detector.
- proton or antiproton beam remnants and beam losses in very forward regions.

3.9.1 $\vec{\not\!\!E}_T$ corrections

The $\not\!\!E_T$ measured by the CDF calorimeter ("raw" $\not\!\!E_T$) needs to be corrected for the same reasons that the jet energies do, as described in section 3.8. Hence, the $\not\!\!E_T$ needs to be recomputed using the corrected values of the jet energies. The event $\not\!\!E_T$ is corrected using the corrected jet E_T^{Corr} values with the following formula:

$$\not\!\!E_x{}^{Corr} = \not\!\!E_x{}^{Raw} - \sum_{i=1}^{N_{jets}} E_x^{Corr,i} - E_x^{Raw,i} \tag{3.11}$$

$$\not\!\!E_y^{\,Corr} = \not\!\!E_y^{\,Raw} - \sum_{i=1}^{N_{jets}} E_y^{Corr,i} - E_y^{Raw,i} \tag{3.12}$$

where $\not\!\!E_x^{\,Raw}$ and $\not\!\!E_y^{\,Raw}$ are calculated with equations 3.7 3.8. The azimuthal direction of the corrected $\not\!\!E_T$ is recomputed in this way:

$$\phi_{\not\!\!E_T{}^{Corr}} = \tan^{-1}\left(\frac{\not\!\!E_y^{\,Corr}}{\not\!\!E_x^{\,Corr}}\right) \tag{3.13}$$

The $\not\!\!E_T$ energy used everywhere in this analysis is the corrected $\not\!\!E_T$

3.10 Missing transverse momentum

Similarly to missing transverse energy $\vec{\not\!\!E}_T$, it is possible to define a missing transverse momentum $\vec{\not\!\!p}_T$ using the spectrometer, as the negative vector sum of the charged particles momenta.

$$\vec{\not\!\!p}_T = -\sum_{tracks} \vec{p}_T \tag{3.14}$$

In events where only charged particles and undetected particles are produced, the $\vec{\not\!\!p}_T$ is highly correlated in module and direction to the undetected particle(s) momentum, and thus provide a way to measure their energy which potentially has better resolution than $\vec{\not\!\!E}_T$. In the search described here, the presence of quarks in the final state complicates the picture.

In the radiation and hadronization process forming a jet, most particles produced are pions (with a 10-20% of kaons). Due to isospin symmetry, roughly 2/3 of the energy of a jet will be carried by charged pions, which will be measured with both calorimeter and tracking chamber. The $\not\!\!p_T$ underestimates the undetected particle's energy because it does not take into account the energy carried by the neutral components of the jets, and for the same reason has a worse angular resolution. Also, the fact that in the CDF detector the tracking efficiency drops quickly to zero when tracks η become larger than 1 has to be taken into account. For these reasons, $\vec{\not\!\!p}_T$ cannot substitute $\vec{\not\!\!E}_T$ as a measurement of energy or direction of undetected

particles; still, the $\vec{\not{p}}_T$ can provide informations complementary to those obtained $\vec{\not{E}}_T$. As we will see in Chapter 5, angular correlations between $\vec{\not{E}}_T$ and $\vec{\not{p}}_T$ will be very useful in the implementation of a QCD data-driven model.

For $\vec{\not{p}}_T$ calculation we use the same track quality criteria required for the tracks used in the calculation of the position of the primary vertex. The selections are listed in Tab. 3.4 and only tracks satisfying these conditions are used when calculating $\vec{\not{p}}_T$. The quantities used are:

- N_{COT}: The number of hits in the COT detector
- N_{SVX}: The number of hits in the SVX detector
- N_{COT}^{Ax} ($\geq$ 5 hits) and N_{COT}^{St} ($\geq$ 5 hits): the number of axial and stereo COT layers that have at least 5 hits
- N_{SVX}^{Ax} and N_{SVX}^{St}: the number of hits in axial and stereo SVX layers
- χ^2: the χ^2 of the track fit
- χ^2_{SVX}: the χ^2 of the track fit, using only informations from the SVX

Also, only tracks with 0.5 GeV $< p_T <$ 200 GeV, $|\eta| < 1.5$ and $|Z_{vtx}| < 2$ cm are used.

Table 3.4: Quality requirements for tracks used in $\not\!\!\vec{p}_T$ calculation. The categories are listed in the order the selections are checked. If the first categoriy requirements fail, we next check the second category, etc.

Category	***Requirements***
Strong COT	$N_{COT} > 0$ $(\chi^2 - \chi^2_{SVX})/(N_{COT} - 5) < 4.0$ $N^{St}_{COT}(\geq 5 \text{ hits}) \geq 3;\ N^{Ax}_{COT}(\geq 5 \text{ hits}) \geq 3$
Weak COT, good χ^2	$N_{COT} > 0$ $(\chi^2 - \chi^2_{SVX})/(N_{COT} - 5) < 4.0$ $N^{St}_{COT}(\geq 5 \text{ hits}) \geq 2;\ N^{Ax}_{COT}(\geq 5 \text{ hits}) \geq 2$ $N^{Ax}_{SVX} \geq 4; N^{St}_{SVX} \geq 3$ $\chi^2_{SVX}/(N_{SVX} - 5) < 8.0$
Very weak COT, good χ^2	$N_{COT} > 0$ $(\chi^2 - \chi^2_{SVX})/(N_{COT} - 5) < 4.0$ $N^{Ax}_{SVX} \geq 5; N^{St}_{SVX} \geq 3$ $\chi^2_{SVX}/(N_{SVX} - 5) < 8.0$
Inside-Out or Silicon stand-alone	$N_{COT} > 0$ $N^{Ax}_{SVX} \geq 5; N^{St}_{SVX} \geq 3$ $\chi^2_{SVX}/(N_{SVX} - 5) < 8.0$

3.11 Trigger selection

For our analysis, we used the MET_DIJET trigger [57, 58, 59], which was originally conceived for the Higgs boson search. Here we describe how objects are reconstructed at the various trigger levels and the requirements we make to select multijet events suitable to be good candidates for our analysis. In particular, the jet reconstruction techniques are described here, since they are different for different trigger levels.

Level 1

The Level 1 trigger system is based on custom electronics designed for fast decision making, and allows the reconstruction of basic physics object that can be triggered on.

The calorimeter trigger towers have a width $\Delta\eta = 0.2$ and $\Delta\phi = 15$, and the entire calorimeter can thus be represented by a 24×24 trigger tower map. The trigger tower energy information is sent to both the Level 1 and Level 2 trigger with a 10-bit energy resolution, with a least significant count of 125 MeV and a resulting full scale of 128 GeV. To reduce the complexity and the processing time, the Level 1 trigger uses only a 8-bit trigger tower energy information, by dropping the least significant bit and the most significant bit. Level 1 does not perform any clustering, and it simply select events on the number of trigger towers above programmable transverse energy thresholds or on the values of the computed total E_T and $\not\!\!E_T$ The transverse projection of the tower energies are calculated with the assumption that the event primary vertex is located at $z = 0$, and the missing energy at L1 is calculated as a vector sum of trigger tower pairs. The missing energy at L1 has poor resolution, due to a limited available information and the need to make a fast decision, and is usually underestimated. Therefore, the $\not\!\!E_T$ threshold at L1 is chosen to be as low as possible in order to maximize efficiency.

At Level 1, the MET DIJET trigger requires at least one trigger tower with $E_T \geq 10$ GeV, and a $\not\!\!E_T \geq 28$ GeV.

Table 3.5: Summary of the MET DIJET trigger requirements.

Trigger Level	***Requirements***
Level 1	≥ 1 trigger tower with $E_T \geq 10$ GeV
	$\not\!\!E_T \geq 28$ GeV
Level 2	≥ 2 clusters with $E_T \geq 3$ GeV
Level 3	$\not\!\!E_T \geq 30$ GeV

Level 2

At level 2, MET_DIJET trigger requires at least 2 clusters with transverse energy above 3 GeV and $|\eta| \leq 3.6$. The $\not\!\!E_T$ is recomputed with the additional available informations. Again,$\not\!\!E_T \geq 28$ GeV is required.

Level 3

The information provided by the cluster finding algorithm at the trigger level can be considered as a first-order jet reconstruction. At Level 3, a looser time constraint enables to exploit the full detector segmentation for a better jet energy and direction determination.

At L3, our trigger requires the newly-computed $\not\!\!E_T$ to be greater than 30 GeV. A summary of the trigger level requirements can be found in Table 3.5. The efficiency of this trigger is greater than 99% for all our signal hypothesis, after requiring the kinematical cuts described in chpt. 6 [59]. Systematical errors due to this acceptance are negligible compared to other systematics (see chpt. 7).

Chapter 4

Experimental Overview

4.1 Previous searches in the semileptonic channel

Traditionally, the "golden channel" for measurements and searches involving top quarks is the semileptonic one. The BR is lower than in the hadronic channel, but the lepton in the final state allows to easily trigger, substantially reducing the QCD background. The first search for a fourth generation T' quark decaying in top quark and invisible particles, performed at CDF [52] and, more recently, at ATLAS [53] has been made in this channel. In this analysis, the main backgrounds are top pair production and W plus jets events, which both give rise to real missing energy through a leptonic W decay. In the CDF analysis, based on a 4.8 fb$^-$1 data sample, the separation between backgrounds and signal is performed using the reconstructed transverse mass of the leptonically decaying W candidate:

$$m_T^W = m_T(\vec{p}_T^{\,l}, \vec{\not{E}}_T) = \sqrt{2|\vec{p}_T^{\,l}|\vec{\not{E}}_T|(1 - \cos\Delta\phi(\vec{p}_T^{\,l}, \vec{\not{E}}_T))} \tag{4.1}$$

where $\vec{p}_T^{\,l}$ is the lepton transverse momentum and $\vec{\not{E}}_T$ is the missing transverse energy vector. In $t\bar{t}$ and W plus jets events the $\not{E}_T$ originates primarly from the neutrino in $W \to l\nu$ decay, so that m_T^W will show a strong jacobian peak at the W boson mass. On the other hand, in signal $T'T' \to tt + XX$ events there is additional $\not{E}_T$ coming from the undetected X particles. As a result of this, the m_T^W

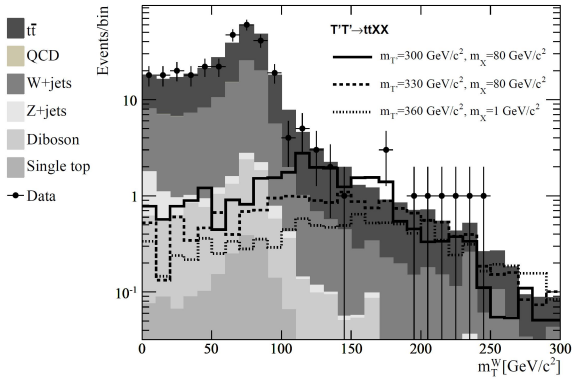


Figure 4.1: Reconstructed W mass for data, backgrounds and three different signals. From [52]

variable does not reconstruct the correct W boson mass in signal events, and shows a substantally higher tails at high values (see Fig. 4.1).

The possible presence of signal is thus investigated performing a binned maximum-likelihood fit in the m_T^W distribution, using the expected SM and signal templates. In this manner, upper limits on the signal cross sections are derived, and then converted in an exclusion plot in the $(m_{T'}, m_X)$ plane. As can be seen in Fig. 4.2, the results exclude at 95% confidence level the existence of T' in this model up to $m_{T'} = 360 GeV/c^2$, for $m_{T'} \leq 100 GeV/c^2$.

The ATLAS analysis, which has been performed using a sample of 1.34 fb$^-$1, was completed a few weeks after this work, and basically consisted of a counting experiment. Its results will be described and compared with our result in the Conclusions.

4.2 The advantages of the hadronic channel

The large QCD multijet production cross sections results in a very large background for signatures in which only jets are expected. Nevertheless, the model we want to investigate predict also a large $\not{E}_T$ in the final state. This feature allows us to investigate our signal hypothesis despite of this apparently overwhelming background. Appropriate kinematic cuts, mainly based on $\not{E}_T$-correlated variables, can reduce the

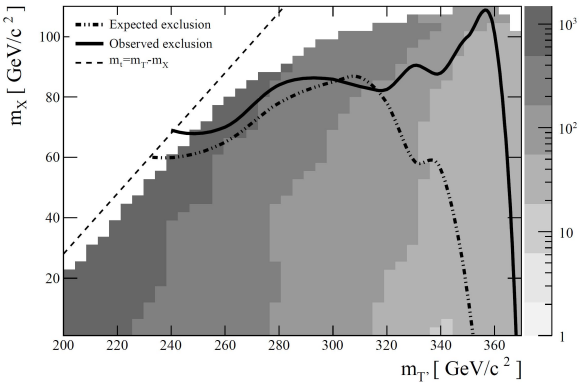


Figure 4.2: Excluded region at 95% C.L. (line) and cross section limits (shaded areas) in the semileptonic channel. From [52].

QCD multijet background to a level comparable to the signal. Also, the different characteristics betweeen signal and QCD multijet events with respect to $\not\!\!E_T$-related variables allow us to define a data-driven QCD background model. The $\simeq 50\%$ larger branching ratio results in better expected sensitivity for the hadronic channel respect to the semileptonic, as we have demonstrated with this work. In a recent study [27], pre-existing to both analysis, simulations of a search for $T'T' \to ttXX$ are made. As can be seen in 4.3, the expected hadronic sensitivity is higher. It's worth to point out that the methods used in [27] are different either from those used in our analysis and from those of the semileptonic analysis, since they simply consist in strict kinematic cuts to reduce the background to negligible levels, followed by a counting experiment. Nevertheless, as can be seen comparing Fig. 4.2 and Fig. 4.3, the results for the semileptonic channel turned out to be similar to those of the published semileptonic CDF analysis.

4.3 Signal characterization

The process we want to investigate consists of the following subprocesses:

- The production of a fourth generation T' pair via quark-antiquark annihilation or gluon-gluon fusion. As in the case of top pair production, the quark-antiquark annihilation is expected to give the dominant contribution because

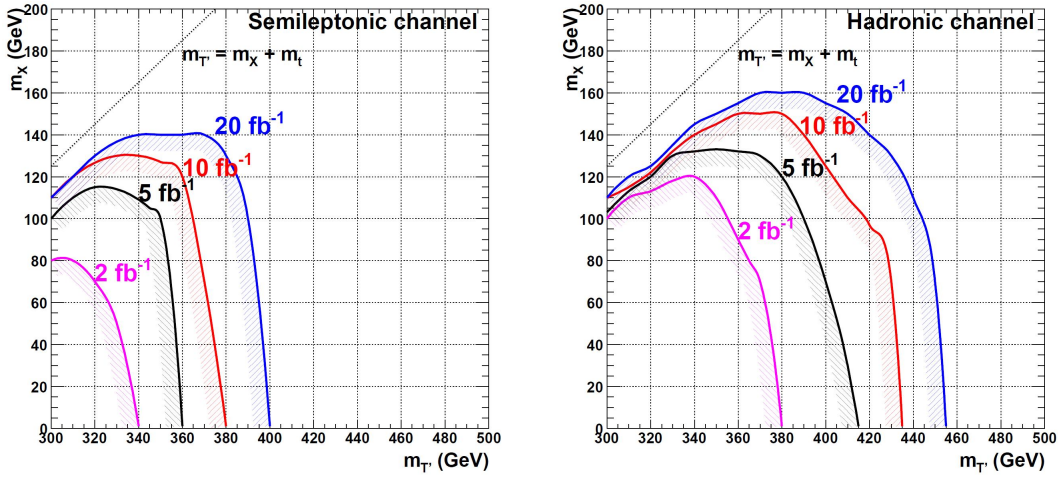


Figure 4.3: Expected excluded region at 95% C.L. for the semileptonic and hadronic channel at the Tevatron. From [27].

of the higher average momentum carried by the valence quarks/antiquarks. The production cross sections at the Tevatron for different T' masses have been calculated from Matteo Cacciari [31], using the same framework of the study described in [54] with CT10 PDF [55].

- The decay, before any type of hadronization, of each T' quarks in a top quark plus a DM particle. We assume a 100% branching fraction of T' quarks to these products.

- The SM decay of the two top quarks in a W boson plus a b quark.

- The SM decay of both W bosons in a quark-antiquark pair.

A feynman diagram of the process is shown in Fig. 4.4.

From the final product of the overall process we expect a nominal signature of six hadronic jets from the six quarks and large missing transverse energy from the two DM particles. The detailed distribution of measured variables we expect depends largely on the values of $m_{T'}$ and m_X.

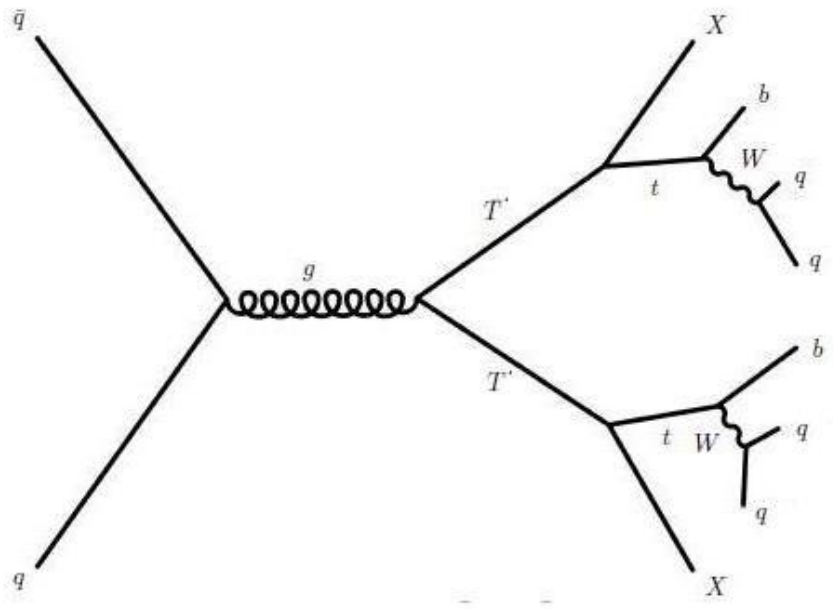


Figure 4.4: Feynman diagram of the process searched for in this analysis.

4.4 Main expected backgrounds

The above-mantioned signature can be mimicked by three different categories of processes:

4.4.1 QCD multijet production

The generic production of many jets by quantum chromodynamics interactions between quarks and gluons is the most probable source of hadronic jets in the particle colliders. Because of this, it is also the greatest source of backgrounds for searches which do not require leptons in the final state. Although the cross section decrease as $(\alpha_s)^n$ with the increasing jet multiplicity, energetic events with six or more jets with $E_T > 30GeV$ have still cross sections of 30-40 pb [56], i.e. more tan 10 times larger than our signal largest expected cross section. Ideally, this would not be a problem, because our signature consist also of large missing transverse energy. Since QCD multijet processes don't include electroweak production of neutrinos (apart from small contribution of heavy quarks' semileptonic decays), missing transverse

energy in this type of events, if exactly measured, is expected to be zero. Nevertheless, statistical fluctuations and non linearity in the calorimeter response have to be taken into account. Also, the detector is not completely hermetic: there are non-instrumented regions ("cracks") like the separation zone between central and forward calorimeter, or the "chimney", a hole in the calorimeter that host cryogenic and instrumental connections to the inner detector; if a particle travel along a boundary between two towers its energy is not precisely measured as well. Because of this, QCD multijet events have a smeared non-zero $\not\!\!E_T$ distribution. The probability to have large $\not\!\!E_T$ due to detector effects is low, as we will see in Chpt. 7, but due to the high cross section a significant amount of such background events events is present.

4.4.2 Top pair production

The top pair production process has a cross section of about 7 pb, which is much less than QCD multijet but still larger than our signal, expecially in the high $m_{T'}$ hypothesis. Paradoxically, for our measure the contribution of the hadronic channel is much less problematic than the semileptonic. In fact, $\not\!\!E_T$ in hadronic $t\bar{t}$ events is due to detector effects, like for QCD multijet, and so tends to be small. On the other hand, a semileptonic events with $\not\!\!E_T$ coming from the neutrino and 1-2 additional jets from initial and/or final state radiation (see Fig. 4.5) can give rise to 5-6 jets plus large $\not\!\!E_T$ events. $t\bar{t}$ background will be the most limiting for our analysis.

4.4.3 Electroweak processes

We separately considerate contributions from:

- W + jets. The main background from this category come from leptonic W decays (see for example Fig. 4.5)
- Z + jets. Events with 5-6 jets and $Z \to \nu\nu$ cannot be reduced with lepton veto.
- WW, WZ and ZZ ("diboson" production) plus jets.

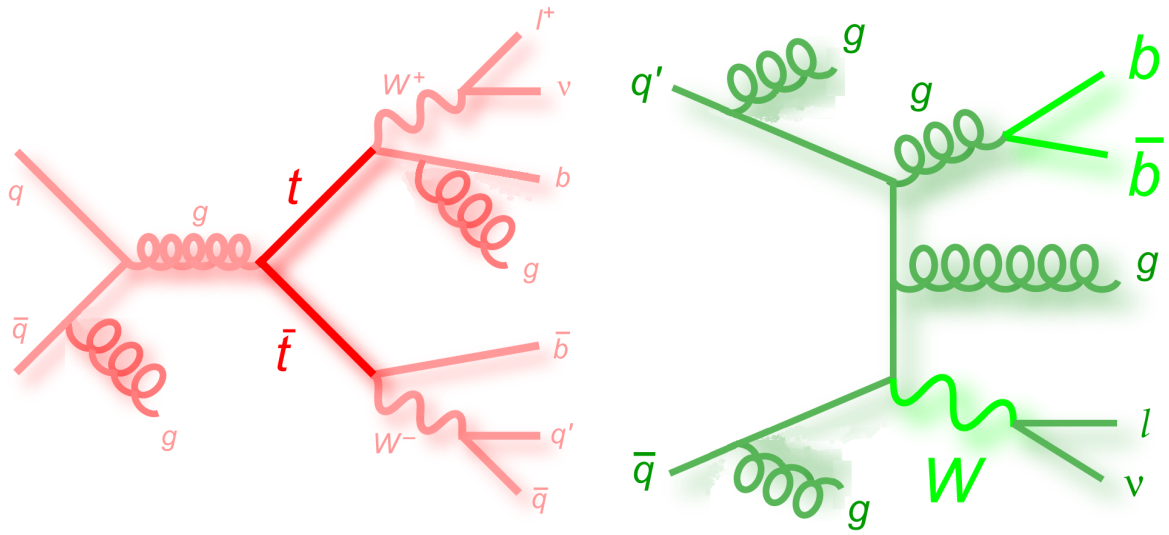


Figure 4.5: Feynman diagram of two background processes for this analysis: (left) top pair production with additional jets from initial and final state radiation and semileptonic decay mode, (right) W plus jets production.

4.5 Analysis strategy

Our analysis procedure basically consist of three steps:

- **Signal and background modeling**

  The first step consists in evaluating the expected distributions in the variables of interest for the signal and for all the SM backgrounds. In this kind of analysis this is usually done using Monte Carlo (MC) programs, *i.e.* complex numerical simulations built up from the theoretical knowledge of the process under examination. In this analysis we use MC to model all the processes except QCD multijet, for which simulation approach is not feasible, requiring a data-driven method. All these procedures will be described in detail in the next chapter.

- **Event selection**

  Once we have models for signal and backgrounds is possible to define a procedure to optimize the signal over background ratio. More precisely, the figure of merit describing the sensitivity in a counting experiment is $S/\sqrt{S+B}$, where S and B are respectively the expected number of signal and background events

in the current configuration. This quantity approximate the significance of observing signal S against the poissonian fluctuation $\sqrt{S+B}$ in the signal plus background hypothesis. In order to maximize this quantity, a series of kinematical cuts are made. These cuts are kept general enough to be useful in all $(m_{T'}, m_X)$ hypothesis, and a reference point for future searches of this kind of signature. The detailed event selection is described in chpt. 6.

- **Limits extraction**

  To increase sensitivity above the mere counting experiment threshold, we choose the distribution in $\not\!\!E_T$ significance, which is defined as the $\not\!\!E_T$ divided by the square root of the sum of the energy collected in all the calorimetric towers, to perform a binned maximum-likelihood fit. The detailed procedure is described in chpt. 7.

Chapter 5

Signal and Background Modeling

To model our signal events, and the possible SM sources of backgrounds, we use different methods: $t\bar{t}$ and electroweak processes, as well as signal events, are modeled with different Monte Carlo programs; QCD multijet production is modeled with a novel data-driven technique.

5.1 Monte Carlo tools

We model the production and decay of T' pairs with MADGRAPH [60]. Additional radiation, hadronization and showering are described by PYTHIA [61]. The dominant SM background is the $t\bar{t}$ production. We model this background using PYTHIA with top mass equal to 172.5 GeV/c^2, which is consistent with the current world best estimate of this parameter [16], normalizing it to the $t\bar{t}$ NLO cross section [62]. The second dominant SM background process is the associated production of a W or Z boson and jets. Samples of simulated W/Z+jets events with light- and heavy-flavor jets are generated using the ALPGEN [63] program, interfaced with parton-shower model from PYTHIA. A matching scheme is applied to avoid double counting of partonic event configurations [64]. The samples are normalized to the inclusive cross sections [65], scaled by 1.4 to account to next-to-leading-order (NLO) corrections. Due to the limited reliability of PYTHIA to simulate very high multiplicity events, a 40% uncertainty [66] is assigned for the extracted yields of events

with a W or Z boson and jets. We also model contribution from $WW/WZ/ZZ$ (diboson) [67] with PYTHIA. Single top production [68] is simulated with MADGRAPH interfaced with PYTHIA. The detector response for all simulated samples is modeled by the official CDF detector simulation [69].

5.2 QCD data-driven modeling

The most significant background at the first stage of the analysis is the QCD multijet production. Although these processes generally do not produce neutrinos, mismeasured jet energies do result in imbalance in the measured transverse energy. As a result, QCD events can pass the basic requirements on $\not\!\!E_T$. Because of the high production rate for QCD at a hadron collider and the large statistics needed in order to describe this process adequately in an analysis looking for a very small signal, the Monte Carlo simulation of an acceptable amount of QCD events is prohibitive. More importantly, the systematic uncertainties associated with the Monte Carlo simulation of QCD jet production are high. For these reasons, we estimate the QCD background solely from data.

5.2.1 $\vec{\not\!\!E}_T$-$\vec{\not\!\!p}_T$ angular correlations

For events with real undetected particles, $\vec{\not\!\!E}_T$ is a good measurement of the vectorial sum of those particles four-momenta , and so is the $\vec{\not\!\!p}_T$ which thus will tend to be parallel to the $\vec{\not\!\!E}_T$. However, in QCD events the nature of these two variables is very different.

First, we consider here events with only 2 jets. For QCD events, the energy conservation requires that the vector sum of the jet transverse energies amounts to zero. According to the energy conservation law, the 2 jets will have the same magnitude of transverse energy, and will come out back-to-back in azimuthal space. The mismeasurement of jet energies makes the $\vec{\not\!\!E}_T$ align to the jet with less measured energy. $\vec{\not\!\!p}_T$ could be present in some amount too. However, it will be generated for very different reasons. In the jet fragmentation and hadronization processes a certain number of charged particles are produced inside each jet, and they will be detected

by the tracker. The intrinsic fluctuations inherent to the parton shower process will result in large fluctuations in the fraction of energy carried by charged particles inside a jet. The fraction of charged particles for a jet is completely independent from the measured jet energy in the calorimeter. In QCD dijet events the $\vec{\not\!\!E}_T$ comes from mismeasurement of one of the jet energy, so it will always be aligned to the under-measured jet. The $\vec{\not\!\!p}_T$ comes from random fluctuations in the charged fraction of the jet energies, so it will mainly point to the direction of jet with less energy being carried by reconstructed charged particles. As a net result, the $\vec{\not\!\!E}_T$ and $\vec{\not\!\!p}_T$ direction will be mainly correlated or anticorrelated (see Fig. 5.1. This is reflected in the distribution of the azimuthal distance between the two vectors, $\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T)$. As can be seen in Fig. 5.2, events containing undetected particles like neutrinos are concentrated near 0, while QCD multijet events have almost equal chance to populate the region around 0 or around π. This peculiar distribution allows to effectively suppress the QCD contribution in sample with $\not\!\!E_T$ and jets, as testified by the several publications that take advantage of it [70, 71, 72, 73, 74].

5.2.2 A $\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T)$-based data-driven QCD model

The characterization of $\vec{\not\!\!p}_T$ has been performed [75] in detailed studies of the $\not\!\!E_T$+2/3 jets sample comparing CDF data to a Monte Carlo simulation of QCD events using the PYTHIA program. In that study are investigated the general $\not\!\!p_T$ properties and the causes of the peculiar shape of $\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T)$ distribution reported above. Noteworthy, a similar shape is obtained in events with 3 hadronic jets, as can be seen in Fig. 5.2. The aim of the study was in particular to check the absence of correlation between $\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T)$ and the other kinematic characteristics of a QCD event, which is expected on the basis of the consideration made above. If this is the case, events in the $\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T) > \pi/2$ could be used to model QCD events in the complementary region, as has been done in [73, 74] in dijet signature. The comparison of many kinematic variables distribution in the two regions for MC QCD 2- and 3-jets events confirmed the uncorrelation assumption, confirming the validity of a $\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T)$-based QCD model even in 3-jets events. Our signature is much more complicated, including 5 or more hadronic jets. However, energetic QCD multijet

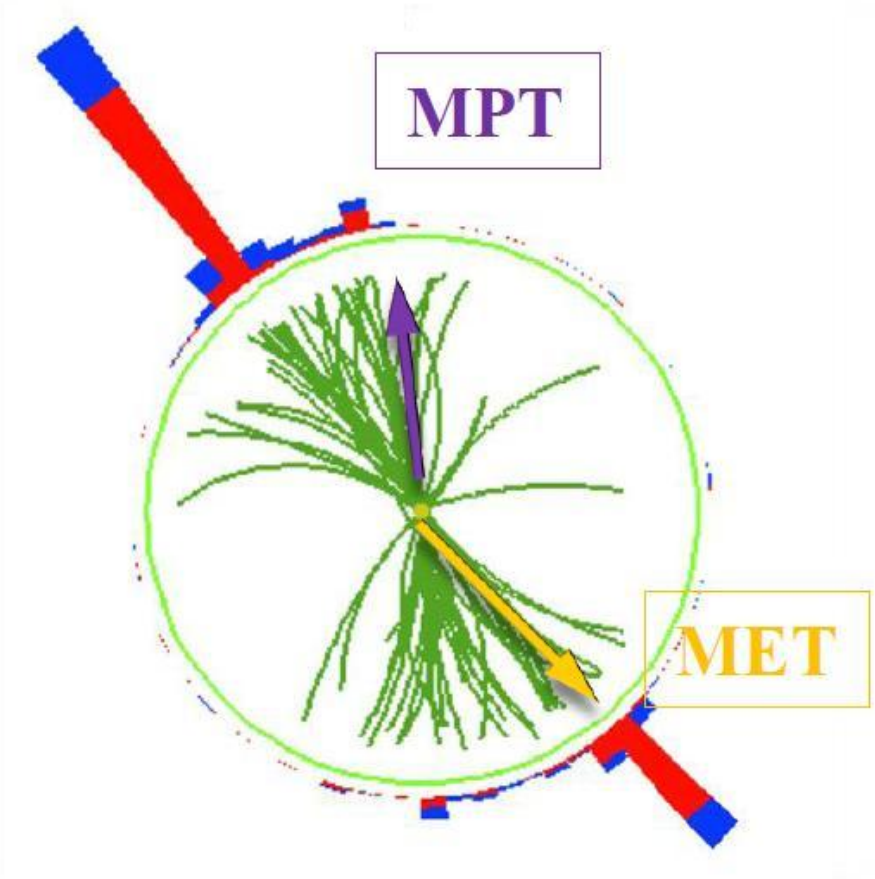


Figure 5.1: Schematical representation of $\vec{\not{E}}_T$ (MET) and $\vec{\not{p}}_T$ (MPT) in a QCD multijet event, where $\vec{\not{E}}_T$ is due to calorimetric mesmeasurement and $\vec{\not{p}}_T$ is due to statistical fluctuation in the number of charged particles in a jet.

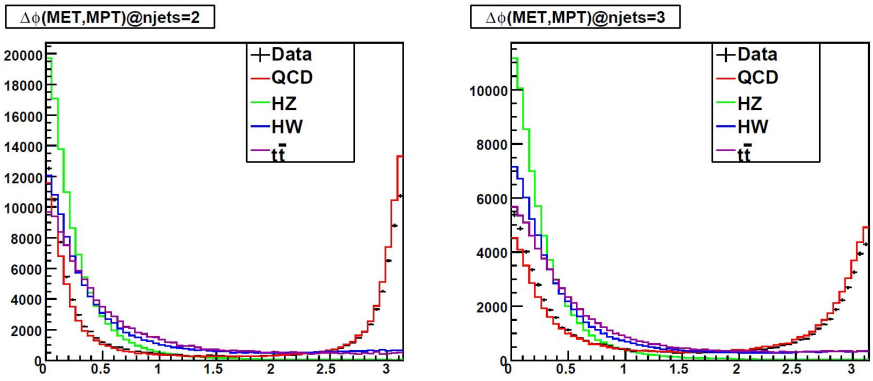


Figure 5.2: The $\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T$ distribution in events with 2 (left) or 3 (right) jets. Lines represents QCD, Higgs+W/Z and $t\bar{t}$ MC, while dots represents data collected with the MET_DIJET trigger. Plots are normalized to data. From [75].

events are often dijet-like, i.e. large part of the energy is carried by two highly energetic, nearly back-to-back jets. For this reason, all the above considerations are expected to hold. This is confirmed by the $\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T)$ distribution shape for preselected events, namely events containing at least five jets, no leptons and high $\not\!\!E_T$. As we will see in the next chapter, this sample consist of QCD multijet for more than 95%. As can be seen in Fig. 5.3, the data shows the expected saddle-like distribution, while events producing undetectable particles, in this case the $T'T' -> tt + XX$ signal, the $\vec{\not\!\!E}_T$ and $\vec{\not\!\!p}_T$ are aligned to each other. All non-QCD backgrounds of this analysis, apart from a small contribution of hadronic $t\bar{t}$ (see chapter 5), produce neutrinos. Because of this, even after QCD-rejecting selection cuts, we expect the $\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T) > \pi/2$ region to be an almost pure QCD sample. So, we reject these events, and use them to model the QCD in the signal region. Since the distribution is not exactly symmetric, we need to apply a normalization factor to the QCD background sample. We get this normalization from three signal-depleted control regions, described in details in the next chapter. Two of these control regions have enough statistic to check also the correct modeling of kinematic variables distribution by QCD region events.

This is the first analysis in which a $\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T)$-based QCD modeling method is tested for events with many (> 3) jets in the final state.

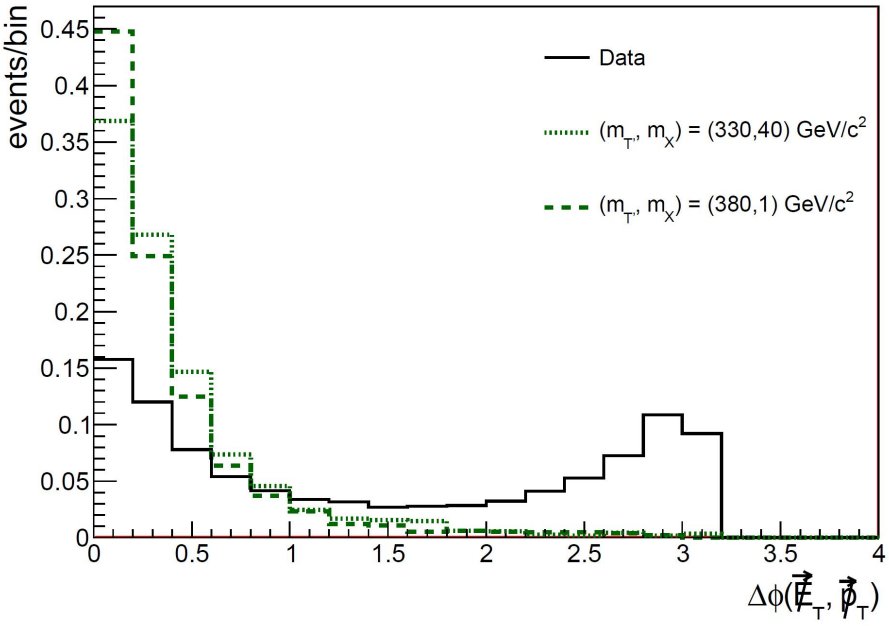


Figure 5.3: The $\Delta\phi(\vec{\not{E}}_T, \vec{\not{p}}_T)$ distribution in two signal hypothesis, and in collisions requiring five or more jets, large $\not{E}_T$, and no leptons in the final state. All histograms are normalized to the unit area. A QCD model is obtained from data with $\Delta\phi(\vec{\not{E}}_T, \vec{\not{p}}_T) > \pi/2$.

Chapter 6

Event Selection

Using signal and background models discussed in the previous Chapter, we develop a background rejection procedure taking advantage of the expected kinematic variables distributions. This is done in two different steps. First, we choose a basic selection in addition to the trigger requirements. Then, comparing signal and background distribution shapes, we optimize a set of kinematic cuts in order to maximize $S/\sqrt{S+B}$ and, at the same time, to keep enough statistic for a reliable QCD data-driven model and for an efficient shape analysis. Finally, we discuss three control regions, orthogonal to the signal region, to perform cross-check on the quality of the QCD background estimate and to obtain the normalization factor.

6.1 Preselection requirements

The starting data sample consists of an integrated luminosity of 5.7 fb^{-1} collected with the MET_DIJET trigger. We require the run to be a good one (*i.e.* all detector parts needed for the analysis are in working conditions) and to have at least one well-reconstructed primary vertex (as defned in chpt. 3) which lies inside the luminous region $Z_V \leq 60$ cm (this requirement is satisfied for $\simeq 97\%$ of the interactions). The trigger, described in section 3.11, asks for $\not\!\!E_T > 30$ GeV, one hadronic jets with transverse energy $E_T \geq 10$ GeV and $|\eta| \leq 2.4$, and at least another hadronic jet with $E_T \geq 3$ GeV and $|\eta| \leq 2.4$.

After offline reconstruction, we evaluate the number of jets with $E_T \geq 15$ GeV and $|\eta| \leq 2.4$, and require it to be at least five; also, we require $E_T \geq 30$ GeV for the first two jets and $E_T \geq 20$ for the third, fourth and fifth jet. The N_{jets} lower cut is set to 5 instead of 6 because, with large jet multiplicity, the probability to have one jet lost along the beamline or to have the merging of two jets is large (see the N_{jets} distribution in Fig. 6.6). We also veto events containing at least one lepton, where leptons are identified as described in section 3.6. This cut allows to substantially reduce all backgrounds apart from QCD (non-QCD background), with the exception of $Z \to \nu\nu$ + jets. In particular, this cut greatly reduce semileptonic $t\bar{t}$; residual background coming from this process arise only if the lepton fails identification.

After applying these cuts on the data, we are left with about 50000 events, while expected contribution from non-QCD backgrounds accounts for about 2000 events, as is shown in Tab. 6.1. The preselected sample is dominated for more than 95% by QCD multijet. We show in Fig. 6.1 - 6.7 a comprehensive list of kinematic variables for the data, the $t\bar{t}$ MC background, and for two signal hypothesis. The differences in all these variables distributions allow to perform kinematic cuts to isolate signal from background. From the plots in Fig. 6.1 - 6.7 we can see the shapes differences between the two largest source of backgrounds, *i.e.* QCD multijet (which, at this point, is dominating in data) and $t\bar{t}$, and two representative points in the $(m_{T'}, m_X)$ plane for our $T'\bar{T}'$ signal.

Table 6.1: Number of predicted and observed events in the preselection region. The expected number of signal events is shown for two representative points in the $(m_{T'}, m_X)$ parameters space. The uncertainty in the predicted number of events is due to the theoretical cross section uncertainty and to the uncertainty on luminosity.

Process	Events
$t\bar{t}$	1566 ± 210
W+ jets	395,7 ± 160.1
Z+ jets	98.9 ± 40.0
$WW/WZ/ZZ$	80.0 ± 10.0
Single top	7.2 ± 1.0
Total MC	2148 ± 267
Observed	49979
$(m_{T'} = 330\,GeV/c^2, m_X = 40\,GeV/c^2)$	91.5 ± 12.3
$(m_{T'} = 380\,GeV/c^2, m_X = 1\,GeV/c^2)$	35.2± 4.7

- The azimuthal distances between $\vec{\not\!\!E}_T$ and the 5 more energetic jets, $\Delta\phi(\vec{\not\!\!E}_T, \vec{J}_i)$
- The azimuthal distance between $\vec{\not\!\!E}_T$ and $\vec{\not\!\!p}_T$, $\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T)$ (see chpt. 5)
- The transverse energies of the jets, $J_i E_T$
- The scalar sum of the transverse energies of all the hadronic jets, H_T
- The scalar sum of the transverse energies collected in each of the calorimetric towers, $Sum\, E_T$
- The number of jets, N_{jets}
- The $\not\!\!E_T$
- The $\not\!\!p_T$
- The $\not\!\!E_T$ divided by the square root of the $Sum\, E_T$, $\not\!\!E_T$ significance

The study of these distributions highlights the differences between signal and backgrounds in the $\not\!\!E_T$-related variables. These differences are stronger for high $m_{T'}$ and/or low m_X, because in these conditions the momentum carried by the decay products of the T' is larger.

6.2 Selection cut optimization

Since the preselected data sample is more than 95% QCD multijet, it can be used as a starting approximation of a QCD multijet sample. The choice of the selection cuts start from a visual inspection of the distribution in the preselection region, and then are subsequentially optimized in order to maximize $S/\sqrt{S+B}$, but also keeping enough statistic to avoid excessive fluctuations in the QCD data-driven model. The final selection cuts are:

- $\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T) < \pi/2$
- $\Delta\phi(\vec{\not\!\!E}_T, \vec{J}_i) > 0.4$ (for i = 1,2,3); > 0.2 (for i = 4,5)

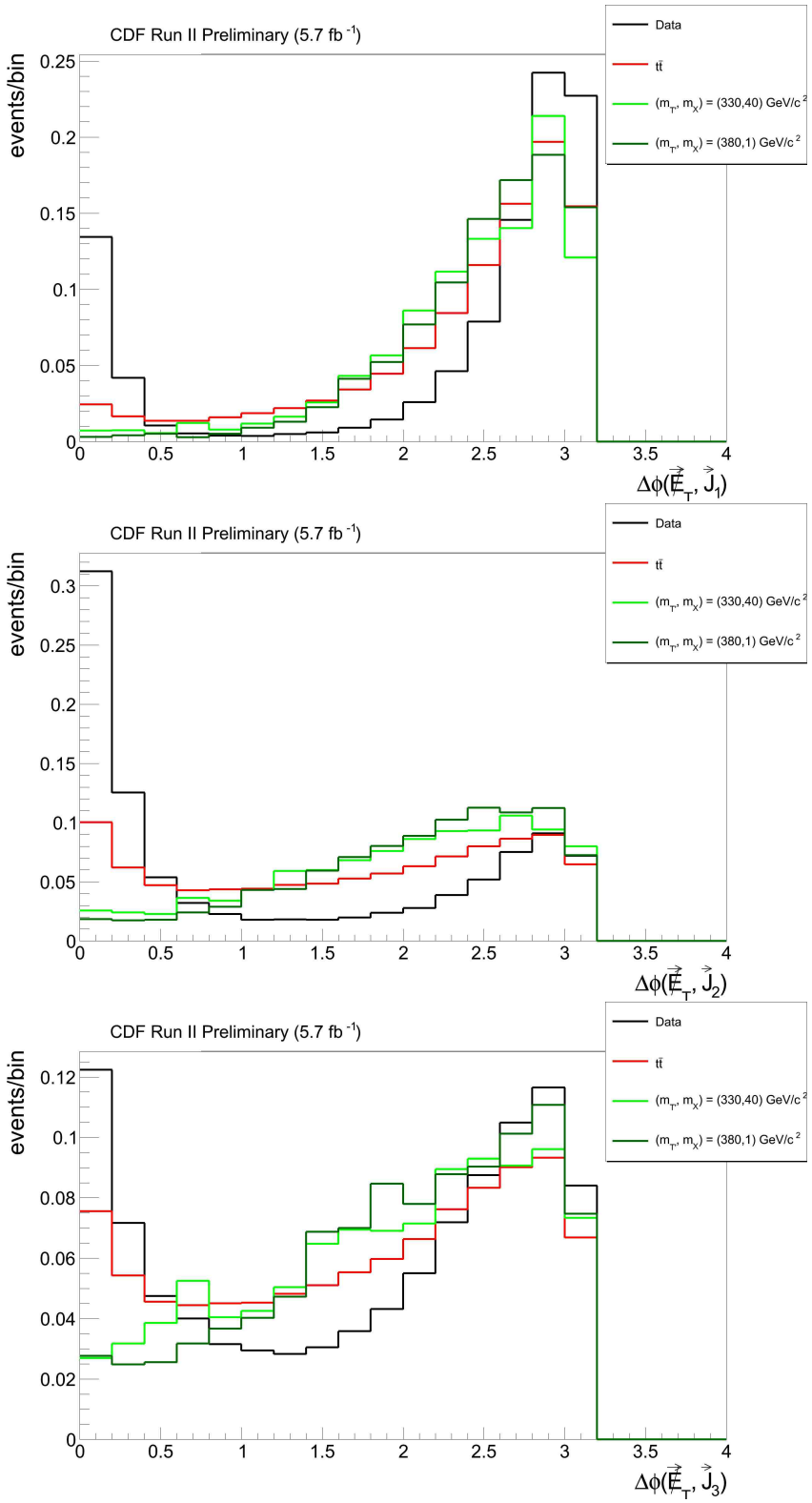


Figure 6.1: Distribution of the azimuthal distances between $\vec{\not{E}}_T$ and the 3 more energetic jets for the preselected data sample, the $t\bar{t}$ MC sample and two signal hypothesis. All histograms are normalized to the unit area.

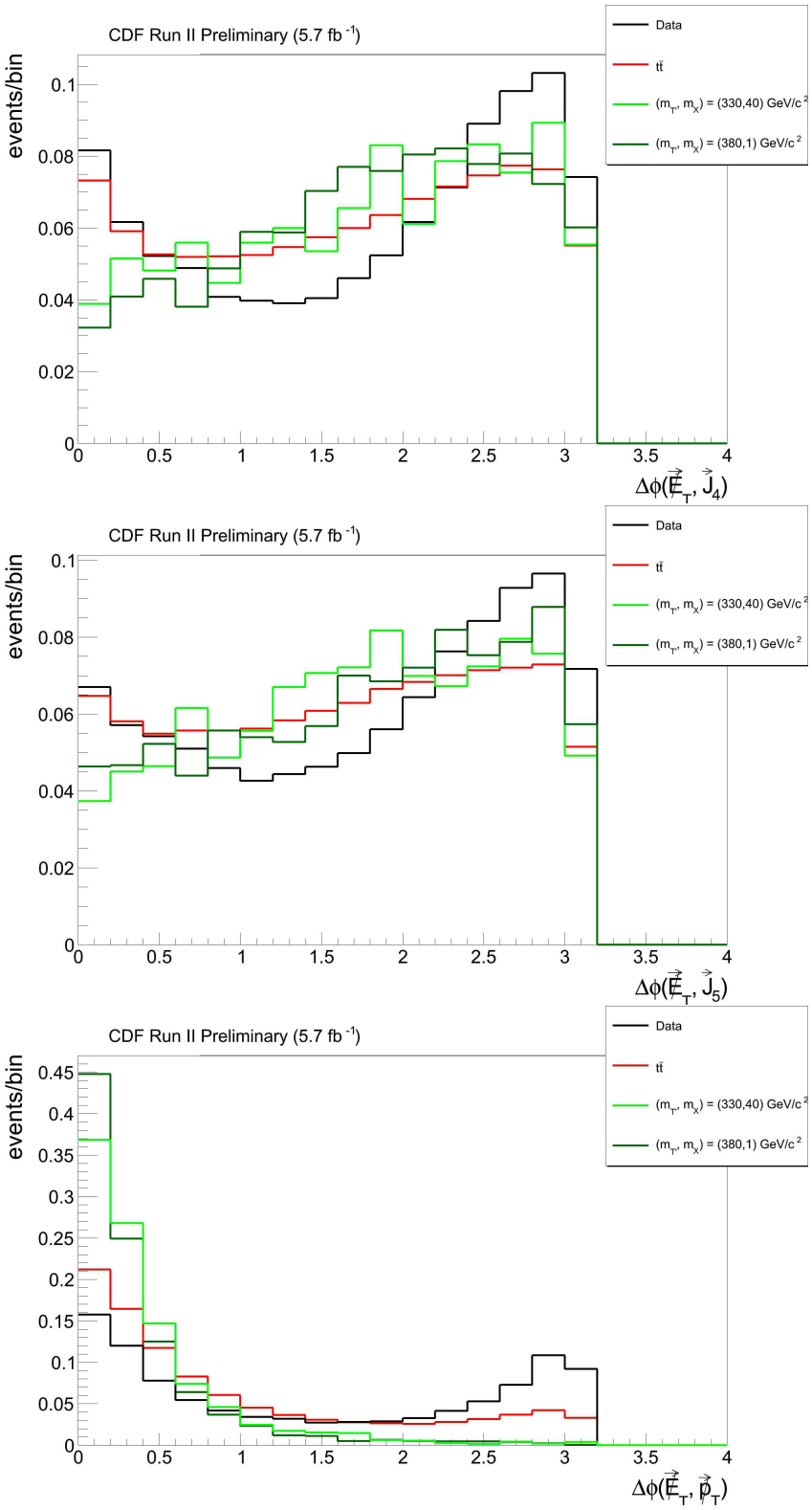


Figure 6.2: Distribution of the azimuthal distance between $\vec{\not\!\!E}_T$ and the fourth jet, the fifth jet and the $\vec{\not\!\!p}_T$ for the preselected data sample, the $t\bar{t}$ MC sample and two signal hypothesis. All histograms are normalized to the unit area.

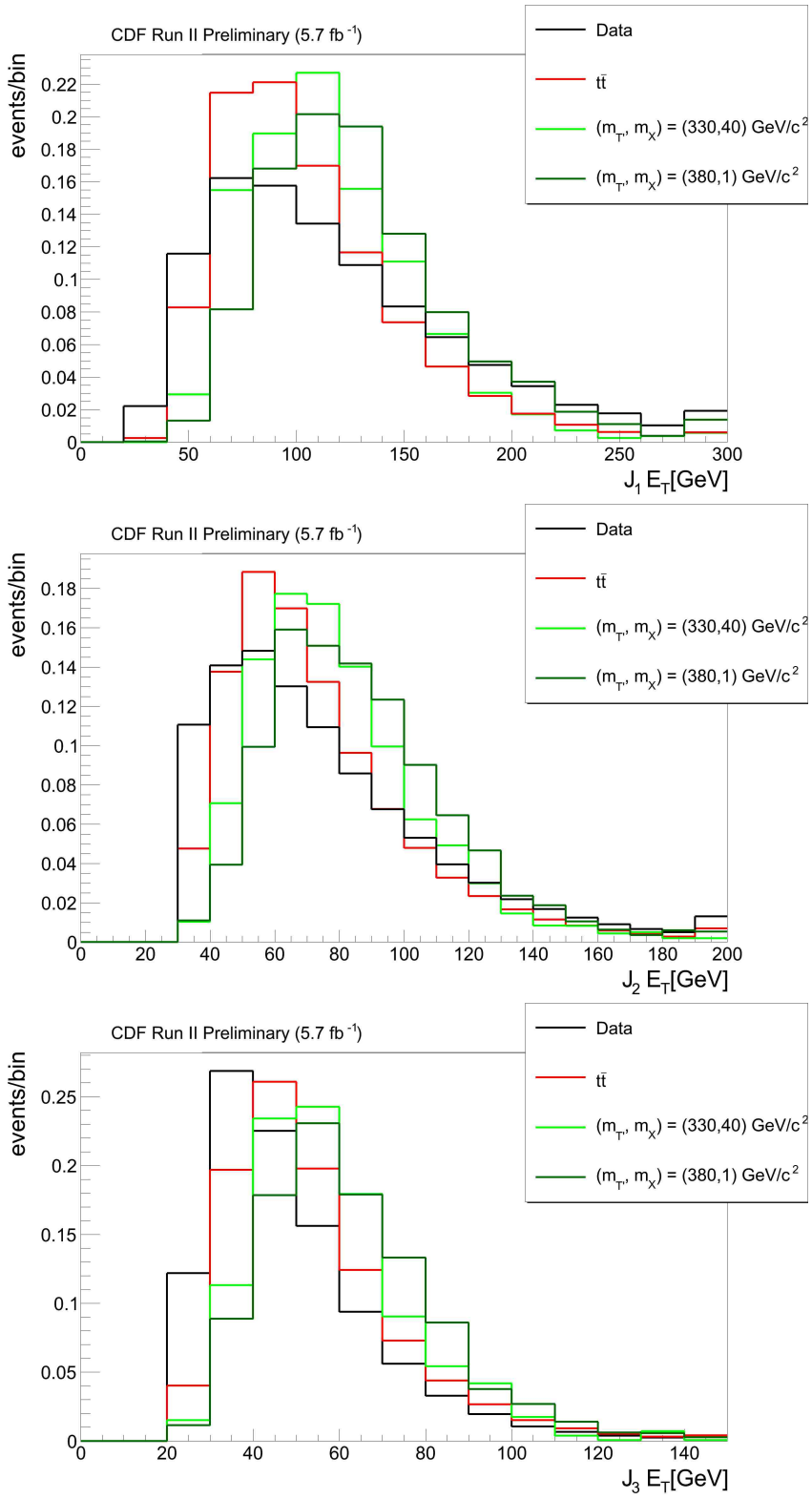


Figure 6.3: Distribution of the first 3 jets transverse energies for the preselected data sample, the $t\bar{t}$ MC sample and two signal hypothesis. All histograms are normalized to the unit area.

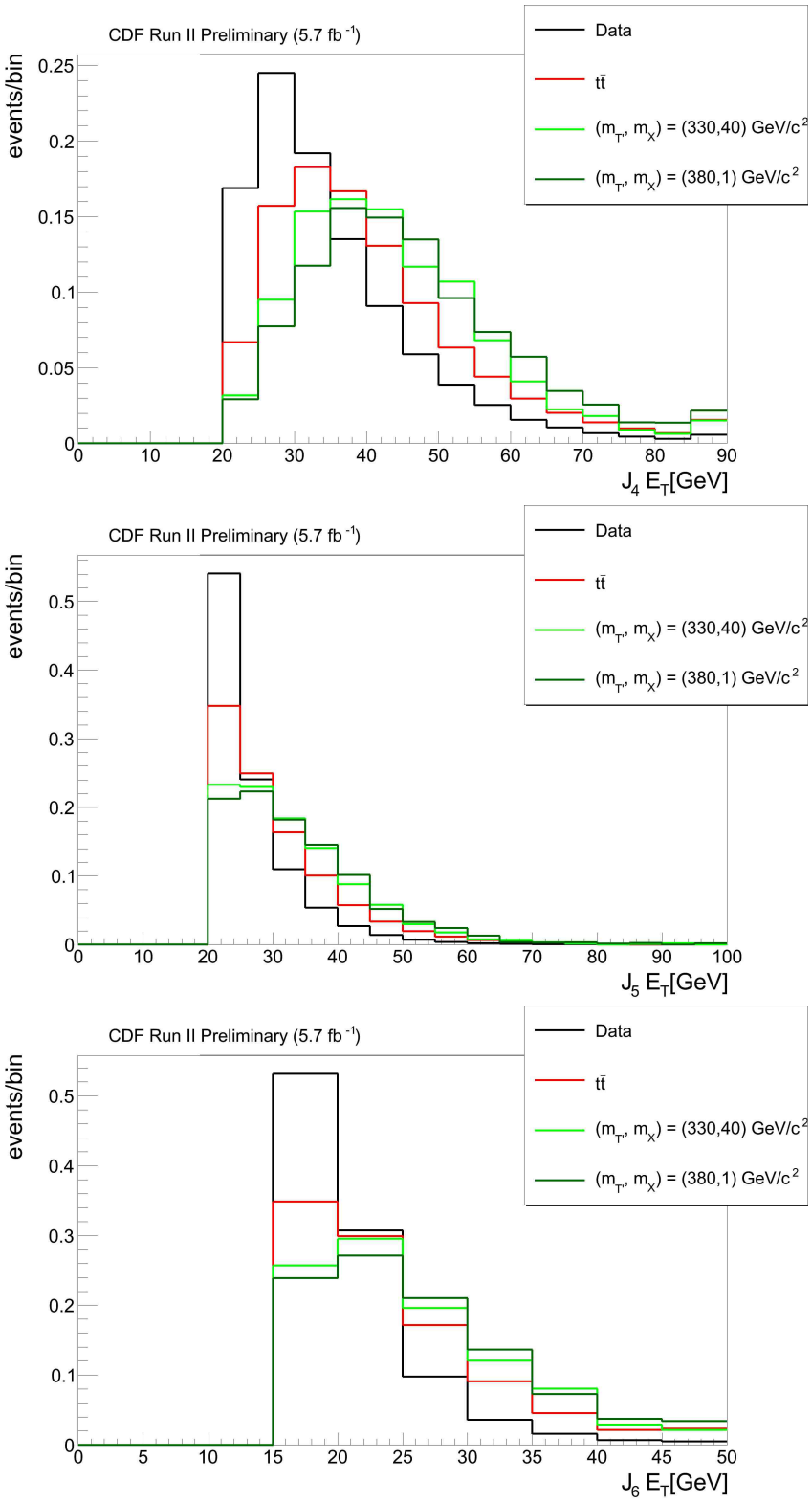


Figure 6.4: Distribution of the fourth, fifth and sixth jet's transverse energies for the preselected data sample, the $t\bar{t}$ MC sample and two signal hypothesis. All histograms are normalized to the unit area.

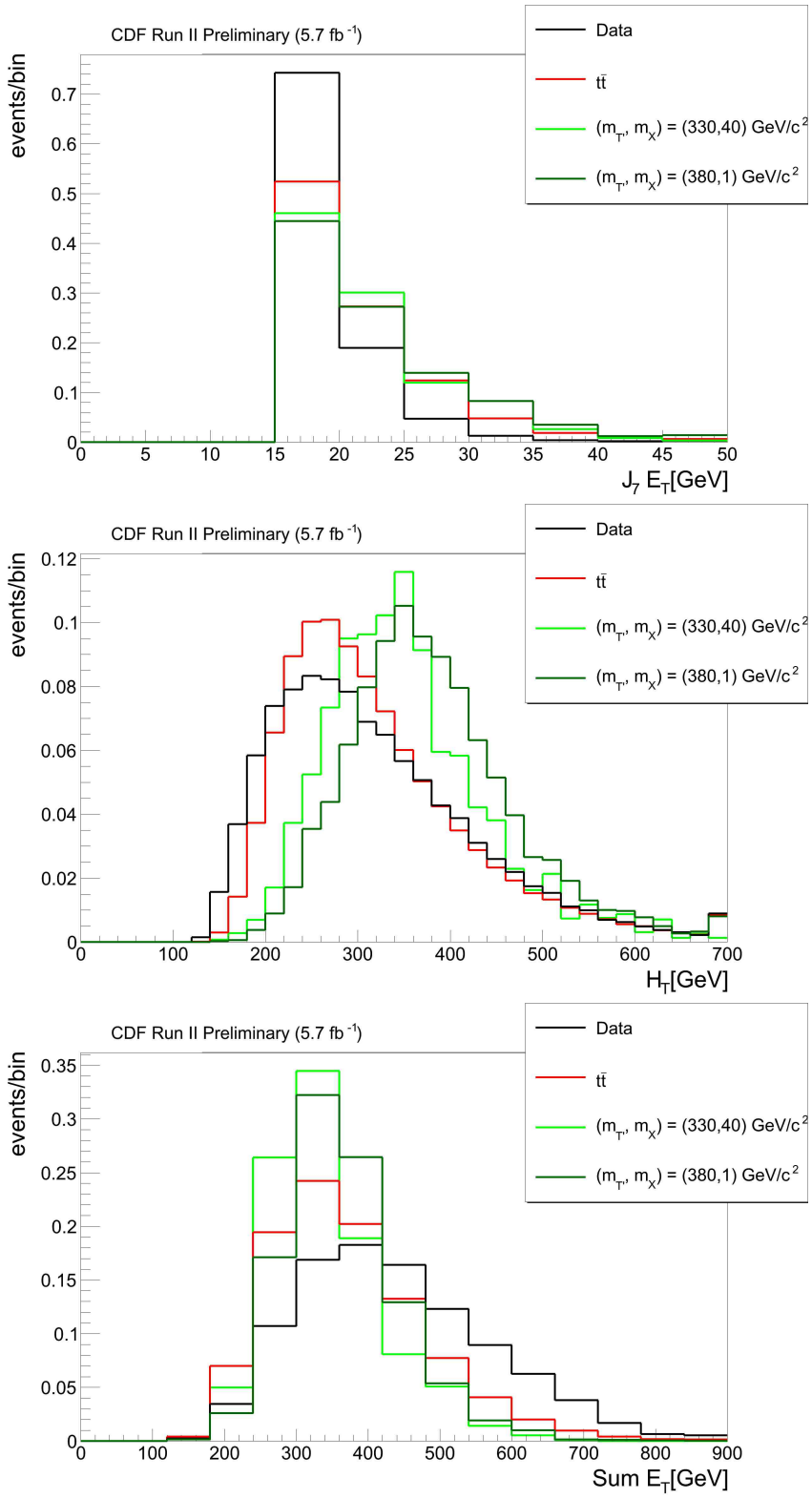


Figure 6.5: Distribution of the seventh jet's transverse energy, the scalar sum of the jes transverse energies and the total transverse energy collected in all the calorimeter for the preselected data sample, the $t\bar{t}$ MC sample and two signal hypothesis. All histograms are normalized to the unit area.

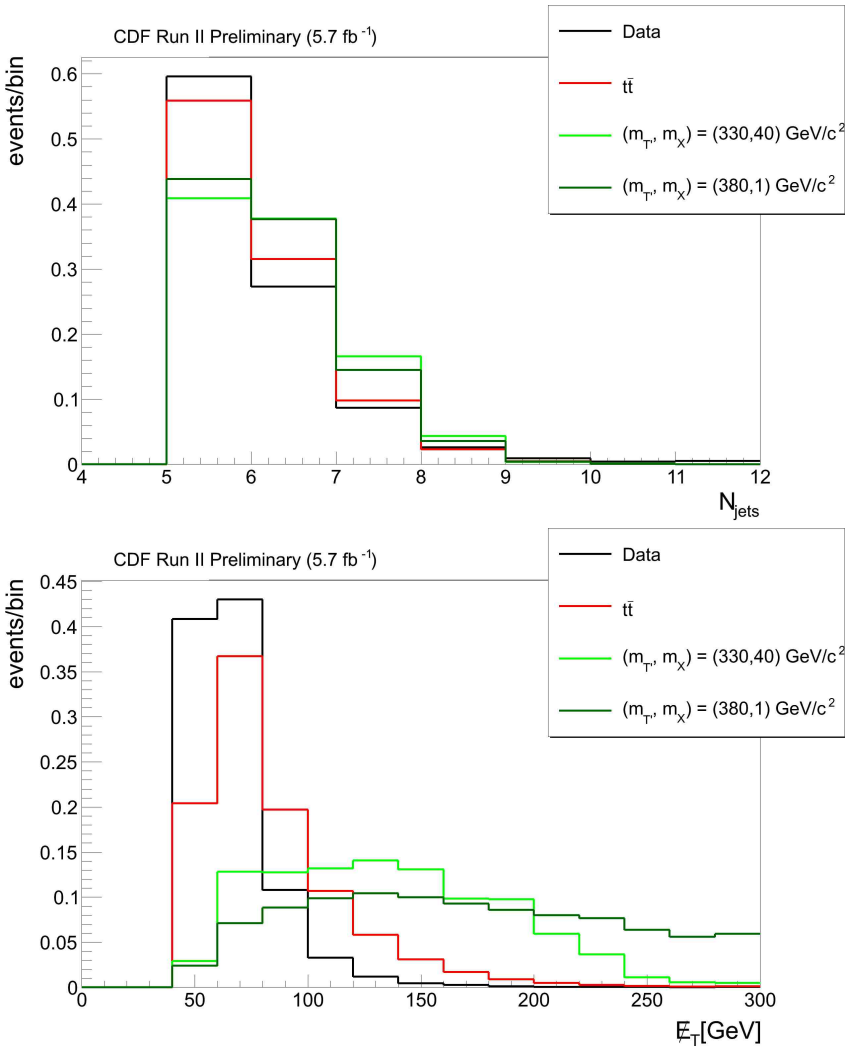


Figure 6.6: Distribution of the number of jets and the $\not\!\!E_T$ for the preselected data sample, the $t\bar{t}$ MC sample and two signal hypothesis. All histograms are normalized to the unit area.

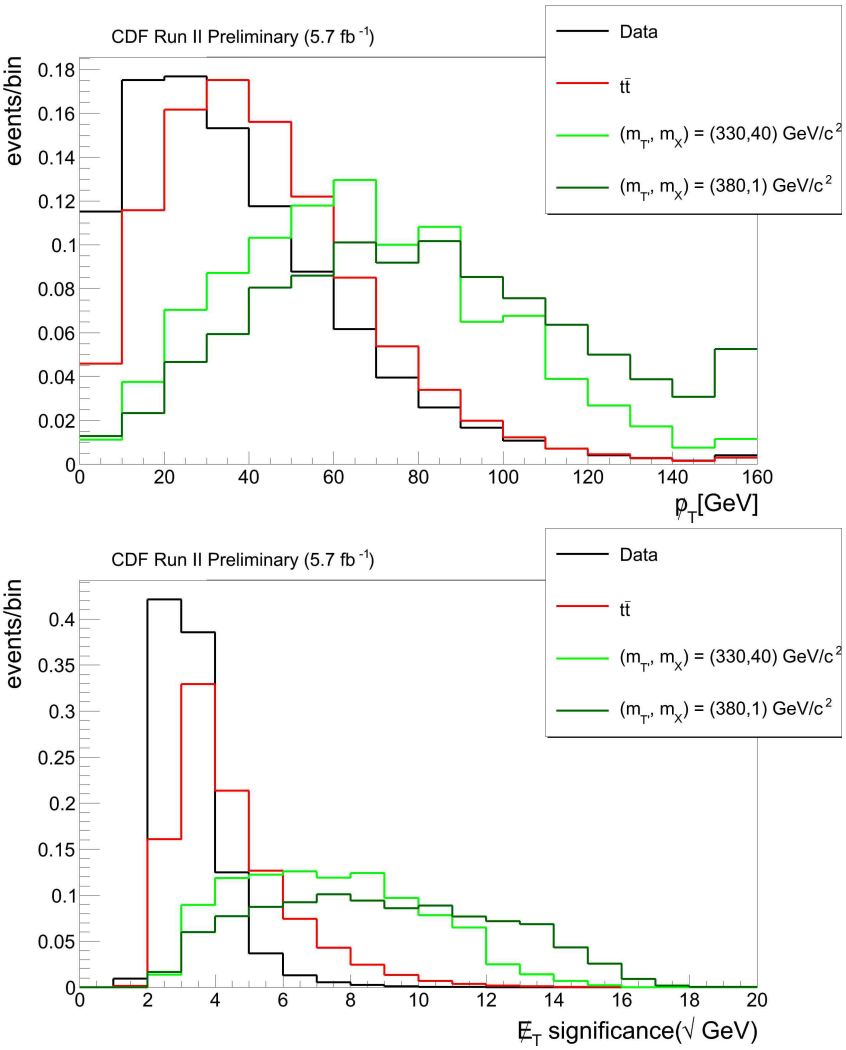


Figure 6.7: Distribution of the number of the $\not\!\!p_T$ and the $\not\!\!E_T$ significance for the preselected data sample, the $t\bar{t}$ MC sample and two signal hypothesis. All histograms are normalized to the unit area.

Most of the QCD multijet events have the $\vec{\not{E}}_T$ aligned with one of the subleading jets. This happens because of mismeasurement of that jet's energy. In dijet-like events, the most likely case is when we have two jet of similar energy, and one of them have its energy underestimated, so that $\vec{\not{E}}_T$ will be aligned with the second jet. As can be seen in Fig. 6.1, this is also our case.

- $H_T > 220$ GeV

  Top quarks produced by massive T' decays are expected to be more boosted, on average, than top quarks from SM production. Because of this, the hadronic jets are expected to be more energetic for our signal. This cut allow to reduce all the SM backgrounds.

- $\not{p}_T > 20$ GeV

  This is complementary to the $\not{E}_T$ cut. Since $\not{p}_T$ underestimates the undetected particle's energy because it does not take into account the energy carried by the neutral components of the jets (see chpt. 6) the optimal cut is less strict than the one on $\not{E}_T$.

- $\not{E}_T significance > 3$ GeV$^{\frac{1}{2}}$

  Since the calorimeter resolution on a jet's energy measurement is proportional to $E_{jet}{}^{\frac{1}{2}}$, $\sqrt{Sum\, E_T} = \sqrt{\sum_{towers} E_T}$ is an estimate of the uncertainty on $\not{E}_T$. This cut is equivalent to require a $\not{E}_T$ different from 0 at 3 σ.

- $N_{jets}(E_T > 15$ GeV, $|\eta| < 2.4) \leq 10$

  The signal we are looking for contains on average 6 jets. Due to jet merging and radiation, this number is actually the mean of a relatively symmetric distribution. The probability to have more than 10 jets is naively ${\alpha_S}^4$ (actually a bit larger due to the fact that there are many quarks that can radiate). On the other hand, events with very large jet multiplicity arise naturally from pileup. Also, the more the jets in the final state, the larger the probability that the energy of at least one of those is mismeasured, giving thus rise to large $\not{E}_T$. We confirmed that these events are actually coming from pile up by isolating them and plotting the number of vertexes (N_{vtx}) distribution and the

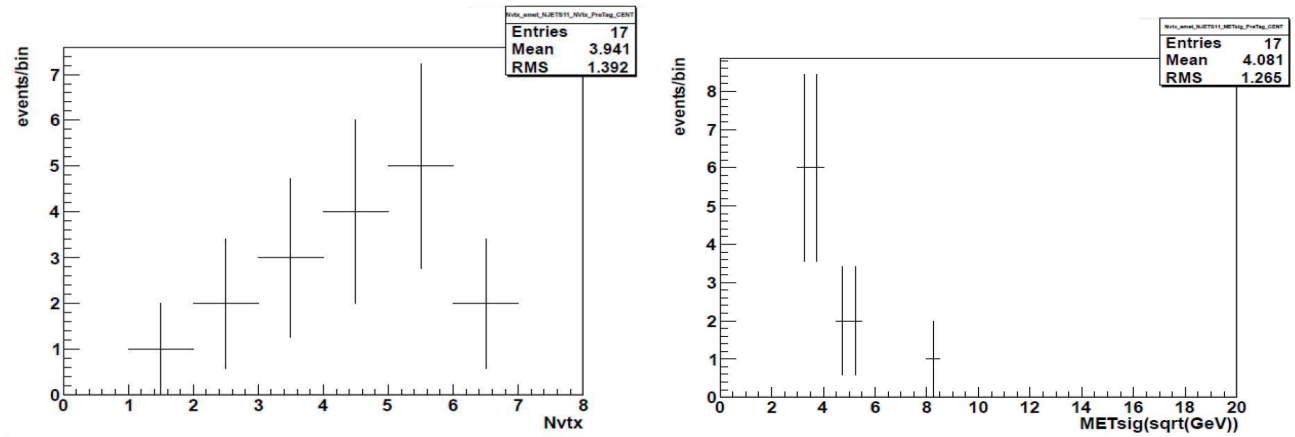


Figure 6.8: Distributions of number of reconstructed primary vertexes (left) and $\not\!\!E_T$ significance (right) for events passing the selection cuts which have more than 10 hadronic jets.

$\not\!\!E_T$ significance distribution (see Fig. 6.8). There are 17 events with more than 10 jets. The number of reconstructed primary vertexes distribution peaks at 4, to be compared to a peak of about 2 in our signal region. The $\not\!\!E_T$ significance distribution is steeply falling as expected from QCD. Our conclusion is that the events with $N_{jets} > 10$ in the preselected sample is QCD multijet with significantly higher pile-up than average.

We also checked the effects of including a b-tagging requirement (see section 3.5). Though reducing QCD multijet and W/Z processes, b-tagging requirements reduce acceptance on signal without reducing $t\bar{t}$, which is the largest background in the high $\not\!\!E_T$ significance region, *i. e.* the region with higher sensitivity (see section 6.4). Therefore, we don't include b-tagging requirements in our selection. The final selection, including preselection requirements, is summarized in Tab. A.1.

These selection cuts, togheter with the preselection requirements, define our signal region. The efficiency on the signal is 70-90%, depending on the $m_{T'}$ and m_X we are assuming (higher for high $m_{T'}$ and/or low m_X). The selection allows to reduce QCD background by two order of magnitude, and the $t\bar{t}$ and W/Z+jets backgrounds by a factor of 3, as can be seen comparing values in Tab. 6.1 and 6.4. It is important to point out that the above cuts are not designed to reject as much QCD multijet as possible, searching for an absolute maximum of $S/\sqrt{S+B}$. In a counting experiment this would be the right strategy, but since we want to

Table 6.2: Selection cuts defining the signal region.

Signal Region
$\not{E}_T > 50$ GeV
$N_{jets}(E_T > 30$ GeV, $\lvert\eta\rvert < 2.4) \geq 2$
$N_{jets}(E_T > 20$ GeV, $\lvert\eta\rvert < 2.4) \geq 5$

$\Delta\phi(\vec{\not{E}}_T, \vec{\not{p}}_T) < \pi/2$
$\Delta\phi(\vec{\not{E}}_T, \vec{J}_i) > 0.4$ (for i = 1,2,3)
$\Delta\phi(\vec{\not{E}}_T, \vec{J}_i) > 0.2$ (for i = 4,5)
$H_T > 220$ GeV
$\not{p}_T > 20$ GeV
$\not{E}_T$ significance > 3 GeV$^{\frac{1}{2}}$
$N_{jets}(E_T > 15$ GeV, $\lvert\eta\rvert < 2.4) \leq 10$

do a shape analysis we need enough statistic to avoid excessive fluctuations in the selected data. Moreover, our QCD sample needs enough statistic to be a reliable QCD model. Therefore, our selection cuts are a compromise between enhancing signal and keeping high statistic.

6.3 Control regions

The QCD sample for our data-driven model is made of events that pass all these selection cuts, inverting the $\Delta\phi(\vec{\not{E}}_T, \vec{\not{p}}_T)$ requirement. To perform an unbiased measure of the normalization factor we apply to our QCD sample, we define three signal-depleted control regions. We do this inverting one of the selection cuts for each control region we want to define, keeping others unchanged. The three control regions are defined by one of the following cuts:

- $N_{jets} = 4$
- $\not{E}_T sig < 3$ GeV$^{\frac{1}{2}}$
- $\not{p}_T < 20\,GeV$

For each of these regions, we will have a correspondent QCD region with inverted $\Delta\phi(\vec{\not{E}}_T, \vec{\not{p}}_T)$ requirement. We calculate a QCD normalization factor for each control

region as the one to exactly match the number of observed data events:

$$R_{QCD} = \frac{N_{obs} - N_{t\bar{t}} - N_{W/Z+jets} - N_{diboson} - N_{single\,top}}{N_{QCD}} \tag{6.1}$$

where N_{obs} is the number of observed events in the control region, $N_{t\bar{t}}$, $N_{W/Z+jets}$, $N_{diboson}$ and $N_{single\,top}$ are the expected number of non-QCD background events in the control region, and N_{QCD} is the number of observed events with $\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T) > \pi/2$ in the control region (QCD region). Results are shown inTab. 6.3. The average value we get from the three normalization factor is 1.2, and we assign a ± 20% uncertainty over this value. This uncertainty cover the spread we observe in the three measured values.

Table 6.3: Number of predicted and observed events in the three control regions. The uncertainty in the predicted number of MC events is due to the uncertainty of theoretical cross section and luminosity. The QCD scale factor is calculated to match the QCD events with the difference between observed events and the sum of SM MC predictions

	$N_{jets} = 4$	$\not\!\!E_T$ significance < 3 GeV$^{\frac{1}{2}}$	$\not\!\!p_T < 20$ GeV
Observed	2467	1219	289
QCD (scaled)	1482 ± 140	1151 ± 39	222.4 ± 18.9
$t\bar{t}$	580.6 ± 77.9	52.1 ± 7.0	50.2 ± 6.7
$W+$ jets	263.5 ± 106.6	12.2 ± 4.9	12.0 ± 4.8
$Z+$ jets	103.1 ± 41.7	2.1 ± 0.8	2.4 ± 1.0
$WW/WZ/ZZ$	31.9 ± 4.0	1.6 ± 0.2	1.8 ± 0.2
Single top	10.8 ± 1.5	0.3 ± 0.04	0.30 ± 0.04
Observed scale factor	1.35 ± 0.12	1.18 ± 0.04	1.03 ± 0.09

To test the QCD background modeling, we check the agreement of the distributions between data and total background in the control regions, where the total background include non-QCD background and QCD background normalized as described before. We can do this in the two regions with more statistics, $N_{jets} = 4$ and $\not\!\!E_T significance < 3$ GeV$^{\frac{1}{2}}$. In the first control region we have representative samples of all SM backgrounds, while the second is mainly QCD multijet. The distributions show good agreement between data and expectations, as can be seen in Fig. 6.9. Other kinematic variables' distributions in control regions are shown in

Appendix B.

6.4 Signal region

In Tab. 6.4 we show the number of predicted and observed events in the signal region, with QCD normalization extracted from control regions. Fig. 6.10 shows the $\not\!\!E_T$ significance distribution in the signal region. The $\not\!\!E_T$ significance is the most discriminating variable between signal and background, ando thus will be extensively used in the limits setting procedure, as we will see in the next Chapter. As can be seen, the data shows good agreement with the SM expectations. Other kinematic variables' distributions are shown in Appendix A.

Table 6.4: Number of predicted and observed events in the signal region. The expected number of signal events is shown for two representative points in the $(m_{T'}, m_X)$ parameters space. The uncertainty in the predicted number of MC events is due to the uncertainty of theoretical cross section and luminosity.

Process	Events
QCD	745.4 ± 124.3
$t\bar{t}$	498.2 ± 66.8
W+ jets	119.7 ± 48.4
Z+ jets	39.4 ± 15.9
$WW/WZ/ZZ$	17.9 ± 2.2
Single top	5.3 ± 0.8
Total	1423 ± 150
Observed	1507
$(m_{T'} = 330\,GeV/c^2, m_X = 40\,GeV/c^2)$	66.7 ± 3.0
$(m_{T'} = 380\,GeV/c^2, m_X = 1\,GeV/c^2)$	27.3 ± 1.9

At this point, we have a reliable background modeling, and the selection cuts have higly enhanced $S/\sqrt{S+B}$; as can be seen in Tab. 6.5, the most sensitive region is the one with high $\not\!\!E_T$ significance (≥ 10 GeV$^{\frac{1}{2}}$).

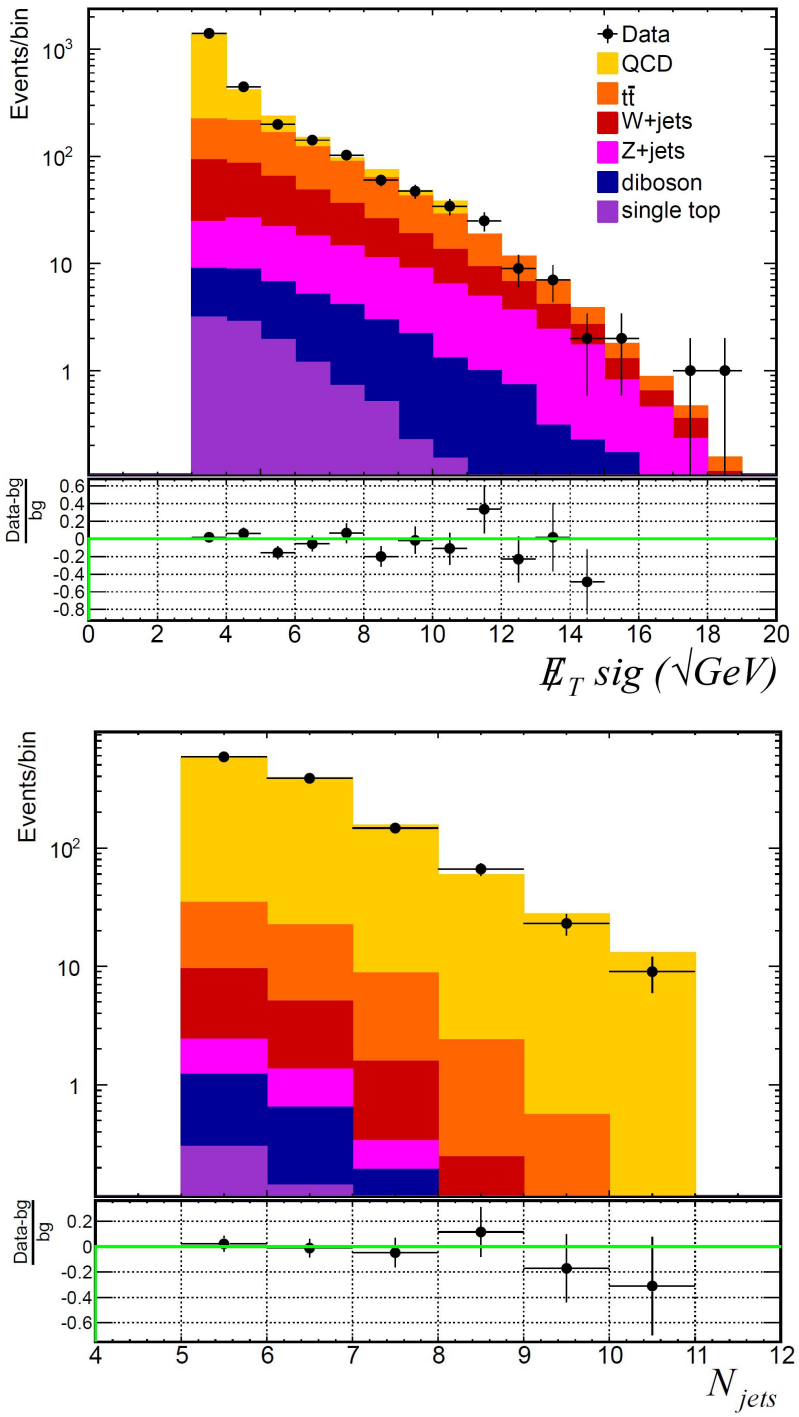


Figure 6.9: Distributions of $\not\!\!E_T$ significance in the $N_{jets} = 4$ control region (top) and of N_{jets} in the $\not\!\!E_T$ significance <3 control region for data and expected backgrounds. Each process is normalized to the number of expected events. The QCD contribution is normalized to the number of observed events minus the $t\bar{t}$, W/Z+jets, diboson and single top expected contributions.

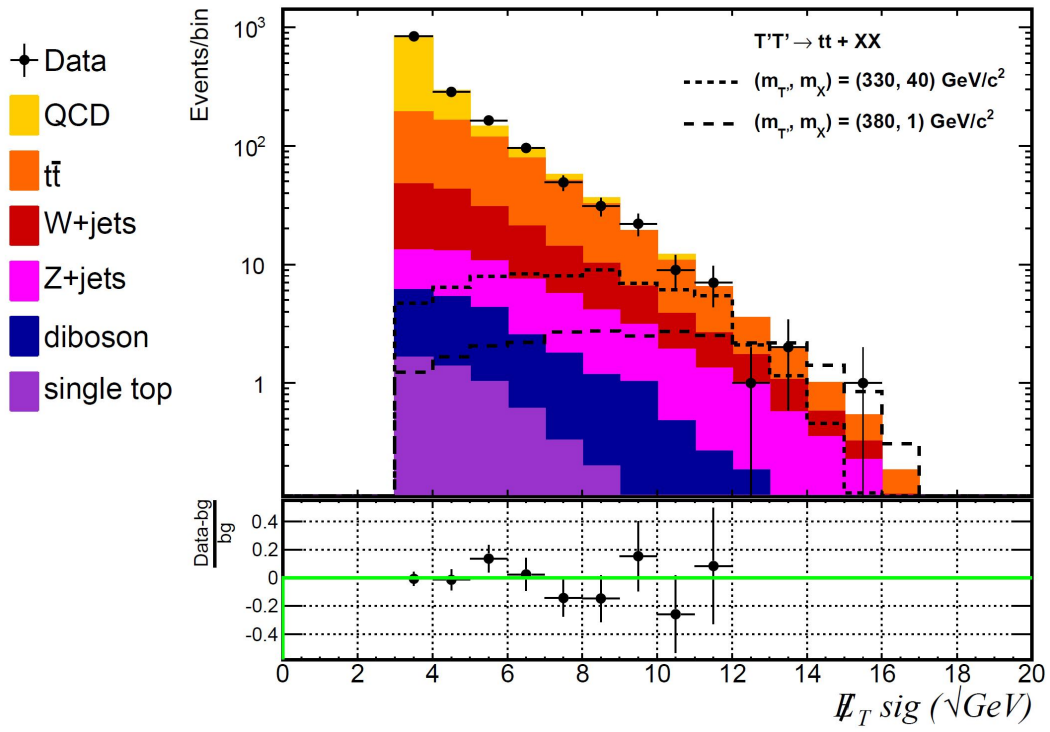


Figure 6.10: Distributions of $\not\!\!E_T$ significance in the signal region for data, expected backgrounds and two signal hypothesis. Each process is normalized to the number of expected events. The QCD contribution is normalized to the number of observed events minus the $t\bar{t}$, W/Z+jets, diboson and single top expected contributions.

Table 6.5: Values of $S/\sqrt{S+B}$, for preselection region, signal region and a subsample of the signal region with high $\not\!\!E_T$ significance. Values are computed for two representative points in the $(m_{T'}, m_X)$ parameters space: S_1 ($m_{T'} = 330\,GeV/c^2, m_X = 40\,GeV/c^2$) and S_2 ($m_{T'} = 380\,GeV/c^2, m_X = 1\,GeV/c^2$).

	$S_1/\sqrt{S_1+B}$	$S_2/\sqrt{S_2+B}$
Preselection Region	0.4	0.2
Signal Region	1.7	0.6
Signal Region ($\not\!\!E_T$ significance ≥ 10 GeV$^{\frac{1}{2}}$)	2.5	2.0

Chapter 7

Limits Setting

As shown in previous chapter, the observed $\not\!\!E_T$ significance distribution show no significant excess above SM predictions. Therefore, we place upper limits on the productions of T' quark pairs decaying in top quark and undetected particles. In order to calculate the upper limits, we analyze the $\not\!\!E_T$ significance distribution with a binned maximum-likelihood fit. The limits are computed using the Bayesian likelihood method [76, 77] with flat prior probability for the signal cross section and Gaussian priors for the uncertainties on acceptance and backgrounds. The software package we use for limit computation is called MCLIMIT [78] and contains an implementation of a multichannel Bayesian limit calculator [79], that allows to consider the correlations between various acceptance and background priors.

7.1 The Bayesian approach

In the simplest case of a counting experiment and a single source of background, Bayesian approach to calculate the limit is the following. Let us assume an experiment is conducted, and n events (Poisson distributed) are observed, with a mean expectation $s\epsilon + b$, where s is the number of expected signal events, ϵ is the acceptance (product of branching fraction, detector efficiency, luminosity factor, etc.) and b is the number of background events we expect. We now assume that the background is known precisely, and the signal acceptance is known with a precision

σ_ϵ from other measurements. The Bayes theorem then states that the posterior $p(s, \epsilon|n)$ for s is:

$$p(s, \epsilon|n) = \frac{P(n, s|\,\epsilon)\pi(s)\pi(\epsilon)}{\int\int P(n, s|\,\epsilon)\pi(s)\pi(\epsilon)\, ds\, d\epsilon} \tag{7.1}$$

where $P(n, s|\,\epsilon)$ is the probability of observing n events in presence of a signal s with acceptance ϵ. while $\pi(s)$ and $\pi(\epsilon)$ are respectively the prior probability densities for s and for ϵ. Since the number of observed events in data is given by the Poisson statistic, we have:

$$P(n, s|\,\epsilon) = \frac{(s\epsilon\, +\, b)^n}{n!} e^{-(s\epsilon + b)} \tag{7.2}$$

Then we can obtain the posterior for s integrating over ϵ:

$$p(s|n) = \int_0^\infty p(s, \epsilon|n)\, d\epsilon \tag{7.3}$$

Finally, to obtain the limit on s, the posterior density function $p(s|n)$ is integrated until the desired confidence level (C.L.) is reached. Therefore, the 95% C.L. upper limit s_u can be obtained by the following equation:

$$\int_0^{s_u} p(s|n)\, ds = 0.95 \tag{7.4}$$

7.2 Multichannel Bayesian fit procedure

The method described above refers to a single counting experiment. In the case of a binned distribution of a discriminant variable, with a total number of bins N, we can treat each bin as a statistically independent counting search. Therefore, for the k^{th} bin we will have n_k observed events and $s\epsilon_k + b_k$ expected events, where s is the total number of expected signal events, while ϵ_k and b_k are the signal acceptance and the amount of expected background in the k^{th} bin. All of the ϵ_k and b_k have uncertainties and are considered "nuisance parameters". Similar to the procedure above, they are assigned priors, that may be correlated (for example, as in the case of luminosity uncertainty). We can write the overall prior as:

$$\pi(\epsilon_1,\, b_1, ...,\, \epsilon_n, b_n) \tag{7.5}$$

so that the posterior can be written, as a function of s, in the following way:

$$p(s|n) \,=\, \pi(s) \int_{...2N...} \int \pi(\epsilon_1,\, b_1, ...,\, \epsilon_n, b_n) \left[\prod_{k=1}^{N} \frac{e^{-(s\epsilon_k + b_k)} (s\epsilon_k + b_k)^{n_k}}{n_k!} \right] d\epsilon_1\, db_1 \,...\, d\epsilon_N\, db_N \tag{7.6}$$

The above integral is calculated with Monte Carlo integration, i.e. generating M random $(\epsilon_1,\, b_1, ...,\, \epsilon_n, b_n)$ vectors ("ensemble") according to their priors, and then averaging over M (in our analysis, $M = 10000$). The software we used takes into account the correlations between nuisance parameters when generating the random vectors. In this analysis, we use a a flat prior for the signal cross-section, and integrate over Gaussian priors for the systematic uncertainties. We distinguish between rate and shape uncertainties. Rate systematics are uncertainty on the expected number of signal or background events. Shape systematics arise when a systematic effects introduce a significant variation in the distribution of interest (in our case, the $\not\!\!E_T$ significance); in this case, the histograms are interpolated within their shape uncertainties, when generating the random sets. Shape uncertainties are provided as inputs to the MCLIMIT package, in addition to the nominal histogram shapes.

7.3 Systematic uncertainties

Systematic uncertainties are uncertainties that cannot be reduced simply increasing the sample size. They arise from incomplete knowledge of detector effects, like mechanical misalignement, miscalibration or detector noise, or from limited theoretical knowledge of a process in exam. Some categories of systematics, such as the luminosity uncertainties, are the same for all the processes; some others, such as cross sections uncertainties, affects each process in a different amount; finally, some theoretical uncertainties are typical of a single process.

Since the main background in the high sensitivity region is $t\bar{t}$ production, is necessary to carefully evaluate the systematics for this process. Moreover, since this

is basically a shape analysis, we have to check for possible shape uncertainties in $\not\!\!E_T$ significance distribution, togheter with the rate uncertainties, generated by the systematics for this background. The rate systematics we take into account in this analysis for all the processes are:

- **Luminosity**

  This systematic source accounts for the uncertainty in the $p\bar{p}$ inelastic cross section and for the uncertainty in the acceptance of the luminosity monitor of CDF to inelastic $p\bar{p}$ collision events [80] and it is estimated as 6% [81]. This uncertainty is applied to the rate predictions based on Monte Carlo simulation, *i.e.* all processes apart from multijet production.

- **Theoretical cross sections**

  For all physics processes modeled by Monte Carlo simulation, we normalize to the most up-to-date theoretical computation of the cross section, and corresponding uncertainties. We use 12% uncertainty for top quark pair production,[62], 40% uncertainty for the W and Z background processes,[82], 11% for the diboson prediction,[67], and 15% for the signal prediction [31].

For the $t\bar{t}$, we also take into account the following systematics:

- **Jet Energy Scale**

  The Jet Energy Scale (JES) is a series of corrections applied to the raw energy measured from the calorimeters (see Chapter 4.2.2) in order to estimate the true energy of the jet [88]. These corrections are necessary to account for various detector effects, such as calibration, η and ϕ dependence, and so on. The uncertainties of each correction are convoluted together to yield an overall uncertainty on the energy of the jet. The Jet Energy Scale (JES) rate uncertainty for $t\bar{t}$ is determined by varying the jet energy correction factor by plus and minus one standard deviation while correcting the transverse energy of the jets. The rate variation in the $t\bar{t}$ prediction, taken into account in MCLIMIT, is $\pm$ 10 %. The JES variation propagates to the $\not\!\!E_T$ and $SumE_T$ reconstruction, and thus to the $\not\!\!E_T$ significance shape, as is shown in 7.2. Thus, JES is included also as a shape systematic.

- **Monte Carlo generators**

  A source of systematic effect arise from uncertainty in the Monte Carlo modeling of the hard interaction. PYTHIA's and HERWIG's accuracy has been tested in many different studies. In both programs, the partonic interactions are generated using leading-order QCD matrix elements, including initial- and final-state parton showers. CTEQ5L [83] parton distribution functions are used for the proton and the antiproton. In the case of PYTHIA, fragmentation into hadrons is carried out using the string model [84] as implemented in JETSET [85]. We want to check for systematic effects on $t\bar{t}$ $\not\!\!E_T$ significance shape comparing it with the same distribution generated by HERWIG, in which the model used for fragmentation is the cluster mode [86]. The differences appeared to be negligible as can be seen in Fig 7.1. The rate variation, taken into account in MCLIMIT, is $\pm 9\%$.

- **Initial/Final State Radiation**

  Extra jets originating from the incoming partons and/or outgoing partons affect the measurement. Since effect of the initial state radiation (ISR) and final state radiation (FSR) is correlated, it is studied simultaneously. The model used by PYTHIA for gluon radiation from partons emitted from the hard-scattering interaction has been tuned with LEP data. Monte Carlo simulated samples are generated for single top quark signals and $t\bar{t}$ with $\Lambda_{\rm QCD}$ doubled (more ISR) or divided in half (less ISR) and with the initial transverse momentum scale and the hard scattering of the shower both multiplied (more ISR) or divided (less ISR) by four. The parameters for the final-state showering are also adjusted in PYTHIA, except for the hard-scattering scale. The uncertainties are then computed by comparing the efficiencies and kinematics of the varied ISR/FSR events to the nominal ones. For our analysis, we want to check for ISR/FSR systematic effects on $t\bar{t}$ $\not\!\!E_T$ significance shape. To do this, we compare the nominal distribution shape with the ones obtained enhancing or reducing ISR contribute by 1σ. Variations appeared to be negligible as can be seen in Fig 7.1.n The rate variation, taken into account in MCLIMIT, is $\pm 6\%$.

- **Color reconnection effect**

  Due to the fact that hadrons originating from $t\bar{t}$ decay overlap in space and are created in time almost simoultaneously, it is natural to expect that there are correlations between these hadrons due to color reconnection [87]. Pythia take in account such effect, and we check for systematic effects on $t\bar{t}$ $\not\!\!E_T$ significance shape varying the magnitude of color reconnection by $\pm 1\sigma$. Even in this case, effects on the shape seems to be negligible as can be seen in Fig 7.1. The rate variation, taken into account in MCLIMIT, is $\pm 3\%$.

Finally, we take in account the following systematics:

- Rate uncertainties on signal prediction due to JES range from $\pm 2\%$ $\pm 4\%$. We assign a $\pm 3\%$ to all the signal samples.

- We assign a $\pm 20\%$ rate uncertainty to the QCD multijet background, due to the uncertainty on the normailzation factor we get from the control regions (see section 6.3).

A summary of all the systematics is given in Tab. 7.1.

Table 7.1: Summary of the systematic uncertainties. For every process, we also take into account a $\pm 6\%$ rate systematic due to luminosity uncertainty.

Process	Systematic source	Rate uncertainty	Shape systematic
$t\bar{t}$	Cross section	$\pm 12\%$	no
	JES	$\pm 10\%$	yes
	MC Generator	$\pm 9\%$	no
	Color Reconnection	$\pm 3\%$	no
W/Z+jets	Cross section	$\pm 40\%$	no
Diboson	Cross section	$\pm 11\%$	no
Single top	Cross section	$\pm 13\%$	no
Signal	Cross section	$\pm 12\%$	no
	JES	$\pm 3\%$	yes
QCD multijet	Normalization	$\pm 20\%$	no

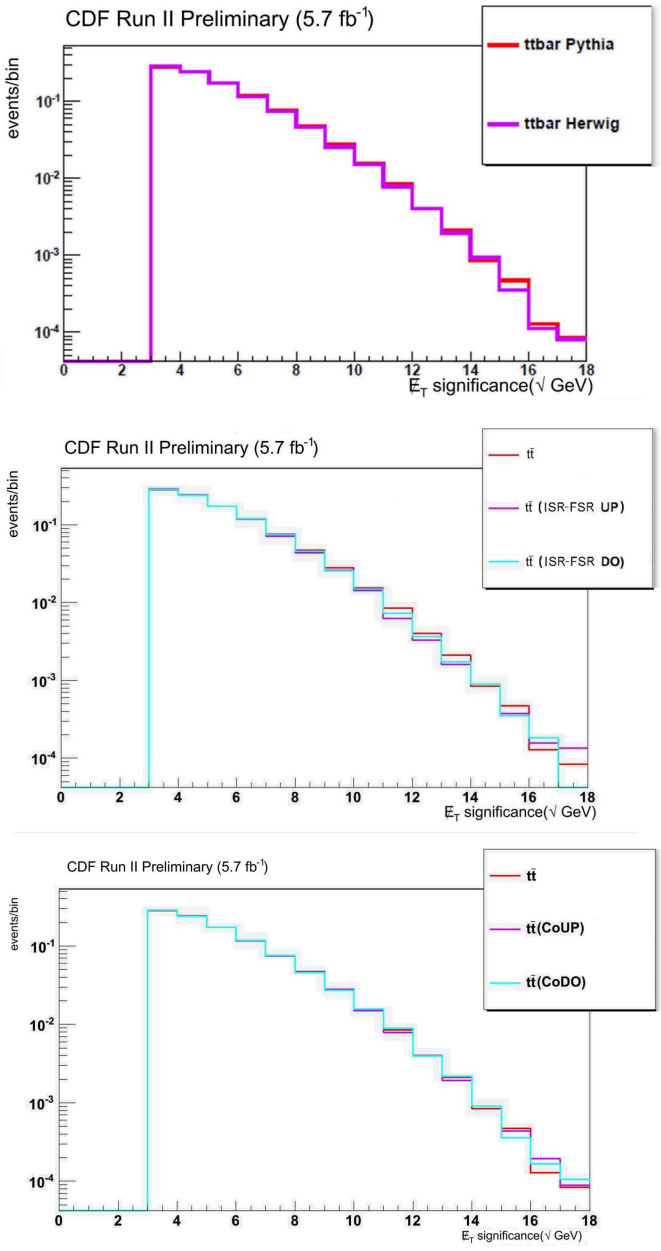


Figure 7.1: Comparisons of the $\not\!\!E_T$ significance distributions of the $t\bar{t}$ (top) for different MC generators, (center) varying initial and final state radiation effect by $\pm 1\sigma$ and (bottom) varying color reconnection effect by $\pm 1\sigma$. All distributions are normalized to the unit area.

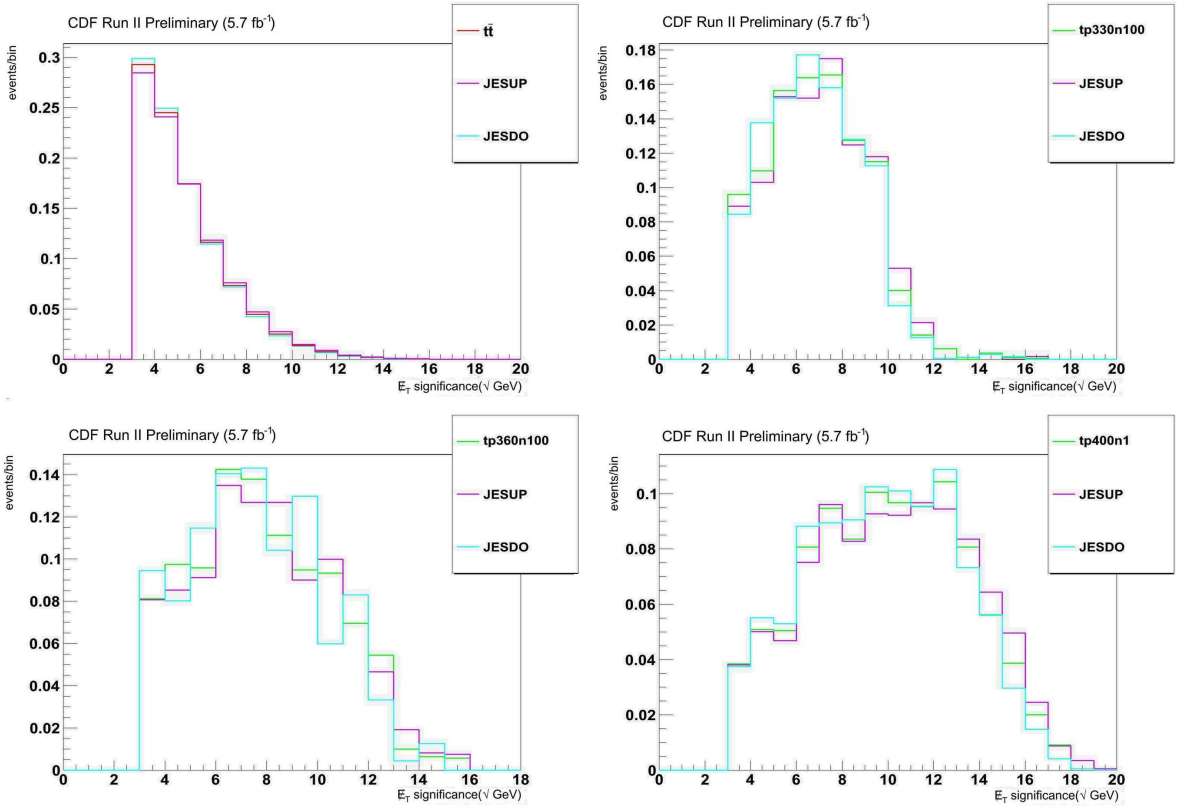


Figure 7.2: Comparisons of the $\not\!\!E_T$ significance distributions for $t\bar{t}$ (top left) and three different signals (top right: ($m_{T'} = 330$ GeV/c^2, $m_X = 100$ GeV/c^2); bottom left: ($m_{T'} = 360$ GeV/c^2, $m_X = 100$ GeV/c^2); bottom right: ($m_{T'} = 400$ GeV/c^2, $m_X = 1$ GeV/c^2) varying JES by $\pm 1\sigma$. All distributions are normalized to the unit area.

7.4 Results

To obtain the upper limit s_u, the marginalized posterior is integrated over s using numerical integration. To obtain expected limits, the procedure is repeated N times, generating a new ensemble for each repetition, and averaged over n (N = 10000 for the results presented in this analysis). We perform this procedure for a grid of points in the $(m_{T'}, m_X)$ plane. The expected and observed results are shown in Tab. 7.2 for representative signal points. We convert the observed upper limits on the pair-production cross sections to an exclusion curve in mass parameters space $(m_{T'}, m_X)$; the results are shown in Fig. 7.3. We exclude fourth generation T' quarks decaying into t quarks and undetected particles X for $m_{T'} \leq 400$ GeV/c^2 and $m_{T'} \leq 70$ GeV/c^2, thus significantly extending limits obtained from the previous analysis in the semileptonic channel. As expected, the largest sensitivity enhancement is reached for high $m_{T'}$ and medium-low m_X, which results in higher momentum carried from the T' decay products, and thus in higher $\not\!\!E_T$.

Table 7.2: Expected 95% C.L. upper limits on $T'T'$ production, where the uncertainty range covers 68% of the pseudoexperiments, and observed exclusion limits for representative signal points.

(mT', mX)GeV/c^2	σ 95% C.L. exp. exclusion (pb)	σ 95% C.L. obs. exclusion (pb)
(200,40)	2.02 ± 0.65	1.90
(220,40)	2.14 ± 0.75	3.00
(260,1)	0.23 ± 0.08	0.18
(280,1)	0.15 ± 0.05	0.12
(280,40)	0.18 ± 0.07	0.15
(300,1)	0.09 ± 0.03	0.09
(300,80)	0.20 ± 0.06	0.16
(300,100)	0.29 ± 0.09	0.38
(330,1)	0.05 ± 0.02	0.03
(330,100)	0.13 ± 0.04	0.18
(360,1)	0.03 ± 0.01	0.02
(360,100)	0.06 ± 0.02	0.04
(380,100)	0.06 ± 0.02	0.05
(400,1)	0.023 ± 0.008	0.016

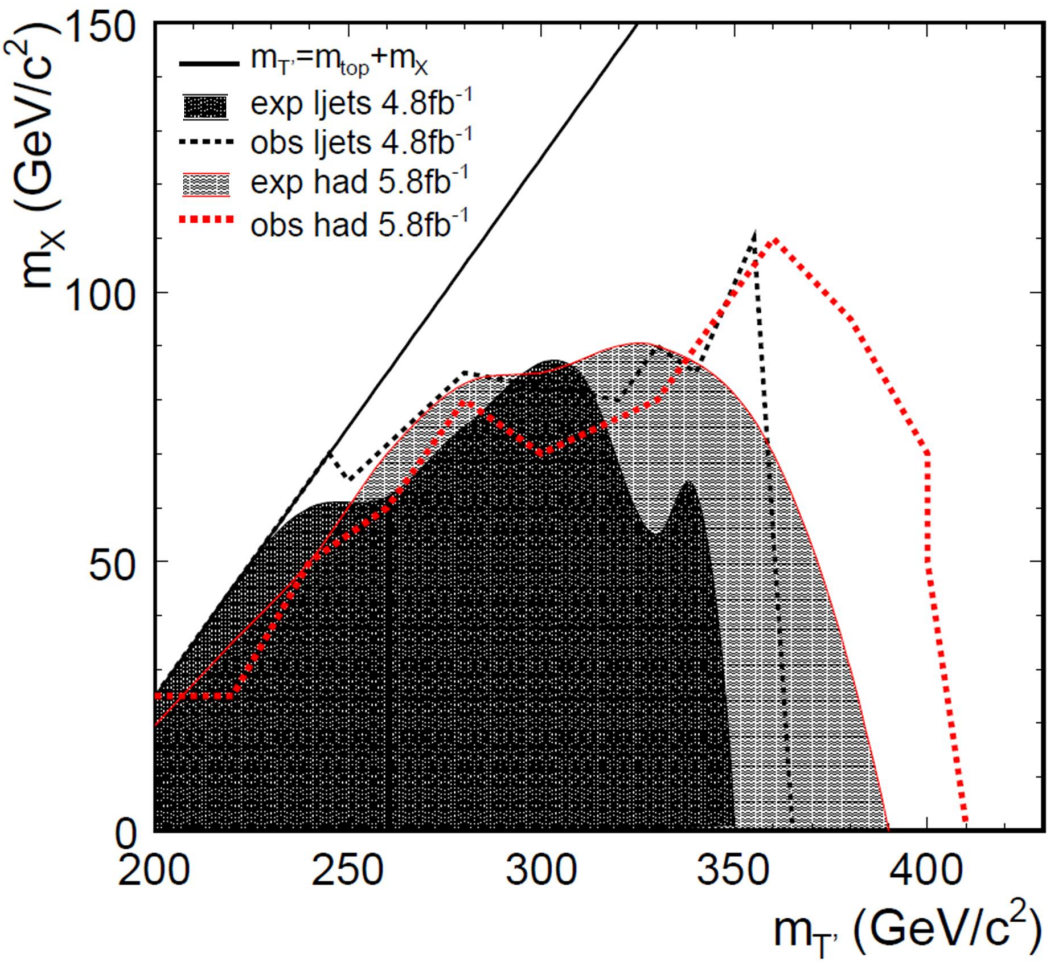


Figure 7.3: Expected (exp) and observed (obs) 95% C.L. exclusion region in the $(m_{T'}, m_X)$ parameters space. Results from previous search in semileptonic channel are also shown.

Conclusions

In summary, we have performed a direct search for a massive fourth generation T' quark decaying into top quarks and dark matter using data corresponding to an integrated luminosity of 5.7 fb^{-1} accumulated in Run II of the Tevatron using the CDF II detector. We performed the first search of these kind of processes in the full-hadronic channel. The observed events' $\not\!\!E_T$ significance show no significant excess above SM predictions, and 95% C.L. upper limits are placed on the T' production cross-section. A novel data-driven method to determine the multi-jet background and a novel event selection for searches of these kind of processes in the hadronic channel have been developed. As a result of this original work we have extended the previous limits set by searches in the semileptonic channel, despite of difficulties in dealing with QCD multijet background. The results, reported in [89], have recently been accepted for publication in *Physical Review Letters*. After the completion of this work, the ATLAS collaboration has performed a similar analysis in the semileptonic channel [53]. The new limits are shown in Fig. 7.4.

This study has demonstrated the higher sensitivity of the hadronic channel in this kind of searches. The improvement of the techniques described in this work, and their adaptation to the LHC environment, togheter with the iincrease in the collected data sample, will further contribute to the exploration of new physics scenarios, first of all the production of supersymmetric scalar top decaying into top quarks and neutralinos.

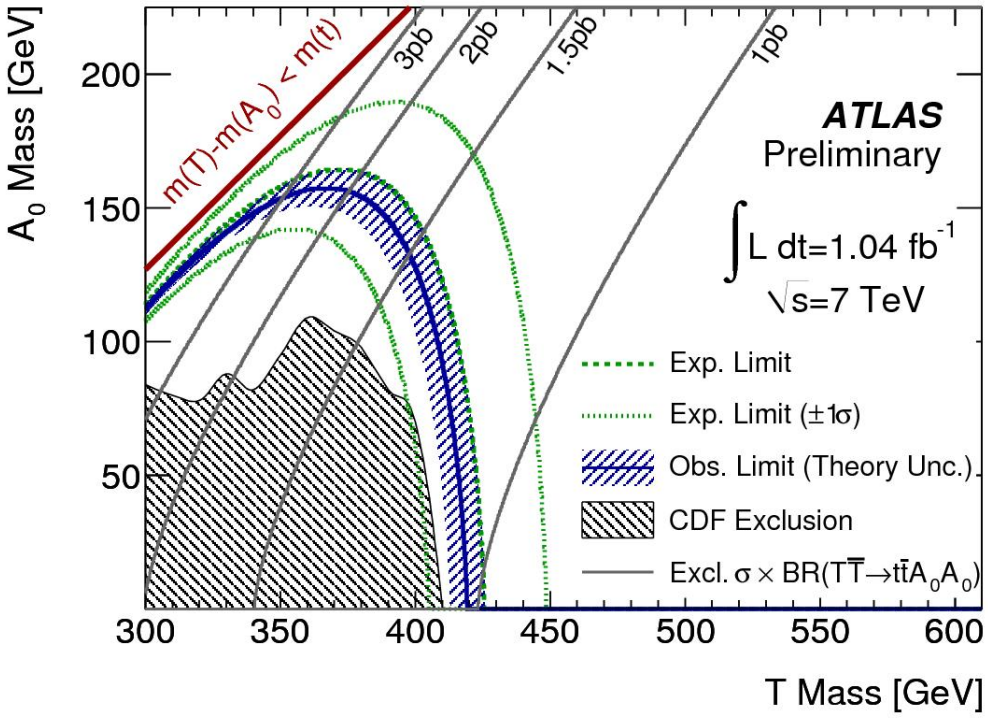


Figure 7.4: Limits set in the $(m_{T'}, m_X)$ plane from the ATLAS analysis. The CDF exclusion is the result of the present analysis. From [53].

Bibliography

[1] F. Zwicky, Helv. Phys. Acta 6 (1933) 110

[2] S. Smith, Astrophys. J. 83 (1936) 23

[3] H.W. Babcock, 1939, Lick Obs. Bull., 19, 41 (No. 498)

[4] R. W. Peebles, The large-scale structure of the universe (Princeton University, Princeton, N.J.) 1980.

[5] A. A. Penzias, R. W. Wilson, A Measurement of Excess Antenna Temperature at 4080 Mc/s, Astrophysical Journal Letters (1965).

[6] P. de Bernardis et al. [Boomerang Collaboration], Nature (2000) 404, 955

[7] D. N. Spergel *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **148** (2003) 175-194. [astro-ph/0302209].

[8] D. N. Spergel *et al* 2007, ApJS, 170, 377

[9] A.K. Drukier, K. Freese, D.N. Spergel, Phys. Rev. D33, 3495 (1986)

[10] K. Freese, J. Frieman, A. Gould, Phys. Rev. D37, 3388 (1988)

[11] R. Bernabei *et al.*, arXiv:1007.0595 [astro-ph.CO].

[12] C. E. Aalseth *et al.*, arXiv:1106.0650 [astro-ph.CO].

[13] M. T. Frandsen, F. Kahlhoefer, J. March-Russell, C. McCabe, M. McCullough and K. Schmidt-Hoberg, arXiv:1105.3734 [hep-ph].

[14] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **74** (1995) 2626 [arXiv:hep-ex/9503002].

[15] S. Abachi *et al.* [D0 Collaboration], Phys. Rev. Lett. **74**, 2632 (1995) [arXiv:hep-ex/9503003].

[16] TEVWG and CDF and D0 Collaborations, [arXiv:1107.3178]; A. B. Galtieri, F. Margaroli and I. Volobuev [arXiv:1109.2163 [hep-ex]].

[17] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. D **83** (2011) 112003 [arXiv:1101.0034 [hep-ex]].

[18] T. Aaltonen et al. [CDF Collaboration], public note no. 10395, March 2011.

[19] F. Deliot and D. A. Glenzinski, arXiv:1010.1202 [hep-ex].

[20] G. Isidori and J. F. Kamenik, Phys. Lett. B **700** (2011) 145 [arXiv:1103.0016 [hep-ph]].

[21] E. Aprile *et al.* [XENON100 Collaboration], arXiv:1104.2549 [astro-ph.CO].

[22] E. Armengaud *et al.* [EDELWEISS Collaboration], Phys. Lett. B **702** (2011) 329 [arXiv:1103.4070 [astro-ph.CO]].

[23] G. Angloher *et al.*, arXiv:1109.0702 [astro-ph.CO].

[24] H. J. He, N. Polonsky and S. f. Su, Phys. Rev. D 64, 053004 (2001).

[25] G. D. Kribs, T. Plehn, M. Spannowsky and T. M. P. Tait, Phys. Rev. D 76, 075016 (2007).

[26] B. Holdom, W. S. Hou, T. Hurth, M. L. Mangano, S. Sultansoy and G. Unel, PMC Phys. A **3** (2009) 4 [arXiv:0904.4698 [hep-ph]].

[27] J. Alwall, J. L. Feng, J. Kumar and S. Su, Phys. Rev. D **81** (2010) 114027 [arXiv:1002.3366 [hep-ph]].

[28] J. L. Feng and J. Kumar, Phys. Rev. Lett. **101** (2008) 231301 [arXiv:0803.4196 [hep-ph]].

[29] J. L. Feng, J. Kumar and L. E. Strigari, Phys. Lett. B **670** (2008) 37 [arXiv:0806.3746 [hep-ph]].

[30] A. G. Ivanov [CDF Collaboration], arXiv:0811.0788 [hep-ex].

[31] M. Cacciari, private communication.

[32] Prospino2, http://www.thphys.uni-heidelberg.de/ plehn/prospino/.

[33] H. C. Cheng, J. L. Feng and K. T. Matchev, Phys. Rev. Lett. **89**, 211301 (2002) [arXiv:hep-ph/0207125].

[34] L. Randall and R. Sundrum, Phys. Rev. Lett. **83** (1999) 3370 [arXiv:hep-ph/9905221].

[35] K. Agashe, A. Falkowski, I. Low and G. Servant, JHEP **0804** (2008) 027 [arXiv:0712.2455 [hep-ph]].

[36] H. C. Cheng and I. Low, JHEP **0408** (2004) 061 [arXiv:hep-ph/0405243].

[37] S. R. Elliott, P. Vogel, Ann. Rev. Nucl. Part. Sci. **52**, 115-151 (2002). [arXiv:hep-ph/0202264 [hep-ph]].

[38] P. Nath and P. Fileviez Perez, Phys. Rept. **441**, 191 (2007) [arXiv:hep-ph/0601023].

[39] P. Fileviez Perez and M. B. Wise, Phys. Rev. D **82** (2010) 011901 [Erratum-ibid. D **82** (2010) 079901] [arXiv:1002.1754 [hep-ph]].

[40] L. Balka *et al.* [CDF Collaboration], Nucl. Instrum. Meth. A **267**, 272 (1988).

[41] S. Bertolucci *et al.* [CDF Collaboration], Nucl. Instrum. Meth. A **267** (1988) 301.

[42] S. Aota *et al.*, Nucl. Instrum. Meth. A **352**, 557 (1995).

[43] G. Apollinari, K. Goulianos, P. Melese and M. Lindgren, Nucl. Instrum. Meth. A **412** (1998) 515.

[44] J. Elias *et al.*, Nucl. Instrum. Meth. A **441**, 366 (2000).

[45] S. Klimenko, J. Konigsberg and T. M. Liss, *"Averaging of the inelastic cross sections measured by the CDF and the E811 experiments"*, FERMILAB-FN-0741, December 2003.

[46] F.D. Snider. Tracking at CDF: Algorithms and experience from Run I and Run II. Nucl. Instrum. Meth. A, 566(1):133 - 141, 2006.

[47] C. Hays et al. Inside-out tracking at CDF. Nucl. Instrum. Meth. A, 538(1-3):249 - 254, 2005.

[48] U. Spreitzer *et al.*, Electron Identification in Offline Release 6.1.2, CDF Note 7950, CDF/DOC/ELECTRON/CDFR/7950

[49] U. Grundler *et al.*, High-p_T muons reccomended cuts and efficiencies for Summer 2006, CDF Note 8262, CDF/ANAL/TOP/CDFR/8262

[50] F. Abe *et al. Phys. Rev.*, D45:14481458, 1992.

[51] F. Abe *et al.* [CDF Collaboration], Phys. Rev. D **45**, 1448 (1992).

[52] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. Lett. **106** (2011) 191801 [arXiv:1103.2482 [hep-ex]].

[53] G. Aad *et al.* [ATLAS Collaboration], arXiv:1109.4725 [hep-ex].

[54] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, JHEP **0809** (2008) 127 [arXiv:0804.2800 [hep-ph]].

[55] H. L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin and C. P. Yuan, Phys. Rev. D **82** (2010) 074024 [arXiv:1007.2241 [hep-ph]].

[56] T. Hebbeker, "What to do with multijet events?", D0 Note 4465

[57] S.Amerio *et al.*, "Study of a MET_DIJET trigger for WH to $l\nu bb$", CDF Note 8826

[58] M. Rossi et al., "Studies of the new MET_DIJET acceptance for WH to $l\nu bb$ and overlap with currently used triggers", CDF Note 9178

[59] O. Gonzalez *et al.*, "Using Neural Network Parametrization of the Trigger Efficiency for the MET+JET Sample", CDF Note 10538, CDF/PHYS/MISSING_ET/CDFR/10538.

[60] J. Alwall et al. J. High Energy Phys. 0709 (2007) 028.

[61] T. Sjostrand et al., Comput. Phys. Commun. 238 135 (2001), version 6.422.

[62] S. Moch and P. Uwer, Nucl. Phys. Proc. Suppl. **183** (2008) 75 [arXiv:0807.2794 [hep-ph]].

[63] M. Mangano et al., J. High Energy Phys. 0307 (2003) 001.

[64] M. L. Mangano, M. Moretti, F. Piccinini and M. Treccani, J. High Energy Phys. 0701 (2007) 013, arXiv:hep-ph/0611129.

[65] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (1999).

[66] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. Lett. **100**, 211801 (2008).

[67] Z. Sullivan, Phys. Rev. D 70, 114012 (2004).

[68] J. M. Campbell, R. Frederix, F. Maltoni and F. Tramontano, Phys. Rev. Lett. **102** (2009) 182003 [arXiv:0903.0005 [hep-ph]].

[69] R. Brun *et al.* GEANT 3 manual, CERN Program Library Long Writeup. 1994.

[70] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. Lett. **104**, 141801 (2010) [arXiv:0911.3935 [hep-ex]].

[71] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. Lett. **103** (2009) 092002 [arXiv:0903.0885 [hep-ex]].

[72] T. Aaltonen *et al.* [The CDF Collaboration], Phys. Rev. D **81** (2010) 072003 [arXiv:1001.4577 [hep-ex]].

[73] V. M. Abazov *et al.* [D0 Collaboration], Phys. Rev. Lett. **104** (2010) 071801 [arXiv:0912.5285 [hep-ex]].

[74] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. Lett. **103**, 091803 (2009) [arXiv:0905.4714 [hep-ex]].

[75] Q. Liu, F. Margaroli [The CDF collaboration], Everything you wanted to know on the missing transverse momentum CDF note 9843 (2009).

[76] A. L. Read. Presentation of search results: the CLs technique. Journal of Physics G, 28(10):2693-2704, 2002.

[77] J. Heinrich et al. Interval estimation in the presence of nuisance parameters. 1. Bayesian approach. arXiv:physics/0409129, 2004.

[78] Thomas Junk. Confidence level computation for combining searches with small statistics. Nucl. Instrum. Meth. A, 434(2-3):435 - 443, 1999.

[79] J. Heinrich. Bayesian limit software: multi-channel with correlated backgrounds and efficiencies. CDF Public Note 7587.

[80] D. Acosta *et al.*, Nucl. Instrum. Methods A **494**, 57 (2002).

[81] D. Acosta et al. The performance of the CDF luminosity monitor. Nucl. Instrum. Meth. A, 494(1-3):57 - 62, 2002.

[82] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **74**, 032008 (2006), D. Acosta *et al.* (CDF Collaboration), Phys. Rev. Lett. **94**, 091803 (2005).

[83] J. Pumplin *et al.*, J. High Energy Phys. 07 (2002) 012.

[84] B. Andersson *et al.*, Phys. Rep.**97**, 31 (1983).

[85] T. Sjö strand, Comput. Phys. Commun. **39**, 347 (1986).

[86] B. R. Webber, Nucl. Phys. **B238**, 492 (1984).

[87] P. Z. Skands, arXiv:0905.3418 [hep-ph].

[88] A. Bhatti *et al.*, Nucl. Instrum. Methods A **566** (2006) 375; arXiv:hep-ex/0510047.

[89] T. Aaltonen *et al.* [CDF Collaboration], arXiv:1107.3574 [hep-ex].

Appendix A

Kinematical Distributions in Signal Region

Comparisons of the background model with the observed data in signal regions was shown in Fig. 6.10. In this appendix we show additional distributions of kinematic variables in the signal region. The QCD contribution is normalized to the number of observed events minus the $t\bar{t}$, W/Z+jets, diboson and single top expected contributions. Cuts defining signal region, including preselection requirements, are summarized in Tab. A.1.

h

Table A.1: Selection cuts defining the signal region.

Signal Region
$\not\!\!E_T > 50$ GeV
$N_{jets}(E_T > 30$ GeV, $\lvert\eta\rvert < 2.4) \geq 2$
$N_{jets}(E_T > 20$ GeV, $\lvert\eta\rvert < 2.4) \geq 5$

$\Delta\phi(\vec{\not\!\!E}_T, \vec{\not\!\!p}_T) < \pi/2$
$\Delta\phi(\vec{\not\!\!E}_T, \vec{J}_i) > 0.4$ (for i = 1,2,3)
$\Delta\phi(\vec{\not\!\!E}_T, \vec{J}_i) > 0.2$ (for i = 4,5)
$H_T > 220$ GeV
$\not\!\!p_T > 20$ GeV
$\not\!\!E_T$ significance > 3 GeV$^{\frac{1}{2}}$
$N_{jets}(E_T > 15$ GeV, $\lvert\eta\rvert < 2.4) \leq 10$

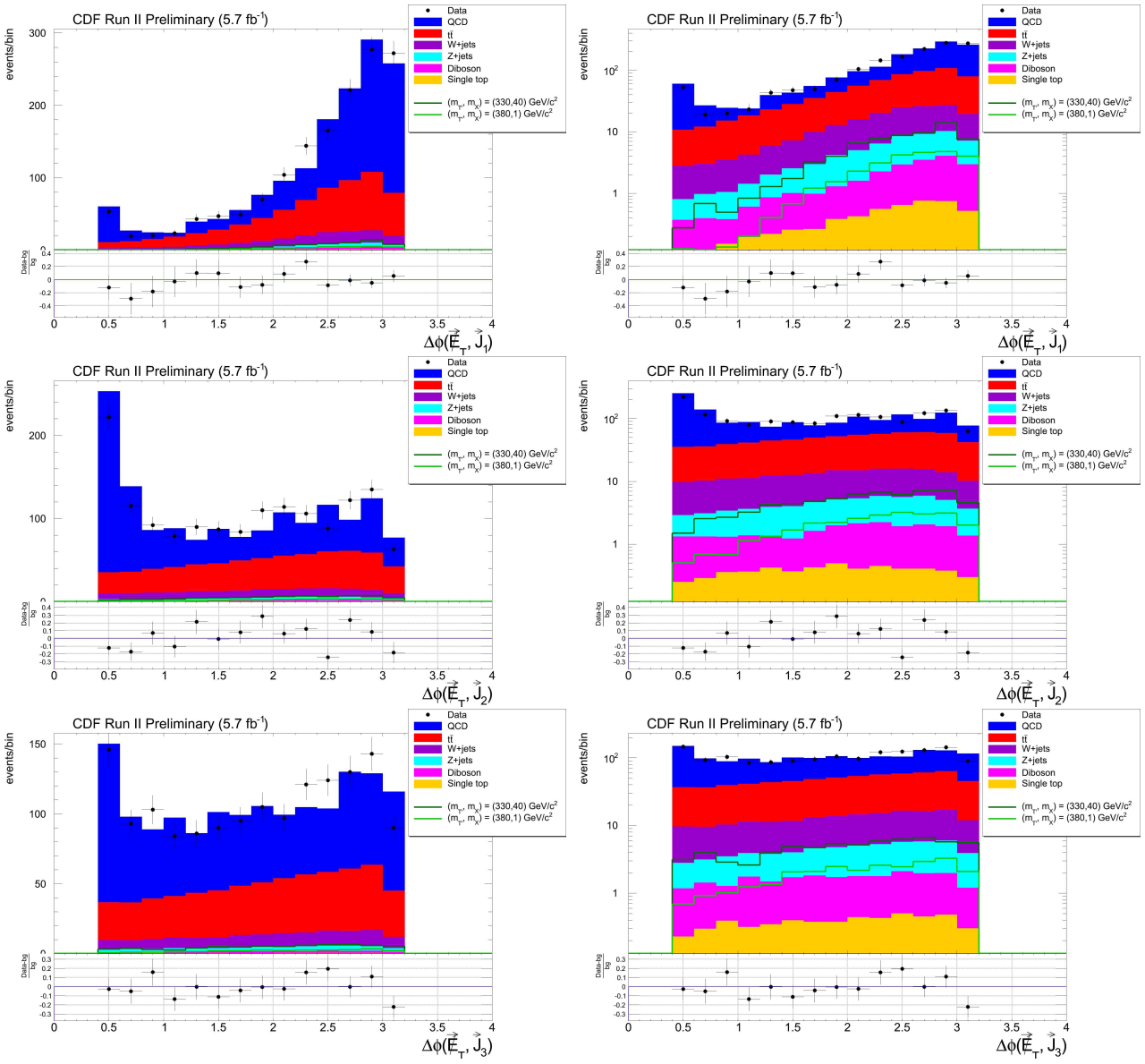


Figure A.1: Distributions of azimuthal distances between $\not\!\!E_T$ and first three jets in the signal region for data, expected backgrounds and two signal hypothesis, in linear and logarithmic scale. Each process is normalized to the number of expected events.

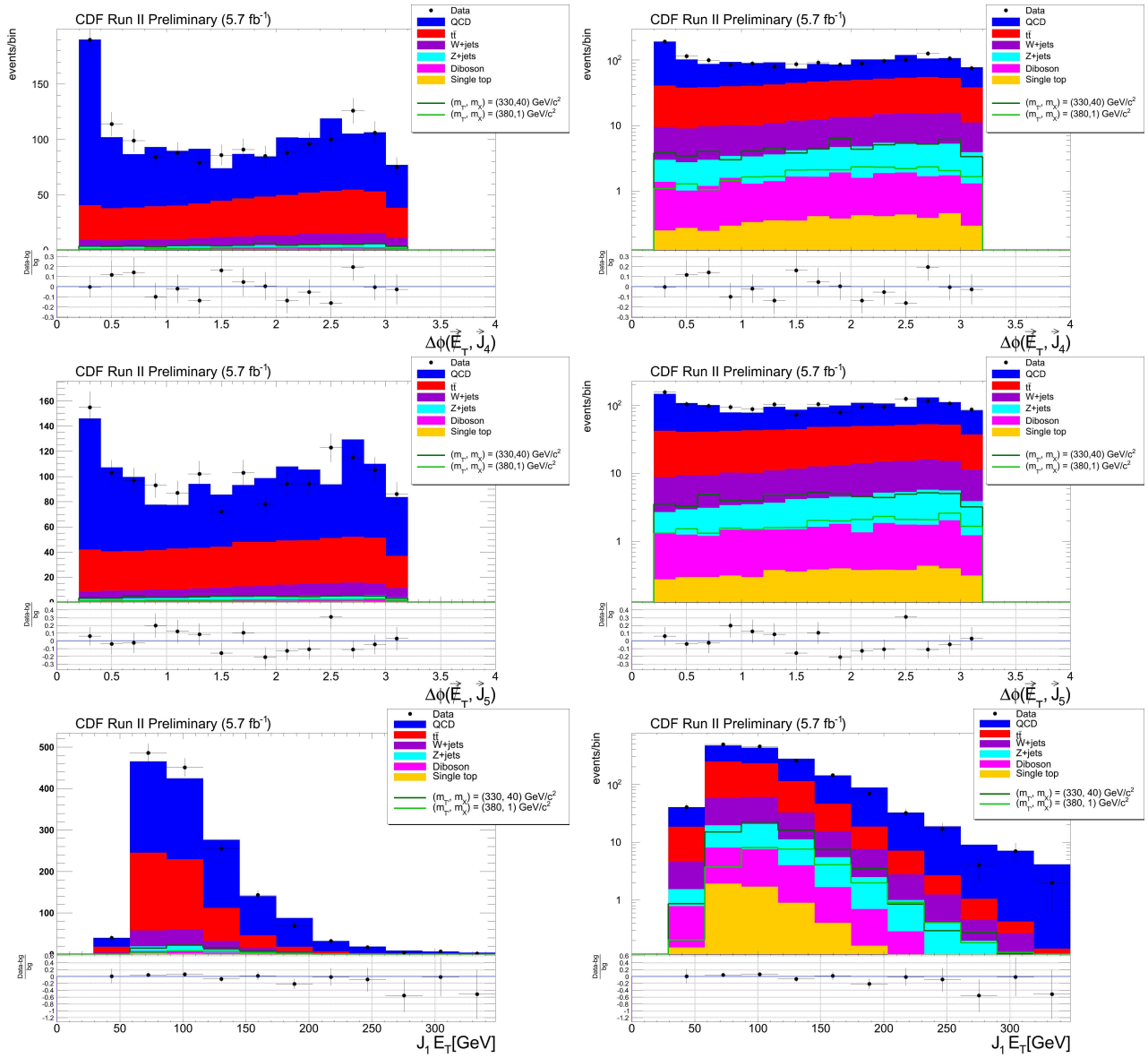


Figure A.2: Distributions of azimuthal distance between $\not\!\!E_T$ and fourth jet, between $\not\!\!E_T$ and fifth jet, and of first jet energy in the signal region for data, expected backgrounds and two signal hypothesis, in linear and logarithmic scale. Each process is normalized to the number of expected events.

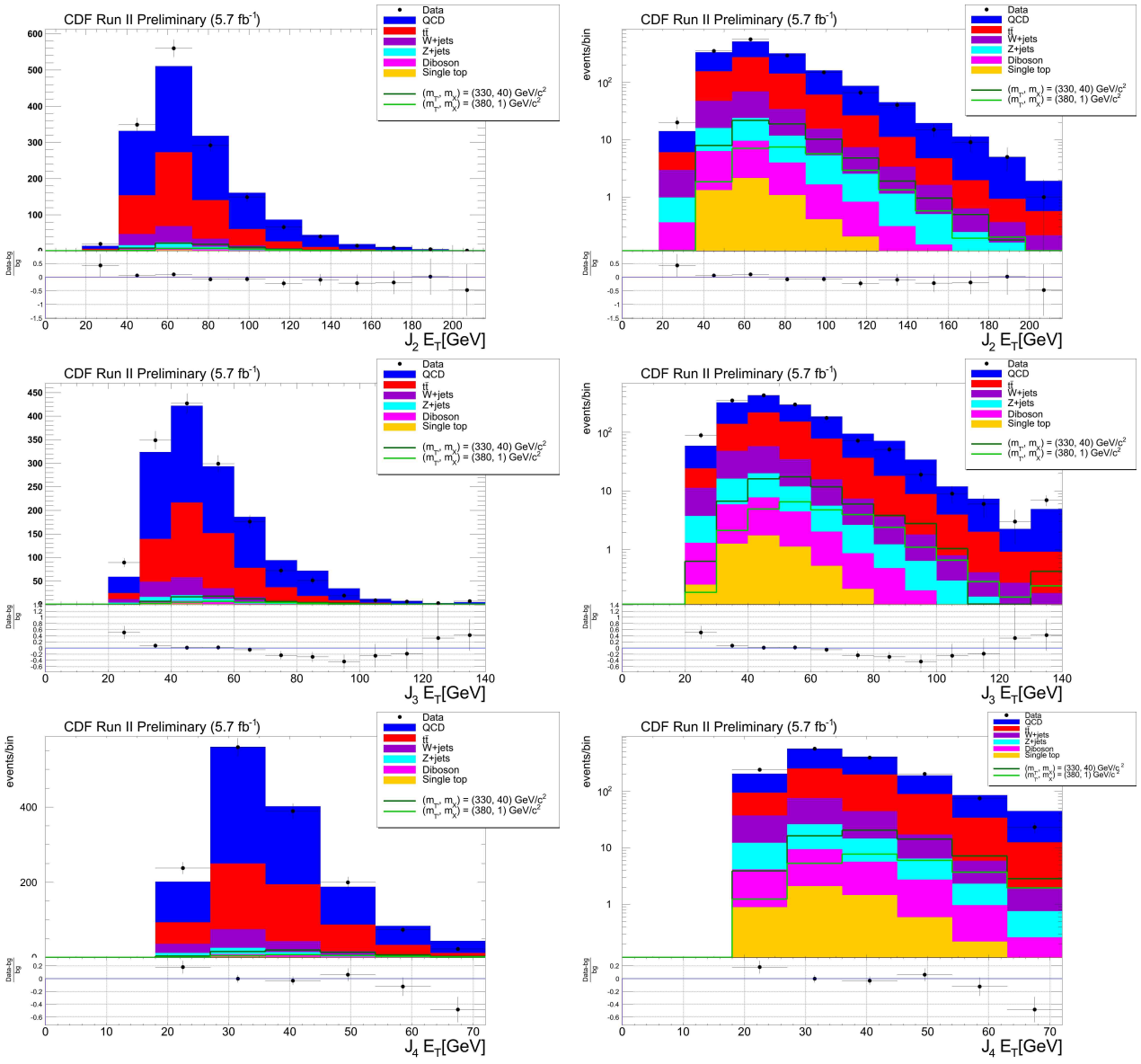


Figure A.3: Distributions of second, third and fourth jet's energies in the signal region for data, expected backgrounds and two signal hypothesis, in linear and logarithmic scale. Each process is normalized to the number of expected events.

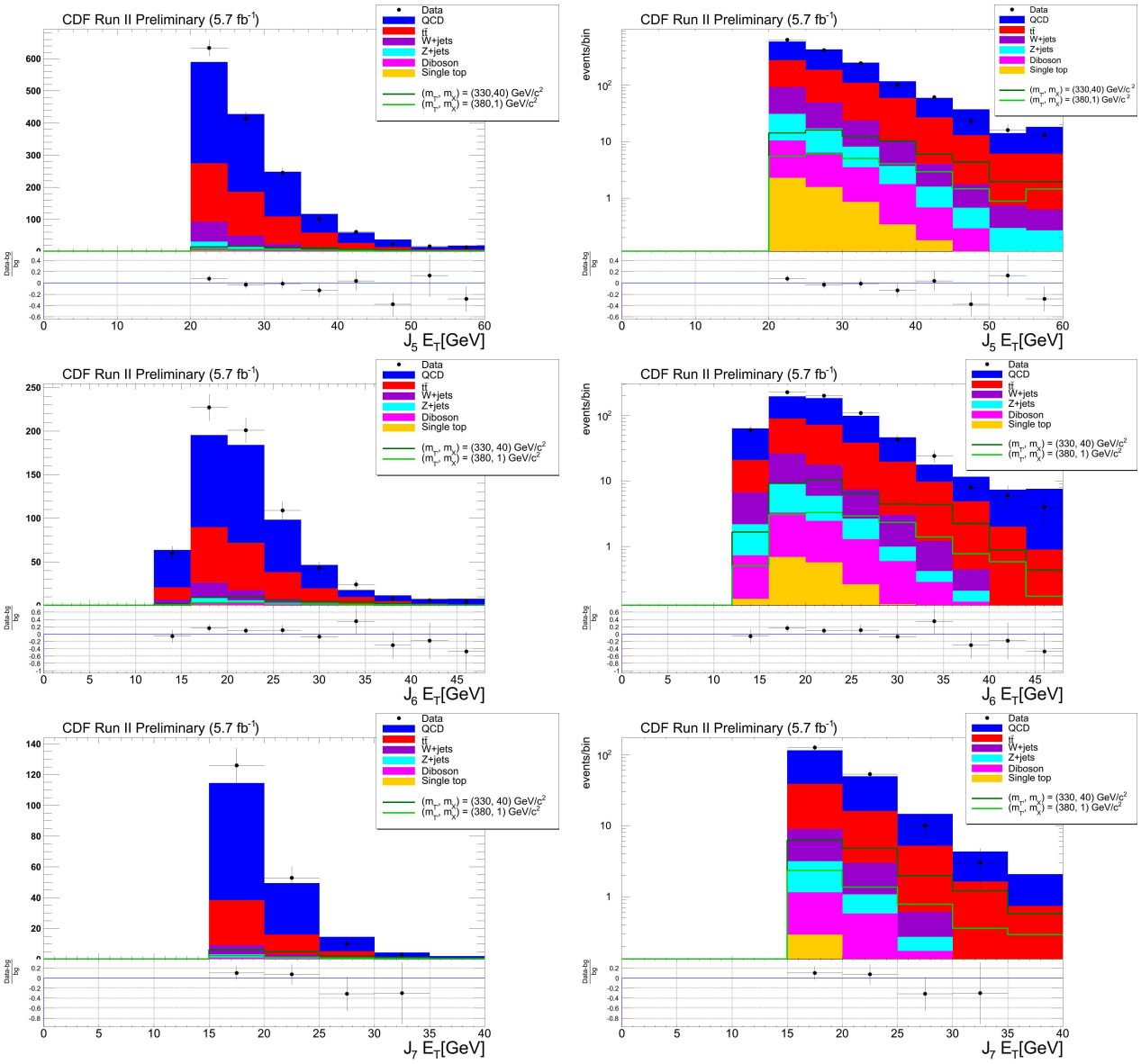


Figure A.4: Distributions of fifth, sisth and seventh jet's energies in the signal region for data, expected backgrounds and two signal hypothesis, in linear and logarithmic scale. Each process is normalized to the number of expected events.

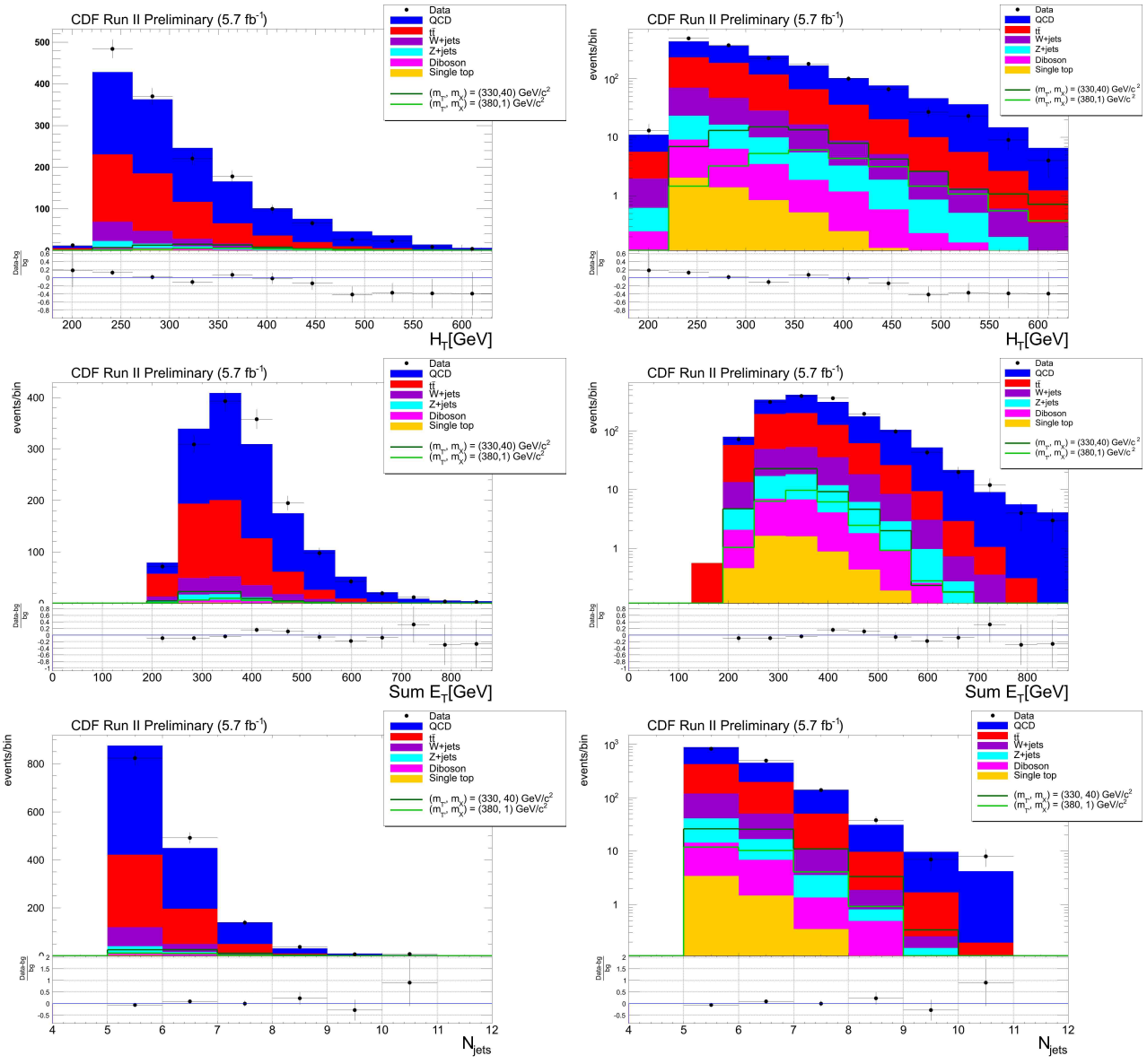


Figure A.5: Distribution of H_T, $SumE_T$ and number of jets in the signal region for data, expected backgrounds and two signal hypothesis, in linear and logarithmic scale. Each process is normalized to the number of expected events.

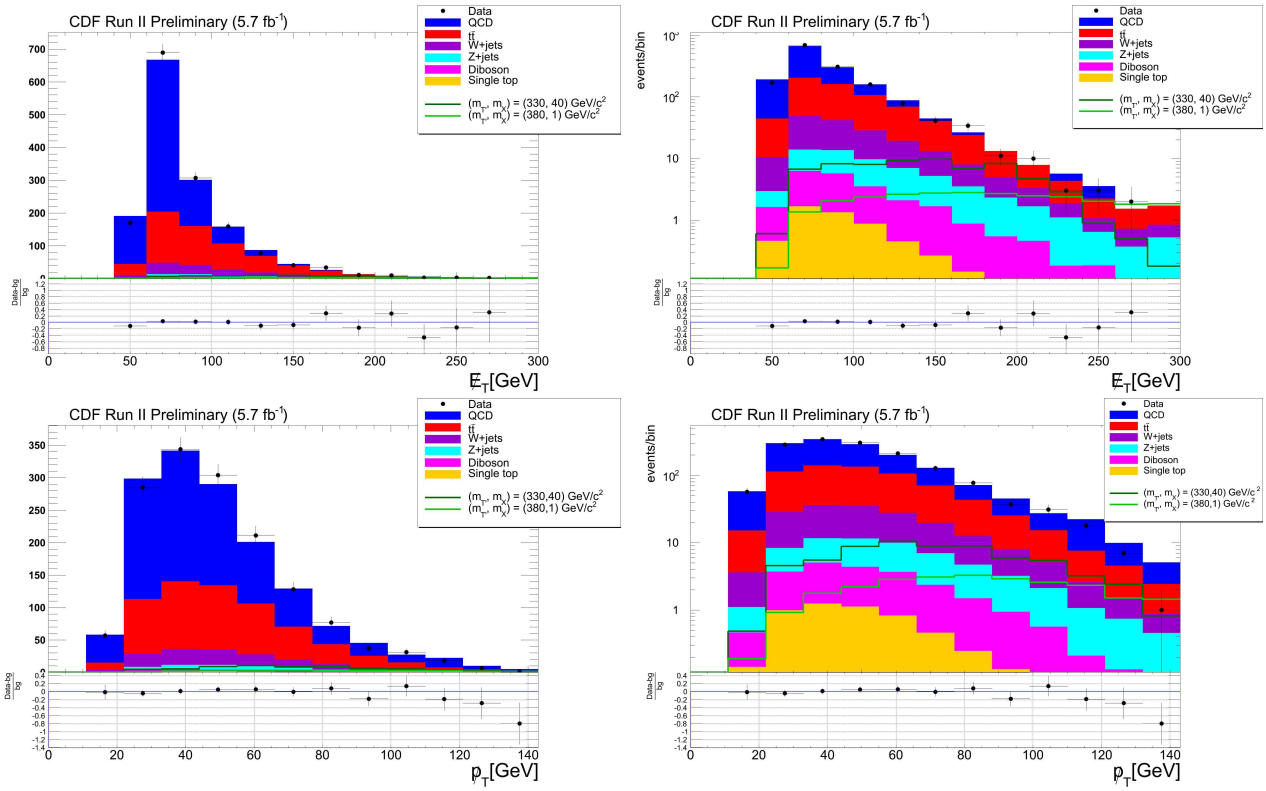


Figure A.6: Distribution of $\not\!\!E_T$ and $\not\!\!p_T$ in the signal region for data, expected backgrounds and two signal hypothesis, in linear and logarithmic scale. Each process is normalized to the number of expected events.

Appendix B

Kinematical Distributions in Control Regions

We utilize control regions to test our ability to model the backgrounds. Two control regions were defined in chpt. 6. The modeling of different types of physics processes is tested in these control regions, allowing us to validate our model of the backgrounds that are expected to contribute to the signal region. Comparisons of the background model with the observed data in control regions were shown in Figures 6.9. In this appendix we show additional distributions of kinematic variables in the control regions.

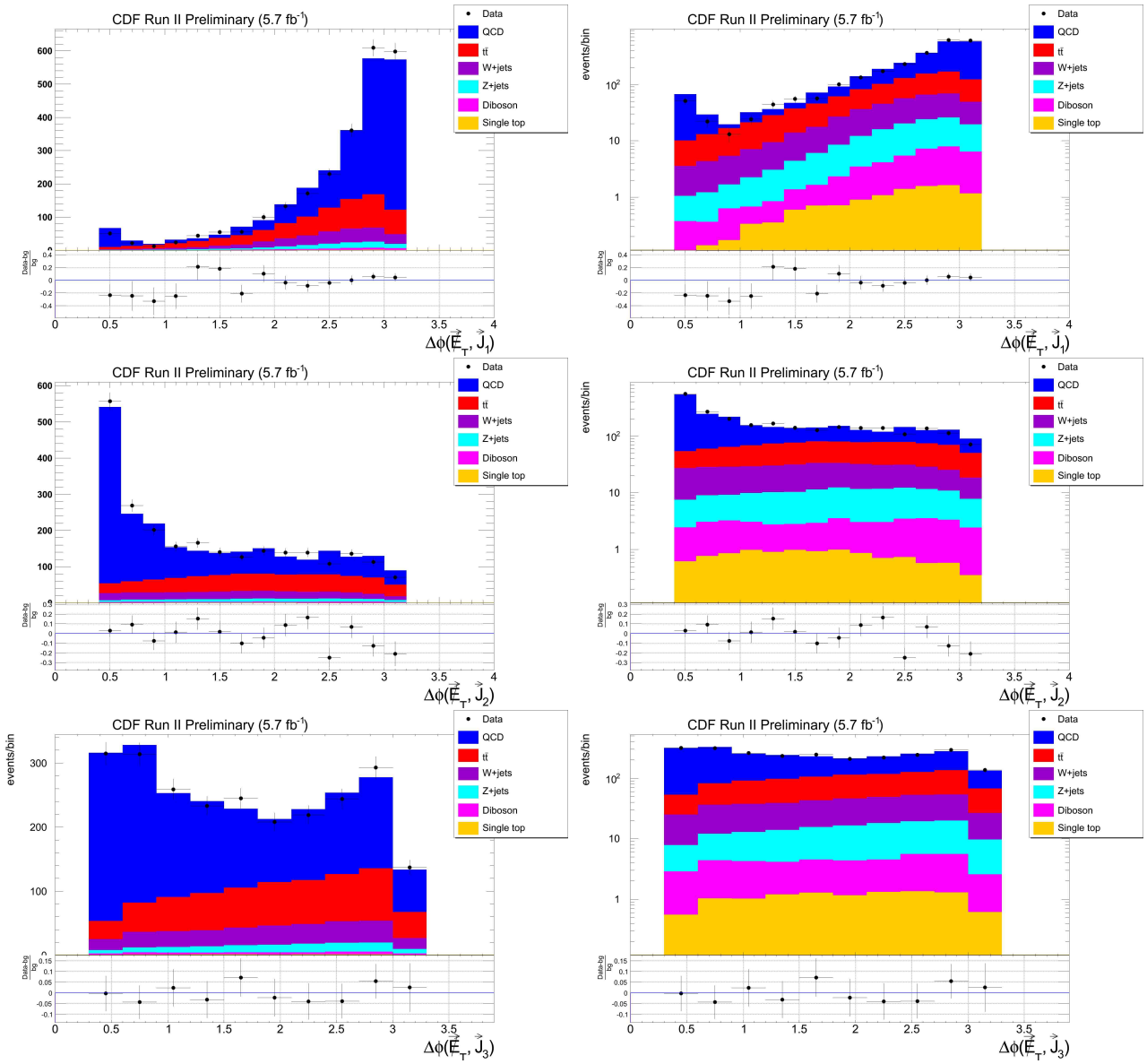


Figure B.1: Distributions of azimuthal distances between $\not\!\!E_T$ and first three jets in the $N_{jets} = 4$ control region for data and expected backgrounds, in linear and logarithmic scale. Each process is normalized to the number of expected events.

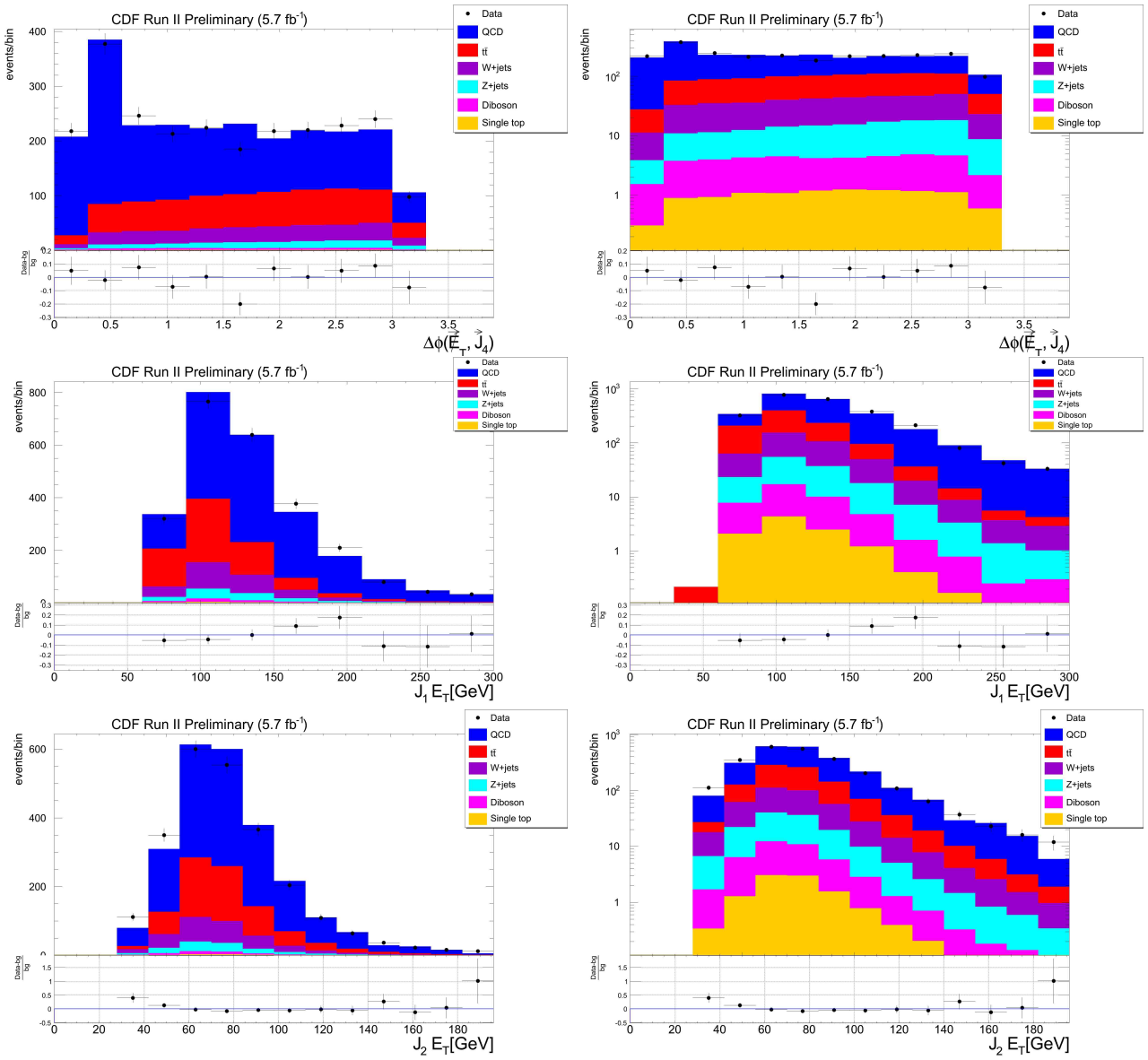


Figure B.2: Distributions of azimuthal distance between $\not\!\!E_T$ and fourth jet, and of first two jets' energies in the $N_{jets} = 4$ control region for data and expected backgrounds, in linear and logarithmic scale. Each process is normalized to the number of expected events.

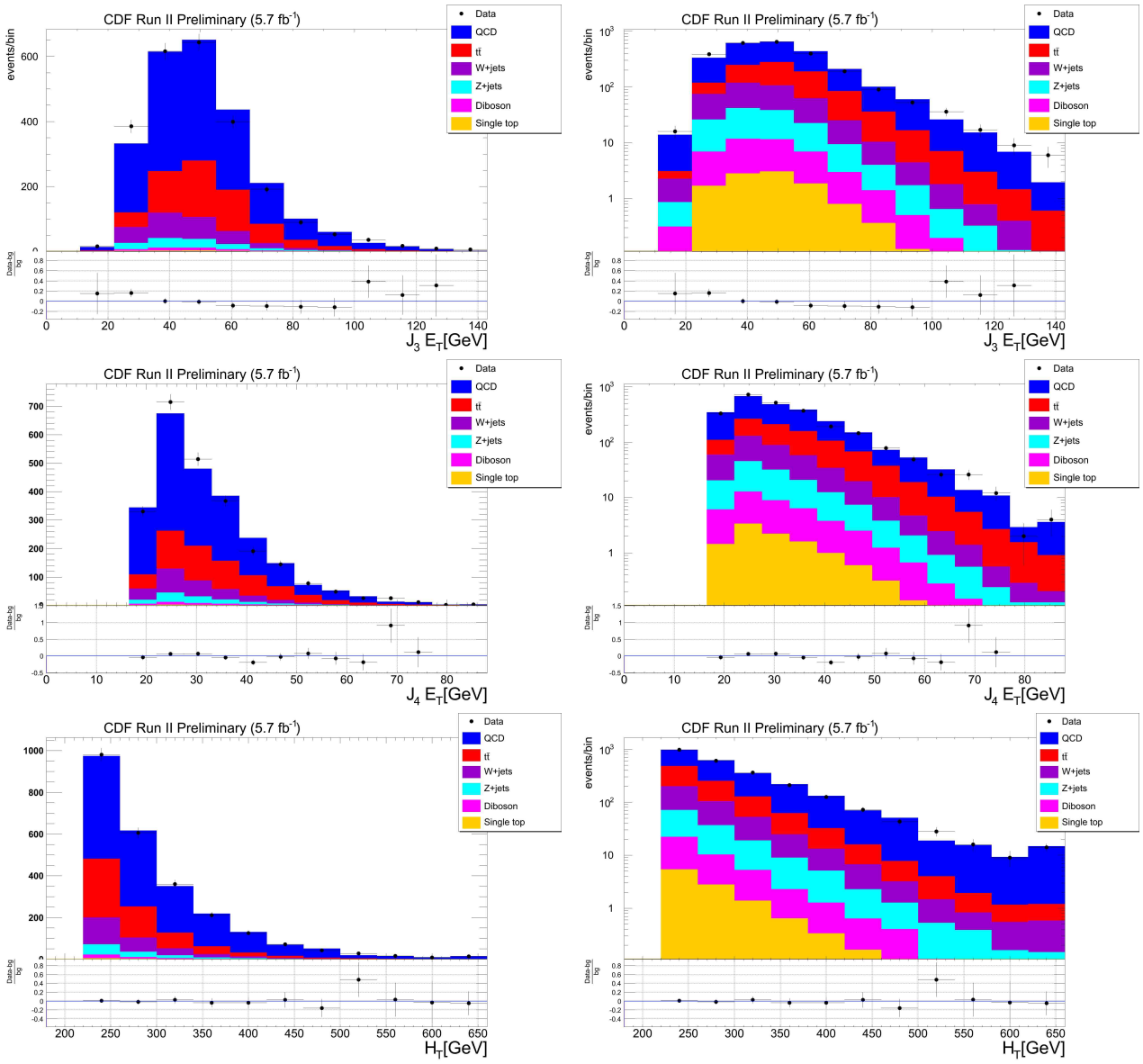


Figure B.3: Distribution of third and fourth jet's energies, and of H_T in the $N_{jets} = 4$ control region for data and expected backgrounds, in linear and logarithmic scale. Each process is normalized to the number of expected events.

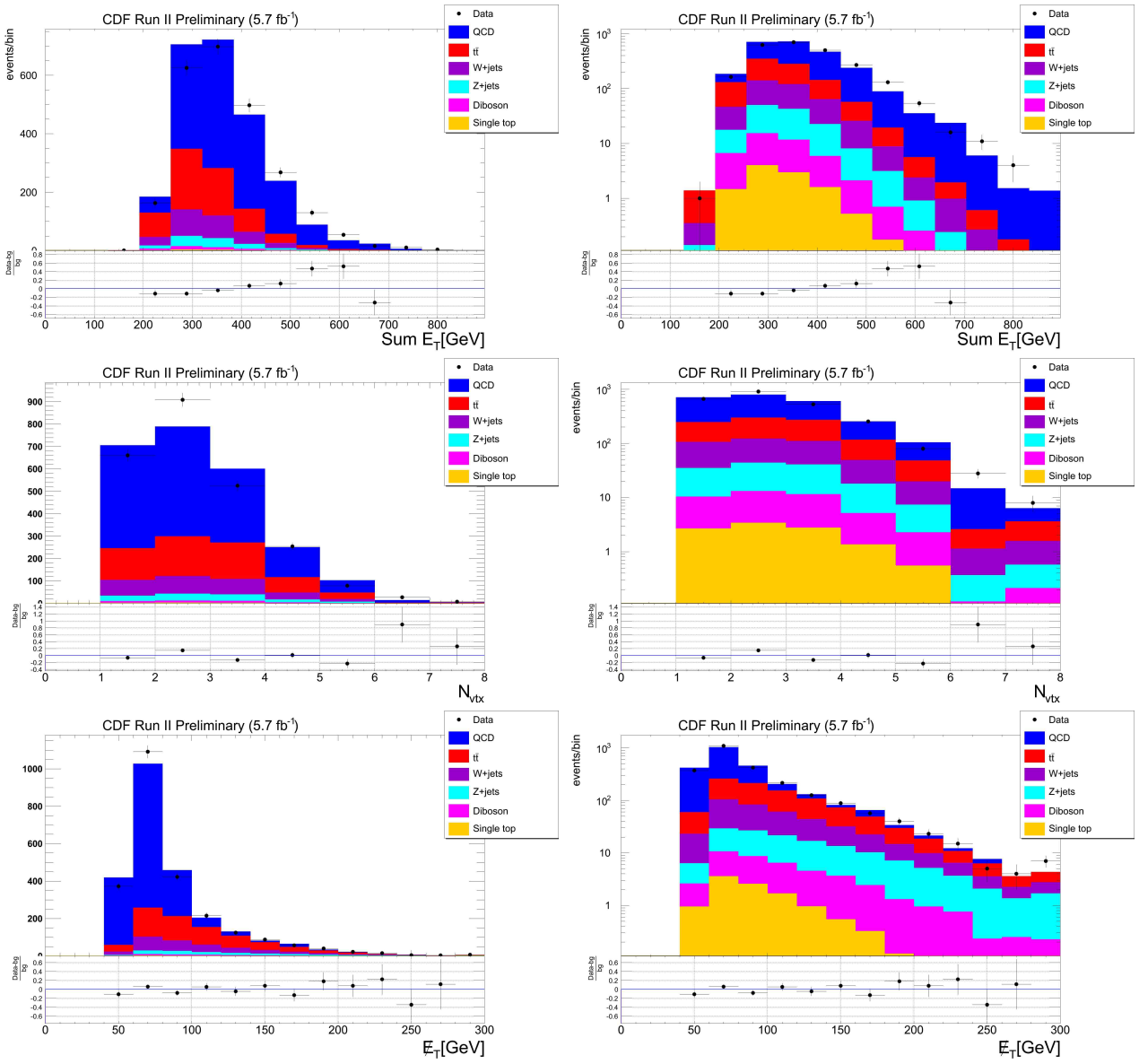


Figure B.4: Distribution of $SumE_T$, number of vertexes and $\not\!\!E_T$ in the $N_{jets} = 4$ control region for data and expected backgrounds, in linear and logarithmic scale. Each process is normalized to the number of expected events.

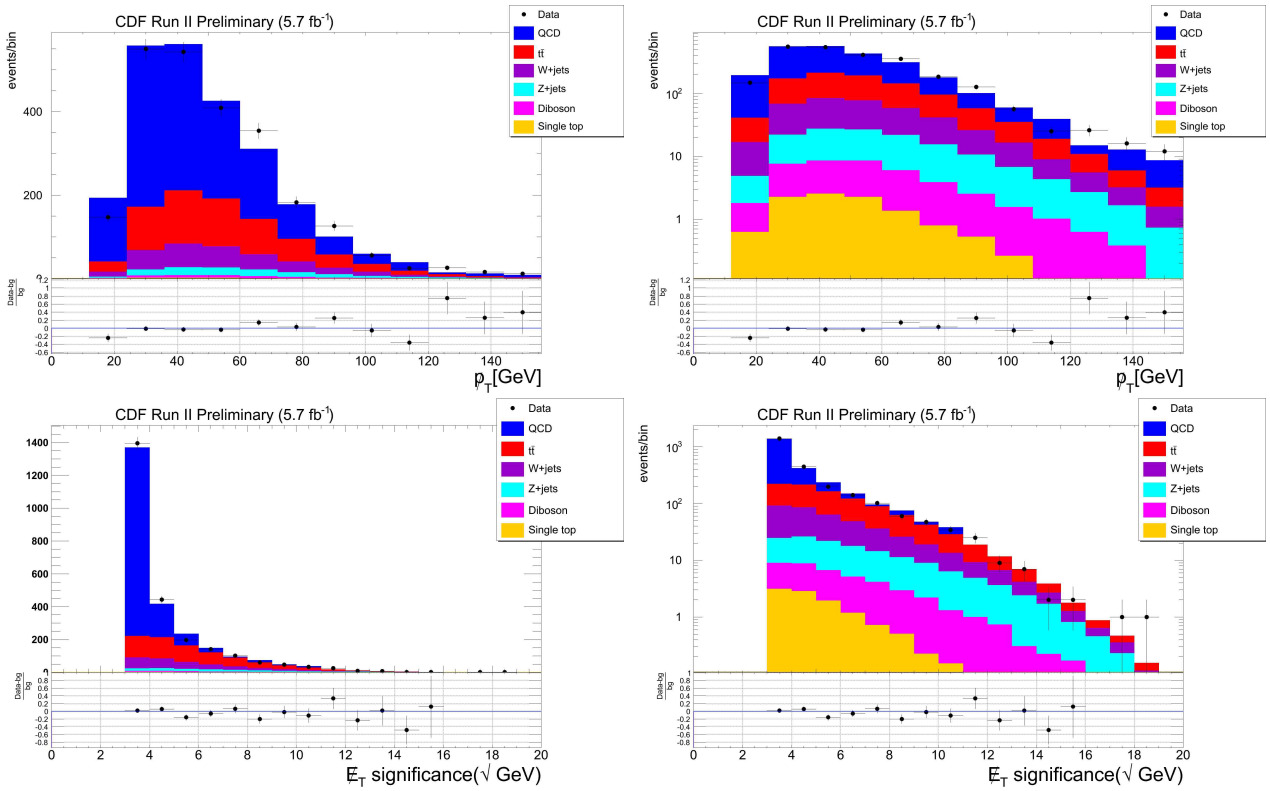


Figure B.5: Distribution of $\not{p}_T$ and $\not{E}_T$ significance in the $N_{jets} = 4$ control region for data and expected backgrounds,, in linear and logarithmic scale. Each process is normalized to the number of expected events.

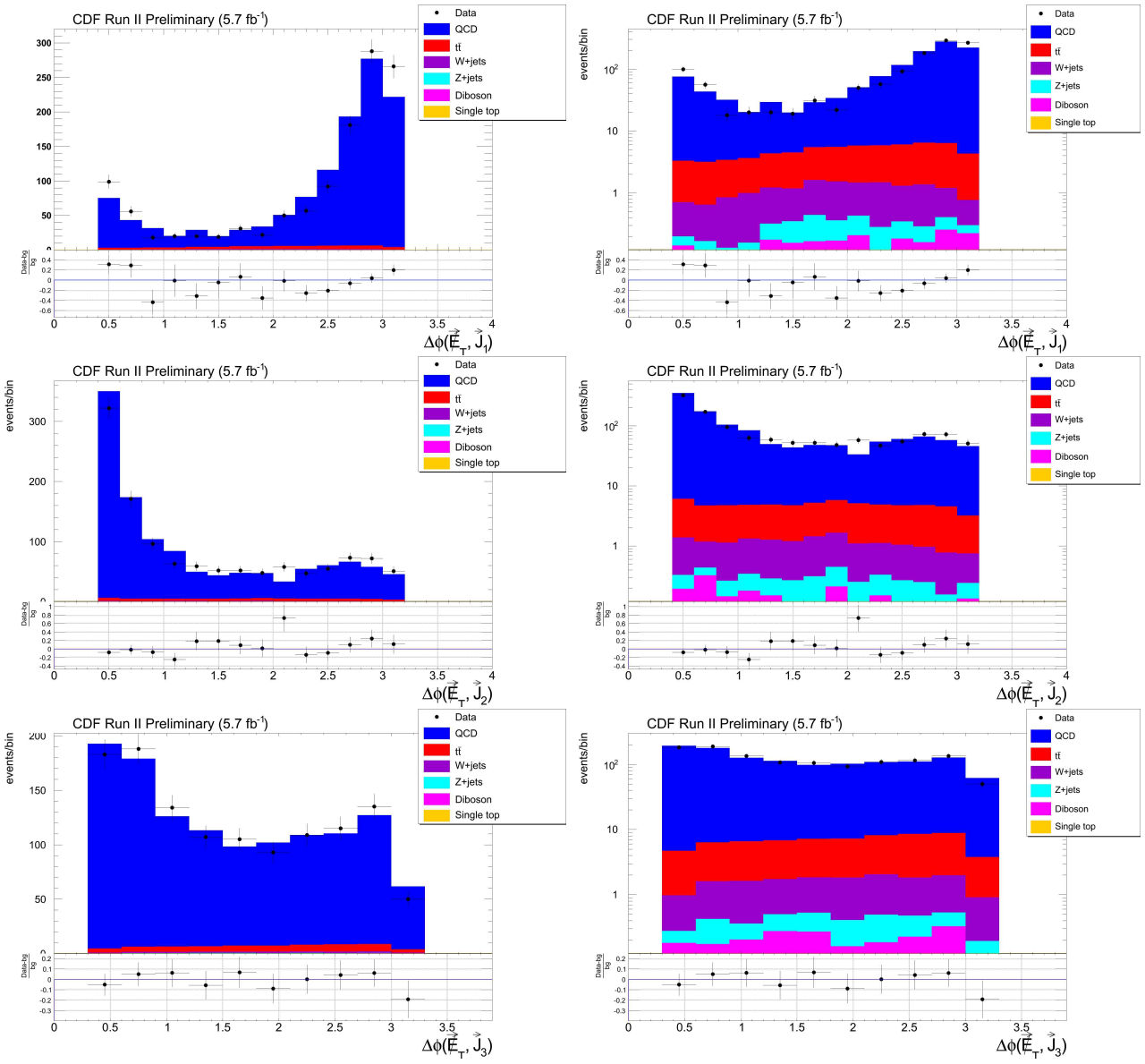


Figure B.6: Distributions of azimuthal distances between $\not\!\!E_T$ and first three jets in the $\not\!\!E_T significance < 3$ GeV$^{\frac{1}{2}}$ control region for data and expected backgrounds, in linear and logarithmic scale. Each process is normalized to the number of expected events.

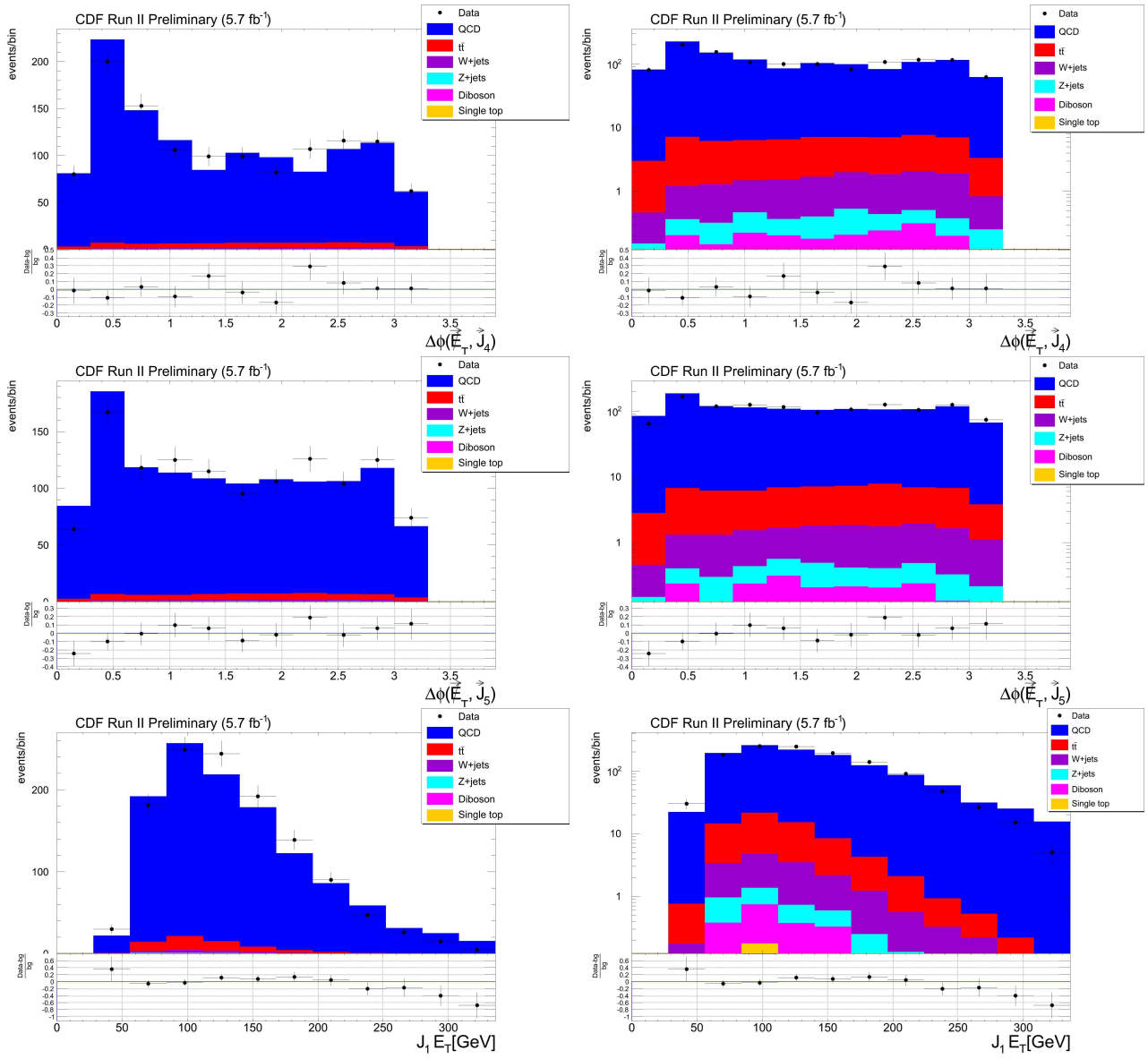


Figure B.7: Distributions of azimuthal distance between $\not\!\!E_T$ and the fourth jet, azimuthal distance between $\not\!\!E_T$ and the fifth jet, and of first jet energy in the $\not\!\!E_T significance < 3$ GeV$^{\frac{1}{2}}$ control region for data and expected backgrounds, in linear and logarithmic scale. Each process is normalized to the number of expected events.

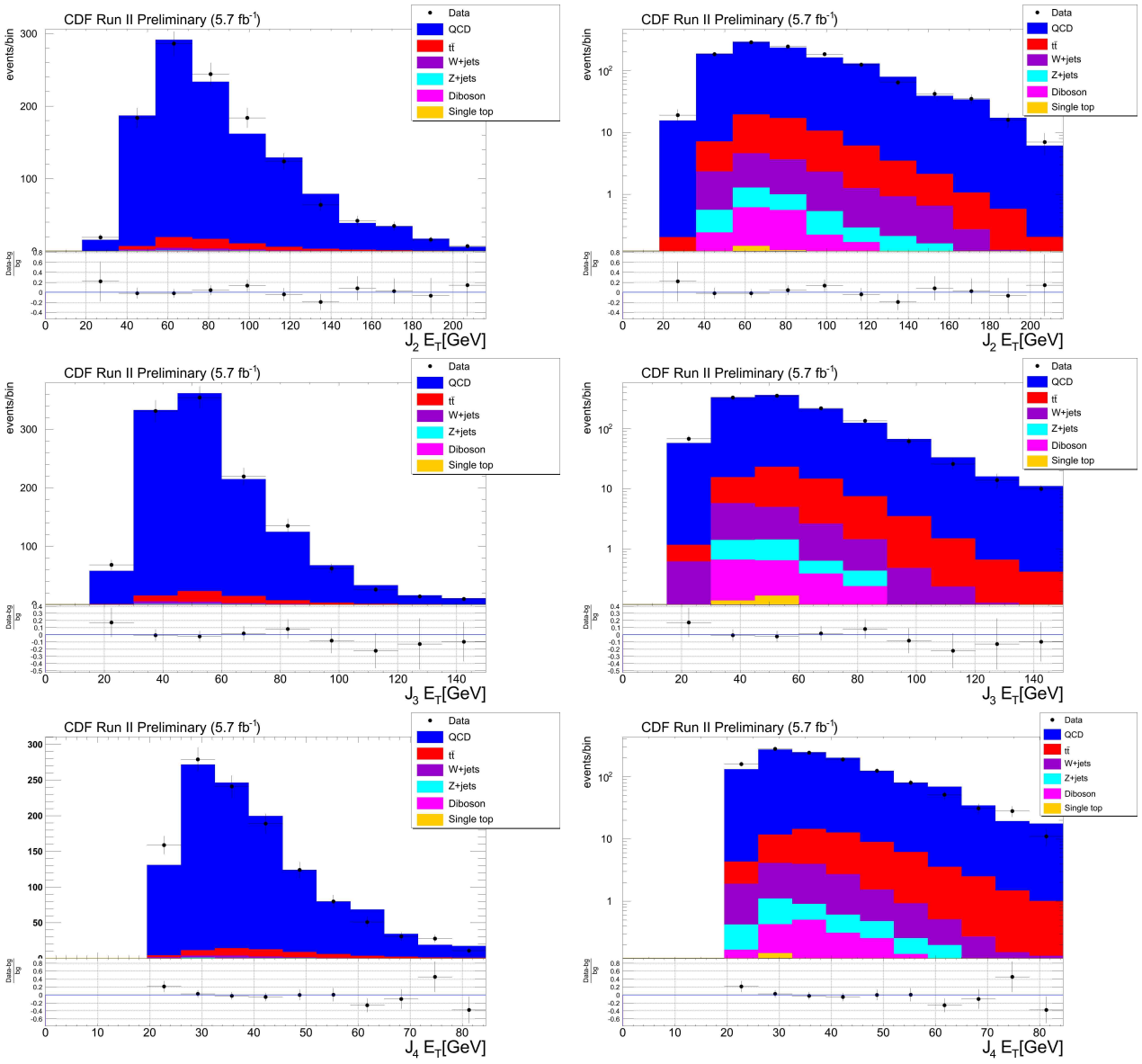


Figure B.8: Distributions of the second, third and fourth jet's energies in the $\not\!\!E_T significance < 3$ GeV$^{\frac{1}{2}}$ control region for data and expected backgrounds, in linear and logarithmic scale. Each process is normalized to the number of expected events.

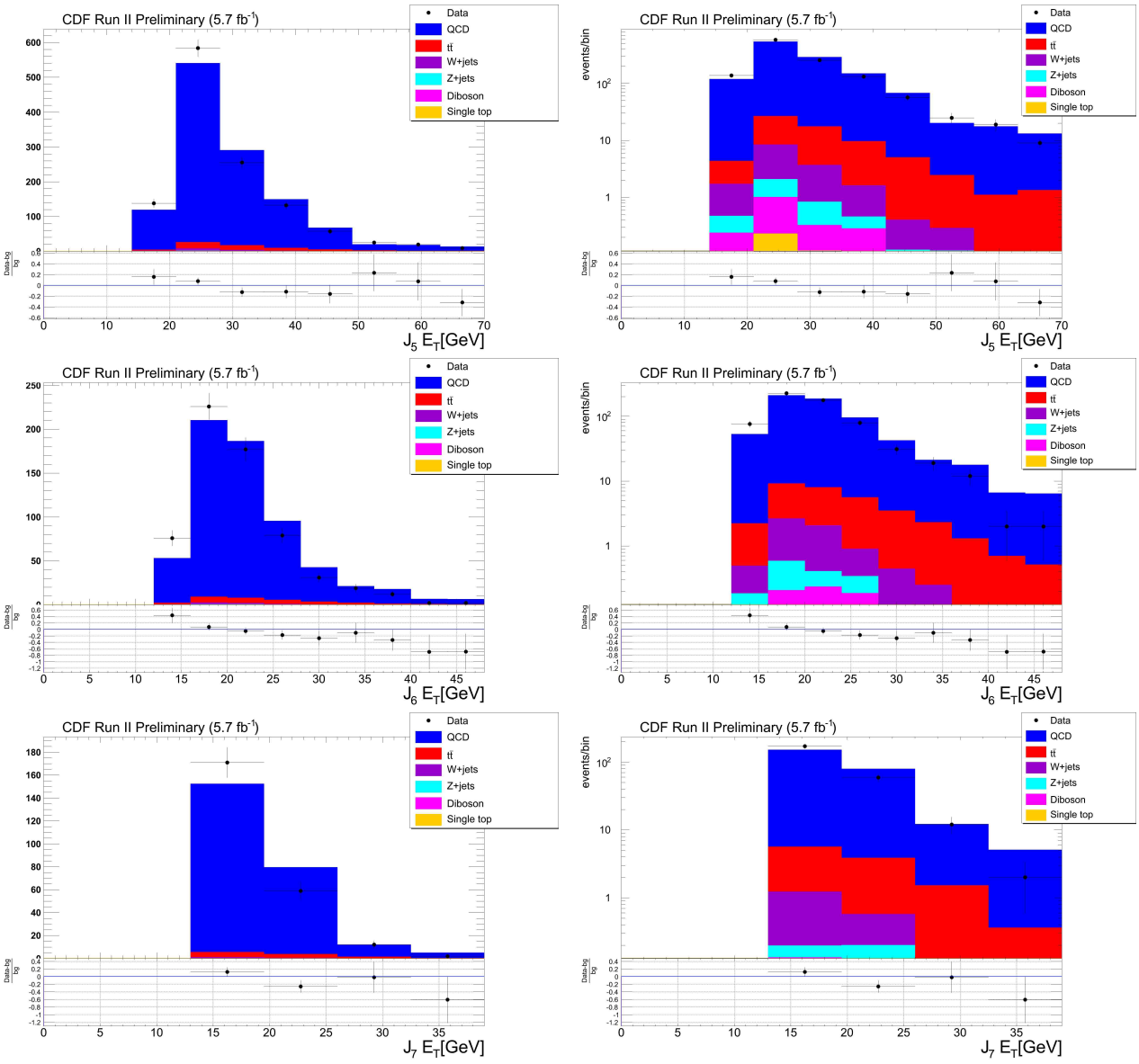


Figure B.9: Distributions of the fifth, sixth and seventh jet's energies in the $\not\!\!E_T significance < 3$ GeV$^{\frac{1}{2}}$ control region for data and expected backgrounds, in linear and logarithmic scale. Each process is normalized to the number of expected events.

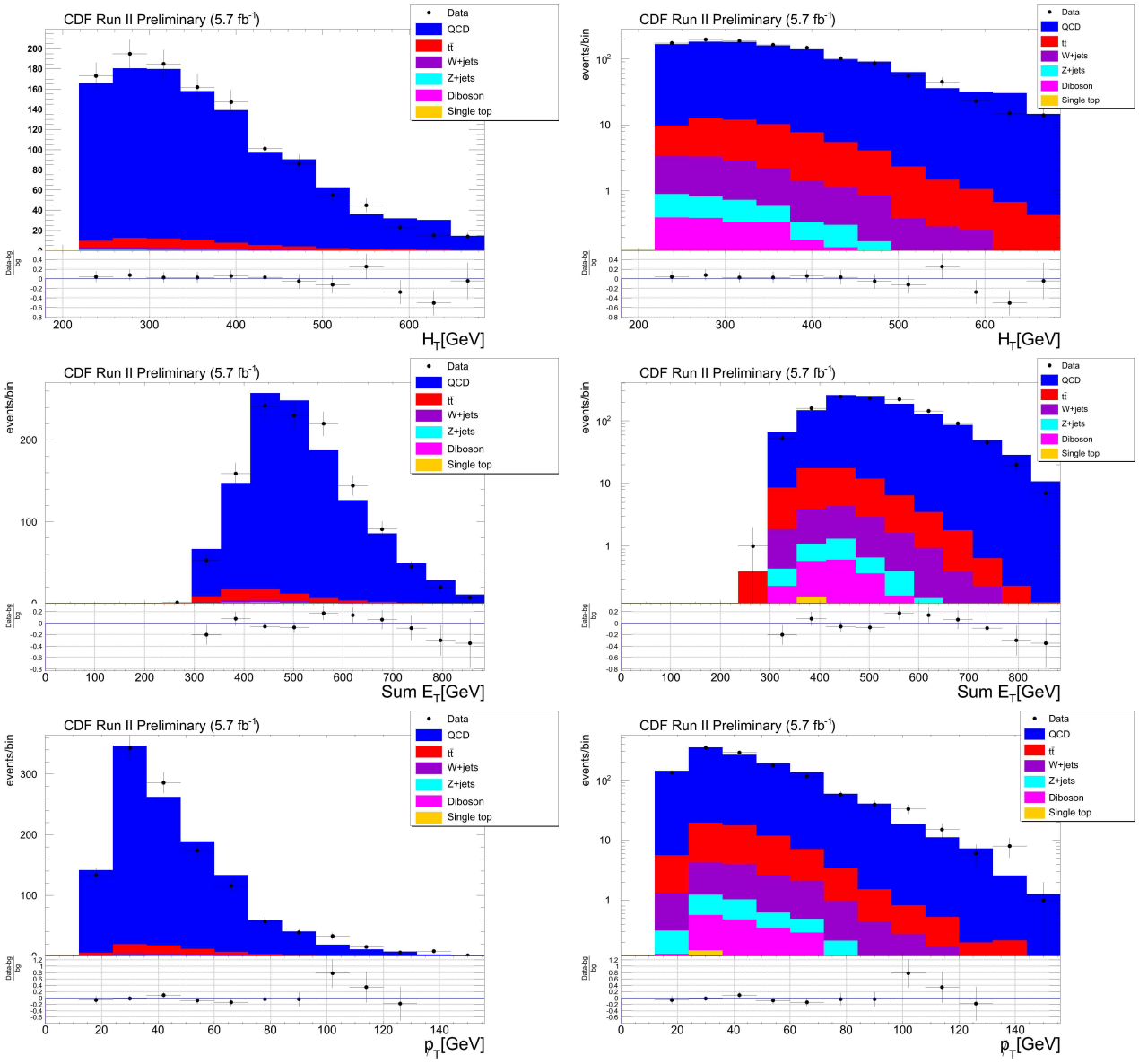


Figure B.10: Distribution of H_T, $SumE_T$ and $\not\!\!p_T$ in the $\not\!\!E_T significance < 3$ GeV$^{\frac{1}{2}}$ control region for data and expected backgrounds, in linear and logarithmic scale. Each process is normalized to the number of expected events.