Università degli Studi di Roma "Tor Vergata"
Facoltà di Scienze Matematiche Fisiche e Naturali

Dottorato di Ricerca in Fisica
XXVIII Ciclo

Measurement of the inclusive isolated prompt photon cross section in $p\bar{p}$ collisions at $\sqrt{s}$ =1.96 TeV, using the full CDF data sample

Relatori:
Prof. Giovanni Carboni
Dr. Ray Culbertson

Coordinatore:
Prof. Maurizio De Crescenzi

Candidato:
Alessandra Lucà

Anno Accademico 2014/15

Contents

Università di Roma "Tor Vergata"

Facoltà di Scienze Matematiche Fisiche e Naturali

Dottorato di Ricerca in Fisica
XXVIII Ciclo

MEASUREMENT OF THE INCLUSIVE ISOLATED PROMPT PHOTON CROSS SECTION IN $p\bar{p}$ COLLISIONS AT $\sqrt{s}$ =1.96 TEV, USING THE FULL CDF DATA SAMPLE

Alessandra LUCÀ

Abstract

The measurement of the cross section for the inclusive production of isolated prompt photons in proton-antiproton collisions at $\sqrt{s}$=1.96 TeV is presented. The data set corresponds to an integrated luminosity of 9.5 fb^{-1}, collected with the Collider Detector at Fermilab in Run II. The measurement is performed as a function of the photon transverse energy (E_T^γ) covering the range of 30 GeV< E_T^γ <500 GeV in the pseudorapidity region $|\eta^\gamma|$ <1.0. To reduce the background coming from the decays of π^0's, η's and other hadrons, photons are required to be isolated in the calorimeter. The output distributions of an Artificial Neural Network are exploited to estimate the remaining contamination from jets faking isolated photons. Results are compared to leading-order and next-to-leading-order perturbative QCD calculations.

Chapter 1

Introduction

The subject of this thesis is the measurement of the inclusive isolated prompt photon cross section with a total integrated luminosity of 9.5 fb^{-1} of data collected with the CDF Run II detector at the Fermilab Tevatron Collider. The prompt photon cross section is a classic test of perturbative quantum chromodynamics (pQCD), probing parton distribution functions (PDFs) and parton-to-photon fragmentation functions (FFs) [1]. In addition, prompt photons can also constitute a background for many other searches such as Higgs boson decays into photon pairs ($H \rightarrow \gamma\gamma$) or SUSY and extra-dimensions with energetic photons in the final state.

The study of prompt photon production also offers some experimental advantages compared to QCD studies using jets. Electromagnetic calorimeters have better energy resolution than hadronic calorimeters, and the systematic uncertainty on the photon absolute energy scale is smaller. Moreover, the reconstruction of the photon kinematics does not require the use of jet algorithms.

However, the measurements with photons in the final state require a good understanding of the background, mainly dominated by light mesons such as $\pi^0$and η, which decay into two or more collinear photons. Since these photons are produced within a jet, they tend to be non-isolated in most of the cases, and can be suppressed by requiring the photon candidates to be isolated in the calorimeter. In the case the hard scattered parton hadronizes leaving most of its energy to the meson, the photon produced in the decay may have surrounding energy deposits small enough to appear as an isolated photon. To further reduce this remaining isolated background, a technique based on a Likelihood fit of the Artificial Neural Network output distributions was developed.

The measured cross section is compared to leading order (LO) and next-to-leading order (NLO) pQCD calculations.

This thesis is organized as follows:

- Chapter 2 contains a brief review of the Standard Model of particle physics, including a short discussion of the perturbative quantum chromodynamics. The phenomenology of prompt photon production follows along with the summary of the most recent measurements.
- Chapter 3 and Chapter 4 contain a description of the Tevatron and the CDF detector, respectively.
- Chapter 5 deals with the reconstruction of the “physical objects” of interest for the analysis.
- Chapter 6 discusses about the currently available predictive tools.
- Chapter 7 describes the different data sets used in the measurement, the trigger, and the event selection requirements.
- Chapter 8 is devoted to the explanation of the background subtraction method and the determination of the photon signal fraction.
- Chapter 9 discusses the final results and the comparison to the theoretical predictions.
- Chapter 10 finally presents the conclusions.

Chapter 2

Theoretical Motivation

This chapter provides the theoretical motivation for the study of prompt photons production at hadron colliders. First, a brief review of the Standard Model of particle physics is presented; then, the physics of hadron colliders is introduced, including a short discussion of perturbative quantum chromodynamics (pQCD). The phenomenology of prompt photon production follows along with the summary of the most recent measurements.

Contents

2.1 Short Review of the Standard Model

The Standard Model of particle physics (SM) is a successful theory for the elementary particles phenomenology [2]. Developed in the second half of the 20th century, it has been tested experimentally in many ways and to high precision during the last decades. The Standard Model is a theory which describes the strong interactions and the electroweak processes through the *Quantum Chromodynamics* (QCD) and the Glashow-Weinberg-Salam (GSW) theory, respectively [3, 4]. Gravity is not incorporated, but in the quantum world of particle interactions the effects of gravity are insignificant. According to the Standard Model, all matter is built from 12 fundamental particles and and their corresponding anti-particles. The basic constituents of matter are 6 *quarks* (q) and 6 *leptons* (ℓ); these subatomic particles are 1/2-spin fermions.

The six leptons are electron e^-, muon μ^-, τ-lepton, and three neutrinos ν_e, ν_μ, ν_τ; the six quarks are *up*, *down*, *strange*, *charm*, *bottom* and *top*, denoted as u, d, s, c, b and t, respectively. These elementary particles are arranged into three “families” or “generations”, as shown in Figure 2.1. Quarks and leptons completely differ in the phenomenology of their interactions, being the former ones, subject to all the three interactions, while the latter ones are not subject to the strong force. Interactions between them are mediated by particles with integer spin, called **gauge bosons**: *photons* (γ), $W^\pm$ and Z bosons, and *gluons* (g) [5]. The Higgs boson (H) is postulated by the electroweak theory primarily to explain the origin of particle masses. Some of the properties of the fundamental particles of the Standard Model are summarized in Figure 2.1 and in Figure 2.2.

The number of bosons for each interaction depends on the symmetry properties of the fermion (particle) field. The Electroweak theory and Quantum Chromodynamics are so-called gauge theories; this translates into the invariance of their equations of motion under a local gauge transformation[1]. In gauge theories the boson (mediator) field appears naturally after requiring the

---

[1] A gauge transformation is a transformation of the fermion wave functions plus a corresponding change of the mediator field, which together leave the Lagrangian unchanged.

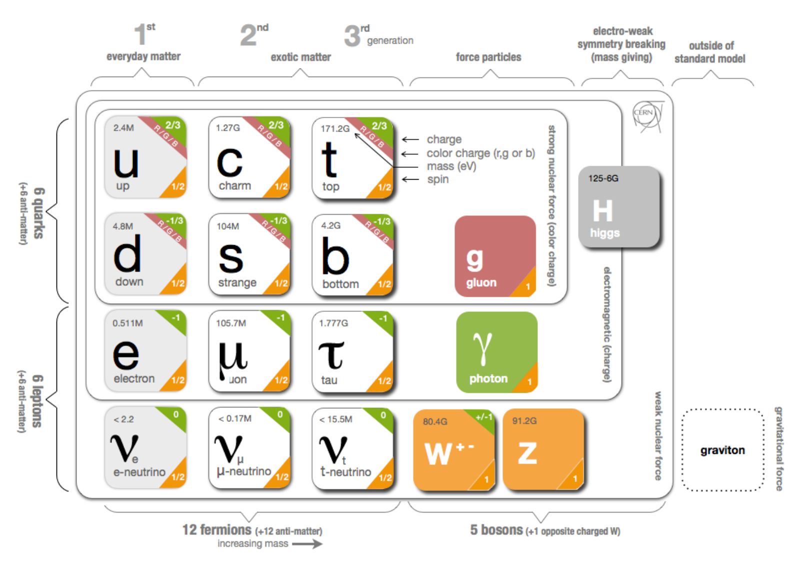


FIGURE 2.1: Fundamental particles and their properties in the Standard Model. Figure taken from Ref. [6].

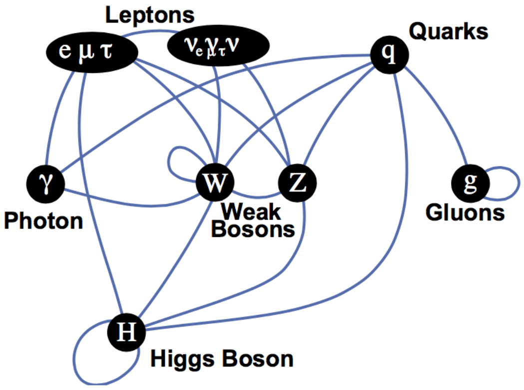


FIGURE 2.2: A diagram showing the tree-level interactions between elementary particles according to the Standard Model. The types of particles in the vertices (darkened circles) are leptons and quarks (top row), the force mediating particles (second row) and the Higgs boson (bottom row). Vertices are connected by edges (blue arcs) if interactions can occur.

Lagrangian invariance under local phase transformation of the fermion fields. A local phase transformation can be seen as an action of the element of some symmetry group; the number of the generators of this group defines the number of mediator bosons.

The interaction of quarks is invariant under $SU(3)_C$ transformations of the quark fields in color space. The $SU(3)_C$ color symmetry is exact and consequently the gluons are massless. The number of group generators is given by $n^2 - 1$, where n is the dimension of the gauge group, so for $SU(3)_C$ it is $3^2-1=8$, which means that there are 8 bosons (gluons) which are carriers of the strong interaction, according to the Quantum Chromodynamics, (see Section 2.2.4).

Quantum Electrodynamics (QED), the theory of electromagnetic interactions, is invariant against global phase transitions of the fermion field; the Fermi theory that describes the weak interaction is not a gauge theory. However, there exists a mechanism to combine these two into a single gauge theory, invariant against $SU(2)_L \times U(1)_Y$ transformations (Electroweak theory), where $SU(2)_L$ is the weak isospin group, acting on left-handed fermions, and $U(1)_Y$ is the hypercharge group. which gives rise to $2^2-1=3$ massive (W^+, W^-, Z) and 1 massless (γ) boson fields. At "low" energy (M $<$ 250 GeV) the $SU(2)_L \otimes U(1)_Y$ symmetry is "spontaneously" broken and the residual group is $U(1)_{emg}$ whose generator is a linear combination of the $U(1)_Y$ generator and a generator of $SU(2)_L$: the corresponding gauge boson is of course the photon and the associated "coupling" is $\alpha \sim \frac{1}{137}$. Symmetry breaking implies that the other gauge bosons acquire a mass: they are the heavy $W^\pm$, Z bosons discovered at CERN in the mid 80's. The symmetry breaking mechanism is associated to the names of R. Brout, F. Englert, P. Higgs, G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble (1964), but it is often simply referred to as the Higgs mechanism [7]. The existence of a massive boson, the Higgs boson, would be associated with the Higgs field. On 4 July 2012, the ATLAS and CMS collaborations at CERN's Large Hadron Collider (LHC) reported independently the existence of a Higgs-like boson [8]. On March 14, 2013 this newly discovered particle was tentatively confirmed as a Higgs boson [9].

2.2 Fundamental Interactions

In this section more details on the fundamentals interactions formalism are given. The Standard Model is a non-abelian local gauge theory. In a gauge theory the fundamental particles are described by quantized fields ψ, characterized by quantum numbers (e.g: spin). The interactions among particles are elegantly described via local *Gauge symmetries*: according to the Noether's theorem [10], every differentiable symmetry of the action $S = \int \mathcal{L}(\psi, \partial_\mu \psi) d^4x$ of a physical system[2] has a corresponding conservation law. The interactions in the SM are reviewed below.

---

[2] $\mathcal{L}$ is the Lagrangian density, which describes the dynamics of the system.

2.2.1 The Electromagnetic Interaction

The Lagrangian density of a free fermion with mass m is:

$$\mathcal{L}_{Dirac}(x) = \overline{\psi}(x)(i\not{\partial} - m)\psi(x) \tag{2.2.1}$$

where ψ is the Dirac spinor, $\not{\partial} \equiv \gamma^\mu\partial_\mu$, γ^μ being the Dirac matrices.

The Lagrangian density yielding the Maxwell equations for the electromagnetic scalar and vectorial potential A^μ, associated to the photon field, can be expressed as:

$$\mathcal{L}_{Maxwell}(x) = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{2.2.2}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor.

Equation 2.2.1 is invariant under global $U(1)$ rotations ($\psi \rightarrow e^{iQ\alpha}\psi$), α being the parameter of the rotation and Q an arbitrary number. If the system is further required to be invariant under a local rotation (i.e.: $U(1)$ with a space-time dependent generator $\alpha(x)$), Equation 2.2.1 has to be re-written as follow:

$$\mathcal{L}_{Dirac}(x) = \overline{\psi}(x)(i\not{D} - m)\psi(x) \tag{2.2.3}$$

where $\not{D} \equiv \gamma^\mu D_\mu$, $D_\mu \equiv \partial_\mu + iQA_\mu$ (D_μ is named "covariant derivative"), and $A_\mu \rightarrow A_\mu - \frac{1}{Q}\partial_\mu\alpha(x)$ under the local aforementioned rotation[3]. Thus, the QED Lagrangian density is derived:

$$\mathcal{L}_{QED}(x) = \overline{\psi}(x)(i\not{\partial}_\mu - m)\psi(x) - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - QeA^\mu\overline{\psi}\gamma_\mu\psi \tag{2.2.4}$$

where the new term $QeA^\mu\overline{\psi}\gamma_\mu\psi$ describes the interaction between a fermion of charge Q in units of the electron charge[4] and the photon field. The corresponding current, which is conserved ($\partial_\mu J_\mu^{QED} = 0$, see Noether's theorem), is defined as:

$$J_\mu^{QED} \equiv Q\overline{\psi}\gamma_\mu\psi \tag{2.2.5}$$

In a classic analogy the conservation of J_0^{QED}, integrated over the space, represents the time-conservation of the electric charge, on which the electromagnetic force depends.

---

[3] It can be shown that equations 2.2.2 and 2.2.3 are respectively invariant under $A_\mu \rightarrow A_\mu - \frac{1}{Q}\partial_\mu\alpha(x)$ and $\psi \rightarrow e^{iQ\alpha}\psi$. The latter invariance happens since $D_\mu\psi(x) \rightarrow e^{iQ\alpha(x)}D_\mu\psi(x)$.

[4] e is the charge of the electron. $e = 1.60217646 \times 10^{-19}$ Coulomb.

Starting from Equation 2.2.5, $\mathcal{L}_{QED}$ in Equation 2.2.4 can be re-written as:

$$\mathcal{L}_{QED}(x) = \overline{\psi}(x)(i\, \not{\partial}_\mu - m)\psi(x) - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - eA^\mu J_\mu \tag{2.2.6}$$

2.2.2 The Electroweak Interactions

Weak interaction phenomena show common properties between doublets of fermions ($\binom{\nu_e}{e}, \ldots, \binom{u}{d}, \ldots$). Because of this reason the weak isospin is defined and the weak interaction Lagrangian density is built in order to be invariant under a local $SU(2)_L$ rotation in the weak isospin space. Such a rotation is given by $\psi(x) \to e^{i\boldsymbol{\alpha}(\boldsymbol{x})\boldsymbol{\tau}/2}$, $\boldsymbol{\alpha}$ being the parameters of the rotations, and $\boldsymbol{\tau} = \boldsymbol{\sigma}/2$ the generators of the rotations, with $\boldsymbol{\sigma} = (\sigma^1, \sigma^3, \sigma^3)/2$ being the Pauli matrices. According to the proposal presented by Glashow in 1961 [4], these two theories, QED and weak, have been unified in the *electroweak (EW) theory*, represented by the group $SU(2)_L \times U(1)_Y$, Y being the hypercharge, as defined by the Gell-Mann-Nishijima formula below:

$$Y = Q - T^3 \tag{2.2.7}$$

$T^3 \equiv \tau^3$ is the third component of the weak isospin. The invariance under such a transformation implies that two currents are conserved. These currents, $\mathbf{J}_\mu^{weak}$ and J_μ^Y, are defined in analogy with Equation 2.2.5:

$$\mathbf{J}_\mu^{weak} \equiv \overline{\chi}\gamma_\mu \frac{\boldsymbol{\sigma}}{2}\chi \tag{2.2.8}$$

$$J_\mu^Y = \overline{\psi}\gamma_\mu Y \psi = J_\mu^{QED} - J_\mu^3 \tag{2.2.9}$$

Once the conserved currents are defined, it is straightforward to write the $SU(2)_L \times U(1)_Y$-invariant Lagrangian density:

$$\mathcal{L}_{EW} = \overline{\psi}(i\, \not{\partial} - m)\psi - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} - \frac{1}{4}\mathbf{W}^{\mu\nu}\mathbf{W}_{\mu\nu} + g\mathbf{J}_\mu^{\mathbf{weak}}\mathbf{W}^\mu + \frac{g^{'}}{2}J_\mu^Y B^\mu \tag{2.2.10}$$

where:

- $\mathbf{W}^{\mu\nu} = \partial_\mu \mathbf{W}_\nu - \partial_\nu \mathbf{W}_\mu + g\mathbf{W}_\mu \times \mathbf{W}_\nu$, $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ describe the propagation of the $\mathbf{W}_\mu = (W_\mu^1, W_\mu^2, W_\mu^3)$ and B_μ electroweak fields. They are defined in analogy with the QED case. The definition of $\mathbf{W}^{\mu\nu}$ contains an extra term with respect to $F_{\mu\nu}$: it accounts for the non-abelianity of the $SU(2)$ group.

- g, $g^{'}$ are two different coupling constants arising from the the fact that $SU(2)$ and $U(1)$ commute. The coupling constants can be expressed as:

$$e = g\sin(\theta_W) = g^{'}\cos(\theta_W) \tag{2.2.11}$$

θ_W being the Weinberg angle[5].

Equation 2.2.10 can be re-written as:

$$\mathcal{L}_{EW} = \overline{\psi}(i \not{D} - m)\psi - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} - \frac{1}{4}\mathbf{W}^{\mu\nu}\mathbf{W}_{\mu\nu} \tag{2.2.12}$$

where:

$$D_\mu = \partial_\mu - ig\mathbf{W}_\mu\frac{\boldsymbol{\sigma}}{2} - ig^{'}YB_\mu \tag{2.2.13}$$

The EW Lagrangian density (Equation 2.2.10) uniquely determines the coupling of the boson fields to the fermions, once the quantum number $(T^+, T^-, T^3) \equiv \frac{1}{2}(\sigma^1 + i\sigma^2, \sigma^1 - i\sigma^2, \sigma^3)$ of the fermion fields are specified. To determine these quantum numbers, one must take into account that the W bosons couple only to left-handed chirality states [15] of quarks and leptons[6]. For simplicity fermions are therefore grouped into doublets and singlets as follows:

$$\chi_L = \begin{pmatrix} \psi_u \\ \psi_d \end{pmatrix}_L \tag{2.2.14}$$

$$\chi_R = \psi_R \tag{2.2.15}$$

where "L" and "R" mean respectively left-handed and right-handed chirality states. "u" and "d" are the labels up and down fermions in the weak isospin doublet. Such a grouping uniquely defines the quantum numbers of the fermions.

The electroweak theory must also take into account the observed *flavor changing charged currents* in the quark sector[7]. $\psi_d^{quark} \equiv (d, s, b) \rightarrow \psi_d^{'quark} \equiv (d^{'}, s^{'}, b^{'})$

---

[5] $\sin^2\theta_W = 0.23116 \pm 0.00012$ (Ref. [11]).

[6] No right-handed neutrinos have been observed yet. This statement is considered valid for quarks as a natural extension in the electroweak unification.

[7] The first evidence came from strangeness changing charged current. An example of this phenomenon is $\Lambda \rightarrow pe\overline{\nu_e}$.

where $(d^{'}, s^{'}, b^{'})$ is expressed according to the following:

$$\begin{pmatrix} d^{'} \\ s^{'} \\ b^{'} \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.2.16}$$

where V_{CKM} is the Cabibbo-Kobayashi-Maskawa quark mixing matrix.

2.2.3 EW Symmetry Breaking: Higgs Model

The $SU(2)$ mentioned earlier would be exact if fermions, W and Z^0 bosons were massless. In order for the Standard Model to be compatible with the large observed masses of W and Z^0 bosons[8], symmetry breaking must occur.

Spontaneous electroweak symmetry breaking (EWSB) can be accomplished by the introduction of a scalar field ϕ which interacts with both fermions and gauge bosons. Since the EW interactions are local and the masses of the particles are different from zero also when non interacting (during free propagation), the scalar field must be always locally present; then the scalar field is usually said to have a non vanishing vacuum expectation value. The Lagrangian density for such a scalar field is:

$$\mathcal{L}_H = |D_\mu \phi|^2 - V(\phi) \tag{2.2.17}$$

where:

- the covariant derivative D_μ (Equation 2.2.13) takes care of the interaction between the Higgs field and the electroweak bosons.
- the potential $V(\phi)$[9], which includes the self interaction of the Higgs field, is given by:

  $$V(\phi)_H = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2 \tag{2.2.18}$$

  with $\mu^2 < 0$ and $\lambda > 0$.

---

[8] The quantum fields of the $W^\pm$ and Z^0 mass eigenstates are defined as $W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp i W^2_\mu)$, $Z^0_\mu = \frac{1}{\sqrt{g^2+g^{'2}}}(gW^3_\mu - g^{'} B_\mu)$, with $\mathbf{W}_\mu$, B_μ, g, g$^{'}$ defined in the last section. $M_{W^\pm} = 80.385 \pm 0.015$ GeV/c^2 and $M_{Z^0} = 91.1876 \pm 0.0021$ GeV/c^2 (Ref. [11]).

[9] In order to preserve the renormalizability of the theory terms up to dimension 4 are kept.

The Higgs mechanism predicts two complex scalar fields, such that:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.2.19}$$

where the field ϕ^+/ϕ^0 are respectively the charged/neutral component of the doublet. The symmetry is broken since the minimum of $V(\phi)$ occurs at $\sqrt{<\phi^\dagger\phi>} = \sqrt{\frac{-\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}} \neq 0$. The gauge freedom allows to conveniently choose the scalar field in its ground state, as follows:

$$<\phi>= \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.2.20}$$

The gauge boson mass terms come directly from substituting Equation 2.2.20 into Equation 2.2.17:

$$M_{W^\pm} = g\frac{v}{2} \tag{2.2.21}$$

$$M_{Z^0} = \frac{v}{2}\sqrt{g^2 + g'^2} \tag{2.2.22}$$

while the photon remains massless (Equation 2.2.20 does not break the $U(1)_{EM}$ symmetry[10]).

From Equation 2.2.11 one can see that the masses of the weak bosons are related by the weak missing angle, as follows:

$$sin^2\theta_W = 1 - \frac{M_W^\pm}{M_{Z^0}} \tag{2.2.23}$$

The ϕ doublet in its ground state can be parameterized also in terms of the Higgs boson H:

$$<\phi(x)>= \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{2.2.24}$$

From Equation 2.2.18 and 2.2.24 it is possible to derive the Higgs mass term:

$$M_H = 2v\sqrt{\lambda} = \mu\sqrt{2} \tag{2.2.25}$$

The fermion masses can also be generated if a Yukawa couplings for the upper (U) and lower (D) components of fermion isospin doublets (Equation 2.2.15) is

---

[10] $SU(2)_L \times U(1)_Y$ is spontaneously broken down to the electromagnetic $U(1)_{EM}$ symmetry

considered:

$$\mathcal{L}_{\text{Yukawa},D} = -g_D[(\overline{\psi}^U_L, \overline{\psi}^D_L)\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix}\psi^D_R + \overline{\psi}^D_R(\phi^{+\dagger}, \phi^{0\dagger})\begin{pmatrix}\psi^U_L\\ \psi^D_L\end{pmatrix}] \tag{2.2.26}$$

and a similar expression for $\mathcal{L}_{\text{Yukawa},U}$. $g_{U/D}$ are the coupling of the fermions to the Higgs boson. Those Lagrangians lead to:

$$M_U = g_U v/\sqrt{2} \tag{2.2.27}$$

$$M_D = g_D v/\sqrt{2} \tag{2.2.28}$$

where the constants $g_{U/D}$ can be extracted by the measurements of the fermion masses.

2.2.4 Quantum-Chromo-Dynamics

In the SM the strong interaction is described by the Quantum-Chromo-Dynamics (QCD). This force is responsible for quarks "sticking" together to form composite particles (hadrons). QCD is described by a non-abelian local $SU(3)_C$ group. The QCD Lagrangian density, invariant under a $SU(3)$ local transformation, is given by:

$$\mathcal{L}_{QCD} = \overline{\psi}(i\,\not{\partial}_\mu - m)\psi - \frac{1}{4}G^{\mu\nu}_a G^a_{\mu\nu} - g_S\overline{\psi}^a\gamma_\mu T^c_S\psi_b A^\mu_c \tag{2.2.29}$$

where:

- a,b,c=1,...,8
- g_S is the QCD coupling constant
- T^a_S are the generators of the $SU(3)$ group;
- $G^{\mu\nu}_a = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_S f_{abc} A^b_\mu A^c_\nu$ describes the propagation of the strong A^a_μ field. It is defined in analogy with the aforementioned $\mathbf{W}^{\mu\nu}$. f_{abc} are the structure constants of the $SU(3)$ group: $[T^a_S, T^b_S] = i2f_{abc}T^c_S$

As for the QED and weak interactions, gauge invariance of the QCD gauge symmetry group implies a new degree of freedom to be conserved: the color. However, unlike QED, the gauge symmetry group is non-abelian, causing gluons (carriers of strong interaction) to possess color charge and interact with each other. The strength of interaction is parametrized by the strong coupling constant $\alpha_s \equiv g_S^2/4\pi$. To a first approximation in Q^2/λ^2 one has:

$$\alpha_s(Q^2) = \frac{12\pi}{(33-2n_f)ln(Q^2/\lambda^2)} \tag{2.2.30}$$

where Q^2 is the interaction momentum transfer scale and n_f is the number of different flavored quarks with mass less than Q^2 [14]. Fits to experimental data set the parameter λ around 200 MeV [10]. In Figure 2.3 measurements of $\alpha_s(Q^2)$ as a function of the energy scale Q are summarized.

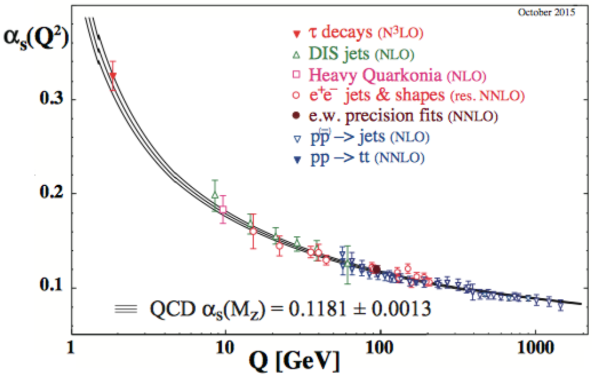


FIGURE 2.3: Summary of measurements of $\alpha_s(Q^2)$ as a function of the energy scale Q. Plot taken from Ref. [21].

QCD features two very important properties:

- **Asymptotic freedom**. $\alpha_s(Q^2)$ becomes small at large Q^2 (see Equation 2.2.30 and Figure 2.3). This means that quarks and gluons interact weakly if they are within a short range or, equivalently, large momentum transfers. This property allows perturbation theory (pQCD) to be used in theoretical calculations to produce experimentally verifiable predictions for hard scattering processes.
- **Confinement**. Colored particles are confined into colorless singlets (hadrons) by an increasing QCD potential with increasing relative distance. If quarks are forced to large relative distances the energy density in the binding color string increases and energy is materialized into colored quark pairs. Thus a hard scattered parton evolves into a shower of partons and finally into hadrons (hadronization). Even if theoretically unproven, confinement is widely believed to be true because it explains the consistent failure of free quark searches.

In the $SU(3)$ representation of QCD, the gluons have an effective color charge that is larger than of the quarks. The effective gluon and quark color charges are proportional to the square roots of C_A and C_F respectively, where $C_A = 3$ and $C_F = 4/3$[11]. The larger effective gluon color charge with respect to the quark one results in a number of properties (e.g: broadness, particle multiplicity with the jet) distinguishing gluon jets from quark jets [17].

2.2.5 Lagrangian of the Standard Model

In summary, QCD and Electroweak Theory are combined into the single Standard Model of particle physics, having the group symmetry structure of:

$$SU(3)_C \otimes (SU(2)_L \otimes U(1)_Y) \tag{2.2.31}$$

i.e. the *direct product* of three simple groups [12]. The Higgs boson permeates the physical vacuum breaking the symmetry of the theory and giving masses to the previously mentioned elementary particles, in a renormalizable theory framework. The SM Lagrangian density, including all the ingredients described in the previous sections, can be found in Appendix A.

---

[11] It can be shown that $C_A = N$ and $C_F = (N^2 - 1)/2N$ for a $SU(N)$ invariant theory, N being a given integer [14].

2.3 Physics at Hadron Colliders

Hadron colliders are powerful tools for studies of QCD dynamics, its parameters and particle structures. The importance of these studies is for both the SM and new physics processes. The later most likely proceed via QCD subprocesses and, at the same time, have as main background QCD subprocesses as well. Therefore, searches of new physics both at the Fermilab Tevatron and the CERN Large Hadron Collider (LHC) require solid knowledge of many aspects of QCD.

QCD provides the formalism to calculate the cross sections for interactions involving hadrons in the initial state and predicts a wide range of final products such as photons, W and Z bosons, jets. The term “jet” refers to a collimated collection of hadrons emerging from a hard-scattering reaction. Jets result from the process of hadronization, wherein the scattered partons are converted to the hadrons appearing in the final state. The four-vector of the jet is closely related to that of the parent parton, so that by studying jet production one can get closer to the underlying parton-level kinematics.

The use of perturbation theory in QCD calculations is made possible by the feature of asymptotic freedom. If a process involves a large momentum transfer, then the running coupling constant α_S may be small enough to justify the use of perturbative techniques. When pQCD can be applied, the factorization theorem states that the cross section of any QCD process can be written as the convolution of basic building blocks such as the quark and gluon distributions in the incoming hadrons, the hard subprocesses describing the large-angle scattering of partons, and the fragmentation functions of quarks and gluons into hadrons.

For example, for the high-p_T process $p\bar{p} \to hX$, where p and $\bar{p}$ are the initial state particles, h is the final state hadron and X represents all other particles, one can calculate perturbatively the cross section as follows:

$$\frac{d\sigma^{p\bar{p}\to hX}}{dP} = \sum_{f_1,f_2,f} \int dx_1 dx_2 dz f_1^p(x_1,\mu^2) f_2^p(x_2,\mu^2)$$

$$\times \left[\frac{d\hat{\sigma}^{f_1 f_2 \to f X'}}{dP}(x_1 p_1, x_2 p_2, p_h, \mu) \times D_f^h(z,\mu^2) \right] \tag{2.3.1}$$

In Equation 2.3.1, P stands for an appropriate set of the kinematic variables of the process, $f_i^p(x_i, \mu^2)$ is the parton distribution function which gives the probability of finding parton i within the initial hadron with momentum fraction x_i (see Section 2.3.2); μ is the factorization scale; p_1 and p_2 are the momenta of the initial hadrons . The $\hat{\sigma}^{f_1 f_2 \to f X'}$ is the hard process cross section for the initial partons f_1 and f_2 which are producing a final state parton f and all other final products X' (see Section 2.3.1). Then, $D_f^h(z, \mu^2)$ is the fragmentation function giving the probability density for finding a hadron h with fraction of momentum z in the final state parton f (see Section 2.3.3).

2.3.1 Partonic cross sections

The partonic cross section $\hat{\sigma}$ can be calculated using pQCD and Feynmann diagrams techniques [18] and can be expressed as:

$$\hat{\sigma} = \hat{\sigma}^{(0)} + \alpha_S \hat{\sigma}^{(1)} + \mathcal{O}(2) \tag{2.3.2}$$

where $\sigma^{(0)}$ is the contribution at leading order (LO), $\sigma^{(1)}$ is the contribution at next to the leading order (NLO). In Appendix B the parton-parton two body scattering differential cross sections are listed.

Calculating high orders terms of the perturbative expansion, two kinds of divergences appear, the ultraviolet and infrared and collinear divergences:

- Ultraviolet divergences (UV), come from the integration over large values of loop momenta, but they are removed after the renormalization of the theory. There isn't an universal scheme for the renormalization method, but several are available: (1) Minimal subtraction scheme, (2) $\overline{MS}$ scheme, (3) on shell scheme. The renormalization of the theory implies the introduction of a scale parameter μ_R, called renormalization scale. The μ_R dependence of the strong coupling constant, $\alpha_S(\mu_R)$, is described by the Callan - Symanzik equation [19].
- Infrared and collinear divergences appear in the calculation of the Feynman diagrams of the real and virtual corrections in the limit of vanishing energy of an emitted parton or when two partons become collinear. In analogy to the renormalization procedure, a factorization scale, μ_F, has to be introduced

for the removal of the infrared and collinear divergences. After this removal, both the parton distribution functions and the partonic cross sections acquire a dependence on the factorization scale.

2.3.2 Parton Distribution Functions

Protons are composed of three *valence* quarks and a number of *sea* quarks and gluons carrying fraction x of the proton momentum. The valence quarks only carry about half of the proton total momentum: the rest is carried by virtual gluons continuatively exchanged by the quarks. These gluons in turn produce virtual $q\bar{q}$ pairs called *sea* quarks. The parton distribution functions [21] $f_i(x, \mu_F^2)$ corresponds to the probability to find a parton[12] i (quark and gluon) carrying a momentum fraction x at a given factorization scale μ_F^2. The factorization scale is an arbitrary parameter introduced to handle singularities in the calculation that cannot be treated perturbatively. These singularities are caused by soft physics effects such as collinear radiation, and are absorbed into the parton distribution functions at a given scale μ_F^2, usually chosen to be of the order of the hard scale probe Q^2.

The evolution of the parton distribution functions (PDFs) with respect to the factorization scale is determined in pQCD, by the DGLAP (Dokshitzer, Gribov, Lipatov, Altarelli, Parisi) [22] equations:

$$\frac{\partial f_i(x, \mu_F^2)}{\partial \ln \mu_F^2} = \sum_{j \in \{q,g\}} \int_x^1 \frac{dz}{z} P_{ij}(z, \alpha_S(\mu_F^2)) f_j(x/z, \mu_F^2)\,, \tag{2.3.3}$$

where $P_{ij}(z, \alpha_S(\mu_F^2))$ are the Altarelli-Parisi splitting functions, describing the probability of the splitting of a parton of kind j into a parton of kind i carrying a fraction z of the longitudinal momentum of j; they have perturbative expansions:

$$P_{ij}(z, \alpha_S(\mu_F^2)) = P_{ij}^{(0)}(z) + \frac{\alpha_S(\mu_F^2)}{2\pi} P_{ij}^{(1)}(z) + \cdots \tag{2.3.4}$$

Expressions for the leading order (LO) and next-to-leading order (NLO) splitting functions can be found in Ref. [23].

---

[12] In 1969 Richard Feynman introduces the name *parton* as a generic description for any particle constituent within the proton and other hadrons [20].

The DGLAP equations determine the μ dependence of the PDFs. The x dependence, on the other hand, is determined by fitting a number of datasets (from deep inelastic scattering and hadron colliders experiments). Several PDFs parametrizations are available[13] and the most widely used come from the CTEQ and MRST/MSTW collaborations [24–27]. The PDFs of valence quarks, gluon and sea quarks from the MSTW group are shown in Figure 2.4 as a function of the momentum fraction for two values of transferred momentum Q^2 at which the proton is probed.

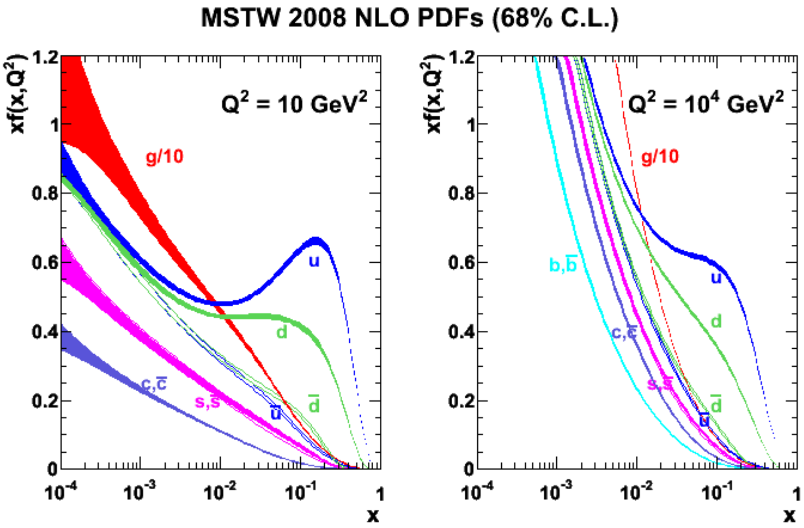


FIGURE 2.4: The MSTW 2008 Proton Parton Distribution Functions (PDFs) at transfer momentum Q^2=10 GeV^2 (Q^2=10^4 GeV^2) on the left (right) [134].

2.3.3 Fragmentation Functions

Physically, an energetic parton fragments ("showers") into many further partons, which then, on later timescales, undergo a transition to hadrons ("hadronization"). Since the parton-hadron transition is non-perturbative, it is not possible to perturbatively calculate quantities such as the energy-spectra of specific hadrons in high-energy collisions [21]. However, one can factorize perturbative and non-perturbative contributions via the concept of fragmentation functions (FFs). These are the final-state analogue of the parton distribution functions that are used for initial-state hadrons. The fragmentation function $D_f^h(z, \mu_f^2)$ gives the probability that parton f will produce the final state

---

[13] See for more information the Les Houches Accord PDFs (LHAPDF): `http://lhapdf.hepforge.org/`.

particle h with momentum fraction z during the fragmentation process; μ_f^2 is the fragmentation scale analogous to the factorization scale μ_F^2. Since the fragmentation functions are non perturbative objects, they cannot be computed from first principles and they need to be extracted by fitting the experimental data of different kind of processes. However, the factorization scale dependence of the fragmentation functions can be obtained in pQCD: like parton distribution functions, they satisfy a DGLAP evolution equation.

Renormalization scale, Factorization scale and Fragmentation scale

A typical QCD calculation involves up to three arbitrary scales: the renormalization scale, the factorization scale and the fragmentation scale. There is no compelling reason why all three scales must be exactly the same. However, the scales should not be chosen to be very different from each other because this would introduce an unphysical hierarchy into the problem. A common choice is to set all three scales equal to the transverse momentum, p_T, of one of the final state objects (particle or jet) which is observed by the experiment [28].

Moreover, if a given calculation were carried out to all orders in perturbation theory, then the final result would not have any dependence on this scale. By truncating the perturbation series at a fixed order predictions exhibit a residual dependence on the scale which implies an uncertainty on the predictions due to the arbitrariness of the scale choice. This uncertainty will be $\mathcal{O}(N+1)$, i.e. of the same order as the neglected terms [21]. For this reason it is customary to use QCD predictions' scale dependence as an estimate of the uncertainties due to neglected terms. Uncertainties are then commonly determined by varying μ by a factor of two up and down around the central scale choice.

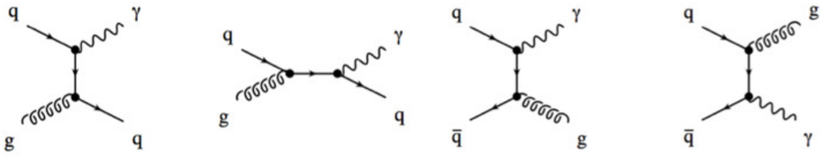
2.4 Prompt Photon Production

The main motivation for the photon measurements is that "direct photon" emerges unaltered from the hard scattering and due to electromagnetic interaction with quarks provides direct probe of the hard scattering dynamics. In addition, photons are free from complications caused by jet fragmentation and systematics caused by their identifications and measurements. Photons may provide direct information on gluon PDF [48], being gluon involved at Leading Order (LO) in contrast to deep inelastic scattering (DIS) processes.

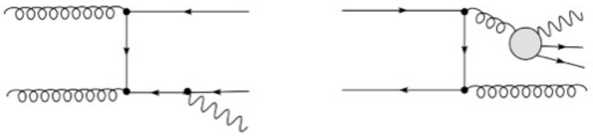
They are also good source for testing gluon resummation techniques and various models of gluon radiation.

In addition, prompt photon production is a major background for important SM processes such as Higgs boson decaying into photon pairs ($H \to \gamma\gamma$) and in searches for new physics with signatures containing photons [29–31].

Prompt photons are defined as photons produced in the beam particle collisions and not originating from particle decays. They include both "direct photons", which originate from the hard interaction, and "fragmentation photons", which arise from the parton fragmentation[14]. Figure 2.5 shows some examples of Feynman diagrams of prompt photon production.



Compton Annihilation



Fragmentation

Figure 2.5: In the upper part, Feynman diagrams of leading order ($O(\alpha\alpha_S)$) *direct photons* production processes in pQCD: (1) quark-gluon Compton scattering $gq(\bar{q}) \to \gamma q(\bar{q})$, (2) quark-antiquark annihilation $q\bar{q} \to g\gamma$. In the bottom part, examples of Feynman diagrams of *fragmentation photons* production. The left one is the point like fragmentation of a quark into a photon, which can be calculated perturbatively for asymptotically large scales. The right one is the non-perturbative fragmentation of a gluon producing a photon.

The term *direct photon* refers to those photons which are produced in the hard-scattering subprocess and are not decay products of some particle. Figure 2.5

---

[14] However, application of photon isolation requirements substantially reduce the contribution of fragmentation photons.

shows the two-body subprocesses which can produce direct photons: the QCD Compton subprocess $gq \to \gamma q$ and the annihilation subprocess $q\bar{q} \to \gamma g$.

These have a Born cross sections of order $\alpha\alpha_S$, because the partonic scattering involves the inelastic emission or absorption of a photon; their cross sections can be written as

$$\frac{d\sigma}{d\hat{t}}(gq \to \gamma q') = -\frac{\pi\alpha\alpha_S}{\hat{s}^2}\frac{e_q^2}{3}(\frac{\hat{u}}{\hat{s}}\frac{\hat{s}}{\hat{u}}) \tag{2.4.1}$$

where $\hat{s} = (q+g)^2$, $\hat{t} = (q-q')^2$, $\hat{u} = (g-q')^2$, and

$$\frac{d\sigma}{d\hat{t}}(q\bar{q} \to \gamma g) = -\frac{\pi\alpha\alpha_S}{\hat{s}^2}\frac{e_q^2}{3}(\frac{\hat{u}}{\hat{s}}\frac{\hat{s}}{\hat{u}}) \tag{2.4.2}$$

where $\hat{s} = (q+\bar{q})^2$, $\hat{t} = (q-g)^2$ and $\hat{u} = (\bar{q}-g)^2$

Depending on the nature of the colliding hadrons and on the values of $\sqrt{s}$ and p_T, either of these two subprocesses can dominate. For photons produced centrally, i.e. $y \sim 0$ in the colliding hadrons' center-of-mass frame, the initial-state partons are probed at a $x \sim x_T = 2p_T/\sqrt{s}$. At medium and large x there is natural hierarchy of parton distributions in the proton, $q \gg g \gg \bar{q}$, while at small x, $g \gg q, \bar{q}$. Thus, for example, in proton-proton collisions the qg Compton process dominates over essentially all the p_T range[15], while in proton-antiproton collisions the Compton process dominates at low p_T and the (valence-valence) $q\bar{(q)}$ annihilation process dominates at high p_T.

Fragmentation photons are produced via the hard radiation from a final-state quark (LO component of the fragmentation contribution) or via the fragmentation of a final-state parton in association with hadronic remnants of fragmentation, including photons produced collinear to the parton momentum (NLO component).

From a technical point of view, the fragmentation contribution comes from the calculation of the higher order corrections in the perturbative expansion in powers of the strong coupling α_S. At higher orders, final state collinear singularities appear in any subprocess where a high-p_T outgoing parton of species k (quark or gluon) undergoes a cascade of successive collinear splittings together with the collinear emission of a photon. The higher order corrections to the cross

---

[15] $q\bar{q}$ annihilation it is suppressed at the LHC as antiquarks must be promoted from the sea.

section can be split into (1) a contribution free from these final state collinear singularities, to be added to the Born term so as to build $d\sigma_{dir}$, and (2) a term $d\sigma_{frag}$ involving these singularities together with accompanying large collinear logarithms. In the fragmentation, the final state collinear singularities and accompanying logarithms can be factorised to all orders in α_S from short distance terms according to the factorisation theorem, and absorbed into fragmentation functions of parton k to a photon $D_k^\gamma(z;\mu_f^2)$. However the splitting of the cross section between $d\sigma_{dir}$ and $d\sigma_{frag}$ is not unique and the $D_k^\gamma(z;\mu_f^2)$ depends on the arbitrary factorization scheme specifying which non-singular parts are factorised together with the collinear singularities; the latter depend in particular on some arbitrary fragmentation scale μ_f. The point-like coupling of the photon to quarks is responsible for the well-known anomalous behaviour of $D_k^\gamma(z,\mu_f)$, roughly as $\alpha/\alpha_S(\mu_f)$, when the fragmentation scale μ_f , chosen of the order of a hard scale of the subprocess, is large compared to $\mathcal{O}(1)$GeV; in this case the photon fragmentation contribution is of the same order $\mathcal{O}(\alpha/\alpha_S(\mu_f))$ as the Born level terms in the direct mechanism.

Schematically, the differential photon cross section in transverse energy E_T and rapidity y can be written as

$$\mathrm{d}\sigma \equiv d\sigma_{dir} + d\sigma_{frag} \sum_{a,b=q,\bar{q},g} \int dx_a dx_b\; f_a(x_a,\mu_F^2) f_b(x_b,\mu_F^2) \times$$

$$\left[d\hat{\sigma}^{\gamma}_{ab}(p_\gamma, x_a, x_b; \mu_R, \mu_F, \mu_f) + \sum_{c=q,\bar{q},g} \int_{z_{min}}^{1} \frac{dz}{z^2} d\hat{\sigma}^{c}_{ab}(p_\gamma, x_a, x_b, z; \mu_R, \mu_F, \mu_f) D_k^\gamma(z;\mu_f^2) \right] \quad (2.4.3)$$

where $f_a(x_a;\mu_F^2)$ is the parton distribution function of parton species a inside the incoming hadrons h at momentum fraction x_a; $D_k^\gamma(z,\mu_f)$ is the fragmentation function of parton k to a photon carrying a fraction z of the parent parton energy (integrated from $z_{min} = x_T \cosh y$, with $x_T = 2E_T/\sqrt{s}$, to 1).

The fragmentation component represents a fraction of the inclusive prompt photon signal which grows with the center-of-mass energy of the collision and it becomes dominant at collider energies. On the other hand, most collider experiments – in particular, the TeV collider experiments CDF and D0 at the Tevatron, ATLAS and CMS at the LHC – measure isolated-photons, because at these energies the inclusive prompt photon signal would be swamped by a large background of secondary photons from hadron decays. The isolation criteria on the hadronic final states of photon candidate events requires that the photon be not accompanied by more than a prescribed amount of hadronic transverse energy in some given cone about the photon (see Section 2.4.1). An analogous, criterion can be implemented in parton level calculations. The isolation cuts do not only suppress the background, they also substantially reduce the fragmentation component.

2.4.1 Isolated Prompt Photons

The main challenge of prompt photon measurements at hadron colliders is the estimation of the signal over a large backgrounds, primarily coming from light mesons (π^0, η, ω) that decay to multiple photons. In order to separate the prompt photon signal from these backgrounds, photon candidates are typically required to be isolated from nearby hadronic activity, which usually accompanies a jet with a leading light meson. One measure of this nearby hadronic activity is the transverse energy in a region around the photon candidate, called the transverse isolation energy ($E_T^{\gamma,iso}$). The isolation region, which set the separation between the photon and the parton, corresponds to the usual definition of a cone or radius $R_0 = \sqrt{\Delta\phi^2 + \Delta\eta^2}$, around the photon, where $\Delta\phi$ and $\Delta\eta$ refer to the difference in azimuthal angle and pseudorapidity from the photon respectively. The transverse isolation energy $E_T^{\gamma,iso}$ corresponds to the sum of the hadronic energy in the transverse direction inside this cone:

$$E_T^{\gamma,iso} = \sum_{had \in R_0} E_T^{had}$$

The isolation cut usually can be defined in different ways: either a cut relative to the photon E_T ($E_T^{\gamma,iso}/E_T^{\gamma} < \epsilon$), or an absolute cut independent of E_T^{γ} ($E_T^{\gamma,iso} < E_{max}$).

Photons produced as part of a fragmentation process also has hadronic remnants. In cases where the cone radius is small and the photon is produced with large z, $E_T^{\gamma,iso}$ can still be small (or zero), and the separation between signal and background is preserved. In cases when the cone radius is large, or when z is small, the value of $E_T^{\gamma,iso}$ will increase. Thus, as already stated the application of an isolation requirement will remove some fraction of the fragmentation component as well as the background.

2.4.2 NLO calculation: the MCFM program

The distinction between the direct and the fragmentation photon production has no physical meaning beyond LO. In fact, from a theoretical point of view, the distinction is defined by an arbitrary choice, which follows from the necessity of factorizing the final state collinear singularities and absorbing them into the fragmentation functions. Several theoretical tools provide predictions for the prompt-photon production calculation such as the JETPHOX [49] and MCFM [50] next-to-leading order Monte Carlo programs. The main advantage of the these MC programs is that one can easily account for any kind of experimental cuts (e.g. on kinematics and/or isolation) implementable at partonic level. In addition, one can easily match the binning of experimental data by histogramming of the generated partonic configurations.

The present analysis predictions relies on the implementation of the photon NLO calculation of both $d\sigma_{dir}$ and $d\sigma_{frag}$ in the MCFM program. MCFM is a **M**onte **C**arlo program that computes the theoretical cross sections for a variety of **F**e**M**tobarn level processes at hadron colliders. Full documentation for the program is available at Ref. [50].

In the MCFM MC, prompt photon production processes are encoded with two options:

- 280 (NLO+F): $f(p_1) + f(p_2) \to \gamma(p_3) + f(p_4)$
- 282 (LO): $f(p_1) + f(p_2) \to \gamma(p_3) + f(p_4) + f(p_5)$

where $f(p_i)$ is a generic partonic jet. Processes denoted as "LO" may only be calculated in the Born approximation. For photon processes, "NLO+F" signifies that the calculation may be performed both at NLO and also including the effects of photon fragmentation and experimental isolation. For the prediction presented in this thesis the process `280` has been used.

Since this process includes a real photon, the cross section diverges when the photon is very soft or in the direction of the beam. In order to get a well-defined cross section, the input file must supply the values of $p_T^{\min}$ and $|y|^{\max}$ for any photons produced in the process; they are specified by the `ptmin_photon` and `etamax_photon` parameters.

The isolation cut is defined by the cone radius R_0, specified with the `cone_photon` and `epsilon_h` parameters. If `epsilon_h` < 1 then the photon is isolated using $\sum_{\in R_0} E_T(\text{had}) < \epsilon_h\, p_T^{\gamma}$. Otherwise `epsilon_h` > 1 sets E_T^{max} in $\sum_{\in R_0} E_T(\text{had}) < E_T^{max}$.

For this analysis the final state is defined by the following basic set of cuts on the photons,

- $p_T^{\min} = 30$ GeV
- $|y|^{\max}| = 1.0$
- $R_0 = 0.4$
- $E_T^{max} = 2$ GeV

The user can also select the desired option for the process:

- `virt`. Virtual (loop) contributions to the next-to-leading order result are calculated (+counterterms to make them finite), including also the lowest order contribution.
- `real`. In addition to the loop diagrams calculated by `virt`, the full next-to-leading order results must include contributions from diagrams involving real gluon emission (-counterterms to make them finite). Note that only the sum of the `real` and the `virt` contributions is physical.
- `tota`. This option runs the *virtual* and *real* real terms in series before performing a sum to obtain the full next-to-leading order result. For photon processes that include fragmentation, it also includes the calculation of the fragmentation contributions.

The option `tota` has been used.

The value of $\alpha_S(M_Z)$ is hardwired with the parton distribution. Furthermore, the parton distribution specifies the number of loops that should be used in the running of α_S. The default mode of operation can be choose from a collection of modern parton distribution functions that are included with MCFM. Some PDF sets, together with their associated $\alpha_S(M_Z)$ values, are given in Table 2.1.

TABLE 2.1: Some example of the available MRS- and CTEQ- type pdf sets and their corresponding values of $\alpha_S(M_Z)$.

PDF set	$\alpha_S(M_Z)$
`mstw8lo`	0.1394
`mstw8nl`	0.1202
`ct10.00`	0.118
`cteq66m`	0.118

The NLO predictions are obtained using MSTW08NL PDFs, for which $\alpha_S(M_Z) =$ 0.12018. Although each PDF set uses a slightly different reference value, the resulting differences in the γ cross sections are very small. This can be seen in Figure 2.6 showing the ratio data/theory calculated with two choices of PDFs: CTEQ6.6M and MSTW08NL. On average the CTEQ6.6M prediction is a few per cent higher than the MSTW08NL distribution.

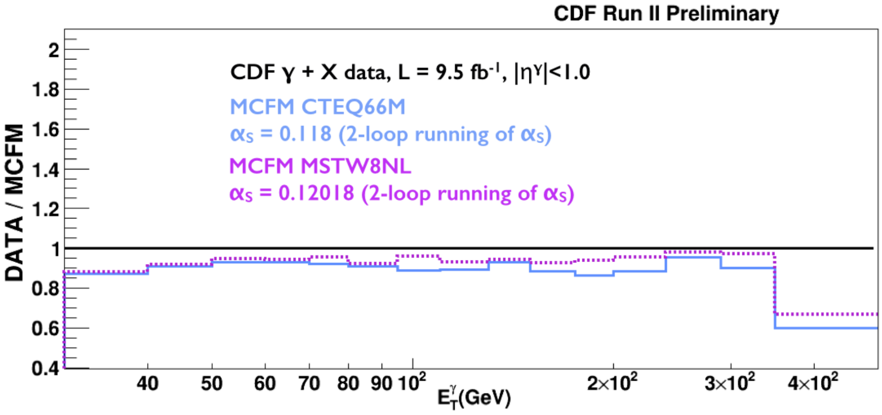


FIGURE 2.6: Ratio of isolated photon spectra in measured in this analysis over the MCFM predictions for two PDFs set: MSTW08NL (default choice for this analysis) and CTEQ6.6M. Although each PDFs set uses a slightly different reference value, i.e. $\alpha_S(M_Z)$ =0.12018 for MSTW08NL and $\alpha_S(M_Z) = 0.118$ for CTEQ6.6M), the resulting differences in the obtained γ cross sections is very small: on average the CTEQ6.6M prediction is a few per cent higher than the MSTW08NL distribution.

The MCFM program has several fragmentation functions sets implemented: Bourhis, Fontannaz and Guillet FFs (`BFGSet_I`, `BFGSetII`) [51] and those proposed by Gerhrman-De Ridder and Glover (`GdRG_LO`) [52]. Gerhrman-De Ridder and Glover calculated the fragmentation functions using a fixed order expansion[16].

Figure 2.7 shows the contributions of each one of the three subprocesses (Compton, annihilation and fragmentation) to prompt photon production at the Tevatron as a function of the photon E_T for the central pseudorapidity region ($|\eta^\gamma| < 1.0$). At the Tevatron collider, quark-gluon scattering is the dominant component up to $\sim$ 120 GeV, beyond which the annihilation of valence (anti)quarks from the (anti)proton beams plays a preeminent role. The single contributions have been obtained by selecting the corresponding Feynman diagrams at NLO with the MCFM Monte Carlo, setting all scales to the photon transverse energy ($\mu_R = \mu_F = \mu_f = E_T^\gamma$) and using the MSTW08NL parton densities and the GdRG parton-to-photon fragmentation functions.

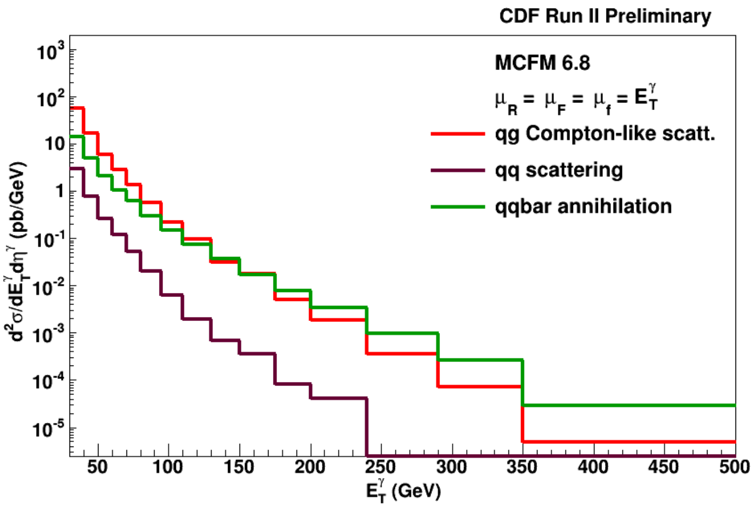


FIGURE 2.7: Contributions of the quark-gluon Compton, $q\bar{q}$ annihilation and fragmentation subprocesses in NLO isolated photon production at the Tevatron, obtained with the MCFM calculation ($\mu = E_T^\gamma$, MSTW08NL PDFs and GrDG FFs).

---

[16] Using the current isolation conditions the LO FFs are sufficient to remove the collinear singularity.

2.4.3 Previous measurements

In hadron colliders, the first direct photon data came from the CERN Intersecting Storage Rings (ISR) pp collider, followed by the $Sp\bar{p}S$ collider [32–35]. More recent prompt photon measurements have been performed at the Fermilab Tevatron collider, by CDF and D0 using $p\bar{p}$ collisions data collected at a center of mass energy $\sqrt{s}$ = 1.8 TeV and $\sqrt{s}$ = 1.96 TeV [36–38] and at the CERN Large Hadron Collider, by ATLAS and CMS using pp collisions data at $\sqrt{s}$ = 7 TeV [39–43].

Figure 2.8 shows the results from D0 and CDF Collaborations at a center-of-mass energy $\sqrt{s}$ = 1.96 TeV, compared with the theoretical predictions from Jetphox:

- D0 measured cross section is based on an integrated luminosity of 326 pb^{-1} and covers a pseudorapidity range $|\eta^{\gamma}|$ <0.9 within the transverse energy interval 23< E_T^{γ}(GeV) <300.
- CDF measured cross section is based on an integrated luminosity of 2.5 fb^{-1} and covers a pseudorapidity range $|\eta^{\gamma}|$ <1.0 within the transverse energy interval 30< E_T^{γ}(GeV) <400.

Figure 2.9 shows the results from the CMS and ATLAS Collaborations:

- CMS results are based on an integrated luminosity of 36 pb^{-1}, covering the pseudorapidity range $|\eta^{\gamma}|$ <2.5 and the transverse energy range 25< E_T^{γ}(GeV) <400. Comparison is made with respect to the Jetphox predictions.
- ATLAS results are based on an integrated luminosity of 4.6 fb^{-1} made measurements in the pseudorapidity regions $|\eta^{\gamma}|$ <1.37 and 1.52≤ $|\eta^{\gamma}|$ <2.37 in the kinematic range 100≤ E_T^{γ}(GeV) <1000. A comparison of the measured cross section is made with respect to Pythia, Herwig and Jetphox theoretical predictions.

In the CDF and CMS measurements, Jetphox predictions are corrected for a scale factor in order to take into account underlying event and parton fragmentation, which give additional contributions to the energy in the isolation cone at the particle level. The scale factors are computed using Pythia Monte Carlo samples with multiple-parton interactions and/or hadronization turned off, by comparing the measured cross section in such samples with the cross section measured in samples simulated under nominal conditions.

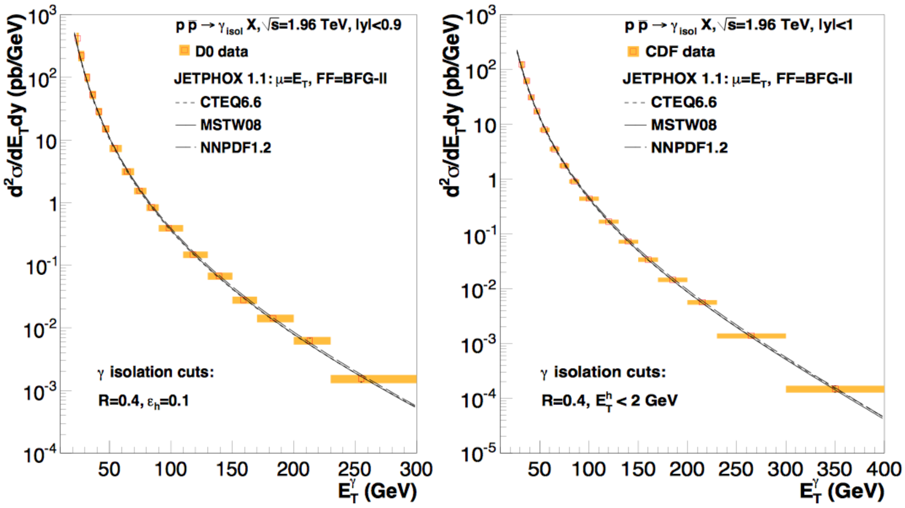


Figure 2.8: Isolated photon E_T-differential cross section at mid-rapidity in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV measured by D0 (left) and CDF (right) compared to Jetphox predictions.

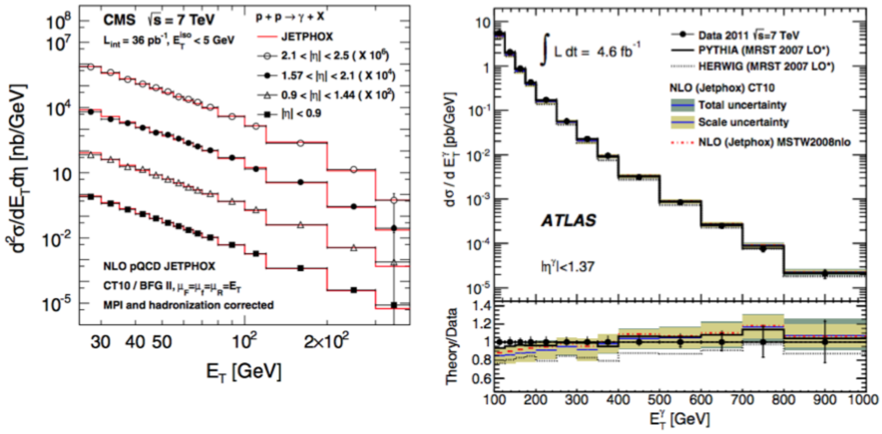


Figure 2.9: Isolated photon E_T-differential cross section pp collisions at $\sqrt{s} = 7$ TeV measured by CMS (left) and ATLAS (right) compared to theoretical predictions.

Chapter 3

Accelerator Complex

The Tevatron in Batavia, Illinois, US was the first large-scale superconducting synchrotron in the world. Originally named the Energy Doubler since, as a proton synchrotron, it was reaching twice the energy of the original Fermilab facility (the "Main Ring"), it began operating in 1983 in fixed target mode and in 1985 as a proton-antiproton collider. From 1985 to 2011 periods of colliding protons and anti-proton alternated with periods of inactivity (shut downs) for upgrading the machine. From 2001 to end of the operations on September 30, 2011, the $p\overline{p}$ *energy in center of mass system was 1.96 TeV.*

In this chapter the proton and anti-proton production is briefly presented, along with the accelerator complex eventually bringing protons and anti-protons to collide at the above center of mass energy. Finally, the accelerator-complex performances are highlighted.

Contents

3.1 Proton production and first step in the acceleration

Hydrogen gas is introduced in a container, where strong ionization induced by pulsed electric field in presence of a containing magnetic field produces occasionally a number of negative ions. These are used for accelerating beam up to GeV energy with negligible electron-capture losses. A pulsed electrostatic extractor then accelerates the negative ions out of the source at a repetition rate up to 15 Hz and at a energy of 15-22 KeV . After that, ions are further accelerated by a *Cockroft-Walton electrostatic accelerator* to an energy of 750 KeV. Ions, segmented into bunches, are fed into the *Linear Accelerator (Linac, [53])*. The Linac (Figure 3.1) is approximately 150 m long and comprises two sections. In the first one, five accelerating cavities with a drift tube inner core, fed by a single RF generator resonating at 201.25 MHz, accelerate ions to approximately 116 MeV. The second one, comprising 7 RF cavities at 805 MHz, fed by a set of Klystron amplifiers, ramps ions to 401.5 MeV. At the Linac exit the negative ion beam strikes a thin carbon target and turns into a proton beam by electron stripping.

Stripped protons enter the *Booster* (Figure 3.1), a 8 GeV synchrotron whose diameter is about 150 m. To maintain a constant circular orbit the dipole magnetic field in the Booster increases from 0.74 Tesla to 7 Tesla during acceleration. To reach 8 GeV the injected protons circulate in the booster for about 33 milliseconds.

3.2 Anti-proton production and accumulation

The *Main Injector* ([54]) is the next link in the accelerator chain. In accumulation mode, a pulse of 8 $\times 10^{12}$ protons is extracted from the Booster and injected every 2.2 seconds. The main injector is a 53.1 MHz circular synchrotron of a 528.5 m radius, with 18 accelerating cavities and conventional magnets. The protons are accelerated to 120 GeV and then directed to the *anti-proton station*, which is a rotating 7 cm-thick target made of nickel alloys containing chromium, iron and other metals. The resulting particles spray contains some anti-protons with a broad momentum and wide-spread spatial distribution. A cylindrical lithium lens

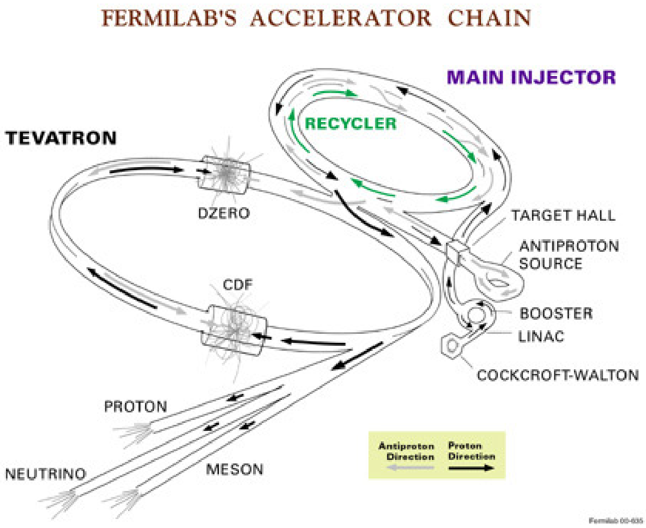


Figure 3.1: *The accelerator complex of the Fermi National Accelerator Laboratory.*

($760\ T/m$) focuses the particles produced around the forward direction. Negative particles in a 35 mrad cone about the forward direction are selected by a 1.5 T pulsed dipole magnet, focused by strong magnetic lenses and injected in the Debuncher Storage Ring. Typically, 21 anti-protons per 106 protons on target are collected. Anti-protons are at 8 GeV energy. The stacking rate is approximately $10 - 20$ mA/hour.

In the *Debuncher ring*, a rounded triangular-shaped synchrotron with mean radius of 90 meters stochastic cooling and bunch phase rotation are used to reduce momentum spread while increasing time spread. After each beam pulse the Debuncher is emptied. The anti-proton bunches (8 ± 0.018 GeV) are transferred with a 60%-70% efficiency to the *Anti-proton Accumulator*, a 75 m mean radius storage ring of larger acceptance housed in the same tunnel as the Debuncher (see a sketch in Figure 3.2). In the Accumulator multiple beam pulses are stacked and $\overline{p}$ are further cooled to increase the anti-proton phase space density.

Besides the small anti-protons production cross section, problems in anti-proton collection, cooling and stacking are among the main causes limiting the final Tevatron luminosity. A further improvement of the anti-proton source is the *Recycler* (see Figure 3.1), a post-accumulator storage ring of constant 8 GeV energy, located in the Main-Injector enclosure and composed of permanent magnets. Because of the larger acceptance of the Recycler (it can store an anti-proton current up to over 2.5 Amps, much larger than the Accumulator), its

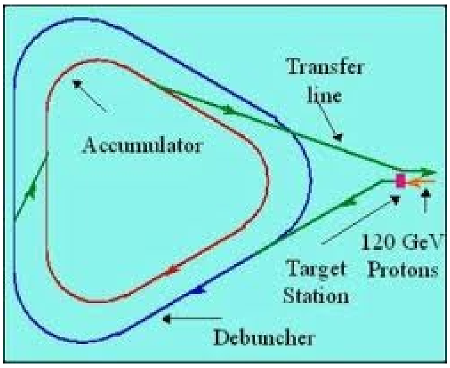


Figure 3.2: *Fermilab Debuncher, Accumulator: a zoom of Figure 3.1.*

role is to store anti-protons, which are periodically transferred from the Accumulator (95% transfer efficiency). Anti-protons are also further cooled to increase the storing capacity of the recycler and the beam transfer efficiency to the Tevatron.

3.3 Injections and collisions

In normal conditions every 10-20 hours the recycler was fully loaded and anti-proton accumulation was stopped. Protons from the Booster were injected into the Main Injector, accelerated to 150 GeV, coalesced into single bunches of about 300×10^9 protons, and then injected into the *Tevatron*, a large synchrotron of 1 Km radius. The entire process is repeated until 36 bunches are transferred with a timing separation of 396 ns from each other. Typically, the transferring efficiency from the Main Injector to the Tevatron is about 65%.

After the protons are loaded, 7-11 anti-proton bunches are extracted from the either Accumulator or Recycler to the Main Injector, accelerated to 150 GeV, coalesced into four 30×10^9 $\overline{p}$ separated by 396 ns, and then injected into the Tevatron. The anti-proton injection process is repeated until 36 anti-proton bunches circulate in the Tevatron.

Protons and anti-protons circulate in the same vacuum pipe. Electrostatic separators reduce to a negligible amount the unwanted interactions, by keeping the beams away from each other at all points in the orbit helix[1], except at the collision points. Protons and anti-protons are accelerated to 980 GeV . A tour of

---

[1] Intra-beam distance is typically 5 times the sum of the beam widths (in a Gaussian approximation)

the Tevatron takes about 21 μs. About one minute is needed to reach the final beam energy[2].

High-gradient focusing quadrupole magnets ("low-β squeezers") reduce the transverse spatial spread to minimize the beam section at the interaction regions and therefore maximize the collision rate. Interactions regions are located where the D0 and the CDF II detector are placed. The resulting transverse beam distributions are approximated by 2D Gaussian functions, with $\sigma = 30\ \mu m$. The typical longitudinal dimension of a bunch is 60-70 cm. The event source is roughly distributed longitudinally as a Gaussian with $\sigma_z = 28$ cm[3].

The 36 bunches of protons or anti-protons are collections of buckets adding up to 1113 in three equispaced trains (see Figure 3.3). Within a train the inter-bunch time is 396 ns (21 buckets) while inter-train time is 2.6 μs (139 buckets). The intra-train empty sectors allow for anti-proton injection without perturbing the orbiting protons and allow enough time for fast kicker magnets to abort the beam into a dump before the arrival of the next train in case of emergency. As a consequence of this configuration, the average bunch crossing rate is 1.7 MHz.

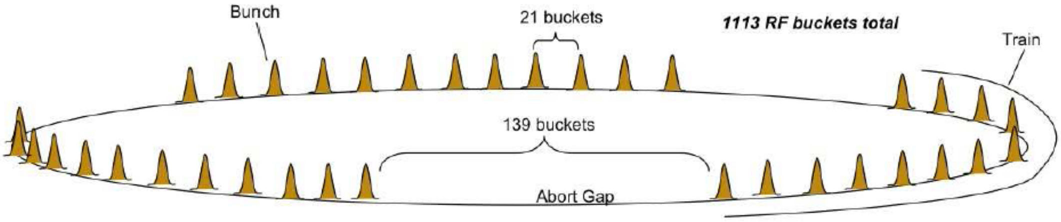


FIGURE 3.3: *Bunch structure of the Tevatron beams in Run II*

The transverse beam profile is shaped by a number of collimators to avoid detector damages from the tails of the proton or anti-protons interacting with the beam pipe (e.g: beam halo). When the beam profile is narrow and the condition are stable, the detectors are powered and data taking can start.

---

[2] The Tevatron comprises about 1000 superconducting magnets including 772 dipoles. Each dipole is approximately 6 m in length and 4 tons in weight. The superconducting coils are made up of niobium-titanium wires embedded in copper. A 4400 A current in the dipoles provides a 4.2 T magnetic field. All superconducting magnets are kept at 4 K temperature

[3] The interaction region is approximately Gaussian in z with $\sigma \sim 28$ cm, as determined by the overlap of the two approximately longitudinally Gaussian bunches.

3.4 Performance

Performance of the Tevatron is qualified primarily by two parameters:

- center of mass energy ($\sqrt{s}$);
- integrated luminosity ($\int \mathcal{L}$dt) in a given time period.

The center of mass energy determines the reach in search for new beyond-the-standard-model phenomena. The latter is directly proportional to the number of events (N) of a given process with a cross section σ

$$N[events] = \int \mathcal{L} dt[cm^{-2}] \times \sigma[cm^2] \tag{3.4.1}$$

In the absence of a crossing angle or position offset, the luminosity in the Tevatron is given by the expression below ([55]):

$$\mathcal{L} = \frac{fBN_pN_{\overline{p}}}{2\pi(\sigma_p^2 + \sigma_{\overline{p}}^2)} F(\sigma_l/\beta^*) \tag{3.4.2}$$

where f is the revolution frequency, B is the number of bunches in each beam, N_p ($N_{\overline{p}}$) is the number of protons (anti-protons) in a bunch, σ_p ($\sigma_{\overline{p}}$) is the rms proton (anti-proton) beam size at the interaction point, and F is a form factor that depends on the ratio of the bunch length, σ_l, to the beta function at the interaction point, β^*. The values of the above parameters are reported in Ref. [55]. It can be shown that fundamental limitations on the Tevatron luminosity are due to N_p/ϵ_{N_p} and $BN_{\overline{p}}$, ϵ_{N_p} being the normalized transverse emittance containing 95% of the proton/anti-proton beam ([55]). Higher luminosities were achieved over time thanks to the anti-proton stack rate.

$$\mathcal{L} \propto \frac{3\gamma f\xi(BN_{\overline{p}})}{\beta^* N_{IR}(1 + \frac{\epsilon_{N_{\overline{p}}}}{\epsilon_{N_p}})} F(\sigma_l/\beta^*) \tag{3.4.3}$$

where:

- $\epsilon_{N_p/N_{\overline{p}}}$ represents the normalized transverse emittance containing 95% of the proton/anti-proton beam;

- N_{IR} is the number of interacting regions;

- ξ is the total head-on beam-beam tune shift seen by the anti-protons[4]:

$$\xi = \frac{r_0}{4\pi}\frac{N_p}{\epsilon_{N_p}}N_{IR} = .000733\frac{Np}{\epsilon_{N_p}}N_{IR} \tag{3.4.4}$$

r_0 being the classical radius of the proton.

Figures 3.4, 3.5 show respectively the instantaneous luminosity at the beginning of each data taking period[5], and the integrated luminosity ($\mathcal{L}dt$), as a function of time: the constant progress in the performances of the machine leads to increased instantaneous and integrated luminosities. Blank parts in the figures correspond to periods of time when the Tevatron was not running. In total Tevatron delivered $\sim$ 12 fb^{-1} to the CDF and D0 experiments. CDF acquired about 85% of it ($\sim$ 10 fb^{-1}), because of inefficiencies in the detector, and dead time due to a number of reasons related to operations (e.g: detector calibrations, etc.).

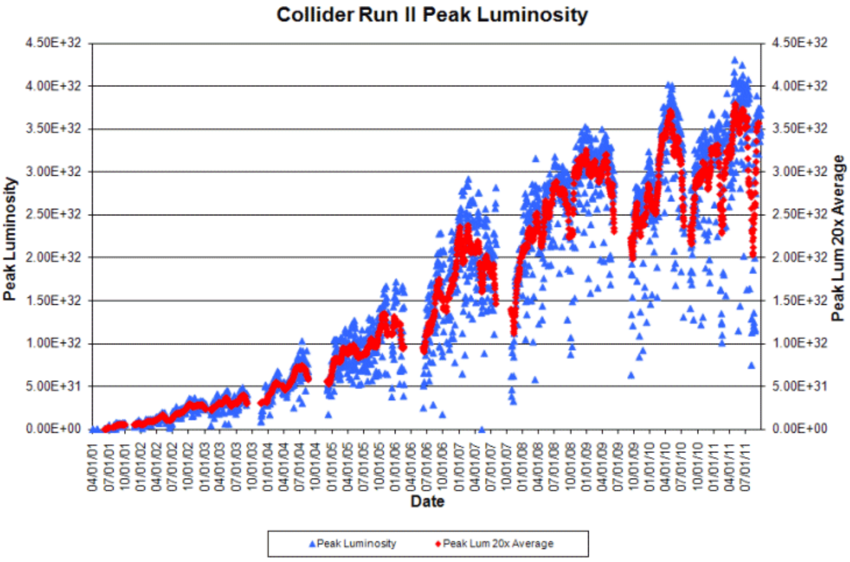


FIGURE 3.4: *The peak instantaneous luminosity during each store (blue triangles) and as average of 20 subsequent stores (red diamonds) over time in the Run II of the Tevatron.*

---

[4] The numerical expression on the right of Equation 3.4.4 is evaluated with N_p in units of 10^9 and ϵ_{N_p} in units of π mm-mrad.

[5] A continuous period of collider operation using the same collection of protons and anti-protons is named store

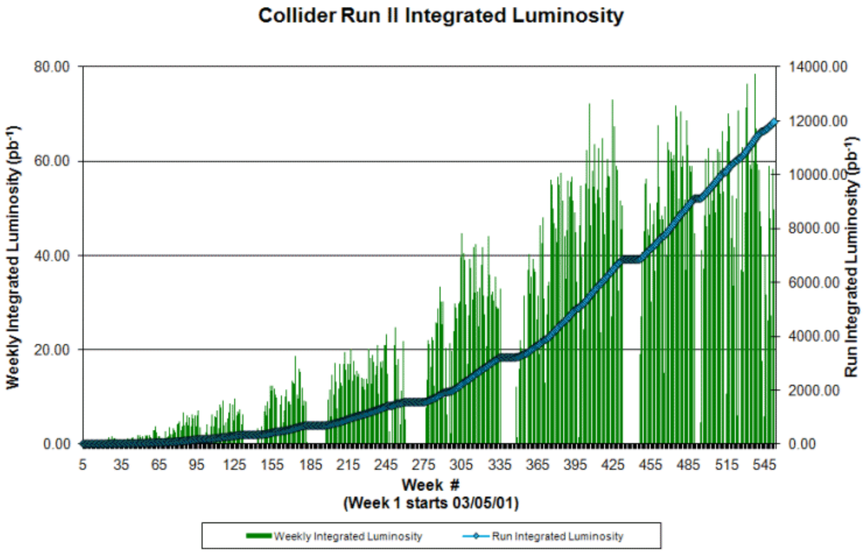


Figure 3.5: *The luminosity integrated during each week (green bars) and in total (cyan diamonds) during the Run II of the Tevatron.*

Chapter 4

The Collider Detector at the Fermilab Tevatron

The Collider Detector at Fermilab (CDF) was designed to study $p\bar{p}$ collisions at the Tevatron. Commissioned in 1985 it was upgraded in 1989 and again in 2001 in order to operate at the expected increased instantaneous luminosity of the accelerator. Each upgrade is considered a "Run". Run 0 was the run before any upgrades, Run I was after the first upgrade and Run II was after the second upgrade. In this section the CDF II upgraded detector is described. Further details are available at Ref. [56].

Contents

4.1 Coordinates

A right-handed Cartesian coordinate system with origin in the B0 nominal interaction point is used for the CDF II detector. The positive z-axis is parallel to the nominal beam line and points toward the proton direction (east). The y-axis points vertically upward, while the x-axis lies in the same plane as the Tevatron and points radially outward with respect to the center of beam (see Figure 4.2). Due to the geometry of the infrastructure (CDF II detector plus the beams from the Tevatron) there is an obvious cylindrical symmetry around the beam line axis. A cylindrical coordinate system ($r \equiv \sqrt{x^2 + y^2}$, $\phi \equiv \tan^{-1} x/y$, z) is used to locate a point on the CDF detector. In this coordinate system the z direction and a direction in the (r, ϕ) plane will labeled as "longitudinal" and "transverse" respectively. Sometimes when describing the particle trajectory, rather than using z, it is convenient to use the polar angle θ. The polar angle is defined relative to the z-axis, the azimuthal angle to the x-axis.

Since hadrons are composite particles, hard interactions at the Tevatron happen with an unknown center of mass energy. Therefore, the overall longitudinal momentum in the initial state is unknown on a event-by-event basis. In the transverse plane the interacting partons are almost at rest since the beams are collimated along the z direction. For this reason it is convenient to use variables which are invariant under boosts along the longitudinal direction. Therefore, rather than using the polar angle $\theta \equiv \tan^{-1} r/z$, the *pseudorapidity* is introduced

$$\eta = -\ln(\tan(\theta/2)) \tag{4.1.1}$$

which is the approximate expression of the rapidity in the ultra-relativistic limit. The rapidity is defined as follows:

$$y = \frac{1}{2}\ln(\frac{E + p\cos(\theta)}{E - p\cos(\theta)}) \tag{4.1.2}$$

$$\tag{4.1.3}$$

where E and p are respectively energy and momentum of the considered particle. It can be shown that under a boost to an inertial frame with velocity β_z along the z direction $y \to y^{'} = y + \tanh^{-1}(\beta_z)$, therefore $dy = dy^{'}$, meaning that the rapidity (and thus the pseudorapidity) difference between two physical systems is invariant under a boost along the longitudinal direction.

Since the longitudinal position of event vertex is distributed around the CDF geometrical center with a $\sim$ 28 cm r.m.s width, it is useful to distinguish the detector pseudorapidity, η_{det}, measured with respect to the geometrical center, from the actual pseudorapidity, η, measured with respect to the real interaction point.

Other common variables which are invariant under a boost along the longitudinal direction are the following:

$$\begin{aligned} P_T &= P\sin\theta \\ E_T &= E\sin\theta \\ \Delta R &= \sqrt{\Delta\eta^2 + \Delta\phi^2} \end{aligned} \tag{4.1.4}$$

$$\tag{4.1.5}$$

P_T and E_T are respectively the transverse momentum and transverse energy. ΔR can be thought as the separation between two particles in the (η, ϕ) plane.

4.2 Detector Overview

The Run II Detector (see Figures 4.1, 4.2) is composed of several components, each optimized for a specific task.

Starting from the interaction point and following the path of an outgoing particle within acceptance there are:

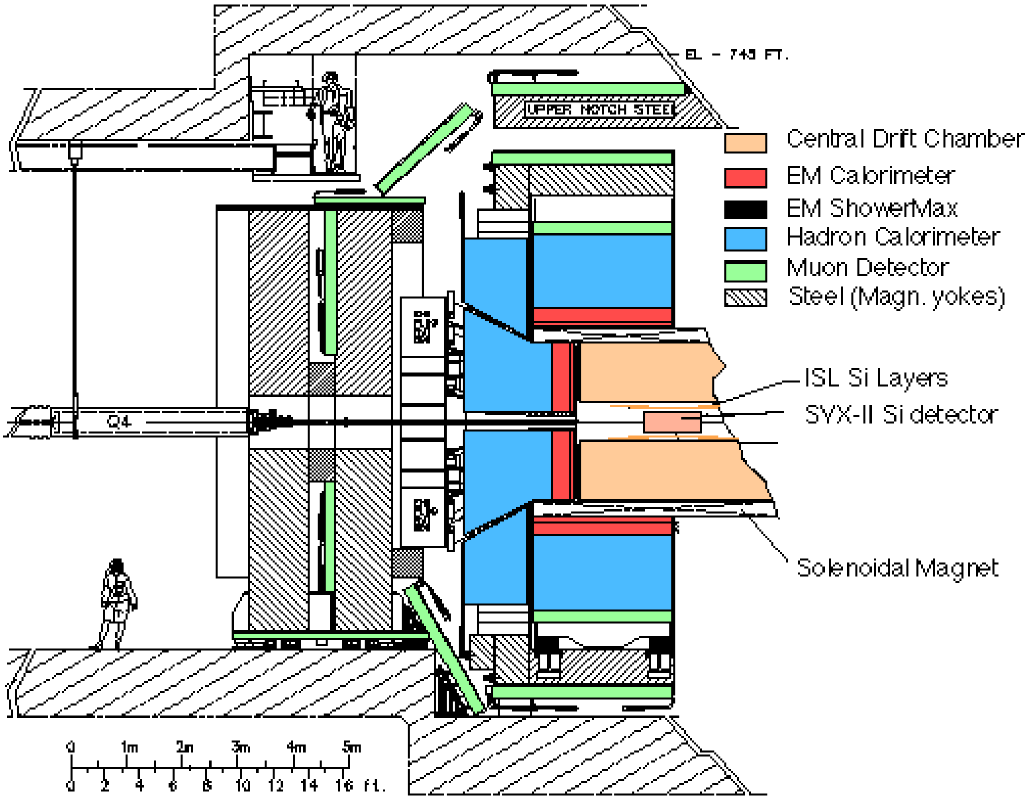


FIGURE 4.1: *Elevation view of the CDF Run II detector.*

- a tracking system enclosed by a superconducting solenoid (1.5 m in radius and 4.8 m in length), which generates 1.4 T magnetic field parallel to the beam axis. The magnetic field is nearly uniform within the tracking volume.
- Electromagnetic and hadronic plastic scintillator calorimeters split into projective towers.
- planar drift chambers backed by scintillation counters (muon detectors).

In the next sections a number of CDF II sub-systems will be discussed. Some of the components (the time-of-flight detector, etc.) of the detector have been neglected since they are not directly related with the topic of this thesis. A detailed description of the upgraded detector can be found in [56].

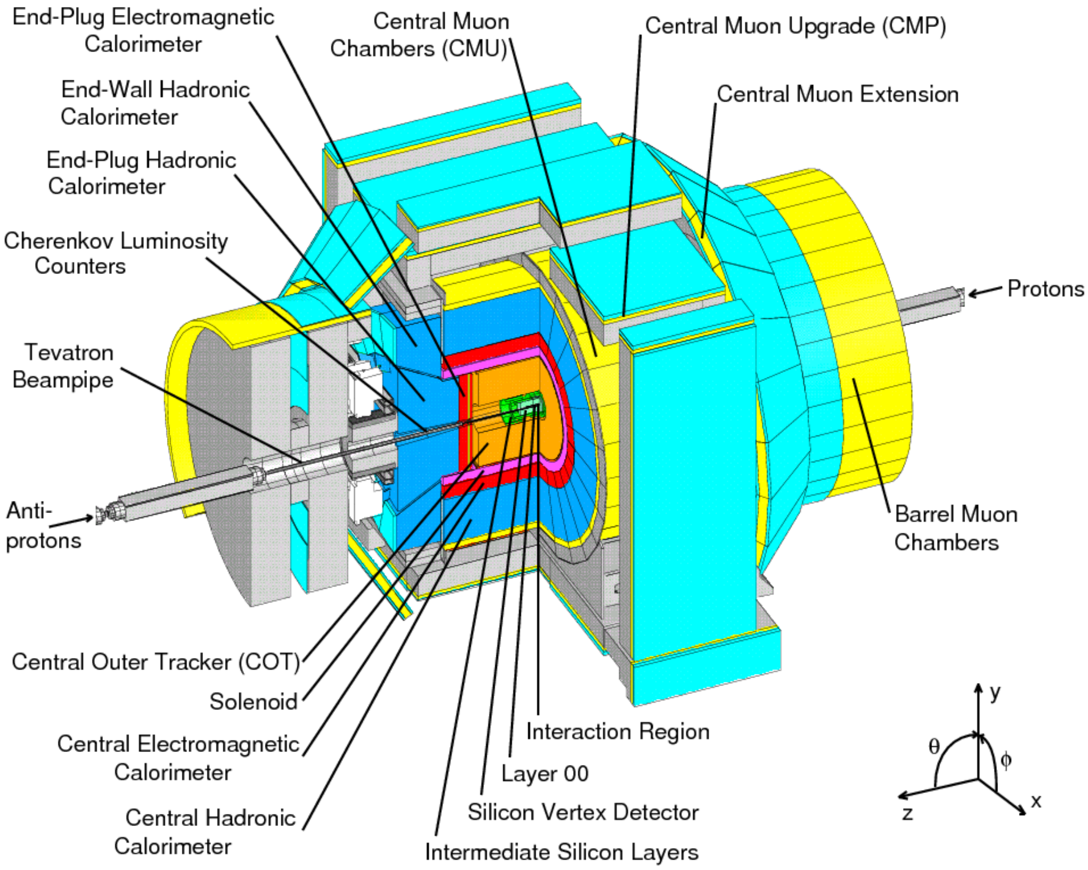


FIGURE 4.2: *Artistic view of the CDF Run II detector.*

4.3 Tracking System

Direction and momentum of charged particles can be measured through a tracking system consisting of three silicon sub-detectors and a large outer drift-chamber (Figure 4.3). These sub-detectors are contained in a superconducting solenoid, which creates a magnetic field of about 1.4 T over a 1.4 m radial distance. The coil is 0.85 electron radiation lengths thick[1].

The tracker is a two component system, comprising a silicon detector at small radii, and a large open-cell drift chamber at larger radii. Although this system is similar to the one used in Run I, it was upgraded in order to overcome some previous limitations (e.g: length not enough to cover the luminous region, non-optimal number of layers, no stereo or weak stereo layers). It was to some extent over-designed in order to face the expected deterioration with the

---

[1] High-energy electrons predominantly lose energy in matter by bremsstrahlung. The mean distance per unity of density material over which an electron loose all its energy but $1/e$ is called radiation length X_0. X_0 and can be approximated by $\frac{716.4 \cdot A}{Z(Z+1) ln(287/\sqrt{Z})} [g \cdot cm^{-2}]$, where A, Z are the mass and atomic numbers of the nuclei composing the material traversed by the electron.

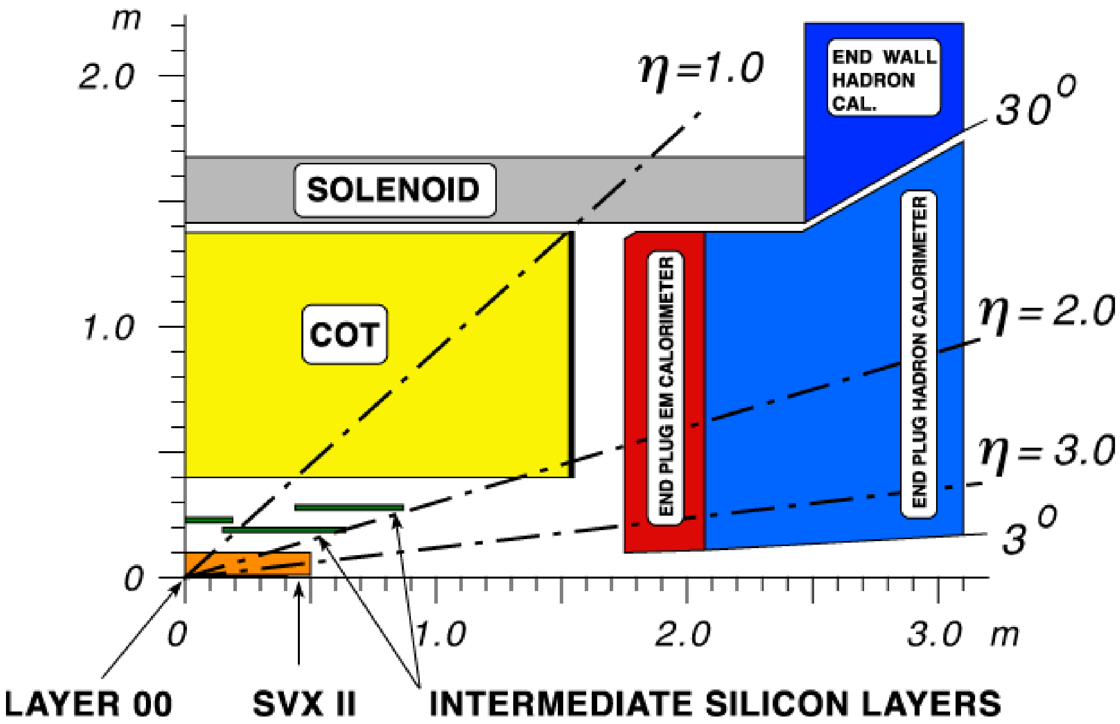


FIGURE 4.3: *Cut-away view along the beam of one quadrant of the Tracking System of the CDF II detector.*

accumulated radiation dose at higher luminosities. The inner detector provides excellent impact parameter, and z resolutions, while the outer detector provides excellent resolution on the curvature, and pseudorapidity. Together they provide an accurate measurement of the azimuthal angle. The two components of the tracking system are described below.

4.3.1 Inner Tracker

With an inner tracker made of silicon, CDF has established the viability and excellent performance of silicon tracking in hadron colliders. The inner-most tracker is also called "vertex detector" in CDF jargon, since its primary role is to provide precise tracking information near the interaction in order to identify displaced tracks, and reconstruct displaced decay vertices produced by long life-time particles[2].

The Inner Tracker is composed of eight layers (seven at $\theta = 90$ of silicon sensors arranged in approximately cylindrical sub-systems coaxial with the beam-pipe:

---

[2] The displaced track identification played a major role in the discovery of the top quark by CDF (see Ref. [57]).

Layer 00 (L00), the Silicon Vertex Detector (SVXII), and the Intermediate Silicon Layers (ISL). Figure 4.4 zooms on the Inner Tracker from two different point of views. All silicon microstrip sensors have a space resolution of 12 μm in the direction transverse to the beam. SVXII and ISL also provide z-measurements with reduced accuracy.

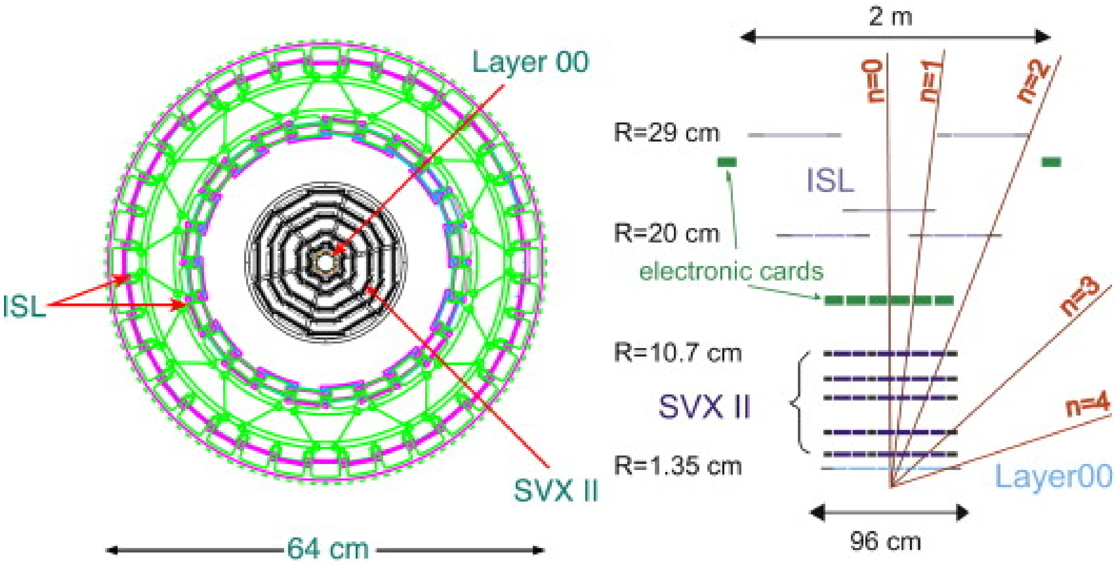


FIGURE 4.4: *Left: cutaway transverse to the beam of the three inner tracker sub-systems. Right: sketch of the silicon detector in a x/y projection.*

- **Layer 00** (see Ref. [58])

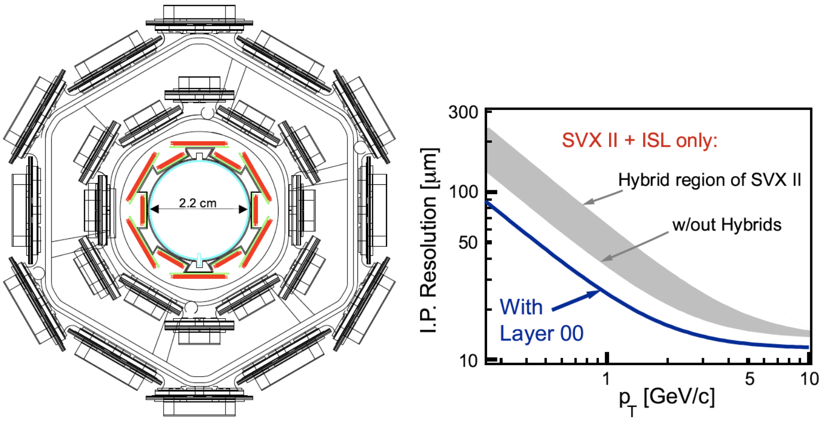
  Layer 00 (L00) is an innovative detector made of a single-sided castellated layer built directly onto the beam pipe. The innermost (128 strip) and outermost (256 strip) sensor layers are located at radii of 1.35 cm and 1.62 cm (Figure 4.5a). The strips are parallel to the beam axis providing samples of the tracks in the (r, ϕ) plane. The implant (readout) pitch is 25 (50) μm. There are 12 sensors along the beam line for a total length of 94 cm. L00 provides full azimuthal coverage and $|z| \lesssim 47$ cm longitudinal coverage. The front-end electronics is located outside the tracking volume to minimize the multiple scattering from inactive material.

  L00 was added in 2001 to the inner tracking system for two reasons.

  - Extend the lifetime of the silicon system: the inner layer of SVXII were expected to have a limited lifetime because of radiation damage. L00 uses radiation hard silicon, which by standing a significantly higher bias voltage can operate at a higher absorbed radiation dose, thus compensating for the SVXII damaged layers.

- Further improve the impact resolution of the tracking system (Figure 4.5b): SVXII readout electronics is located inside the tracking volume, thus degrading the impact resolution because of multiple scattering. This effect is more enhanced for low-momentum particles. Having a minimal silicon material at smaller radii provides a precision measurement on the impact parameter which helps recovering the lost resolution.

During data taking significant noise was observed in L00. Such a noise was characterized by non-uniform pedestals across the strip sensors. The noise was varying event-by-event. It was decided to employ an offline event-by-event pedestal subtraction. Such a procedure was tested in Monte Carlo $p\bar{p}$ and it was found a 95% efficiency with a 95% purity. However, because of the needed pedestal subtraction and of the slow readout (∼30 KHz), L00 could not be used in the online triggers.



(A) (B)

FIGURE 4.5: *End-view of the layer 00 of the CDF inner detector (left). This layer is mounted directly on the beam pipe (cyan). The expected impact parameter resolution of the reconstructed track with and without the layer 00 is also shown as a function of the track P_T (right).*

▶ Silicon Vertex Detector [56], [59]

The silicon vertex detector (SVXII) is built in three cylindrical barrels (Figure 4.6) with a total length of 96 cm, providing coverage over $|\eta_{det}| \leq 2.0$ from the detector center. Each barrel supports five layers of double-sided microstrip detectors at radii between 2.4 *cm* to 10.7 *cm*. SVXII has a cylindrical geometry coaxial with the beam: twelve 30°

azimuthal sectors (“wedges”) provide full coverage over ϕ. A small overlap between adjacent wedges is present (Figure 4.6).

All five layers provide $r - \phi$ measurement on one side. Three out of the five layers combine such a measurement with a 90° stereo measurement on the reverse side. The remaining two layers provide a small angle stereo at 1.2° on the reverse side. All layers are arranged into twelve concentric independent readout units (“ladders”). A total of 405,504 channels are used for SVXII. Water, cooling the channels, is carried between the ladders.

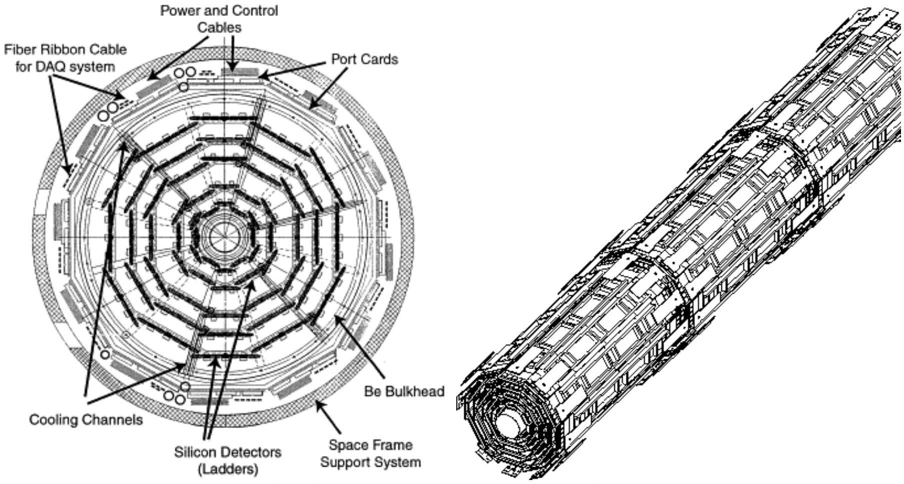


FIGURE 4.6: *End-view (left) and isometric (right) of the CDF Silicon Vertex Detector.*

This sub-detector has a 12 μm resolution on the single hit in the direction transverse to the beam, and provides also some dE/dx information.

With respect to the progenitor SVX in Run I [60], SVXII features an increased length along the z direction. The increase length allows covering 2.5 σ of the luminous region, thus increasing the geometrical and angular acceptances for single tracks. This improvement was designed to enhance the efficiency in the reconstruction of displaced tracks from heavy meson decays.

- **ISL** [61]

  The main ISL purpose is to compensate for incomplete coverage of other sub-detectors in the region $|\eta_{det}| > 1$ by providing precision tracking at $1 < |\eta_{det}| < 2$. Together with SVXII ISL provides 3D tracking information. ISL is placed between SVXII and the drift chamber (Figure 4.4) at radii 20-28 cm. ISL is composed of five barrels in total, one central and of two

> inner and outer barrels in the backward/forward regions. Each barrel has twelve 30° azimuthal sectors to match the SVXII segmentation.
>
> The basic ISL readout is a "half ladder" module. Such a module is made of three silicon sensors ganged together forming a single electrical unit. Sensors are, as in the case of SVX II, double-sided AC coupled microstrip detectors. Each sensor has 512 readout strips along the beamline of 112 μm pitch on one side, and 1.2°-tilted 146 μm pitch on the other side. The total number of channels is 268,800.

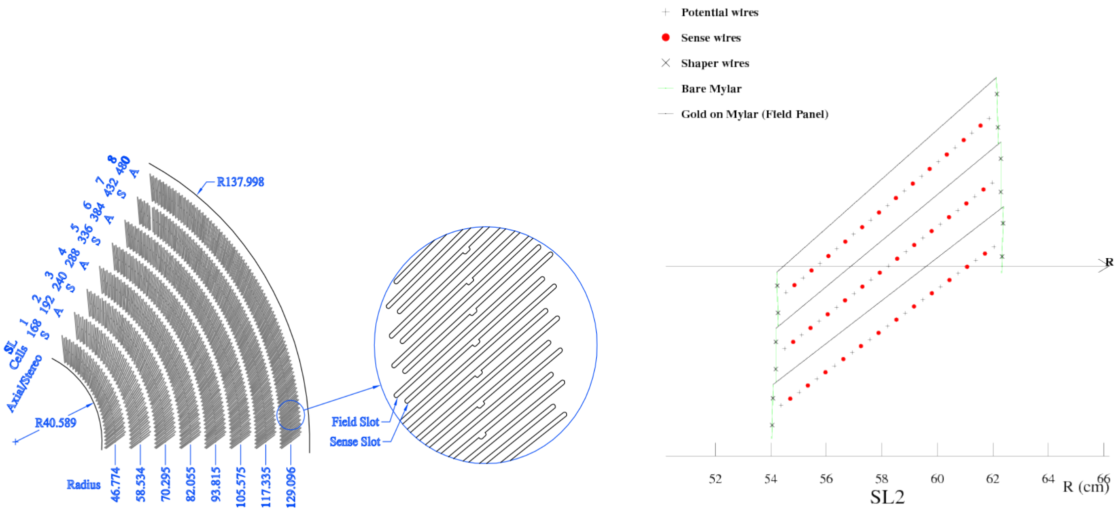
The total amount of material in the silicon system (averaged over azimuthal angle and z) is roughly $10\%/sin\theta$ times the electron radiation length. Therefore in the forward/backward region the average material traversed by particles increases by roughly twice with respect the central region.

The combined resolution of the CDF inner trackers for high momentum tracks is 40 μm in impact parameter and 70 μm along the z direction [62].

4.3.2 Central Outer Tracker

The main tracker at CDF II is the Central Outer Tracker (COT). The COT has a cylindrical shape and is radially right outside the ISL. Its active volume spans from radii ~44 to ~132 cm and $z \lesssim 155$ cm. COT provides full tracking in the central region ($|\eta_{det}| < 1$). With reduced acceptance, its coverage extends to $|\eta_{det}| < 2$ (see Figure 4.3). Radially, the COT is arranged into 4 axial and 4 stereo superlayers, containing 96 planes of wires (see Figure 4.7a). Each superlayer is composed of azimuthal cells. Each cell has alternated sense and field shaping wires (see Figure 4.7b). The latter control the gain on the sense wires optimizing the electric field intensity. Axial superlayers employ sense-wires parallel to the beam axis, while stereo superlayers have the wires alternatively tilted at ±2° with respect to the beam-line. Axial superlayers provide measurements of the hit coordinate in the (r, ϕ) plane, while the stereo superlayer also measure the hit coordinate along the z axis.

Within a cell the ionization released by charged tracks is sampled 12 times (every 0.583 cm) by sense wires. Inside the solenoid magnetic field, the drifting electrons experience a Lorentz force which rotates their path. The cells are tilted



(A) (B)

FIGURE 4.7: *Slots housing the wire-holding fixtures of a 1/6 section of the COT endplate (left). For each super-layer the total number of cells, the wire orientation (stereo r axial), and the average radius [cm] is given. Drift and field wires in three cells (right). The horizontal arrow shows the radial direction.*

by

$35 (Figure\ 4.7b) withrespecttotheradialdirectioninordertomaketheelectronsdriftingperpendicu$

This translates into an overall resolution $\sigma(P_T)/P_T = 0.17\%\ P_T\ [GeV/c]$ [63], P_T being the transverse momentum of the tracked particle.

The main parameters of the tracking system are summarized in table 4.1. Detailed studies on the inner detector performance and aging are described in [65].

COT	
Radial Coverage	44 to 132 cm
Number of superlayers (SLs)	8
Readout coordinates of SLs	+3° 0 -3° 0 +3° 0 -3 0°
Maximum drift distance	0.88 cm
Resolution per measurement	140 μm
Rapidity coverage	$\lvert\eta_{det} \leq 1.0\rvert$
Number of channels	30,240
Material thickness at 90°	1.3% X_0
SVX II	
Radial Coverage	2.4 to 10.7 cm, staggered quadrants
Number of layers	5
Readout coordinates	$r - \phi$ on one side of all layers
Stereo side	r-z, r-z, r-uv, r-z, r-uv (uv ≡ 1.2° stereo)
Readout pitch	60-65 μm r-ϕ; 60-150 μm stereo
Resolution per measurement	12 μm (axial)
Total length	96.0 cm
Rapidity coverage	$\lvert\eta_{det}\rvert \leq 2.0$
Number of channels	405,504
Material thickness at 90°	3.5% X_0
ISL	
Radial Coverage	20 to 28 cm
Number of layers	one for $\lvert\eta_{det}\rvert$ <1; two for 1< $\lvert\eta_{det}\rvert$ <2
Readout coordinates	r-ϕ and r-uv (1.2° stereo) (all layers)
Stereo side	r-z, r-z, r-uv, r-z, r-uv (uv ≡ 1.2° stereo)
Readout pitch	110 μm (axial); 146 μm (stereo)
Resolution per measurement	16 μm (axial)
Total length	174 cm
Rapidity coverage	$\lvert\eta_{det}\rvert \leq 1.9$
Number of channels	268,800
Material thickness at 90°	2% X_0

TABLE 4.1: Design parameters of the baseline tracking system (no L00) [56].

4.4 Calorimeters

The CDF calorimeter measures the particle energy by absorbing their total energy and providing a signal proportional to it. Calorimeter information is also used to estimate the transverse energy of weakly interacting particles, such as neutrinos, by computing the imbalance in the total transverse momentum. CDF uses scintillator sampling calorimeters divided into a front electromagnetic and a rear hadronic compartment. Both calorimeters are segmented into projective towers. Each tower consists of alternating layers of passive absorber material (lead in the front and iron in the rear compartment) and plastic scintillator for shower sampling. The light from the scintillator plates is read out through wavelength-shifting bars or plates and light guides by photo-multiplier tubes (see Figure 4.8).

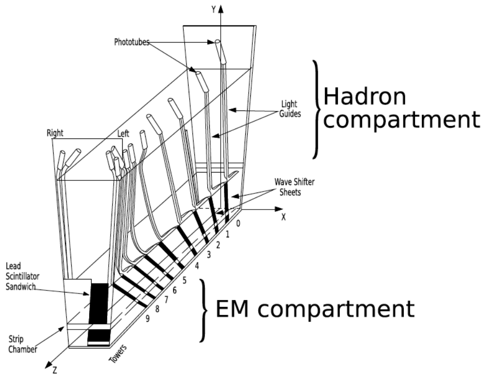


FIGURE 4.8: *Light-shifter plates connected to light guides and to photomultipliers of the front electromagnetic compartment of a central calorimeter wedge.*

High energy electrons and photons generate an electromagnetic shower, which is mostly absorbed in the front calorimeter compartment.

Hadrons interact with the detector matter mostly through inelastic collisions with nuclei of the absorbing medium. Particles produced in the nuclear interactions can loose their energy by ionization and secondary nuclear interactions. Mixed electromagnetic and hadron showers that originate in this process are largely absorbed in the entire (front + rear compartments) calorimeter. The energy-dependent rear leakage is of the order of a few %.

The coverage of the CDF calorimeter extends up to $|\eta_{det}| = 3.6$ and is complete in azimuth (see Figure 4.2). They include the Central Electromagnetic Calorimeter (CEM) and the Hadronic Calorimeter (CHA) in the $|\eta_{det}| < 0.9$ region, the Endwall Hadronic Calorimeter (WHA) at $0.9 < |\eta_{det}| < 1.3$ and the electromagnetic and hadronic plug calorimeters (PEM, PHA) at $1.1 < |\eta_{det}| < 3.6$ (see Figures 4.1, 4.3).

4.4.0.1 Central Calorimeters

The central calorimeters, CEM, CHA and WHA are composed of two parts joining on the median plane of the detector at $z = 0$[3]. Central calorimeters are azimuthally divided into 24 wedges, covering 15° in ϕ each. Each wedge is divided into projective towers of size $\delta\eta_{det} = 0.11$.

---

[3] In this zone, $\eta_{det} = 0$, there is an un-instrumented area about 20 cm thick in the z-direction ("crack")

The **CEM** calorimeter (see Ref. [66]) is made of 31 alternate layers of 0.5 cm thick plastic scintillator plates and 0.32 cm thick lead absorbers: the total amount of material is $18 \cdot X_0$. The CEM energy resolution is[4]:

$$\sigma_{E_T}/E_T = \frac{13.5\%}{\sqrt{E_T[GeV]}} \oplus 1.5\% \tag{4.4.1}$$

the stochastic factor 13.5 is determined primarily by the sampling structure of the calorimeter. CEM also includes two additional specialized detector: the Central Electron Strip Chambers (CES) and the Central Preshower (CPR, [67]). CES is a combined strip/wire gas proportional chamber embedded in CEM at about $6 \cdot X_0$[5]. The CES purpose is to measure the position and the shape of electromagnetic showers in the transverse plane. CES resolution is about 1 cm in z and 1 mm in $r - \phi$. CPR is a set of scintillator tiles located in front of the calorimeter wedges which help distinguishing electrons from charged hadrons because of their probability of showering in the detector material prior to entering the calorimeter.

The **CHA** calorimeter (see Ref. [68]), surrounding the CEM, is composed of 32 alternate layers of 1 cm thick plastic scintillator and 2.5 cm thick steel. The **WHA** calorimeter employs the same technology as CHA, except for the smaller number of layers (15) and the larger thickness of the radiator plates (5 cm). The total calorimeter thickness at normal incidence is 4.7 λ_0 (λ_0 is the charged pion absorption length) for both CHA and WHA.

Resolutions of CHA and WHA for perpendicular particle entrance are approximately:

$$CHA : \sigma_{E_T}/E_T = \frac{50.0\%}{\sqrt{E_T[GeV]}} \oplus 3.0\% \tag{4.4.2}$$

$$WHA : \sigma_{E_T}/E_T = \frac{75.0\%}{\sqrt{E_T[GeV]}} \oplus 4.0\% \tag{4.4.3}$$

---

[4] The energy resolution is expressed as a function of E_T rather than E to account for the larger amount of absorber material encountered at higher θ.

[5] The maximum energy density in the longitudinal development of the electromagnetic shower is expected at about $6 \cdot X_0$

4.4.0.2 Plug Calorimeter

The **PEM** calorimeters (see Figure 4.9) have the same tower segmentation as CEM in η_{det}, and ϕ for $2.11 < |\eta_{det}| < 3.6$ ($15°$ wide ϕ bins). The segmentation in ϕ for $|\eta_{det}| < 2.11$ is finer than CEM ($7.5°$ wide ϕ bins). PEM is composed of 22 layers of 4.5 mm thick lead alternate with 22 layers of 4 mm thick scintillator (see Ref. [69]). The total thickness is about 21 X_0. The PEM energy resolution is:

$$\sigma_{E_T}/E_T = \frac{16.0\%}{\sqrt{E_T[GeV]}} \oplus 1.0\% \tag{4.4.4}$$

As for CEM, PEM is equipped with a shower maximum detector (PES). PES, as well as CES, measures the shower profile to distinguish electrons from charged hadrons. PES consists of two layers of 200 scintillating bars each. Bars are oriented at crossed relative angles of 45°. The position of the shower is measured with an accuracy of about 1 mm.. Further details are described in [70].

PHA, surrounding PEM, has its same tower segmentation. The technology is the same as for CHA, with 23 layers of 2 cm thick steel absorber alternating with 6 mm thick scintillator. The total amount of material corresponds to 4.7 λ_0. PHA resolution is:

$$\sigma_{E_T}/E_T = \frac{80.0\%}{\sqrt{E_T[GeV]}} \oplus 5.0\% \tag{4.4.5}$$

4.4.1 Timing system

Timing readout for the electromagnetic (EM) calorimeters was installed as part of an upgrade to the Run II version of the Collider Detector at Fermilab (CDF). Figure 4.10 schematically shows this system, called *EMTiming*.

Particles from the collision that deposit energy in the EM calorimeter create light within the scintillators [1, 2] that can be used for a timing measurement. Photo-multiplier tubes (PMTs) collect this light and convert it into an electrical analog signal. The EMTiming system routes a copy of the PMT signal to a passive Transition Board (TB) and an Amplifier-Shaper-Discriminator board (ASD) that, in turn, converts analog signal into digital and sends it to a

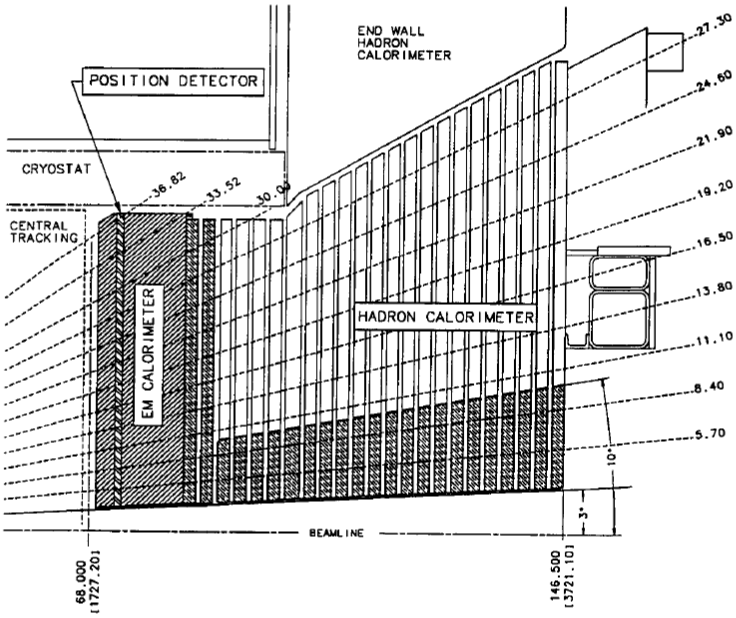


FIGURE 4.9: *Longitudinal view of Plug Calorimeters.*

Time-to- Digital Converter (TDC) board for a timing measurement. The exact system specifications can be found in [98].

It has a resolution of less than a nanosecond and covers the central (CEM, $|\eta|$ 1.1) and plug (PEM, 1.1 $|\eta| < 2.1$) regions of the calorimeter.

In the region $|\eta| < 1$, used in this analysis, photo-multiplier tubes (PMTs) on opposite azimuthal sides of the calorimeter tower convert the scintillation light generated by the shower into an analog electric signal. The energy measurement integrates the charge over a 132 ns timing window around the collision time from 20 ns before thE collision until ~ 110 ns afterwards. New electronics inductively branches off ~ 15% of the energy of the anode signal and sends it to a discriminator. If the signal for a tower is above 2 mV (3–4 GeV energy deposit), a digital pulse is sent to a time-to-digital converter (TDC) that records the photon arrival time and is read out for each event by the data-acquisition system.

4.5 Muon System

Although muons interact mostly electromagnetically as the electrons, because of their much larger mass they can cross a much larger amounts of material before

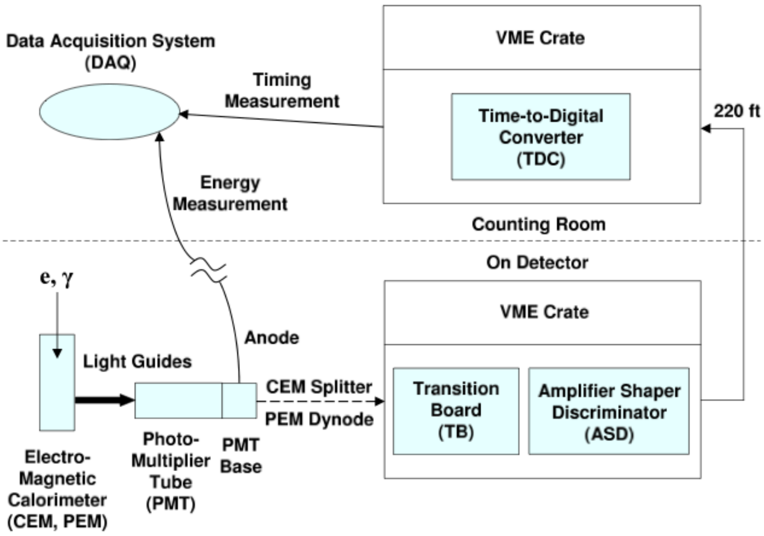


FIGURE 4.10: A diagram of the EMTiming system hardware on the CDF detector.

losing a significant fraction of their energy[6]. For this reason systems dedicated to detect muons are located in the outermost shell of the detector. Muon momenta are measured in the tracker.

Four independent systems are used to detect muons in the $|\eta_{det}| < 1.5$ region: the Central Muon Detectors (CMU), the Central Muon Upgrade Detectors (CMP), the Central Muon Extension (CMX), the Intermediate Muon Detectors (IMU). The (η_{det}, ϕ) coverage of the largest Run II muon detectors is shown in Figure 4.11. Muon detectors share common features (see Ref. [71]). They consist of stacks of

---

[6] At Tevatron energies muons interact in calorimeters as minimum ionizing particles (MIP).

rectangular drift chamber modules[7], composed of single-wire cells. Stacks are four layers deep with laterally displaced cells from layer to layer to compensate for cell edge inefficiencies. The difference in drift-electrons arrival-times between neighbor cells provides a typical resolution of 250 μm for the hit position in the transverse plane. Charge division at the wire ends measures the coordinate along the wire with a 1.2 mm resolution. Chambers are coupled with scintillator counters in order to suppress backgrounds due to secondary interactions in the beam pipe material and to cosmic rays. A muon candidate is reconstructed when a short track segment (stub) in the muon chambers corresponds to the extrapolation of a COT track.

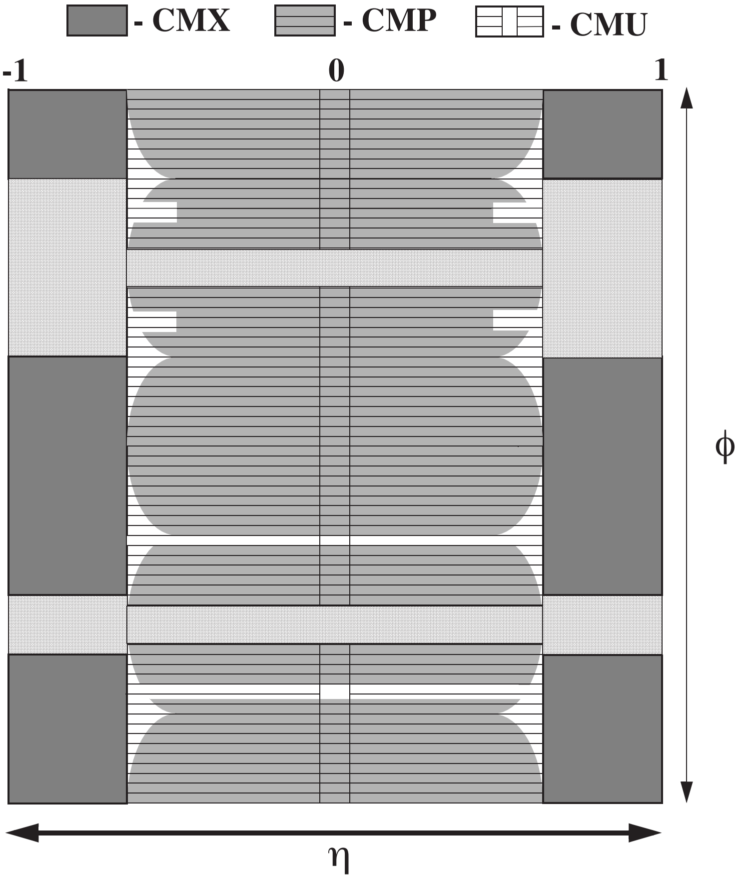


FIGURE 4.11: *Coverage of muon detectors in the* (η, ϕ) *space.* η *is computed with respect to the CDF center.* ϕ *ranges from -*π *to* π *radians.*

The **CMU** detector is behind CHA at a radius of 347 cm from the beam axis and covers the $|\eta_{det}| < 0.6$ region. CMU consists of 144 modules with 16 cells each. The CMU box is arranged in 12.6° wedges. Cells composing CMU are 266 cm long, 2.68 cm thick, and 6.35 wide with a single 50 μm steel wire at their

---

[7] Chambers are filled with a mixture of argon and ethane (50% each)

center (Figure 4.12) parallel to the beam direction. The system is filled with Argon-Ethane gas mixture and alcohol like the COT.

The **CMP** detector is arranged to enclose the $|\eta_{det}| < 0.6$ region in an approximately central box (see Figure 4.11). Scintillator layers (CSP) on the outermost side of the CMP chambers allow identifying bunch crossing. The CMU/CMP system is called CMUP. It detects muons with a minimum energy of about 1.4 GeV.

The **CMX** detector extends the muon identification in the $0.6 < |\eta_{det}| < 1$ region. As for CMP cells are sandwiched to scintillators (CSX).

The forward region of muon system is the **IMU** detector ($1.0 < |\eta_{det}| < 1.5$). The associated scintillator counters sub-system is named BSU.

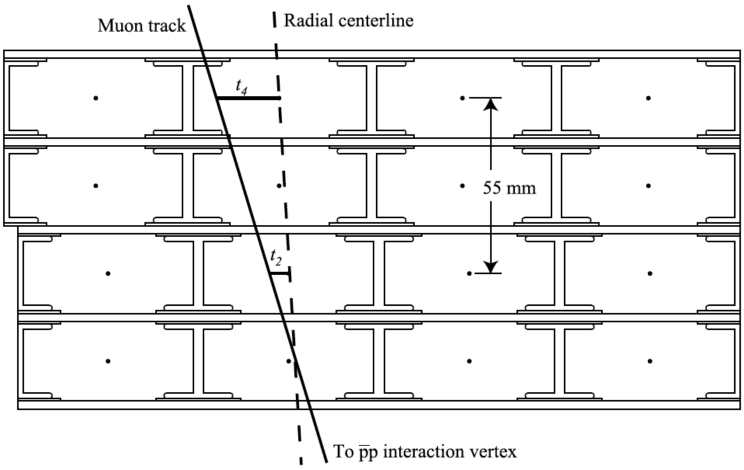


FIGURE 4.12: *Cross section of CMU single wire cells.*

4.6 Cherenkov Luminosity Counters

The Tevatron collider luminosity is estimated by measuring the known interaction rate of inelastic $p\bar{p}$ events.

$$\mathcal{L} = \frac{f\mu}{\sigma_{p\bar{p}}} \tag{4.6.1}$$

where f is the frequency of bunch crossing. A detector consisting of two CLC ("Cherenkov luminosity counters") is used to do so (see Ref. [72]). The two modules, placed inside the CDF end-plug calorimeters in the forward/backward region at $3.7 < |\eta_{det}| < 4.7$, consist of 48 thin, long, gas-filled Cherenkov counters. These counters are grouped in three concentric conical layers (16 counters each) around the beam pipe. Counters point to the center of the interaction region (see Figure 4.13). Counters in the outer two layers are 180 cm long, while the counters in the inner layer are 110 cm long. Counter transverse dimensions range between 2 and 6 cm. This geometry ensures that only charged particles originating from around the nominal interaction point give a full-height Cherenkov signal, and allow distinguishing them from prongs of beam-gas events and from stray machine background.

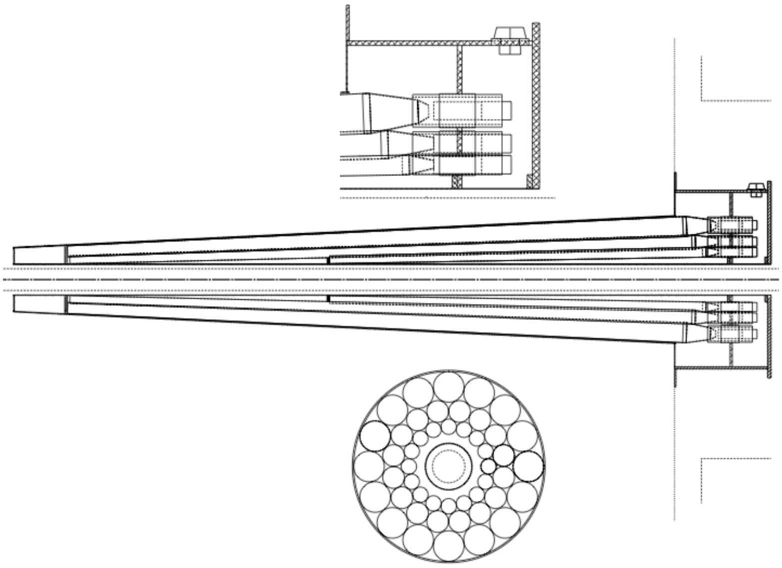


Figure 4.13: *The CDF Cherenkov Luminosity Counters Design.*

The Cherenkov light is collected by conical mirrors at the rear end of the cones and detected by high gain (2×10^6) photomultipliers. The counters are mounted on long aluminum tubes surrounding the beam pipe. Everything is plunged in Isobuthane and kept at the atmospheric pressure. In this condition the Isobuthane refraction index is relatively high ($n = 1.0043$) and the gas is transparent to UV radiation. Around the tube a magnetic field shield absorber is placed in order to reduce the fringe CDF solenoid field. This is done to allow photomultipliers to operate at constant gain.

Prompt particles from the $p\bar{p}$ primary interaction will traverse the full counter length and will generate a large amount of Cherenkov photons. Secondary background particles are mainly electrons or positrons from electromagnetic showers initiated by $\pi^0 \to \gamma\gamma$ decays in the beam pipe or detector material[8]. They may contribute to low-amplitude signals (when they traverse the counters at large angles and with shorter path lengths) or anomalous high-amplitude

---

[8] Beam halo particles often fall under the Cherenkov threshold, thus producing no light.

signals (several particles hit the same counter). Therefore the signal from the secondary particles can be discriminated by setting pulse-height cuts on-line and off-line. The counter time resolution is excellent (< 100 ps) and helps checking the multiple interaction rate as a function of instantaneous luminosity, thus reducing the systematic uncertainties.

The luminosity is measured through the formula below:

$$\mathcal{L} = \frac{R_{p\overline{p}}}{\epsilon_{CLC}\sigma_{in}} \tag{4.6.2}$$

where $R_{p\overline{p}}$ is the rate of the inelastic $p\overline{p}$ events measured with CLC, ϵ_{CLC} is the the CLC acceptance, and σ_{in} is the inelastic cross $p\overline{p}$ section[9].

The CLC acceptance is measured from simulation: $\epsilon_{CLC} = 60.2 \pm 2.6\%$. The inelastic cross section is obtained by extrapolating the CDF, E811 combined measurements at $\sqrt{s} = 1.8$ TeV ($\sigma_{in} = 59.3 \pm 2.3$) [73]. The cross-section extrapolated at 1.96 TeV (assuming a $ln^2 s$ dependence) is 60.7 ± 2.4 mb.

The uncertainty on the measured luminosity is ∼6% and is largely dominated by systematic uncertainties. Main contributors are uncertainties on the detector stability and calibration (≤ 2.5 %), ϵ_{CLC} (4.4%), σ_{in} (4.0%). Further details about the luminosity calculations are described at [63].

4.7 Trigger and Data Acquisition

At Tevatron Run II the interaction rate is typically 2.5 MHz (the bunch crossing frequency), which is much higher than any possible event recording rate (order of 100 Hz). However, $p\overline{p}$ interactions are mostly inelastic, elastic and diffractive with no significant momentum transfer. More interesting events have cross-sections from 10^3 to 10^{12} times smaller than inclusive $p\overline{p}$ cross section (for example $\sigma(p\overline{p} \rightarrow WZ) \sim 4$ pb). The identification of the interesting events is accomplished by dedicated fast online electronics, called the Trigger System, which evaluates the information from the detector and makes an accept/reject decision in real time. The trigger system (see Figure 4.14) is a three-tier system, where each higher level of electronics performs a slower but more accurate event reconstruction and

---

[9] At large angles the rate of elastic events is negligible with respect to the one of inelastic events.

applies a tighter filter with increasing trigger level according to a set of pre-defined conditions.

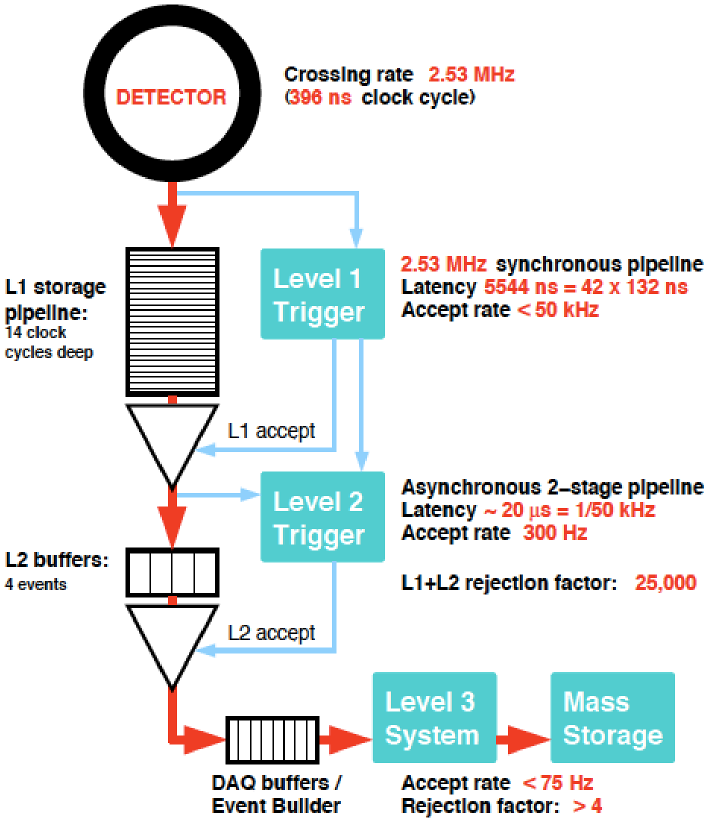


FIGURE 4.14: *CDF trigger block diagram.*

4.7.1 Level 1

The level 1 (L1) is a synchronous pipeline system, where up to 42 subsequent events can be stored for 5.5 μs while custom hardware is making a decision. If no acceptance decision is made up to that time the event is rejected. L1 decisions are made in about 4 μs average time: no dead time is expected from this level. L1 roughly accepts one out of 3×10^4 inelastic events and the typical output rate is 20 KHz.

The L1 decision is generated based on the information below.

- Reconstruction of calorimeter objects (electrons, photons, jets) or calculation of global-event calorimetric observables (e.g: transfer momentum imbalance, total transverse scalar energy, etc.). At this level electron, photon and jet candidates are defined as single-tower ("seed") energy deposit. The nominal CDF geometrical center is used as coordinate origin.

- Track reconstruction[10]: the eXtremely Fast Tracker (XFT, [74]) hardware examines hits on the COT axial superlayers and combine them into track segments. The segments are then linked to form a track candidate with $P_T > 1.5$ GeV/c.
- Muon identification: the observation of hits in the muon detector wire chamber and scintillator systems hits matched to a XFT track, extrapolated to the muon chamber, defines a muon. A loose P_T threshold is applied based on differences in signal arrival times on pairs of wires in the CMU and CMX chambers.

The final L1 decision is made based upon the number of reconstructed objects or upon the calculated observables mentioned above. Events accepted at L1 are stored in one of the four buffers in the front-end readout hardware.

4.7.2 Level 2

The level 2 (L2) trigger system makes use of dedicated hardware to select events. The two main pieces of dedicated hardware are the following:

- cluster finder: add the energy deposited in the towers neighboring the L1 seeds to form a cluster. The cluster energy is an approximate measure of an electron or jet energy
- silicon vertex tracking (SVT, [75]): use the information from the SVX II detector to improve the parameter (e.g: ϕ, P_T, d_0) resolutions of the XFT tracks. Hereby SVT allows to trigger on secondary vertexes from decay of long-lived beauty hadrons.

These two systems work in an asynchronous way since the processing time is highly correlated to the amount of data to be processed.

---

[10] This additional information in the trigger was introduced during the Run II upgrade.

The outputs of these systems is transferred to the global L2 processor, along with L1 data, and with additional information from the CES detector to help in low E_T electron selection. The processor decides whether the event is passed to the next step by exploiting simple selection algorithms. The processor board has been designed to read-in one event while processing another one in order to reduce the deadtime. The maximum L2 acceptance rate is ∼300 Hz (rejection factor $\sim$ 150). The block diagram of L1 and L2 with the involved sub-detectors is schematized in Figure 4.15.

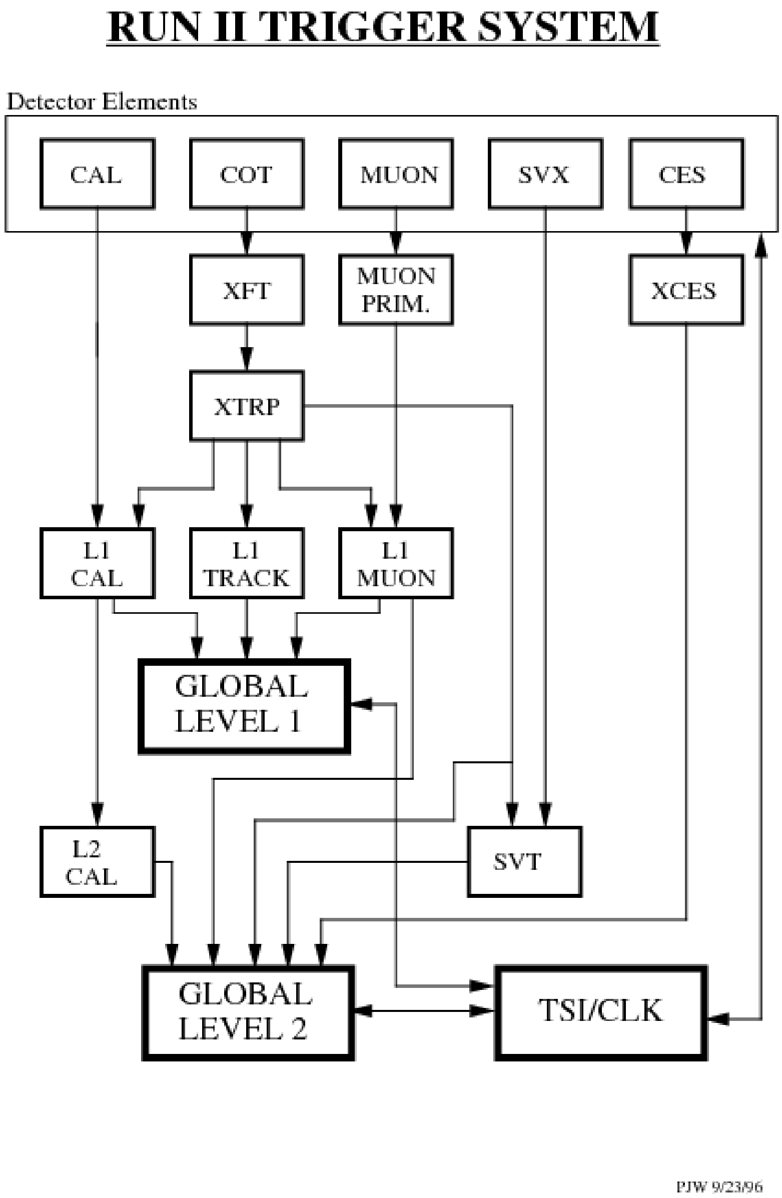


FIGURE 4.15: *Block diagram of level 1 and level 2 triggers. The involved sub-detectors are indicated.*

As the Tevatron luminosity increased during the Run II, the performance of the level 2 trigger started to degrade since:

- the SVT processing time (and therefore the deadtime) increased because of the higher event complexity
- the background rate accepted by the calorimetric trigger increases

In order to cope with the former issue, a faster track fitter and other changes described in [76] were implemented. In order to deal with the latter issues the L2 calorimetric trigger was upgraded such that the full trigger tower information (same accuracy as the offline one) would be available [77].

4.7.3 Level 3

Level 3 (L3) is a software trigger. The L3 farm is made of about 300 commercial dual processor computers. L3 addresses event components delivered by L2 to the Event Builder (EVB), which reconstructs the entire event with the same accuracy as in the offline analysis, which was unavailable to the lower trigger levels. The L3 algorithms makes a full three-dimensional track reconstruction. The final decision to accept an event is made on the basis of a list of required observables indicating candidate events of physical interest (top production events, W/Z events, Drell-Yan events, etc.). Accepted events exit L3 at a rate of up to 100 Hz and are permanently stored on tapes. Store events will be analyzed offline.

A set of requirements to be fulfilled by an event at L1, L2, and L3 is called trigger path. While at CDF II about 150 trigger paths exist, the trigger paths used for this analysis will be described in Chapter 7.

4.8 Data storage and offline processing

The data flow from L3 triggers was stored in real time on fast-access disk. During online acquisition collected data were grouped in run numbers. Each run can contain from few dozen collision events to millions of events depending on the duration of a store. At the end of each run, good run bits for each detector component are set true if they were working properly during the run; portions of runs can also be marked as good. Wether a run is good for physics analysis depends on which components where working and the type of analysis being performed.

Online data are divided into data streams based on similar triggers. For this analysis the CDF's c-stream is used. These sample is structured so that it can be processed using CDF's Stntuple software.

All subsequent data processing is called *offline* data handling. During the offline processing several run numbers are grouped in run periods which integrated luminosity is of the order of ∼100 pb^{-1}. In this thesis the full CDF II dataset collected between February 4^{th} 2002 and September 30^{th} 2011 has been analyzed (from period 0 to period 38), which corresponds to an integrated luminosity of ∼10 fb^{-1}.

Then, stored raw data are unpacked, and physics objects as tracks, vertices, photons are reconstructed. The offline processing is similar to the L3 trigger reconstruction, with the difference that physics objects can be more elaborated and the most up-to-date detector calibrations are used. The next chapter is devoted to physics objects reconstruction.

Chapter 5

Physics Objects Reconstruction

In order to infer the actual physics process occurring at the interaction point, the information on the produced particles is used. As these particles pass through the detector, they interact with it and give rise to electronic signals. From those electronic signals it is possible to determine the location of the interaction point and the particles' properties, such as their energies and momenta.

The process of converting electronic signals recorded by the detector into collections of measurements associated to physical particles is referred to "event reconstruction".

In this chapter, the reconstruction of the "physical objects" of interest for the analysis is discussed. A description of the corrections to the measured photon energy and timing is also given.

Contents

5.1 Track reconstruction

Track reconstruction allows the measurement of charged particle momentum and of the track impact parameter. A charged particle moving in an axial and almost uniform magnetic field ($\overrightarrow{B}$, with $|\overrightarrow{B}| = 1.4$ T for CDF) has a helicoidal trajectory. Such trajectory can be completely described by the following five parameters:

1. C: the half-curvature of the trajectory, defined as $C \equiv 1/(2qr)$ with r as the helix radius and q the charge of the particle. It has the same sign of the particle charge and it is related to the transverse momentum of the track according to the relation:
$$p_T = \frac{cB}{2|C|} \tag{5.1.1}$$
2. d_0: the distance of closest approach in the transverse plane between the helix and the origin, usually referred to as *the impact parameter*. Let x_0 and y_0 be the the coordinates of the center obtained by the projection of the helix on the transverse plane and $r = 1/2C$, the definition of d_0 is
$$d_0 = q\left(\sqrt{x_0^2 + y_0^2} - r\right) \tag{5.1.2}$$
3. λ: the *helix pitch*, i.e. the cotangent of the polar angle between the track and the z-axis ($\cot\theta_0$). The longitudinal component of the momentum corresponds to:
$$p_z = p_T\lambda = p_T \cot\theta_0 \tag{5.1.3}$$
4. z_0: the z position of the track vertex.
5. ϕ_0: the azimuthal angle of the track at its vertex.

Every point along the trajectory satisfies the following set of equations [76]:

$$x(\phi) = r\sin(\phi) - (r + d_0)\sin(\phi_0) \tag{5.1.4}$$
$$y(\phi) = -r\cos(\phi) + (r + d_0)\cos(\phi_0) \tag{5.1.5}$$
$$z(\phi) = z_0 + s\lambda \tag{5.1.6}$$

where s is the length projected along the track, and $\varphi = 2Cs + \varphi_0$.

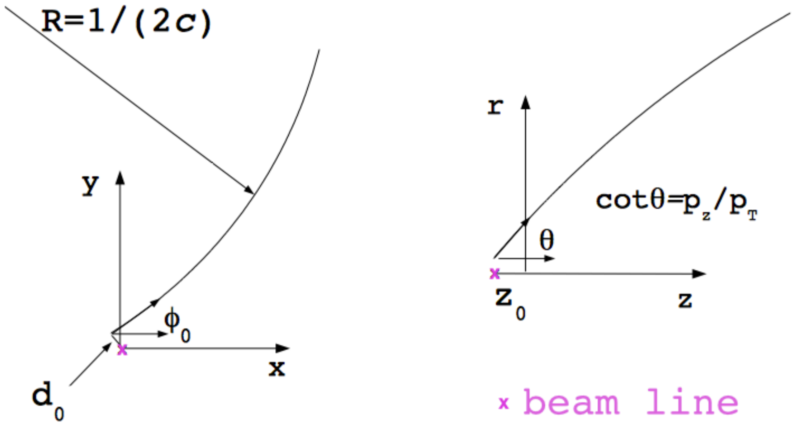


FIGURE 5.1: In a uniform magnetic field charged particles follow a helical trajectory completely defined by 5 track parameters: λ, C, z_0, d_0 and ϕ_0. See text for their definitions.

The track reconstruction is based on a set of signals ("hits") measured in the tracking detectors and associated into clusters. These hits are then fit with an helix functional parametrization which also consider field non-uniformities and multiple scattering effects. At CDF several tracking algorithms have been developed and optimized for different detector regions, in order to derive the previously defined parameters [107]. Using multiple algorithms provides a combined track reconstruction efficiency of about 100% for the central region and up to about 75% in the plug region [49]. The main ones are: the Outside-In algorithm (OI), the Silicon-Stand- Alone (SiSA) algorithm and the Inside-Out (IO) algorithm.

Outside-In Algorithm

For the central region ($|\eta|$ <1), COT tracks are reconstructed with an outside-in (OI) algorithm. Track pattern recognition starts with hits in outer COT superlayers or layers 6 and 7 of the silicon, where there is a lower hit density. First, segments between superlayers are linked by two algorithms ("histogram linking algorithm" and "segment linking algorithm". Then, the COT track is used as a seed to look for a track fit including SVX hits.

Silicon-Stand-Alone Algorithm

Outside the acceptance of the COT, an OI approach is also used: silicon standalone (SiSA) tracks are reconstructed with only silicon hits. The algorithm begins with a $R\phi$ fit; only track candidates with more than 4 hits are considered, given that 3 track parameters (d_0, R, ϕ_0) have to be determined. Then, the Rz hits to be match to the track candidate are searched for. The two remaining longitudinal parameters (λ and z_0) are extracted from a new fit. The transverse parameter are recomputed, as well. Run dependent simulations show that the overall tracking efficiency is $\sim$85% in the region $|\eta| < 0.75$ and $|z_0| < 35$ cm, which is fully covered by SVXII.

Inside-Out Algorithm

Some lower p_T tracks within the inner COT layers fail to be reconstructed by OI algorithms. Inside-out (IO) algorithms improve track reconstruction, also beyond the $|\eta| = 1$ region [49]. The IO algorithm uses SiSA tracks and searches outward to attach COT hits.

5.2 Vertex reconstruction

Many $p\bar{p}$ collisions can occur simultaneously in a single bunch crossing, and be recorded as a single event. Moreover, although only one hard-interaction is expected most of the times, more than one vertex can be reconstructed per event, because of additional min-bias interactions; these remnants of the beam contribute with lower p_T interactions.

A z vertex finding algorithm [101] is used to measure the z position of primary vertices and to count the number of reconstructed vertices. The z position of the primary vertex is exploited in the calorimeter energy reconstruction, as explained in Section 5.4.

COT and/or silicon three-dimensional (stereo) hits are used as input to a histogram algorithm: a histogram is filled with the z position of tracks from these hits and resonances in the histogram give the position of preliminary vertices. Tracks are matched to these seed vertices if their z_0 position lies within a minimum distance (of a few cm) of the preliminary z vertex and if they have a p_T above a given threshold. Then, the vertex z position is recomputed from an error weighted average of the z_0 positions of associated tracks clustered to the vertex:

$$\frac{\sum_i z_0^i/(\Delta z_0^i)^2}{\sum_i 1/(\Delta z_0^i)^2} \tag{5.2.1}$$

Depending on the number of tracks with silicon and/or COT hits, vertices are assigned to a particular class. For example, *Quality 0* refers to all vertices and *Quality* refers to vertices having > 6 tracks with silicon hits and > 1 track with COT hits. Many CDF analyses uses the so-called *Quality 12* class, having vertices with more than 2 tracks with COT hits (tracks with COT hits reduce fake rates). This class of vertices provides a good compromise between true vertex reconstruction efficiency and the number of fake vertices reconstructed. Moreover, there is linear relationship between the number of *Quality 12* reconstructed vertices in an event (N_{vtx}) and instantaneous luminosity [101].

To each reconstructed vertex corresponds a p_T value defined as the sum of the transverse momentum of its associated tracks ($\sum_{tracks} p_T$). The vertex matched to the highest sum p_T of the tracks is defined as the primary vertex of the event, related to the hard interaction. The other reconstructed vertices are typically associated to minimum bias events.

5.3 The Corrected Time

The time of arrival recorded by the EMTiming system TDCs is a “raw” time and is corrected taking into account several effects. The corrected time [98] is given by:

$$t_{corr} = t_{Raw} + C_{Start} + C_{Energy} + C_{Energy\ Asymmetry} - C_{Time\ of\ Flight} - C_{Collision\ Time} \tag{5.3.1}$$

C_{Start} is a constant that takes into account the offset between the average time of arrival at the calorimeter and the TDC start[1]. The main contribution comes from the overall cable lengths, so it can be very different from tower-to-tower.

C_{Energy} corrects for an energy-dependent (*slewing*) effect due to the fixed-threshold discriminators. It is defined with empirically derived constants A_1, A_2 as:

$$C_{Energy} = \frac{A_1}{\ln(x)} + \frac{A_2}{x^2} \tag{5.3.2}$$

where x is the sum of the energies, from the two PMTs, as measured the calorimeter.

$C_{Energy\ Asymmetry}$ within a tower, the PMT energy response differences as well as the location where the particle hits can also affect the measured time of arrival. Both effects are corrected by taking into account the energy asymmetry of the two PMTs, defined as $x = \frac{|E_{PMT1} - E_{PMT2}|}{E_{PMT1} + E_{PMT2}}$. The energy asymmetry correction within the towers is parametrized by the following empirically derived function:

$$C_{Energy\ Asymmetry} = B_0 + B_1 \cdot x + B_2 \cdot x^2 \tag{5.3.3}$$

where B_0, B_1, B_2 are constants.

$C_{Time\ of\ Flight}$ combines the information on the primary interaction position ($\overrightarrow{x_i}$) measured from COT and the calorimeter tower position ($\overrightarrow{x_f}$). The expected time-of-flight is given by:

$$C_{Time\ of\ Flight} = |\overrightarrow{x_f} - \overrightarrow{x_i}|/c \tag{5.3.4}$$

---

[1] The system uses fixed-start TDCs

$C_{Collision\ Time}$ is the measured collision time (given by the tracking chamber); this is subtracted off on an event-by-event basis.

This set of corrections is determined using in-situ data. Electrons from $W \to e\nu$ events have been chosen as an excellent sample for calibrating the detector systems because an electron in the event allows to correctly identify the origin of the event and correctly calibrate the timing systems to this origin. The t_{corr} resolution for electrons from $W \to e\nu$ is 0.64 ns for collision data (see Figure 5.2). Moreover the corrected time distribution is centered at 0 ns.

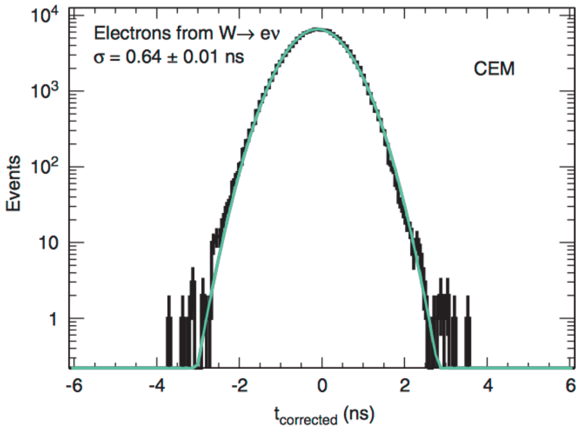


FIGURE 5.2: Timing resolution as estimated from a higher-statistics sample of $W \to e\nu$ events where the time and position of the collision is measured, and corrected. Plot taken from Ref. [111].

A full description of the hardware as well as the correction and calibration procedure can be found in Ref. 16.

5.4 Photon Reconstruction

The measurement presented in this thesis strongly relies on the identification of photons. Some of them convert to an electron-positron pair before reaching the calorimeters. Both photons and electrons detection is based on the energy measurements of the EM calorimeters.

5.4.1 CEM Clustering

This section deals with the reconstruction of such EM objects, based on clustering energy in calorimeter towers [102]. Then, it ends with a description on how the detector coordinates of the cluster are determined. Later chapters discuss the identification of particles from these clusters.

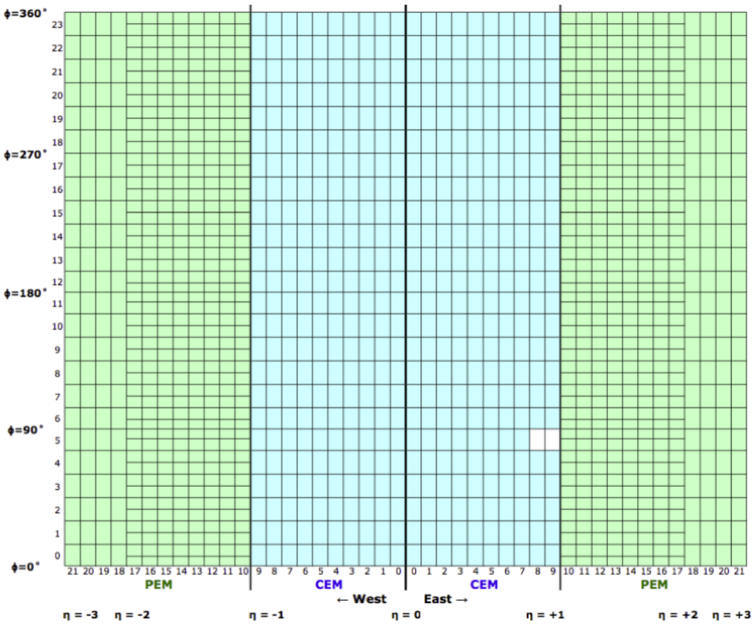


FIGURE 5.3: Calorimeter segmentation in $\eta - \phi$ space for the central (light blue) and plug (green) electromagnetic calorimeters. The x axis shows the west (left) and east (right) η tower numbers (0–21), within the region $0 < |\eta| < 3.6$. Calorimeter wedge numbers (0–23) are shown on the left vertical axis labels; each of them has a coverage of 15deg in ϕ. For a segment of the plug region, each wedge contains two towers such that the tower ϕ segmentation is 7.5deg. The white box, referred to as *the chimney*, is a not instrumented region. It is a gap in the detector used for cables and cryogenic utilities needed for the solenoid.

Calorimeter energy is *clustered* in both the CEM and PEM by looking for all towers with transverse electromagnetic energy $E_T > 3$ GeV, referred to as *seed towers*. The seed towers are then sorted by decreasing E_T such that clustering begins with the largest E_T tower. Nearest neighbours towers, called *daughter towers*, are grouped with the seed tower if they are in the same detector (CEM or PEM) and not already included with another cluster.

In the CEM, daughter towers are defined as border towers with the same ϕ value as the seed tower such that the difference is η andϕ index numbers from the seed

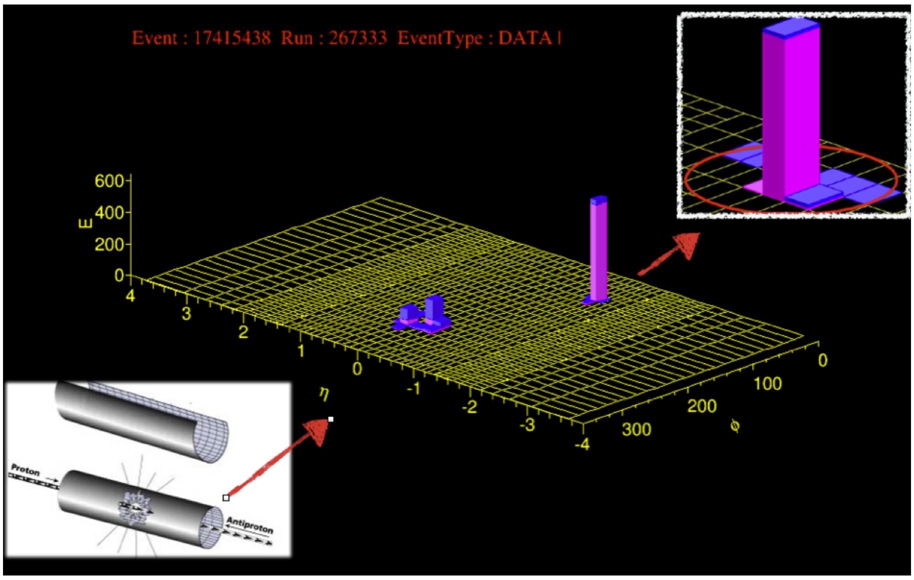


FIGURE 5.4: "Lego" plot of Energy deposit in the calorimeters for a photon plus jet event. The central calorimeter is a cylinder of calorimetry towers (pointing towards the interaction point). To display the event in a easier way, this "cylinder" of energy flow is "unrolled" and made into a flat 2-D plot. The inset on the bottom left shows such a process of unfolding.The yellow grid shows the $\eta\phi$ segmentation and each rectangular region represents a calorimeter tower, with the magnitude of the towers being the total energy deposited.There are two layers: (1) the front EM calorimeter usually displayed in pink (which records almost all the EM showers for electrons and photons), and (2) the hadron calorimeter energies usually displayed in blue (which will record most of the hadronic showers from charged pions, kaons, protons and neutrons). The energy of an EM shower is generally contained within a few CEM towers. Only a small amount of this energy leaks into the hadron towers. The inset on the upper right is a zoom-out of a EM energy deposit. The red circle indicates a region in the $\eta\phi$ space defined by $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

is $\Delta n_\phi = 0$ and $\Delta n_\eta = 1$. Central EM clusters, therefore, contain a minimum of one tower (the seed) and a maximum of three towers (when both neighbors have E_T >100 MeV), all in the same ϕ wedge.

In the plug calorimeters, EM clusters have a square 2 × 2 tower configuration.

EM clusters are treated as massless objects. Their total EM energy (E_{EM}) is the sum of the energies of each EM calorimeter tower included in the cluster. The total amount of hadronic energy (E_{HAD}) of the cluster is obtained from the sum of the associated HAD calorimeter towers. The total energy of the cluster is $E = E_{EM} + E_{HAD}$. Since a real EM object deposits essentially all of its energy in the EM calorimeters, we take the total energy of such a particle to be the measurement from the EM calorimeters alone and often refer to it as just E rather than E_{EM} (or E_T rather than E_T^{EM}).

The transverse energy of the photon is obtained from the energy after correcting it for the position of the photon measured in the CES detector (see next Section) as relative to the hard interaction vertex, given by the polar angle θ:

$$E_T = E\sin(\theta) \tag{5.4.1}$$

The position along the z axis of the primary interaction vertex is used as a reference point to evaluate the polar angle θ_{EM} (θ_{HAD}) of the EM (HAD) calorimeter towers according Eq. 5.4.10.

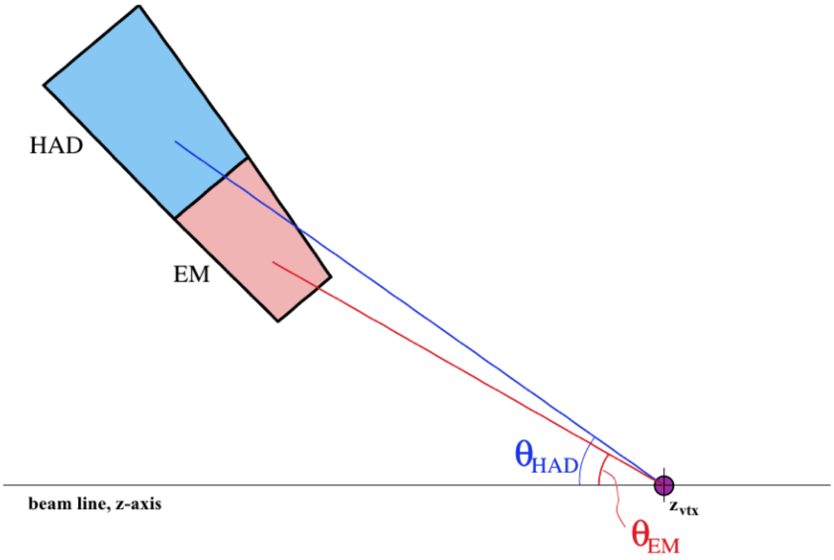


Figure 5.5: The position along the z axis of the primary interaction vertex is used as a reference point to evaluate the polar angle θ_{EM} (θ_{HAD}) of the EM (HAD) calorimeter towers. The transverse energy E_T for an EM (HAD) cluster is calculated from $E_T = E\sin\theta_{EM}$ ($E_T = E\sin\theta_{HAD}$). For trigger towers and during offline calorimeter clustering, the location of the primary vertex is assumed to be z_{vtx} =0. After obtaining a calorimeter cluster, its E_T is computed using the reconstructed primary vertex.

The (η, ϕ) detector coordinates of the cluster are obtained with an energy-weighted method. The EM energy-pondered sum of the η and ϕ detector positions of each tower in the cluster are given by:

$$\eta_{EM} = \frac{\sum_i E^i_{EM} \times \eta^i}{\sum_i E^i_{EM}} \tag{5.4.2}$$

$$\phi_{EM} = \frac{\sum_i E^i_{EM} \times \phi^i}{\sum_i E^i_{EM}} \tag{5.4.3}$$

and, similarly, the hadronic cluster positions are obtained according to:

$$\eta_{HAD} = \frac{\sum_i E^i_{HAD} \times \eta^i}{\sum_i E^i_{HAD}} \tag{5.4.4}$$

$$\phi_{HAD} = \frac{\sum_i E^i_{HAD} \times \phi^i}{\sum_i E^i_{HAD}} \tag{5.4.5}$$

Then, the detector coordinates for the cluster are then obtained from the following calculation

$$\eta = \frac{E_{EM}\eta_{EM} + E_{HAD}\eta_{HAD}}{E} \tag{5.4.6}$$

$$\phi = \frac{E_{EM}\phi_{EM} + E_{HAD}\phi_{HAD}}{E} \tag{5.4.7}$$

5.4.2 CES Clustering

The central Shower Maximum Detector is used to provide a position for the photon within a tower, as already mentioned. The reconstruction of a CES cluster is performed with a "track-based" or a "strip-based" algorithm [103]. The latter is specifically designed for photon identification. This one sorts the strips in decreasing energy and selects the "seeds" from those having an energy above a given threshold. Strips surrounding the seed are collected to form a 1D cluster, and this is also done for the wire layer to form its own 1D shower. For both layers, 11 wires and 11 strips are used to reconstruct a cluster. The two energy showers are then matched to form a 2D CES cluster with the following energy-weighted local CES x and z coordinates:

$$E = \sum_i^N E_i, \quad X = \sum_i^N \frac{X_i E_i}{E} \tag{5.4.8}$$

where E_i and X_i are the pulse height and the position respectively and the index i runs over all the N wires (strips) in the cluster, for the x_{CES} (z_{CES}) coordinate.

Then, a *goodness of fit* χ^2 is derived from the fit of the cluster to a standard set of shower profiles:

$$\chi^2 = \frac{1}{4}\sum_{i}^{N} \frac{[y_i^2 - \hat{y}_i^2)]}{\sigma_i^2} \tag{5.4.9}$$

where, given an i channel, y_i is the measured fraction of energy, $\hat{y}_i$ is the fraction of the standard profile and σ_i is the RMS fluctuations measured in the test beam. A new position is determined from the fit using an iterative correction. Figure 5.6 shows the χ^2 distribution for strip- and wire- clusters for electrons simulated with energies of 50 GeV.

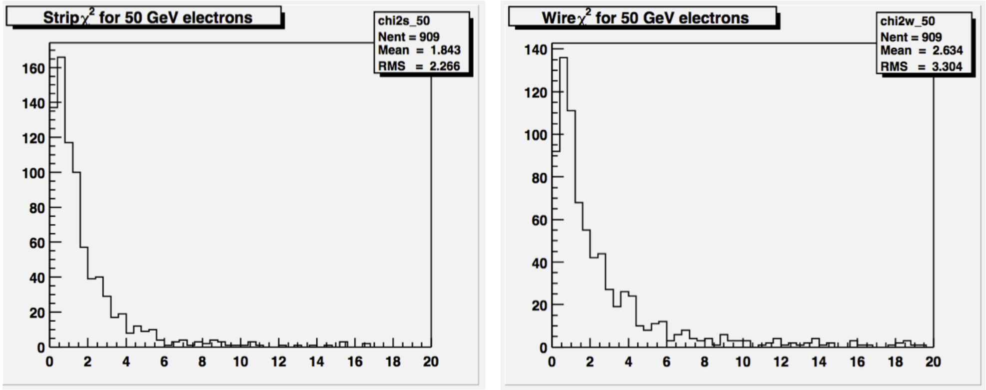


FIGURE 5.6: χ^2 distributions for strip- (left) and wire- (right) clusters for electrons simulated with energies of 50 GeV. Plots taken from Ref. [103].

5.4.3 Photon Energy Corrections

The total energy of an EM cluster is not exactly the energy of the outgoing photon (electron). Therefore, some corrections have to be applied; they are listed below. The first one is made online and is applied only to data. The remaining ones are made offline and are applied both to data and Monte Carlo (MC).

- *Time-dependent gain variations:*
  The gain variations of phototubes for each tower are monitored over time and corrected during run-time in order to detect long-term stability fluctuations. The long-term stability of other component as the light guides, WLS, and the scintillators is also monitored for the measurement of the gradual deterioration of energy resolution and energy scale with time [100]

- *Face (or map) correction in the individual CEM tower:*
  A correction of a few percent is made to the energy offline based on the location where the particle hits the tower. If a shower is initiated near a phototube, there is a higher measurement of energy than if the same shower energy were in the center of a tower. Since towers vary in size, corrections are applied per tower and were evaluated on dedicate 50 GeV electron test beam measurements and cosmic ray studies [104, 105].

- *Wedge–to–wedge correction*:
  The tower clustering is made only in the η direction, thus energy which leaks out of the shower laterally into the neighbouring wedge is lost. The corrections for the lateral shower spread depends on the position of the shower in the CES and energy of the shower.

- *Z-based energy correction*:
  In addition to these corrections, the energy of the photons in data and in Monte Carlo is scaled using electrons from Z decays [108]. These calibrations are performed by plotting the reconstructed Z boson mass from $Z \to e^+e^-$ decays (see Figure 5.7). The energy scale factor is given by the ratio of the Z mass measured from the mean of a fit made to the data, to the accepted PDG value of 91.2 GeV/c^2 [21]. A separate correction is applied to each data-taking period (see Figure 5.8). Figure 5.9 serves as a sanity check: after corrections, the energy scale is flat and consistent with 1.000. A single correction is also applied to MC events. This correction is determined from the ratio of the reconstructed mass obtained from Z MC samples to the PDG.

As previously described, the E_T of the shower is obtained from $E\sin\theta$, where the calorimeter energy includes the energy corrections just discussed. For central and plug offline reconstruction, the polar angle θ is formed relative to the z position of the primary vertex (z_{vtx}) and the position of the cluster in the shower maximum detector (see Figure 5.5). In particular,

$$\sin\theta = \frac{R_{SMX}}{\sqrt{R^2_{SMX} + z^2_{vtx}}} \tag{5.4.10}$$

where R_{SMX} is the radial distance from the z axis to the shower position in the CES or PES detector. The selection of the primary vertex is described in Section 5.2.

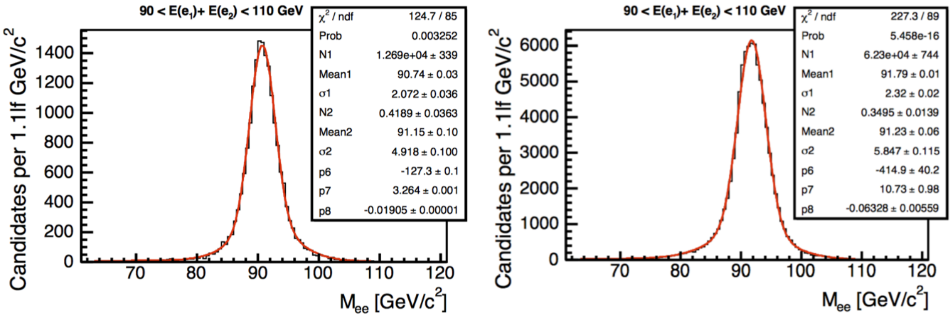


FIGURE 5.7: Example of the reconstructed Z boson mass from $Z \to e^+e^-$ data (left) and MC (right), in one bin of $E(e1) + E(e2)$. The curves indicate results of the fits to double Gaussian and a second-order polynomial background. Plots taken from Ref. [108].

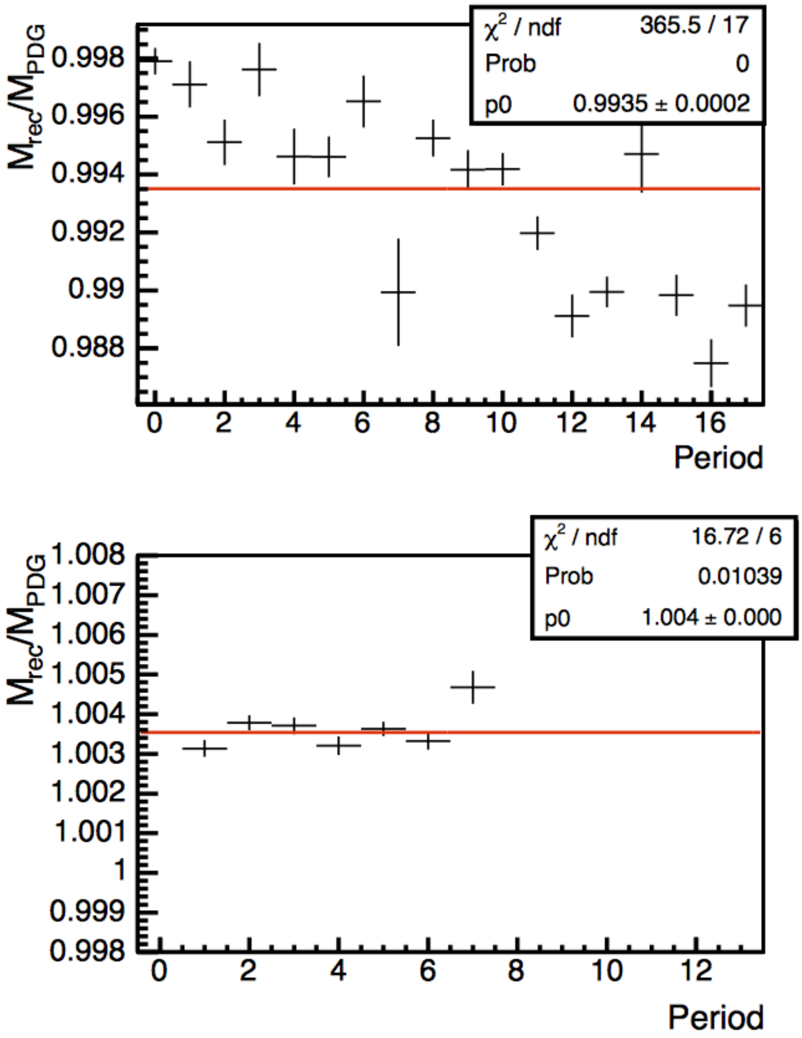


FIGURE 5.8: The energy scale factor in each run period from $Z \to e^+e^-$ data (top) and MC (bottom). Plots taken from Ref. [108].

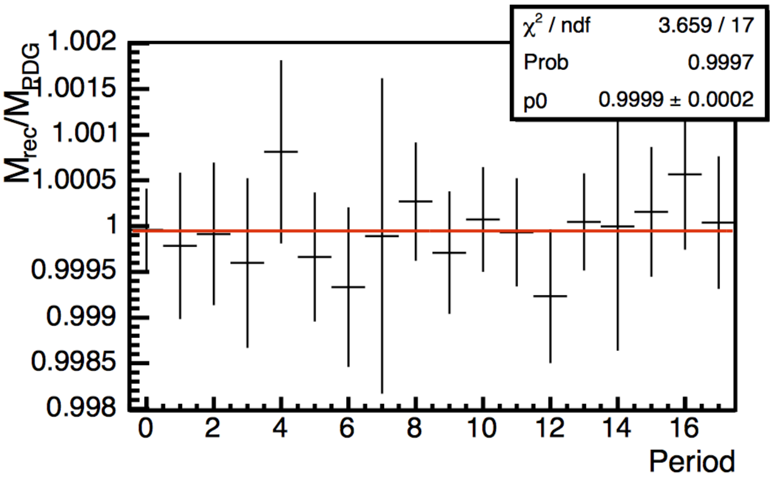


FIGURE 5.9: The energy scale factors in each run period from $Z \to e^+e^-$ data, after applying corrections to the energy of each electron, according to the numbers in Figure 5.8 (top). Plot taken from Ref. [108].

A real photon is massless and, therefore, the total vector momentum of photon candidates is set to the energy obtained from the EM calorimeter energy measurement: $p = E$ and $p_T = E_T$ (with $c = 1$). The photon four-momentum is then

$$P = (E, p_x, p_y, p_z) = (E,\ E_T \cos\phi,\ E_T \sin\phi,\ E\cos\theta) \tag{5.4.11}$$

where ϕ is determined from the azimuthal position of the EM cluster (Equation 5.4.3).

5.5 Missing Transverse Energy

The calorimeter measures both energy and position, thus it can be used to infer the production of non-interacting particles (such as neutrinos) by imposing energy conservation in the plane transverse to the beamline[2]. The calorimeter energy imbalance observed in a given event is referred to as the "missing transverse energy" $\vec{\not{E}}_T$ (or MET) and corresponds to the negative of the vector sum in the transverse plane:

---

[2] Generally the total vectorial summed transverse energy should be the same as at the point of the collision, namely zero.

$$\vec{\not{E}}_T = -(\sum_{towers} E_i \sin\theta_i)[\cos\phi_i, \sin\phi_i] \tag{5.5.1}$$

where E_i is the energy deposit in the i^{th} massless tower, θ_i the polar angle defined from the interaction vertex (Section 5.2) with respect to the tower position, ϕ_i the tower azimuthal angle. Towers included in the sum must be in the pseudorapidity region $|\eta|$ <3.6 and exceed a threshold of 100 MeV in the CEM, CHA, and WHA, 300 MeV in the PEM, 500 MeV in the PHA.

Equation 5.5.1 corresponds to the reconstructed "raw" missing transverse energy[3]. In electroweak processes, where heavy bosons are created, MET is likely to be associated to neutrino's transverse energy. On the other side, large MET can also arise from calorimeter mis-measurements. Both cases represent background sources for in this analysis where the photon candidate is expected to be well balanced to one or more jets.

5.6 Conclusion

This chapter was devoted to a basic description of track, vertex and calorimeter cluster reconstruction. In particular, the reconstruction and identification of photons are critical to the analyses presented in the remainder of this thesis and correction to the energy and timing variables are also discussed. The following chapter focuses on exploiting aspects of the reconstruction ideas introduced here.

---

[3] Some corrections are applied to the raw MET based on the true z vertex position, the difference between the raw and corrected E_T of the jets, the presence of muons in the event.

Chapter 6

Monte Carlo Simulations

In modern experimental particle physics, parton shower Monte Carlo programs have become an indispensable tool for data analysis, since they simulate the final states of high-energy collisions in full detail down to the level of individual stable particles. This chapter briefly outlines how they work. Moreover the predictive tools used to study the dataset under investigation are introduced.

Contents

6.1 Overview on Event generators

Precision QCD calculations and their experimental tests at hadron colliders are challenging. From a theoretical point of view the difficulty is mainly caused by our limited knowledge of parton fragmentation mechanism, diffractive processes, underlying events and particle structure functions. While at the experimental side, limitations come from finite detector or accelerator capabilities.

A parton-shower Monte Carlo (MC) event generator is a computer program designed to give fully exclusive description of QCD events. They are a powerful tool for high-energy collisions at particle accelerator experiments, particularly in data analysis where they are used together with detector simulation [21].

In Chapter 2 a brief overview of some concepts of the Quantum Chromodynamics theory was given. Those concepts are useful to understand the evolution of a hadron collision event. Perturbation theory allows to predict hard scattering cross sections, which corresponds to the first step of the Monte Carlo simulation chain. In particular, the QCD factorization theorem permits to decouple the cross section into different steps.

The MC generation process steps are schematically depicted in Figure 6.1. The first step is the random generation of the considered *hard scattering process* (shown as a black blob in Figure 6.1) at some high scale Q_0. Then, a *parton shower* stage follows (brown). When parton showering stops (at a scale of order 1 GeV), a *hadronization model* (yellow) is used to convert the resulting partons into hadrons. Modeling of a *underlying event* (green) is also needed to deal with the collision between the two hadron remnants; this is usually implemented through additional $2 \rightarrow 2$ scatterings (*multiple parton interactions*) at a scale of a few GeV. Many of the produced hadrons are unstable, so the final stage of event generation is the simulation of the *hadron decays*.

Finally, the products from the generated event are passed through a simulation of the detector geometry and response, and digitized to give an output identical to that of the real detector (Chapter 4). At this point physics objects are reconstructed in the same way for Monte Carlo and real data, as discussed in Chapter 5.

The several stages for a Monte Carlo generation process are discussed in the following sections.

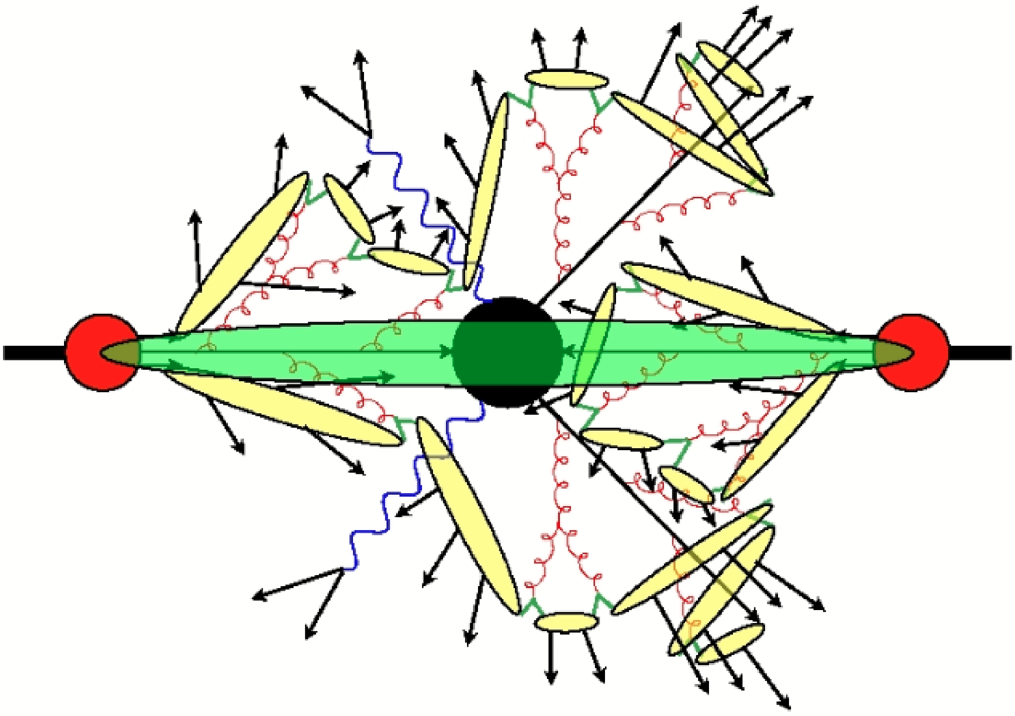


FIGURE 6.1: Schematics of the steps for the simulation of a hadron-hadron collision event, by a parton shower Monte Carlo event generator.

6.2 Hadron collision event

Factorization formula for QCD cross sections

Cross sections for a scattering subprocess $ab \to n$ at hadron colliders can be computed in collinear factorization as:

$$\begin{aligned}\sigma_{\mathrm{ab}\to\mathrm{n}} &= \sum_{a,b} \int \mathrm{d}x_a \mathrm{d}x_b \int f_a(x_a,\mu_F) f_b(x_b,\mu_F) \mathrm{d}\hat{\sigma}_{\mathrm{ab}}(x_a p_a, x_b p_b, \mu_R, \mu_F) \\ &= \sum_{a,b} \int \mathrm{d}x_a \mathrm{d}x_b \int \mathrm{d}\Phi_{\mathrm{n}}\ f_a(x_a,\mu_F) f_b(x_b,\mu_F) \ \times \frac{1}{2x_a x_b s} \left|\mathcal{M}_{\mathrm{ab}}\right|^2 (\Phi_{\mathrm{n}},\mu_R,\mu_F),\end{aligned} \tag{6.2.1}$$

where $f_a(x_a, \mu_F)$ are the parton distribution functions (PDFs) (described in Section 2.3.2), Φ_{n} is the final state phase space, $\frac{1}{2x_a x_b s}$ is the parton flux (with s being the ab center-of-mass energy squared) and $|\mathcal{M}_{\mathrm{ab}}|$ is the matrix element for the event of interest.

Matrix Element Calculations

The calculation of a hard scattering at hadron colliders is based on tree-level Feynman diagrams. Calculations at higher orders, which involve a combination of loop graphs and the soft parts from the QED and QCD radiation, are very difficult and are slowly becoming available. Generation of events is performed in the following way. The matrix element is numerically integrated over the phase-space and a differential cross section is estimated ($d\sigma(\overrightarrow{x})$, $\overrightarrow{x}$ being the degrees of freedom which have not been integrated). Events are generated by the "un-weighting" procedure. The hit-and-miss method is commonly used for the weighting: given $d\sigma(\overrightarrow{x})$, the maximum cross section $d\sigma_{max}$, and a random number g generated uniformly in the interval $(0, 1)$, the event is kept if $d\sigma(\overrightarrow{x})/d\sigma_{max} < g$. This ensures that the accepted events are distributed according to the description provided by the theory, and that all events have the same weight, as in the real data.

6.3 Parton Shower

The parton shower phase of event generators describe what happens to the incoming and outgoing partons involved in the hard collision. The partons involved in the hard process are coloured particles, quarks and gluons. From Quantum Electrodynamics (QED) it is well known that scattered electric charges radiate photons, this is what is called Bremsstrahlung. In the same way, scattered colour charges radiate gluons and this happens for partons on their way in and out of a collision. The main difference to QED is that, due to the non-Abelian structure of $SU(3)$, gluons themselves are coloured and so an emitted gluon can itself trigger new radiation. This leads to an extended shower and the phase space fills up with (mostly) soft gluons. Figure 6.2 shows a diagram of multiple gluon emission off an initial quark line. The showers modeling adds higher-order corrections to the hard subprocess, in an approximative way, since it is not feasible to calculate these corrections exactly: real radiative corrections to any inclusive quantity (like the hard cross section as computed at fixed order in pQCD) are divergent. Instead, an approximation scheme is used, in which the dominant contributions below a cut-off parameter are included iteratively ordered in sequence of, typically, smaller emission angles.

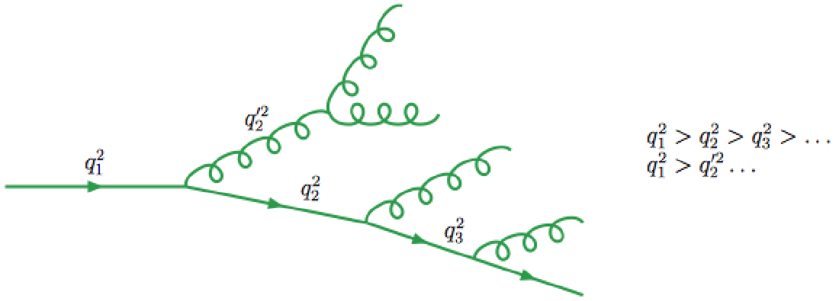


FIGURE 6.2: Diagram showing multiple gluon emission off an initial quark line [133]. The basic idea of parton showers is: accelerated colour charges radiate gluons, but since the gluons themselves are also charged an extended cascade develops. This cascade is modeled as an evolution downward in momentum scale. As we approach the non-perturbative limit, we get more and more radiation and the phase space fills with soft gluons.

These dominant contributions are associated with collinear parton splitting or soft (low-energy) gluon emission.

There are three possible processes for QCD emission: $q \to gq$, $g \to gg$ and $g \to q\bar{q}$. The cross section then factorizes into the product of the parent parton production cross section times a splitting factor. Considering the almost-collinear splitting of a parton of type i into jk, e.g. the $q \to qg$, the n-parton differential cross section before splitting is $d\sigma_n$, then after the splitting it becomes

$$d\sigma_{n+1} \approx d\sigma_n \frac{\alpha_S}{2\pi} \frac{d\theta^2}{\theta^2} dz d\phi P_{i,jk}(z, \phi) \tag{6.3.1}$$

where θ is the opening angle, ϕ is the azimuthal angle of the splitting, z is a parameter generally defined as a ratio the energy of i carried by j .

The function $P_{i,jk}(z, \phi)$ is the Altarelli-Parisi *splitting function* which describes the probability for a quark to emit a gluon, and is the only term that changes in Equation 6.3.1 between $q \to gq$, $g \to gg$ and $g \to q\bar{q}$ splittings. Figure **??**, shows a graphical representation of the splitting with the quark and the gluon (with four-momenta j and k respectively) being emitted at a small angle θ. The virtuality of the splitting parton is given by $q^2 = (j + k)^2$.

By sequential application of Equation 6.3.1, θ, ϕ and z are generated randomly during the Monte Carlo simulation process for each splitting; thus, a parton shower is developed from each coloured parton of the hard subprocess. An important feature of the showering algorithm is the evolution variable. The simplest evolution

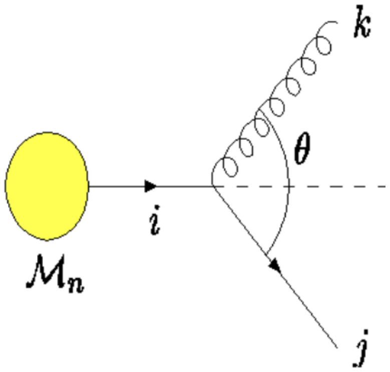


FIGURE 6.3: Graphical representation of the $q \to qg$ splitting. The yellow circle represent the $\mathcal{M}_n$ amplitudes of the tree-level processes.

variable to understand is the virtual mass-squared (the virtuality, q^2) of the partons in the shower. The splitting process $i \to j+k$ cannot take place with all the partons on their mass-shells. The dominant contributions will come from configurations in which the virtualities are strongly ordered, with the parton nearest to the hard subprocess farthest from its mass shell and the virtualities falling sharply as the shower evolves away from it. The upper limit on the initial virtuality is set by some momentum transfer scale Q of the hard subprocess, $q^2 < Q^2$, and the shower is terminated when the virtualities have fallen to the hadronization scale, $q^2 = Q_0^2 \sim$1 GeV^2. At this point, allowing for n splitting processes, the cross section can be written as:

$$\sigma_0 \frac{1}{n!} \alpha_S^n \log^n \frac{Q^2}{\lambda^2}, \tag{6.3.2}$$

where Q is the upper cut-off scale called *annihilation energy* that determines when the showering starts, and λ is the infrared cut-off. This procedure is called “leading log approximation”.

Once the shower is developed, the vertices and lines of the final configuration are assigned weights, which are for each vertex:

$$\theta(q - q_0) \ \frac{\alpha_S(q)}{2\pi} \ \frac{dq}{q} \ P_{i,jl}(z) \, dz \, \frac{d\phi}{2\pi}, \tag{6.3.3}$$

and for each line are the so-called *Sudakov form factors*:

$$\Delta_i(q_1^2, q_2^2) = \exp \left[- \sum_{(jk)} \int_{q_2^2}^{q_1^2} \frac{dq^2}{q^2} \int_0^1 dz \ \ \frac{\alpha_S(q)}{2\pi} \ \ P_{i,jk}(z, \phi) \right] \tag{6.3.4}$$

where q_1 is the value of q at the upstream vertex, and q_2 at the downstream vertex.

At the end of the graph, q_2 is substituted by a cut-off q_0. The Sudakov form factors specify the range of the z parameter for which the splitting is resolvable and represent the probability of *not* splitting.

Thus, the generation of a parton shower thus proceeds as the following. Given the initial scale Q^2 , one solves the equation $\Delta_i(Q^2, q_1^2) = R_1$, where R_1 is a random number uniform on the interval [0,1], for the scale q_1^2 of the first splitting. If $q_1^2 < Q_0^2$, then the splitting was unresolvable and the showering of that parton terminates. Otherwise, if the splitting was $i \to j+k$, then the procedure is repeated on parton j using $\Delta_j(q_1^2, q_2^2) = R_2$ to determine the scale q_2^2 for the splitting of that parton, and similarly for parton k, and so on, until all attempted splittings have fallen below the resolvable scale Q_0^2. At each splitting, the variables z and ϕ are chosen according to the distribution $P_{i,jk}(z, \phi)$ using the Monte Carlo method, with z in the resolvable region specified by the limits of integration in the Sudakov form factor.

Hadronization

In interactions that produce a final-state parton, the escape of the new parton from the local color field causes a process known as *hadronization* or *fragmentation*. The parton effectively begins to radiate gluons, which then decompose into $q\bar{q}$ pairs, which radiate gluons, and so on, creating a shower of partons with progressively smaller momenta. Eventually the partons will cluster into colorless hadrons, the collection of which is called a "jet". While the creation of a final-state parton can be treated within pQCD, the evolution of a parton into a shower of low-energy particles eventually crosses into the non-perturbative regime. The dynamics of this evolution are absorbed into fragmentation functions, which give the probability for a parton to produce a final-state hadron through fragmentation.

Thus, the σ_{ab} cross sections for a scattering subprocess $ab \to n$ at hadron colliders, discussed in Section 6.2 and given by Equation 6.2.1, can be further modified to compute the process $a + b \to C + X$:

$$\sigma_{ab\to C+X} = \int dz_C D_{c_k}(z_C, \mu_f^2)\sigma_{ab\to c_k+X}(\mu_F^2, \mu_R^2)$$

(6.3.5)

where C is a hadron, D_{c_k} is the fragmentation function that defines the probability of a parton c_k fragmenting into C with momentum fraction $z_C = p_C/p_{c_k}$. These fragmentation functions depend on a fragmentation scale, μ_F , and effectively remove the same singularities in the final state as those removed by the factorization theorem in the initial state. As with the PDFs, measurements of the fragmentation functions can be made at a fixed value of Q^2 and evolved through the DGLAP (see Section 2.3.2) equations to extrapolate to larger Q^2. The only physical observables accessible to collider experiments are those measured after hadronization, by which point they are said to be measured at the *particle level*. This means that any cross section computed at the parton level, i.e. before hadronization, is not something that can be observed directly. Monte Carlo programs are typically used to estimate the impact of hadronization on observables to allow theoretical predictions (almost always made at the parton level) to be meaningfully compared with experimental measurements (almost always made at the particle level). In Figure 6.4 the different levels for objects reconstruction can be seen.

The unphysical scales μ_R, μ_F , and μ_f are usually set to be equal (with their common value called μ). In the case of prompt photons, where the photon momentum is usually balanced by a single jet in the transverse direction (see Section 2.3.3), a convenient choice for the scale is the transverse energy of the photon, E_T^γ . This, however, is a somewhat arbitrary choice, and there is some systematic uncertainty associated with the exact choice of scale. This uncertainty can be evaluated by varying the scales around their nominal values and observing the change in value of the observable; a common range for photon measurements is $0.5E_T^\gamma < \mu < 2E_T^\gamma$.

The two main hadronization models in current use are the *string* and *cluster* ones. These are described in the following. The main difference is that the former transforms partonic systems directly into hadrons, while the latter employs an intermediate stage of cluster objects, with a typical mass scale of a few GeV.

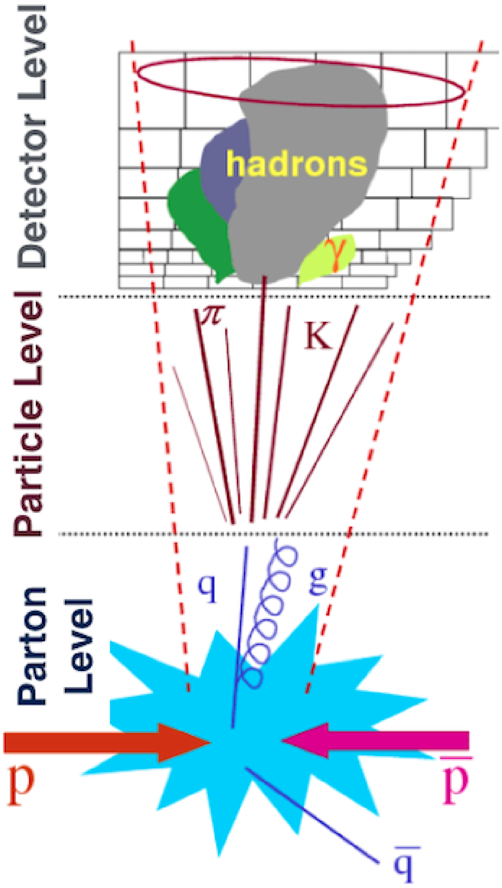


FIGURE 6.4: Measured objects can be defined at: (i) the parton level, (ii) the particle level (the stable particles, i.e., particle with a lifetime of at least 10 ps in events simulated with MC generators), and (iii) the detector level (calorimeter towers).

The string model is based on the assumption of linear confinement.

A quark corresponds to an endpoint of a string, and a gluon to a kink on it, with partons ordered in color along the string. The string offers a very predictive framework for how its space–time motion and breakup translates into an energy–momentum distribution of the primary hadrons. This framework also applies for complicated multiparton configurations, and has been successfully tested in e^+e^- collisions.

The cluster hadronization model is based on the preconfinement property of parton showers, which leads to color-singlet parton clusters with a universal mass distribution at low scales.

Cluster hadronization starts with non-perturbative splitting of gluons into quark-antiquark (and possibly diquark-antidiquark) pairs. Clusters are then formed from color-connected pairs.

Most clusters undergo quasi-two-body sequential phase-space decay. The limited cluster mass spectrum naturally leads to limited transverse momenta and suppression of heavy flavor, strangeness and baryon production.

The decay of heavier clusters requires a more string-like initial stage of anisotropic decay into lighter clusters.

When combined with angular-ordered parton showers, the cluster model gives a fairly good overall description of high-energy collider data, usually slightly less good than the string model but with fewer parameters.

The busy environment of high-energy hadronic collisions could lead to nontrivial collective effects, currently not simulated either in string or in cluster models.

6.4 Initial- and Final-State Radiation

As already stated, colored particles can radiate gluons and charged particles can radiate photons. Initial-state radiation (ISR) refers to the radiation from an initial-state particle before the collision occurs. As a result, an additional object is observed in the detector that do not come from the primary interaction. Similarly, final-state radiation (FSR) is radiation from a final-state particle after the collision; an additional particles will also be reconstructed; this time some of the energy from one of the final-state particles produced is taken out. PYTHIA has a set of parameters that can be adjusted, or *tuned*, to control the amount of ISR/FSR in an event. The default CDF set of values used for generation of the MC samples is called *Tune A* [135].

6.5 The Underlying Event

The preceding steps of hard process, parton shower, hadronization are not sufficient to fully describe the final state of the hard process, in which a high energy parton from each incoming hadron interact to produce an arbitrarily complex final state. This process involves the extraction of a coloured parton from each of the hadrons, which are colourless bound states of many coloured partons. Therefore, in addition one has to consider how the hadron remnants evolve, hadronize and, potentially, interact with each other.

“Underlying event” (UE) refers to the secondary parton interactions at low momentum transfer that occurs together with the main hard process. This effect primarily result as soft particles which give an the additional contribution to the energy deposited in the detector around the interaction point. The underlying event is flavor- and color-connected to the hard scattering and in real data in general can not be separated from the event of interest. It is typically observed as jets of particles close to the direction of the beam and cannot be modeled with perturbative QCD. Phenomenological models are used to simulate the underlying event and are tuned to *minimum bias*[1] data from hadron colliders [44]. Indeed processes involving IR divergences have free parameters that can (must) be phenomenologically constrained. Hadronization models typically have 10 to 30 free parameters, Multi-Parton Interaction (MPI) models usually add at least 5 more free parameters. The Pythia parameters that control the behavior of the underlying event are “tuned” in such a way that the MC simulation better fits measured distributions from Tevatron data, such as the p_T spectrum from jets or the p_T spectrum from lepton pairs produced by Z boson decays [135, 136]. The MC generators used for this analysis use the standard CDF UE tune, Tune A [135].

Pile-up

In-time pile-up events arise from the multiple parton inelastic (MPI) scatterings of protons in the same bunch of the hadron generating the hard process of interest. They mainly consist in soft QCD interactions and are modelled in a similar way as the UE. Pile-up is also tuned to minimum bias events [44].

---

[1] A “bias” in event selection is, in general, any kind of assumption made on the final state and therefore any kind of cut applied in selecting only these events. In minimum bias events only very loose requirements are imposed on the collision data.

6.6 Hadron decay

Hadrons produced during the hadronization are often unstable and decay into the stable particles that are observed in the experiment. A “stable” hadron has ambiguous definition. From the detector point of view, only a hadron sufficiently long-lived to interact with the detector material before decaying can be identified; such a particle is considered as a stable.

The hadron decay modeling uses a set of look up tables containing branching ratios and decay modes. All particles are included with their proper mass distribution and decay properties.

6.7 Generators

Monte Carlo generators can be classified as either *multi-purpose* generators, i.e. capable of performing the full simulation chain described in this chapter, or as *specialized* generators, i.e. devoted to a single aspect of the simulation which they often describe with higher accuracy than typical multi-purpose generators. As should be clear after the discussion about the modeling of hadronic collisions, there are some choices to be made that might result in some generators being better suited to describe particular processes. Therefore the Monte Carlo generators should be chosen according to their performance in modeling the event of interest. The following sections describe the main characteristics of the generators that are used in the analyses that are the subject of this thesis.

6.7.1 PYTHIA

PYTHIA is the most senior and established general-purpose event generator. The `Fortran` version PYTHIA 6 is still very widely used.

Here follows a very brief summary of the PYTHIA 6 program. The physics is documented in the PYTHIA 6 manual [110] and the literature quoted there.

The PYTHIA event generator is a leading-order (LO) parton-shower Monte Carlo program, capable to generate a broad range of QCD, electroweak, and more exotic events. It uses LO Matrix Element calculations to generate hard interactions between partons and it is optimized for 2 → 1 and 2 → 2 processes. It accounts for QED radiation emitted off quarks in the initial state (ISR) and final state (FSR). The underlying event is simulated using the MPI model, and the hadronization of partons is modeled with the Lund string model [28]. The fragmentation component is modeled as FSR, with no knowledge of the fragmentation functions D_k^γ. PYTHIA 6.216 was used for the production of signal and background MC samples and to compute the expected prompt-photon production cross section; of the available PDF datasets, CTEQ5L has been selected.

6.7.2 SHERPA

Likewise PYTHIA, SHERPA , acronym for **S**imulation for **H**igh-**E**nergy **R**eactions of **PA**rticles, is designed to be a multi-purpose event generator for particle production at high-energy colliders [140]. SHERPA is a publicly available code and can be obtained from `https://sherpa.hepforge.org`. It has been written in `C++` from the beginning.

As in all Monte Carlo event generators the simulation of an actual event is split into different steps corresponding to the evolution stages of a scattering process from the high scale of the hard interaction down to scales of order Λ_{QDC} where hadronization sets in and the hadrons seen in the detectors are formed. Within SHERPA different physics modules host the different stages of the event evolution. The most important of them shall be briefly reviewed here.

It is a simulation of higher-order perturbative QCD effects, including NLO corrections to hard processes and resummation as encoded in the parton shower; It contains a very flexible tree-level matrix-element generator for the calculation of hard scattering processes within the Standard Model and various new physics models. The emission of additional QCD partons off the initial and final states is described through a parton-shower model. This technique is known as “Matrix Element plus Parton Shower merging” (ME + PS). A cluster model is used for hadronization. For the underlying event SHERPA uses a multiple-interaction model based on that of PYTHIA but differing in some respects.

Both SHERPA and PYTHIA generators have their own hadron decay modules with extensive tables of particle properties, branching fractions and decay distributions.

While these have much in common, they are not identical since, as explained above, significant physics choices and modelling are involved. For the same reason they are not interchangeable, since these features influence the values of parameters relevant to other parts of the program that are needed to give the best agreement with experimental data.

A simple model of multiple interactions is used to account for underlying events in hadron-hadron collisions.

SHERPA 1.4.1 generator with CT10 PDFs has been used to produce MC signal samples and to compute the expected prompt-photon production cross section.

6.8 Detector Simulation

Once all the stable hadrons are generated, their propagation through the detector material is simulated using the GEANT 3 package [121]. GEANT is a platform for handling:

- Geometry: physical layout of the experiment, including detectors, absorber and how the layout would affect the path of the particles.
- Tracking: mathematical models are used to simulate the passage of a particle through matter. Those models include particles interactions, which cause showering to secondary and tertiary particles, and decays.
- Detector response: simulate how the detector would respond as the particle passes through its volume.
- Detector conditions: the conditions of the detector for each time period can be inputted in order to mimic as close as possible the real data. The status of the different sub-detectors is taken into account.

The modeling of the particle interaction at the passage through matter is computationally intensive. Therefore the simulation of the calorimeter response is performed using parametrized response function tuned to the data. This is done by employing a program called GFLASH [123]. The result is a rapid and accurate response of the individual calorimeter towers to the energy deposited by the outgoing particles. TRIGSIM [122] is used to simulate the trigger data and the corresponding digital response from the detector. Simulated events are analyzed with the same CDF reconstruction software that is used for the collision data, and physics objects are constructed based on vertex, tracking, and calorimeter variables.

Chapter 7

Datasets and Event Selection

In this chapter, the data and MC samples used for the measurement are described. Moreover, the trigger and event selection criteria are discussed. Each cut is designed to separate real prompt photons from photons from hadronic jets, electrons and other backgrounds. Emphasis is put on the definition of the variables used for the photon identification.

Contents

7.1 Datasets

7.1.1 Data samples

For this measurement the full CDF Run II dataset is used. As already mentioned in Chapter 4, data correspond to the so called “cph1 stream” in the Stntuple format, covering the period 0-38 of data taking. These samples contain a trigger bit variable that is set to true for each event that was collected by the dedicated photon triggers (Section 7.3.1).

Not all data in the “cph1” samples were actually stored if the detector was not fully operational. Thus, in conjunction with the triggering system a list of runs, for which all the necessary subsystems were properly working during data taking, is defined. Such a set of suitable runs is based on good run bits (see Section 3.7) saved into a text file, known as “Good Run List”, which is read in and stored in an array. The array is pointed on an event-by-event basis to determine if a particular event should be included in the analysis. The runs included determine the integrated luminosity of the data sample. The Good Run List used for this analysis is the `goodrun_v45_pho_01.txt`, and corresponds to an integrated luminosity of 9.5 fb^{-1}. More details on data samples are provided in Appendix C.

7.1.2 Monte Carlo samples

Monte Carlo (MC) samples are used to calculate the photon fraction in each E_T^γ bin, to unfold the cross section measurement back to the hadron level and to compare results to theoretical predictions. These samples are generated using Gen6 PYTHIA 6.216 Tune A with realistic simulation and SHERPA 1.4.1 with Tune A minbias and realistic simulation. A trigger simulation is used for these events in order to apply the same requirements used for data selection.

More details on Monte Carlo samples can be found in Appendix C.

7.2 Photon selection variables

The photon identification, “photon ID” in the following, is based on a set of detector variables designed to distinguish prompt photons from other particles. The experimental signature of a photon is an energy cluster in the electromagnetic calorimeter (*EM cluster*) with no charged particle tracks pointing to it. Thus, photon ID mainly relies on how isolated the photon is in the calorimeter, on the amount of energy in the EM vs. HAD calorimeters, and the presence of tracks.

The selection aims also to distinguish prompt photons, i.e. photons originating from the hard scatterings (see Chapter 2) and collinear photon pairs from the decay of a neutral meson inside a jet. Thus, additional requirements, called cuts, on each individual detector variable are applied. The selected cut for a particular variable is based on the different spread of values measured that has to be consistent with that of a prompt photon.

All the different variables used to select the photon signal are listed below.

- *CES fiducial*

  The fiducial region is defined by the CES local coordinates $|x| < 21$ cm and 9 cm $< |z| <$ 230 cm. These requirements correspond to the active region of the detector and ensure that the energy cluster is well contained.

- *CES* χ^2 *(Strip+Wires)*

  The CES detector measures the lateral shower profile of the photon candidates in the event. The observed shower shape is compared to the predicted shower shape for a single photon, obtained from test beam. The consistency between these two showers is evaluated with a χ^2 parameter, the "CES χ^2". The "CES χ^2" is calculated for every strip (wire) layer by comparing the energy in the 11 strips (wires) of the CES strip (wire) cluster to what we expect from a single shower obtained from electrons in a test beam experiment.After computing the χ^2 for each CES plane, the χ^2 of the fit is defined as the average $\chi^2 = (\chi^2_{strip} + \chi^2_{wire})/2$. The shower profile cut requires the CES χ^2 to be below 20[1] I

- *Transverse Energy*

  The E_T is the transverse component of total energy deposited by the photon in the EM calorimeter, defined in Section 5.4. Since photons are massless objects, $E_T = p_T$ .The corrections applied to the measured energy are discussed in Section 5.4.3. All photon candidates are required to have corrected $E_T >$30 GeV.

- *Hadronic Leakage*

  The ratio of energy deposited in the hadronic calorimeter[2] to the energy deposited in the electromagnetic calorimeter (Had/EM) for the photon

---

[1] If the photon candidate CES χ^2 is above 20, it may come from a meson decay.

[2] The considered hadronic energy deposition (E_{Had}) lies directly behind the EM cluster

cluster towers is required to be small: for events with photons with E_T^{γ} <200 GeV (E_T^{γ} >200 GeV), each photon is required to have $\frac{E_{Had}}{E_{EM}} < 0.055 + 0.00045 \times E_T$ ($\frac{E_{Had}}{E_{EM}} < 0.125$). This requirement aims to reject background from jets: electromagnetic showers deposit most (typically $> 95\%$) of their energy in the electromagnetic calorimeters, while hadron showers in general deposit energy in both the hadronic and electromagnetic compartments.

- *The calorimeter isolation*

  A well isolated photon occupies 1 or 2 calorimeter towers without leaving additional energy outside of the EM cluster. The calorimeter isolation $E_T^{\gamma,ISO}$ is a measure of the energy surrounding the EM cluster. It is defined as the difference between the energy deposited in a cone of radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} <$ 0.4 around the candidate, in both the EM and Had calorimeters and the energy of the EM cluster:

  $$E_T^{(\gamma,ISO)} = E_T^{R=0.4} - E_T^{cluster} \tag{7.2.1}$$

  Such a cut is very helpful to reject most of the QCD background except for the case when the pion carries almost all the momentum of the jet. This occurs in one out of every 1000 cases, thus since the jet cross section is also approximately that much high, the isolation cut leaves a signal-to-background ratio of approximately 1.

- *Track rejection and track isolation*

  There should not be charged particle track associated with the EM cluster from a photon. However, one track pointing to the EM cluster may remain acceptable in a photon candidate if the p_T is not characteristic of an electron when compared to the deposited energy. Thus the maximum allowed p_T of this track is limited by a constant term but can grow with the corrected energy of the EM cluster E_T. Since jets contain charged particles that leave tracks, there is also a requirement on the *isolation track, SumPt4*, defined as sum of all tracks in a cone around the EM cluster. This selection cut constrains the sum of the transverse momenta of all the tracks within 5 cm of the vertex and $\Delta R < 0.4$ compared to the direction of the cluster.

- *Lateral shower shape*

  The purpose of this variable is to distinguish the photon and electron shower development in the EM calorimeter to that of jets, based on the lateral sharing of energy in towers adjacent to the EM cluster's seed tower. The *Lshr* variable (pronounced *el share*) is a measure of how well the EM shower profile agrees with the expected profile for a single EM shower. Lshr is defined as

$$Lshr = \frac{0.14 \sum_i (E_i - E_i^{exp})}{\sqrt{(0.14\sqrt{E_{EM}})^2 + \sum_i (\Delta E_i^{exp})^2}} \tag{7.2.2}$$

where the sum is over towers in an EM cluster adjacent to the seed tower and in the same wedge as the seed tower (i.e. over either a one or two tower sum). The value E_i is the measured energy in an adjacent tower, E_i^{exp} is the expected energy in the adjacent tower obtained from test beam data, E_{EM} is the total EM measured EM energy of the 1–3 tower cluster, and ΔE_i^{exp} is an estimate of the i uncertainty in E_i^{exp}.

- *CES/CEM*

  This is the ratio of energy measured by the CES to the energy measured by the CEM. This variable helps distinguish prompt photons and photons from neutral meson decays.

- *"PMT Asymmetry", A_P*

  A "Spike Killer" function is applied to rejects events caused by a high-voltage breakdown ("spike") between the PMT photocathode and the surrounding material. It is based on the so-called "PMT Asymmetry" $A_P = \dfrac{|E_{PMT1} - E_{PMT2}|}{E_{PMT1} + E_{PMT2}}$ where E_{PMT1} and E_{PMT2} are the two PMTs response energy [142]. A_P is required to be less than 0.6. More details can be found in Section 8.3.

- *MET/E_T^γ*

  A already stated, the missing transverse energy, (defined in Section 5.5) is helpful to suppress background from electroweak and noncollision processes. High MET events are vetoed by applying the condition MET/E_T^γ <0.8. The reasons for this cut will be discussed in Chapter 8.

- *Photon Timing Variable*

  It corresponds the *corrected time* of arrival [142], t_{corr}, defined in Section 5.3. The photon timing variable is used to separate prompt photons from noncollision backgrounds.

For a prompt photon, with “perfect measurements”, $t_{corr} = 0$ ns. Since the detector is not perfect this measurement has an intrinsic resolution and is thus represented by a Gaussian centered at $t_{corr} = 0$ ns.

Secondly, photon candidates that have nothing to do with the collision and originating from sources external to the detector, typically from “cosmic rays", present another source of large t_{corr} events. These events are discussed in more detail in Section 8.3.1. For now it is sufficient to remark that the signal region is defined between about -6.6 ns and 6.6 ns. Other regions are dominated by wrong vertex, or cosmic rays. Each of which can potentially be measured as a background using data.

Table 7.1 gives a comprehensive summary of *photon identification variables* typically used.

Table 7.1: Set of typical *photon identification variables*.

	Photon identification variables
Fiducial	$\|X_{CES}\| < 21$ cm and 9 cm $< \|Z_{CES}\| < 230$ cm
E_{Had}/E_{EM}	The ratio of the energy in the hadronic calorimeter behind the cluster, to the energy in the clusters as measured in the EM calorimeter
E_{ISO}(GeV)	Energy in a cone of $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.4$ around the object, excluding the cluster energy
$\sum p_T$(GeV/c)	Total p_T of tracks in a cone of $\Delta R = 0.4$ around the cluster
A_P	$\frac{\|E_{PMT1} - E_{PMT2}\|}{E_{PMT1} + E_{PMT2}}$ where E_{PMT1} and E_{PMT2} are the two PMT energies
χ^2_{Strip}	A χ^2_{Strip} comparison of the shower-maximum profile to test beam data expectations
Lshr	A comparison to the energy deposition in adjacent towers, to expectations

7.3 Event selection

The event selection is one of the most important part of the analysis, by which as much as possible background has to be suppressed while keeping the signal events. The selection of the candidate events is a three stages process. The first stage is the online selection when dedicated triggers are applied; this occurs during the data taking, even before the data is written on the storage. The second stage is the $\gamma + X$ “selection”; this takes place offline.

7.3.1 Inclusive Photon Trigger

The data and MC events are required to have passed at least one of the following trigger paths: PHOTON 25 ISO, ULTRA PHOTON 50 or SUPER PHOTON 70 EM. The first path requires one isolated electromagnetic cluster with $E_T > 25$ GeV; the second path requires one electromagnetic cluster with $E_T > 50$ GeV and has no isolation requirement. The SUPER PHOTON 70 EM trigger only applies a loose E_T cut and a loose HAD/EM cut, which prevents a potential inefficiency arising at high E_T where the EM energy becomes saturated causing the HAD/EM to be miscalculated. The transverse energy in the trigger uses $z = 0$ for the vertex.

Details of the trigger specifications are outlined in Table 7.2.

TABLE 7.2: Summary of the requirements implemented in the trigger paths at different levels.

	`PHOTON 25 ISO`	`ULTRA PHOTON 50`	`SUPER PHOTON 70`
Level 1			
Trigger Tower E_T^γ	> 12 GeV	> 12 GeV	> 20 GeV
Level 2			
L2 EM Cluster E_T^γ	> 21 GeV	> 40 GeV	> 70 GeV
L2 EM Cluster E_{HAD}/E_{EM}	< 0.125	< 0.125	< 0.125
L2 EM Cluster E_T^{iso}	$< 3.0 GeV \| \| 0.15 E_T$	-	-
Level 3			
L3 EM Cluster E_T^γ	> 25 GeV	> 50 GeV	> 70 GeV
L3 EM Cluster E_{HAD}/E_{EM}	$< .055 + .00045\ E_T^\gamma$ (for $E_T^\gamma <$200 GeV)	< 0.125	$< 0.2 + 0.001\ E_T^\gamma$
L3 EM Cluster E_T^{Iso}	$< 2 \| \| 0.10\ E_T^\gamma$	-	-
L3 EM Cluster χ^2_{CES}	< 20	-	-

7.3.2 Photon Selection

All events are required to be marked “good” for photons using the goodrun list and have a class 12 vertex. To maintain the projective geometry of the calorimeter towers, all events must have a well reconstructed primary vertex within 60 cm around the center of the detector. Events are required to have at least one photon candidate. The photon candidate is the one passing the “Loose photon ID cuts”, listed in Table 7.3, with energy in the range of 30 GeV $< E_T^\gamma <$ 500 GeV and pseudorapidity $|\eta^\gamma| < 1.0$. If there are multiple photon candidates which pass the η^γ and the E_T^γ cuts, the one with the highest E_T^γ is selected.

Table 7.3: Set of cuts for Loose photons.

Variable	**Loose Cut**
E_T^γ	$>$25 GeV
CES X and Z Fiducial	CES $\|X\|$<21 cm, 9<CES $\|Z\|$<230 cm
E_{HAD}/E_{EM}	$<$0.125
$E_T^{\gamma,ISO}$	$<0.15 E_T^\gamma$ for $E_T^\gamma \leqslant 20$ GeV
	$<$2.0 GeV for $E_T^\gamma >$20 GeV
SumPt4	$<$5.0 GeV/c
Track p_T	$<0.25 \times E_T^\gamma$

In order to remove noncollision backgrounds, a photon candidate with MET>0.8 $_T^\gamma$ is rejected; the spikes rejection is also applied. The photon timing must be consistent to the collision timing: $|t^\gamma|$ <6.6 ns.

Chapter 8

Background Subtraction

The photon selection criteria described in Chapter 7 exploit several variables in order to separate signal from background. However, the background rejection is not perfect and non-prompt photons still remain in the candidates sample. An estimate of this residual contamination is an important part of the measurement of the final cross section. The first part of this chapter deals with the estimation of the main background coming from light mesons decays; this is performed via a statistical technique applied to the outcome of a dedicated Artificial Neural Network. The final section discusses other minor sources of backgrounds including noncollision backgrounds.

Contents

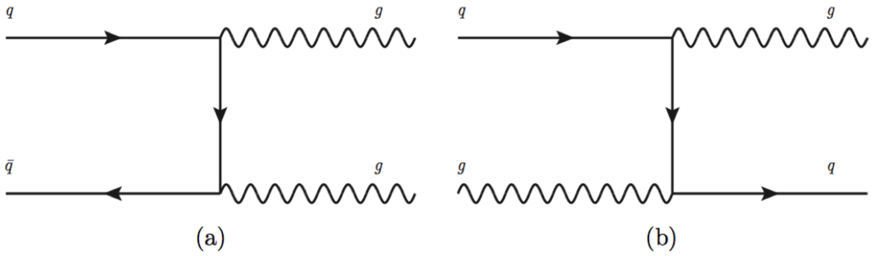
8.1 Introduction

The inclusive photon sample ($\gamma + X$) is contaminated by backgrounds arising from several sources. These can be separated into two different categories: collision and noncollision events.

Collision backgrounds come from SM production, such as strong interaction (QCD) and electroweak processes. Noncollision backgrounds come from photon candidates that are either emitted by cosmic ray muons as they traverse the detector or are from beam related backgrounds producing an energy deposit in the calorimeter that is reconstructed as a photon. In the following Sections each background is discussed and the signal estimation technique is outlined.

8.2 Neutral meson background

The QCD multijet background is the largest source of fake prompt photons, especially at at low E_T. In these events jets from light neutral mesons decay into photons. Often, π^0 and η particles are produced and they decay almost exclusively into several photons. Their dominant decay modes are shown in Table 8.1. As shown in Figure 8.1, some of the Feynman diagrams of jet production are very similar to those of prompt photon production (Figure 2.5), in which the photon is replaced by a quark or a gluon.



(a) (b)

FIGURE 8.1: Example of Feynman diagram for dijets production in the SM. Light mesons from these jets can eventually decay into one or more photons. This background is very significant as the dijet production cross section is much larger than that of $\gamma + jets$.

Some of these photons are rejected with the isolation cuts at trigger and offline levels, and with the CES χ^2 cluster cuts, as explained in Section 7.3. However, when these photons are collinear, they fake a single photon shower in the calorimeter, reducing the effectiveness of the CES based cuts. Since in most of

TABLE 8.1: Dominant decay modes of light mesons (π^0 and η) as a fraction of the full decay width (Γ_i/Γ). Values taken from Ref. [21].

Meson	**Dominant Decay mode**	**Fraction** (Γ_i/Γ)
π^0	$2\ \gamma$	$(98.823 \pm 0.034)\%$
η	$2\ \gamma$	$(39.41 \pm 0.20)\%$
	$3\ \pi^0$	$(32.68 \pm 0.23)\ \%$

the cases photons from mesons will be accompanied by other hadronic particles that will deposit their energy around the photon candidate, the isolation cut provides a highly effective way to remove these contributions. In the case the hadronization of the hard scattered parton results in most of the energy transferred to the mother meson, the resulting photons will be not eliminated by the photon ID cuts. These photons cannot be distinguished from prompt photons on an event by event basis. Instead, this background is removed in a statistical manner.

In the previous measurement [38], the background subtraction method was based on the isolation energy in the calorimeter around the photon candidate. The fraction of the data determined to prompt photon production (signal fraction) was estimated by a χ^2 fit of the isolation distribution in the data to signal and background Monte Carlo isolation templates, for every bin in photon E_T .

In this measurement, the signal fraction is evaluated by means of an Artificial Neural Network, trained to better discriminate between photon signal-like and photon background-like events (Section 8.2.1). The output is fit with a dedicated Likelihood technique (Section 8.2.2).

8.2.1 Artificial Neural Network

An Artificial Neural Network (ANN) is any simulated collection of interconnected neurons, with each neuron providing a particular response for a given set of input signals [145]. By giving an external signal to some input neurons the network is set into a defined state that can be determined from the response of one (or several) neurons. Thus the Artificial Neural Network can be seen as a mapping from a space of input variables $\{x_1, ..., x_n\}$ onto a one-dimensional (for example in case of a signal-versus-background discrimination) or multi-dimensional space of output variables $\{y_1, ..., y_m\}$. The mapping is nonlinear if there is at least one nonlinear response to a neuron input.

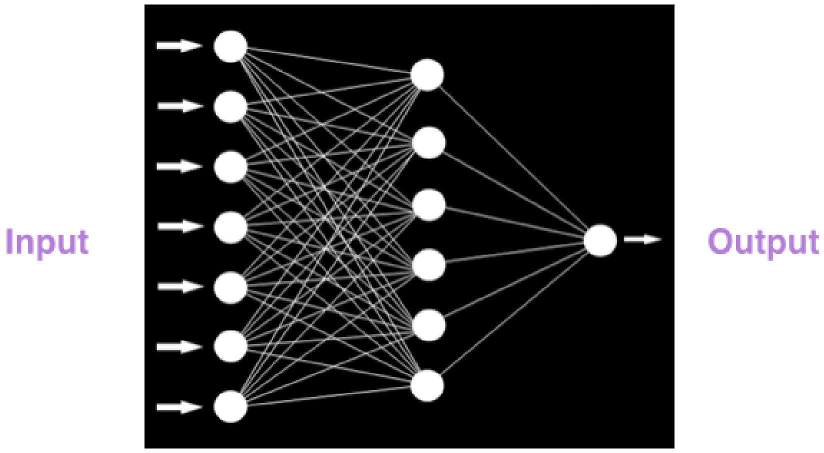


Figure 8.2: Diagram of an Artificial Neural Network (ANN). This multivariate technique is often used in High Energy Physics in order to discriminate a signal from backgrounds. The learning algorithm used to train ANN relies on a set of input variables from which it *learns* the statistical expected behaviour and gives as output a single continuos variable which determine how an event is signal-like.

The Toolkit for Multivariate Analysis (`TMVA`) provides a Root-integrated [144] environment for the processing, of multivariate techniques. All multivariate methods in `TMVA` respond to supervised learning only, i.e., the input information is mapped in feature space to the desired outputs. `TMVA` is specifically designed for the needs of high-energy physics (HEP) applications. The package includes Artificial Neural Networks [145].

The ANN developed for the photon identification in CDF is briefly discussed in the next section.

8.2.1.1 ANN training

Multivariate techniques (MV) require a dedicated signal sample of particles known to be *true* photons and a background sample representative of *fake* photons. Using iterative training algorithms, these techniques learn how to distinguish between the signal and background samples. To achieve the best performance, it is important to carefully choose training samples that match the desired use of the multivariate classifier. As stated above, the most common photon *fakes* are neutral pions or etas from jets. Thus, the algorithm is trained to distinguish harder cases where a jet looks like a real photon.

The signal photons were chosen from a simulated sample of inclusive photon production[1], while *fake* photons were chosen from a sample of simulated jets[2]. The background candidates were checked to ensure that they were not actually photons radiated by a quark. This requires a matching between the generator-level particles in the simulation and the particles reconstructed from the simulated detector output. As a result, the matching process required that any candidate background photon could not match to a generator level radiated photon. All selected photon candidates were required to pass a set of loose photon cuts (see Chapter 7).

Several MV algorithms were *trained* with input variables sensitive to differences between photons and jets. Based on this training [143], an artificial neural network (ANN) was found to provide both the best efficiency for identifying photons (signal efficiency) and efficiency for rejecting jet backgrounds (background rejection). The variables used as inputs for the ANN are described next.

- *Hadronic Leakage,* ***HadEm*** This is the ratio of energies deposited in the hadronic (Had) and electromagnetic (EM) calorimeters. Chosen because it is typically used in the photon selection cuts (see Chapter 7). Photons are expected to deposit mostly in the EM calorimeter.

- *Calorimeter Isolation,* ***EIso4*** The sum of additional energy in a cone with radius 0.4 around the photon. Taken from typical photon ID cuts (see Chapter 7). Photons are expected to be ‘isolated’ without extra particles in the same cone.

- *Track Isolation,* ***SumPt4*** The sum of track momenta in a cone (radius 0.4) around the photon. Taken from typical photon ID cuts (see Chapter 7). Photons are again expected to be ‘isolated’ without extra particles in same cone. Technically a ‘tracking variable’ but useful for electrons and jets as both can increase *SumPt4*.

- ***Chi2Strip*** *and* ***Chi2Wire***, Compares the electromagnetic shower to the expected lateral shape. Taken from typical photon ID cuts (see Chapter 7),

---

[1] dataset **gq0sqd**, inclusive photon production with *minbias*; 3444250 raw events

[2] dataset **q8is01**, dijet production with dijet p_T >40 GeV/c 77093242 raw events

except that Strip and Wire values are used separately so that e.g. a bad Wire shape doesn't throw out a photon.

- *Lateral shower shape,* ***Lshr*** Also uses lateral shower shape, but compares the lateral sharing of energy between towers to expectations. Not a standard cut variable, but is used often for electrons and gives additional information about whether energy deposition is 'photon- like' as photons are expected to have little sharing.
- ***CES/CEM*** The ratio of the energy from the shower maximum detector to the total measured energy. It but should provide some separation from π^0's (see Chapter 7).

As a remark, the *Chi2Strip and Chi2Wire* variable is the same as described in Section 7.2, except that Strip and Wire values are used separately.

Figure 8.3 shows the input variable shapes. Note that regions of high overlap indicate background that isn't easily separated with simple cuts. Figure 8.4 shows the corresponding ANN output distribution. By TMVA convention, signal (background) events accumulate at large (small) classifier output values.

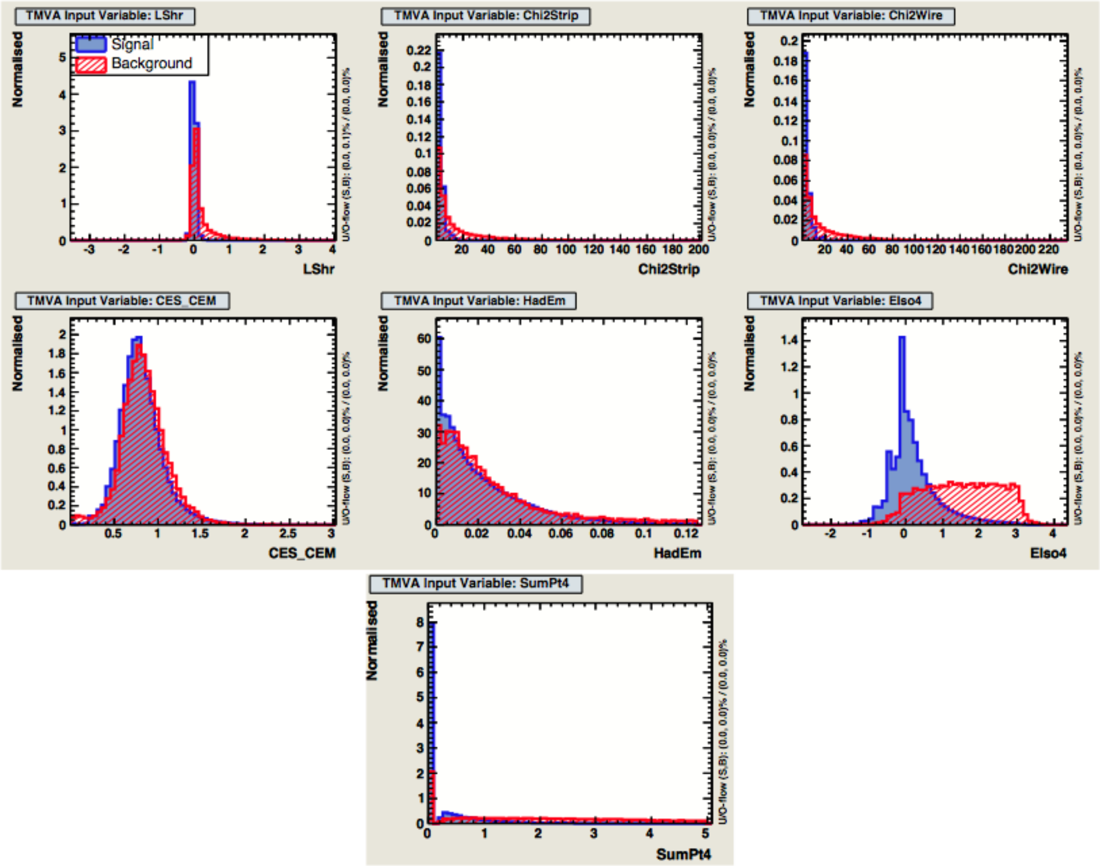


FIGURE 8.3: Normalized distributions of the input training variables for signal (filled area) and background (hatched area). Note that regions of high overlap indicate background that isn't easily separated with simple cuts [143].

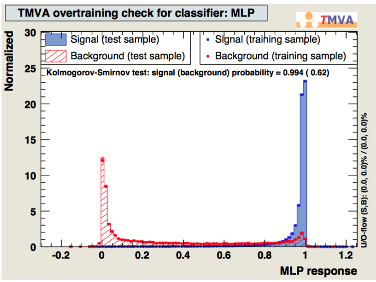


FIGURE 8.4: Output distribution for signal (filled area) and background (hatched area) after the TMVA ANN training. By TMVA convention, signal (background) events accumulate at large (small) classifier output values [143].

8.2.1.2 MC templates

The developed *ANN function* is then applied to the inclusive photon (signal) and QCD (background) MC samples listed in Section C. The corresponding ANN output distributions, the so-called *MC templates*, are then used to perform the fit (8.2.1.2). In the background sample contributions from ISR and FSR are removed as they represent part of the signal. Among the ANN input variables [?], one of the most important is the calorimeter isolation (*EIso4*); in the inclusive photon MC sample is calibrated to data by applying the corrections described in Appendix D. Further, to account effects dependent on luminosity, MC events are reweigthed to have the same $N_{z_{Vertex}}$ as observed in data.

Such a set of weights is obtained by comparing the vertex distributions of data to the MC. The data and MC vertex distribution are normalized to unit area and the ratio of data to MC is taken. A weight matrix is built depending on the number of vertices and on the E_T bin.

$$w(N_{vtx}, E_T^{\gamma}) = \frac{X^{data}(N_{vtx}, E_T^{\gamma})}{X^{MC}(N_{vtx}, E_T^{\gamma})} \tag{8.2.1}$$

where $X^{data}(N_{vtx}, E_T^{\gamma})$ ($X^{MC}(N_{vtx}, E_T^{\gamma})$) is the fraction of data (MC) events in the bin with n vertices and E_T^{γ} .

MC events are required to pass the same cuts as data and a set of cuts for the MC samples simulate the trigger conditions. Figure 8.5 shows the the obtained MC templates, for one E_T^{γ} bin.

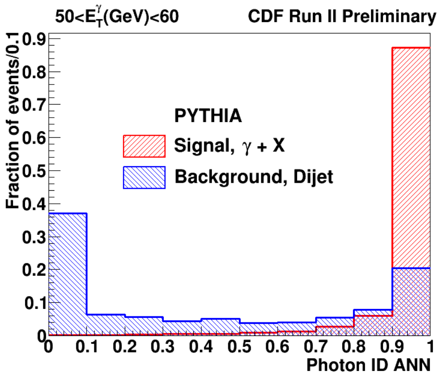


FIGURE 8.5: MC templates: the ANN output distributions, in one E_T^{γ} bin, for the MC signal (red) and background (blue). MC templates are normalized to unity.

8.2.2 Fit procedure

For this analysis the estimation of the composition of the Data sample, relies on Monte Carlo simulations of the different sources (signal and background). Data values are binned and for bins with a small content a Maximum Likelihood technique based on Poisson statistics is more appropriate than a χ^2 minimization. This Section explains how to take into account the statistical fluctuations of MC samples, which have a finite statistics, as well. The fitting method in Section 8.2.2.2 relies on the following algorithm.

Given an experimental dataset, the goal is to evaluate the proportions P_j of its different sources. These sources can not be modeled with an analytic form but only with dedicate Monte Carlo simulations. Thus, the space related to the dataset values is divided into n bins, giving a set d_i numbers which correspond to the i^{th}bin content. Let a_{ji} be the number of events from source j in bin i, and P_j the different sources strength, the predicted number of events in bin i is given by

$$f_i = N_D \sum_{j=1}^{m} P_j a_{ji}/N_j \tag{8.2.2}$$

where N_D (N_j) is the total number in the data sample (MC sample for the source j), i.e. :

$$N_D = \sum_{i=1}^{n} d_i \tag{8.2.3}$$

and

$$N_j = \sum_{i=1}^{n} a_{ji} \tag{8.2.4}$$

The P_j are the actual proportions and sum to one and strength factors can be rewritten as $p_j = N_D P_j / N_j$; thus Equation 8.2.2 is equivalent to

$$f_i = \sum_{j=1}^{m} p_j a_{ji} \tag{8.2.5}$$

The estimation of the p_j values can be performed by minimizing

$$\chi^2 = \sum_{i}^{n} \frac{(d_i - f_i)^2}{d_i} \tag{8.2.6}$$

For such a χ^2 the d_i distribution are assumed to be Gaussian, which is a good approximation to the Poisson for large numbers.

However, because of the finite size of the data sample and the splitting into several bins, many d_i are not large enough and Equation 8.2.6 is not appropriate. Going back to the Poisson distribution, according to which the probability to observe a given d_i value is:

$$e^{-f_i}\frac{f_i^{d_i}}{d_i!} \tag{8.2.7}$$

the proportions p_j are estimated by maximising the total Likelihood or, equivalently, its logarithm[3]:

$$\ln\ \mathscr{L} = \sum_{i=1}^{n} d_i\ \ln\ f_i - f_i \tag{8.2.8}$$

In this way, the small bin content in data events is properly accounted for. Such a method is often called “Binned Maximum Likelihood” fit.

By the way, this does not still account for the finite size of Monte Carlo samples, giving rise to statistical fluctuations in the a_{ji} term.

These can be included in Equation 8.2.6 by modifying the error in the denominator:

$$\chi^2 = \sum_{i}^{n} \frac{(d_i - f_i)^2}{d_i + N_D^2 \sum_j^m a_{ji}/N_j^2} \tag{8.2.9}$$

But this equation uses the incorrect Gaussian approximation. One has to find a procedure equivalent to the binned Maximum Likelihood technique.

First, Equation 8.2.5 changes into

$$f_i = \sum_{j=1}^{m} p_j A_{ji} \tag{8.2.10}$$

where A_{ji} is some unknown expected number of events for the source j in bin i

The total likelihood is given by the combined probability of the observed d_i and the observed a_{ji}:

---

[3] The relation $a^b = e^{b\ \ln\ a}$ is used and the constant factorials is omitted

$$\ln\,\mathcal{L} = \sum_{i=1}^{n} d_i\,\ln\,f_i - f_i\,+\,\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ji}\,\ln\,A_{ji} - A_{ji} \tag{8.2.11}$$

The estimates for the p_j (which we aim to know) and the A_{ji}, (in which we are not really interested) are evaluated by maximising this Likelihood. With this method the MC statistics is correctly incorporated.

In order to solve this maximisation problem of $m \times (n+1)$ unknowns, Equation 8.2.11 (together with Equation 8.2.5) is differentiated and derivatives are set to zero. In such a way two sets of equations are obtained:

$$\sum_{i=1}^{n} \frac{d_i A_{ji}}{f_i} - A_{ji} = 0 \quad \forall j \tag{8.2.12}$$

and

$$\frac{d_i p_i}{f_i} - p_j + \frac{a_{ji}}{A_{ji}} - 1 = 0 \quad \forall i, j \tag{8.2.13}$$

The obtained $m \times (n+1)$ simultaneous equations are nonlinear and coupled (f_i are functions of the p_j and the A_{ji}), and can be further simplified. Equation 8.2.13 is equivalent to

$$1 - \frac{d_i}{f_i} = \frac{1}{p_j}\left(\frac{a_{ji}}{A_{ji}} - 1\right) \quad \forall i, j \tag{8.2.14}$$

Let t_i be the left hand side ($t_i = 1 - d_i/f_i$), Equation 8.2.14 becomes:

$$t_i = \frac{1}{p_j}\left(\frac{a_{ji}}{A_{ji}} - 1\right) \quad \forall i, j \tag{8.2.15}$$

and the right hand side can be written as:

$$A_{ji} = \frac{a_{ji}}{1 + p_j t_i} \tag{8.2.16}$$

With this simplification for a set of p_j, the n unknown A_{ji} are given by the n unknown t_i values. Starting from Equation 8.2.15, if d_i is null then t_i is 1; otherwise, if $d_i \neq 0$,

$$\frac{d_i}{1 - t_i} = f_i = \sum_j p_j A_{ji} = \sum_j \frac{p_j a_{ji}}{1 + p_j t_i} \tag{8.2.17}$$

If these n equations are satisfied, then all the $m \times n$ Equations 8.2.13 are satisfied.

In summary, the solution to a binned Maximum Likelihood fit with finite Monte Carlo statistics is obtained by:

- solving the m Equations 8.2.12 iteratively; in this way the m variable p_i are found;
- solving the corresponding Equations 8.2.17, at every stage; these gives the A_{ji} values.

8.2.2.1 Additional remarks about the solution

From algebra some nice aspect can be highlighted. The Equations 8.2.12 can be simplified as:

$$\sum_{i=1}^{n} t_i A_{ji} = 0 \quad \forall j \tag{8.2.18}$$

or, replacing d_i/f_i by $1 + (A_{ji} - a_{ji})/p_j A_{ji}$

$$\sum_{i=1}^{n} A_{ji} = \sum_{i=1}^{n} a_{ji} \quad \forall j \tag{8.2.19}$$

Equations 8.2.19 shows that in the A_{ji} estimation, some source will change their shape with respect to the distribution of the MC values a_{ji}, but their overall total number will not change.

Moreover, if Equation 8.2.13 is multiplied by A_{ji} and summed over j, it becomes

$$\sum_{j} d_i - p_j A_{ji} + a_{ji} - A_{ji} = 0 \tag{8.2.20}$$

Finally, by summing over i an using Equation 8.2.19,

$$\sum_{i} d_i = \sum_{i} \sum_{j} p_j A_{ji} \tag{8.2.21}$$

and

$$N_D = \sum_{j} p_j N_j \tag{8.2.22}$$

which gives the normalization. In the χ^2 method of Equation 8.2.6 this automatic normalization does not hold anymore. In fact, the return set p_i values gives a fitted number of events generally lower than the actual one, because downward fluctuations have smaller error and have higher weight.

In conclusion, as well as providing an error estimate which includes the effect of finite MC statistics, the new $\ln \mathcal{L}$ technique provides an unbiassed estimate of the fractions of each MC source present in the data, even when there are bins with very few events from either data or Monte Carlo.

8.2.2.2 TFractionFitter

The algorithm described above is implemented in TFractionFitter package of the Root software (the full documentation can be found at [146]). This routine uses minuit minimization package to fit the Monte Carlo distributions to the data distributions and returns the best values for the fraction of each Monte Carlo component present in the data sample; the fit also returns the post-fit template (MC) predictions varied within statistics ($\hat{h}$). Each bin is allowed to vary independently, thus the post-fit templates shape can differ from the shape of the nominal templates in order to better describe the data distribution. The fit result is an histogram (R) given by the sum of the post-fit MC templates scaled to their own rates.

$$R(fit\_Result) = f^S \cdot \hat{h}(signal) + (1 - f^S) \cdot \hat{h}(background) \qquad (8.2.23)$$

Signal and background templates are normalized to the total number of data events. An example of this fit application is given in Figure 8.6.

8.2.3 Fit validation study

A correct behavior of the used fit in not obvious a priori. Thus, it is important to validate the obtained signal fraction as well as its error estimate. In this Section the fit validation study is described. The validation procedure used for this analysis is based on *Toys Monte Carlo*, or *pseudo-experiments*. *Pseudo-experiments* correspond to a large sample of simulated events; each pseudo-experiment was generated as follows:

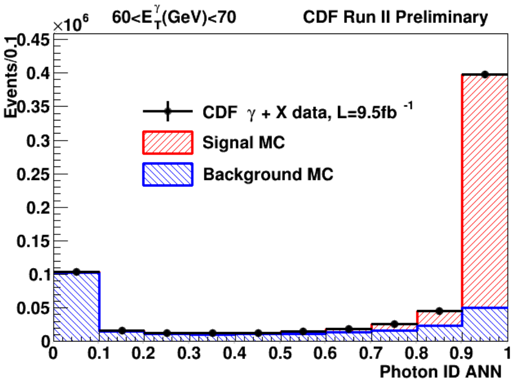


FIGURE 8.6: Fit to Artificial Neural Network (ANN) output distributions for one E_T^γ bin. This figure shows the distribution fit result given by the sum of MC signal (red) and background (blue) templates scaled to TFractionFitter rates. MC templates are normalized to data (black points).

- the ANN data distribution is fit to the MC templates; the fit returns the signal fraction f^S, or f^S_{data};
- then, the bin content of the ANN data and Monte Carlo distributions is varied independently according to Poissonian statistics, which means that every bin is allowed to fluctuate within its statistical error. The *new* ANN distribution obtained from data is referred to as *pseudo-data* distribution.
- the ANN pseudo-data distribution is fit to the *new* MC templates; the fit returns the signal fraction f^S_{pseudo};

These steps are repeated for each simulated event (∼10000).

The f^S_{pseudo} values, are then used to build the residual distribution and pull distribution. One expects both distributions to be Gaussian.

The residual distribution of the signal fraction is defined by the difference between the value from the fit to data (f^S_{data}) and the value from the fit to pseudo-data (f^S_{pseudo});

$$Residual = f^S_{data} - f^S_{pseudo} \tag{8.2.24}$$

If there is no bias in the fit, such a distribution is centered at zero within uncertainties.

The pull distribution is defined as:

$$Pull = \frac{f^S_{data} - f^S_{pseudo}}{\sigma_{pseudo}} \tag{8.2.25}$$

> If the error (σ) quoted by the fit is reliable, this distribution has unit width within uncertainties.

An example of the obtained residual and pull distributions for one E_T^γ bin is shown in Figure 8.7. Both distributions in the figure can be approximated with Gaussian distributions. The residual distribution has a mean value consistent with 0, while the pulls ditribution has a width ∼1. The fit procedure for such a bin is therefore reliable.

The same studies are repeated for all bins of interest. Figure 8.8 shows the obtained values for all the E_T^γ bins: the points and error bars correspond respectively to the mean values of the residual distributions and the widths of the pull distributions. One can easily see that the means values of the residual distributions are around zero and the pull widths are either ∼1 or close to 1. The fit is accurate for all considered bins. All details about the aforementioned distributions are reported in Appendix E.

For the few bins where the pulls widths slightly deviate from 1 (e.g: bin 30$<$ $E_T^\gamma(GeV)$ $<$40) the quoted uncertainties on signal fractions are corrected for such a deviation. Corrections are computed as one over the pull width.

8.2.4 Fit Results and Signal Fraction

As already stated, the ROOT routine TFractionFitter is used to fit the data in each photon E_T bin. Figures 8.9 and 8.10 show the results of the fits in the 15 E_T bins covering the range $30 < E_T^\gamma(GeV) < 500$. At low E_T, the uncertainty of the measured signal fraction (error bars) is mostly caused by low statistics of the di-jet sample while for high E_T it is dominated by the data statistics. The signal fractions determined for each photon E_T are shown in Figure 8.11.

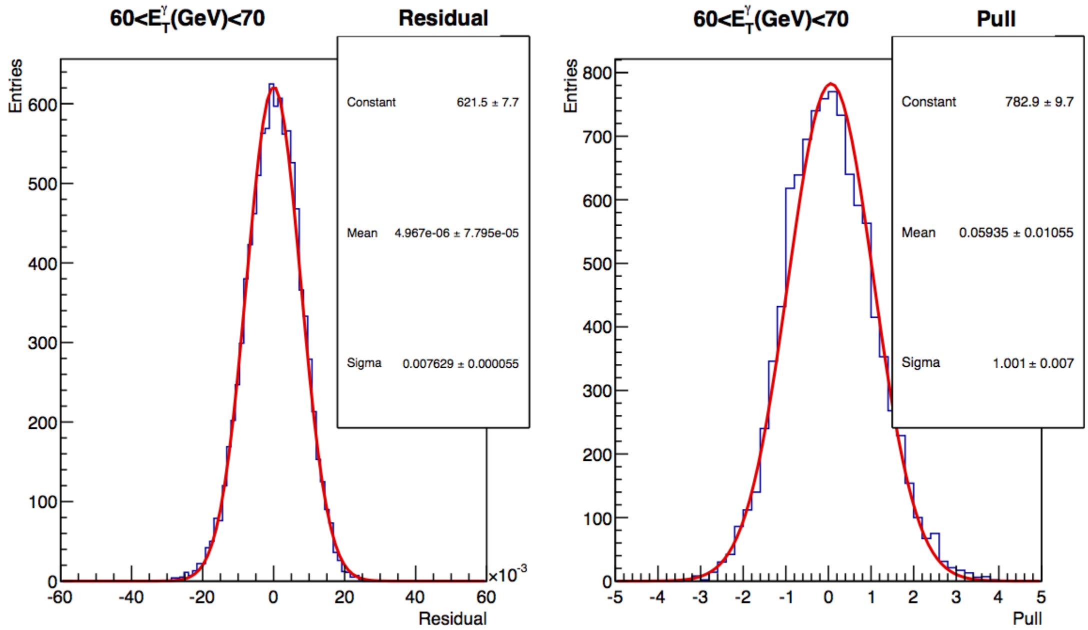


FIGURE 8.7: Residual (left) and pull (right) distributions for pseudo-experiments with N = 10000, in one E_T^γ bins. The residual distribution represents the difference between the expected signal fraction (from Data fit to MC templates) and the "trial" signal fraction (from pseudo-experiments). The width of the pull distributions should be compatible with unity if the statistical error quoted by the fit is reliable. See text for further details

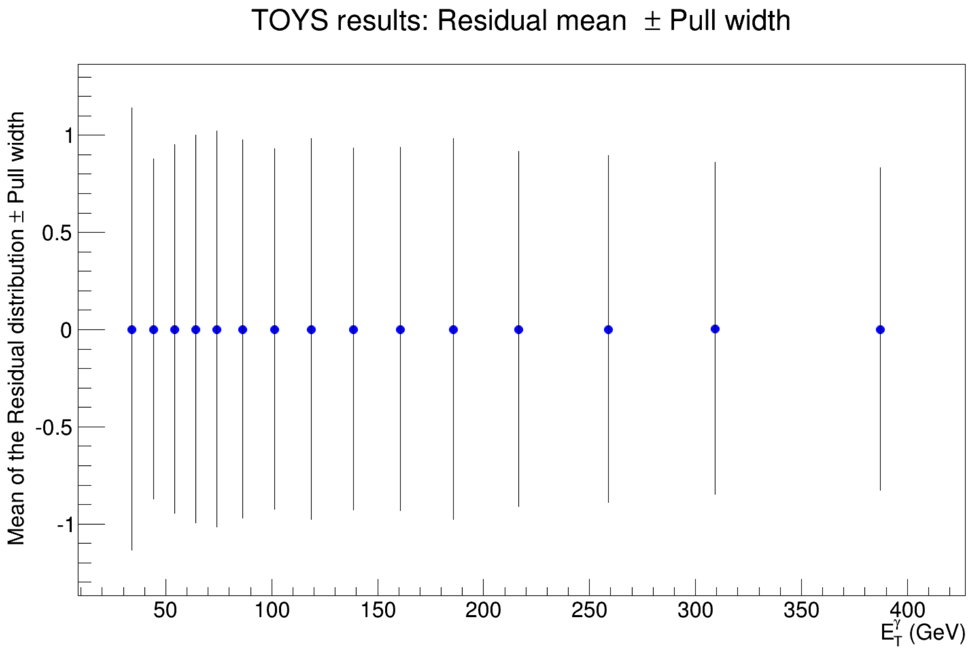


FIGURE 8.8: Mean values of the residual distributions (points) and pull widths (error bars) as a function of E_T^γ. The mean values of the residual distributions are around zero and the pull widths are either ∼1 or close to 1. See text for details.

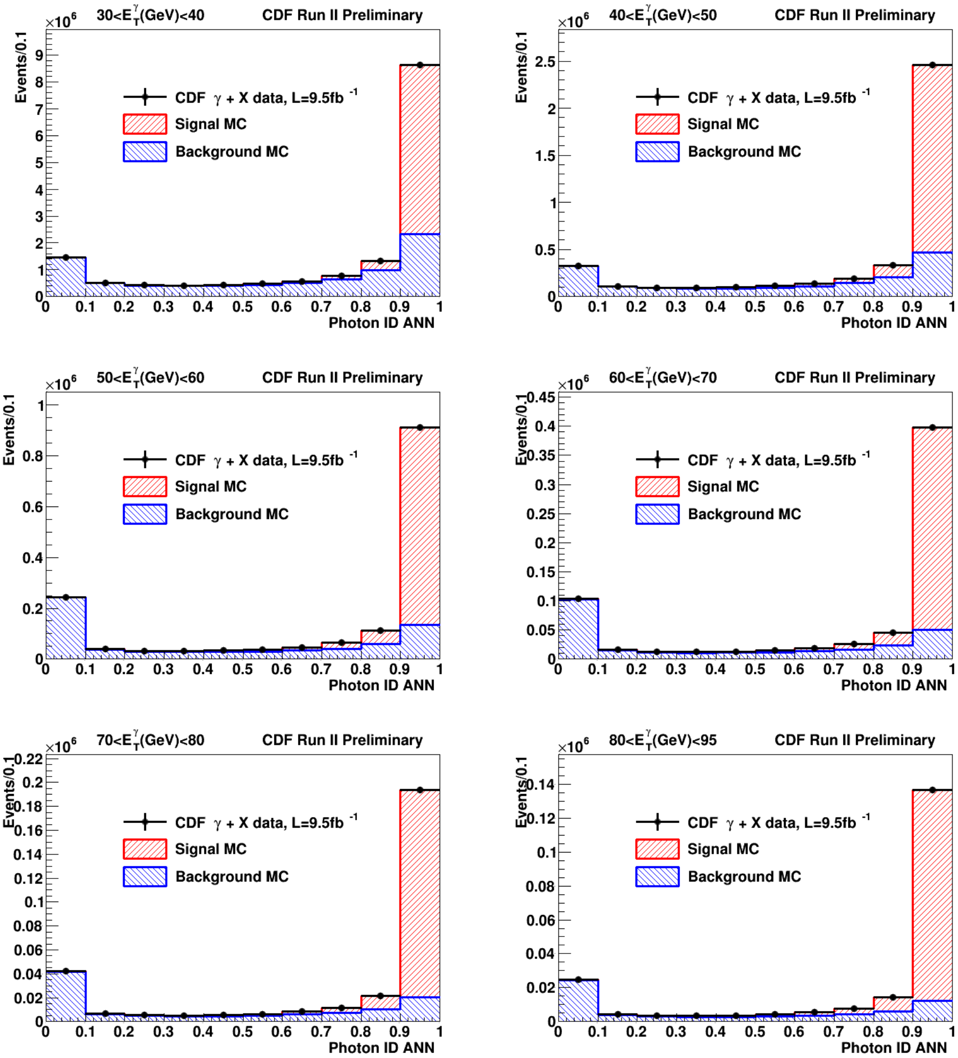


FIGURE 8.9: Fits to the ANN distributions in bins of photon transverse energy for $30 < E_T^\gamma(GeV) < 95$. The distribution fit result is given by the sum of MC signal (red) and background (blue) templates scaled to TFractionFitter rates. MC templates are normalized to data (black points). See text for more details.

8.2.5 Systematic uncertainties in the Signal Fraction

Figure 8.12 shows the total systematic uncertainty and single contributions on the signal fraction as a function of photon E_T. In the whole range, the dominant systematic effect comes from varying the ANN input variable *EIso4*.

The following systematic effects on the signal fraction are evaluated.

Photon energy scale A value of ± 1.5% systematics in the photon energy scale is considered according to the studies in Ref. [108]. This uncertainty takes into account both geometrical and energy dependence differences between

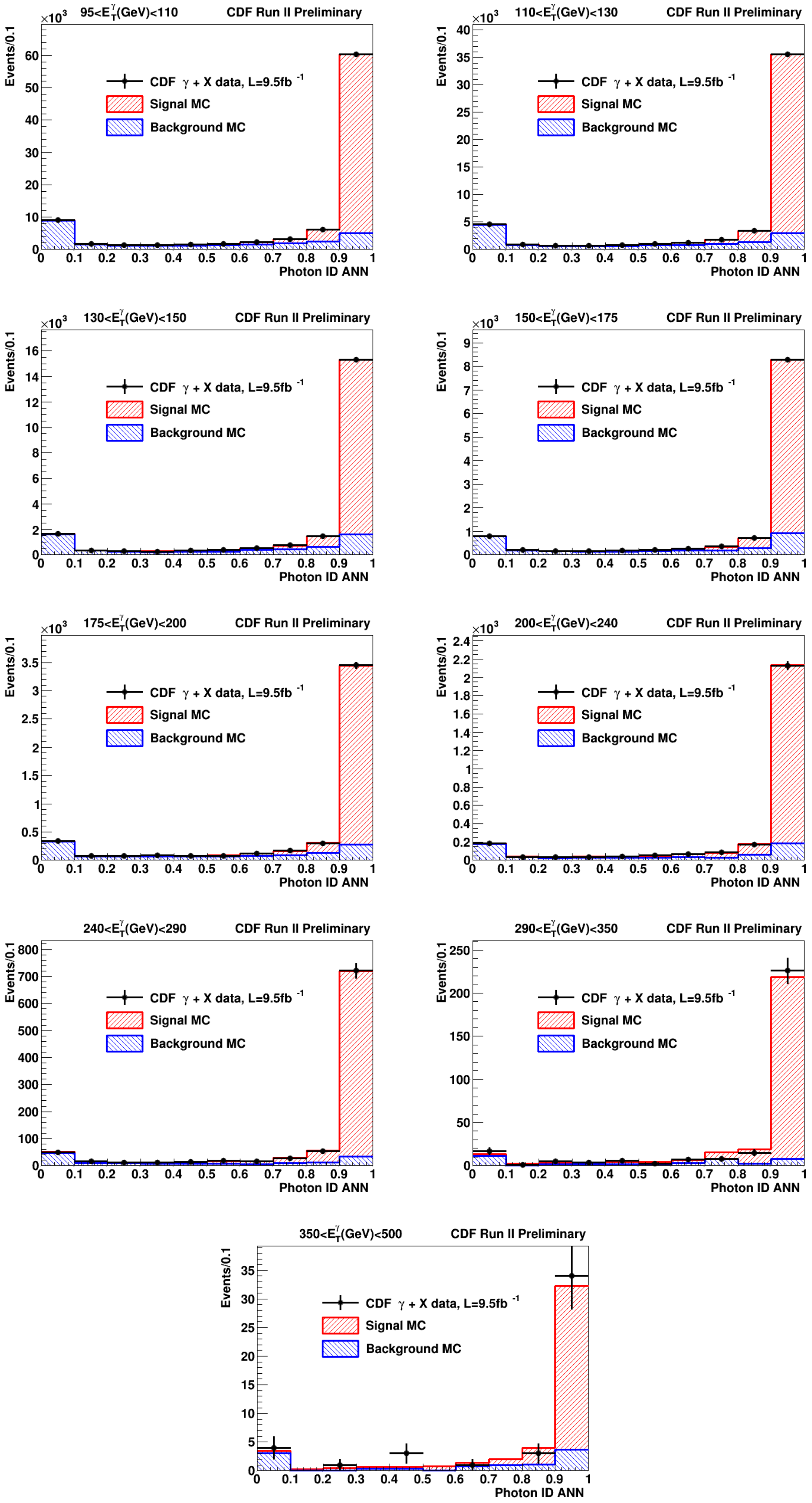


FIGURE 8.10: Fits to the ANN distributions in bins of photon transverse energy, for $95 < E_T^{\gamma}(GeV) < 500$. The distribution fit result is given by the sum of MC signal (red) and background (blue) templates scaled to TFractionFitter rates. MC templates are normalized to data (black points). See text for more details.

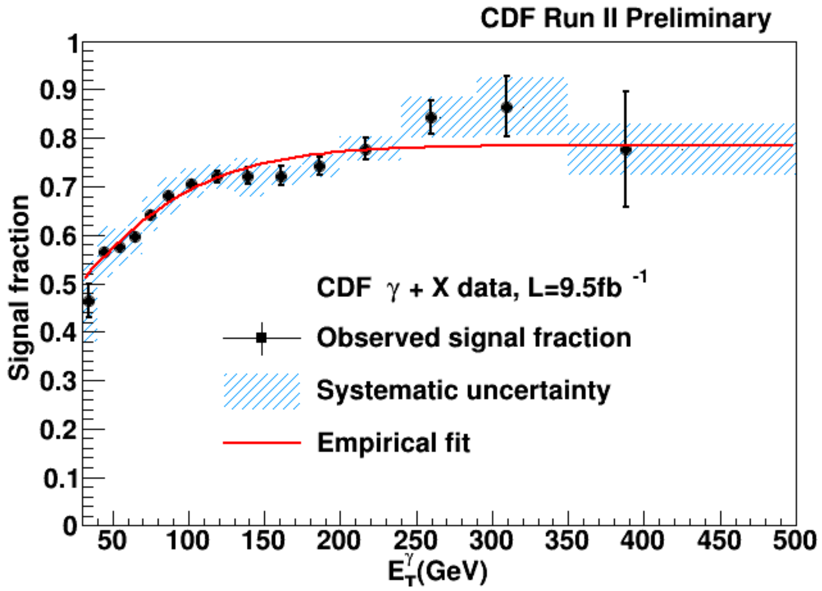


FIGURE 8.11: Signal fraction as a function of the photon transverse energy. The error bars represent the statistical errors and the azure bands represent the systematic errors.

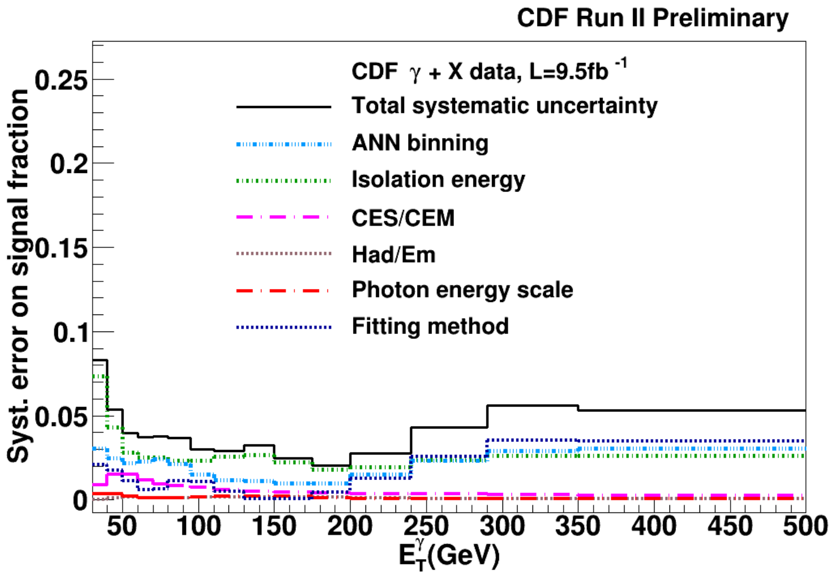


FIGURE 8.12: Total systematic uncertainty and single contributions on the photon fraction as a function of photon E_T

data and MC. This effect is studied by varying the E_T^γ-values by ± 1.5% in MC. This will cause migration effects between E_T^γ bins and change the ANN template shapes. The systematic effect on fit results is less than 1% in the whole E_T^γ range.

ANN variables Some ANN input variables are less than well modeled in the

MC: *EIso4*, *CES/CEM*, and *Had/Em*. They are varied by ±50% based on data and simulation comparisons to study how sensitive the result is to the mismodeling of the ANN variables. The *EIso4* variation turns out to have a large effect systematic effect (∼8%) at the lowest E_T^γ.

Two-bin fit A different fit technique is also used: *two-bin fit*. This is a very simple counting method based on ANN distributions divided in two bins. More details can be found in Appendix F.

ANN binning By default, the ANN histograms are divided into 10 bins from 0.0 to 1.0. Different binnings are used to test sensitivity to shapes. Figure 8.13 shows fit results (signal fractions) as a function of the number of ANN bins, for a particular photon E_T range ($95 < E_T^\gamma(\mathrm{GeV}) < 100$). The half span from the highest and the lowest values is taken as systematic effect.

Using different binnings has an effect of less than 4% on the fit results in the whole E_T^γ range (Figure 8.12) .

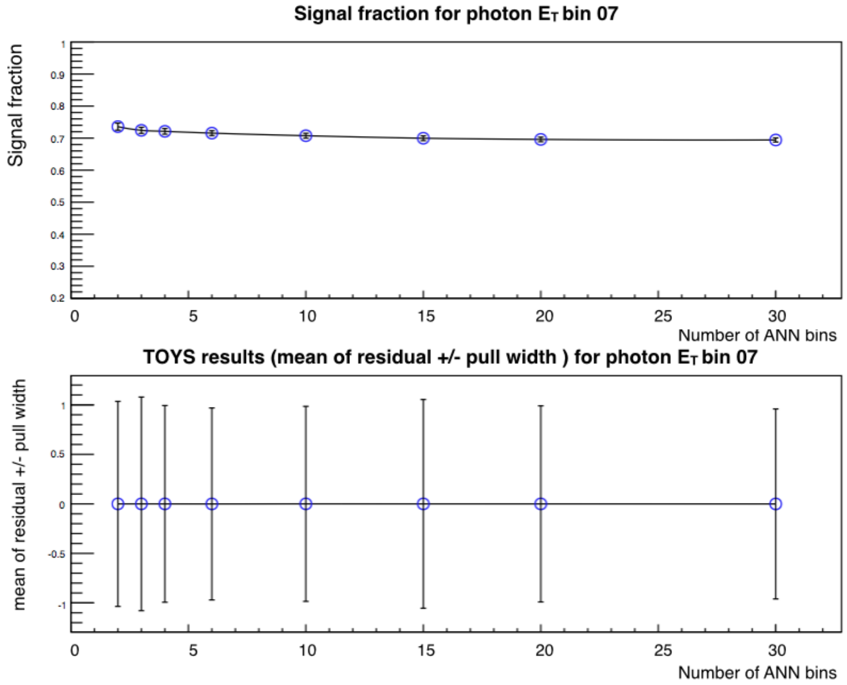


FIGURE 8.13: Both plots refer to a particular photon E_T bin (95< E_T^{γ}(GeV) <100). *Upper plot*: Fit results (signal fractions) as a function of different ANN binnings. By default the number of ANN bins is set to 10. The half span from the highest and the lowest values is taken as systematic effect. *Bottom plot:* Mean of residual distributions (points) and pull distributions widths (error bars) as a function of the different number of ANN bins. For each number of bins, points are centered at zero and the width is close to 1. This implies the reliability of fit results. See text for details.

8.3 Minor Backgrounds

In addition to the SM backgrounds, there are several other sources that can produce fake photons in the detector. They are photomultiplier tube spikes and noncollision photons. In the following, it will be shown that they are minor backgrounds.

PMT Spikes

In the Central Electromagnetic Calorimeter (CEM) an energy deposit is identified from the output of the two photo-multiplier tubes (PMTs) that collect the light from the scintillator. A high voltage breakdown (“spike”) in the PMT is unrelated to an energy deposit in the CEM and can create a false candidate.Since photons that come from the collision will deposit almost the same amount of energy in each PMT, these instrumental effects can be separated by considering the asymmetry of the two energy measurement of the PMT from a tower as shown in Equation 8.3.1.

$$A_P = \frac{|E_{PMT1} - E_{PMT2}|}{E_{PMT1} + E_{PMT2}} \tag{8.3.1}$$

where E_{PMT1} and E_{PMT2} are the two PMT energies. In the event selection the corresponding cut, referred to as *spike killer*, is:

$$A_P < 0.6 \tag{8.3.2}$$

Previous measurements [142] compared the A_P distribution for photon candidates (from both real photons and spikes) to real electrons from $W \rightarrow e\nu$ events, showing that this requirement rejects 100% of all spikes with a minimal loss in efficiency for real photons (see Figure 8.14). Thus, this source is neglected in the background estimate.

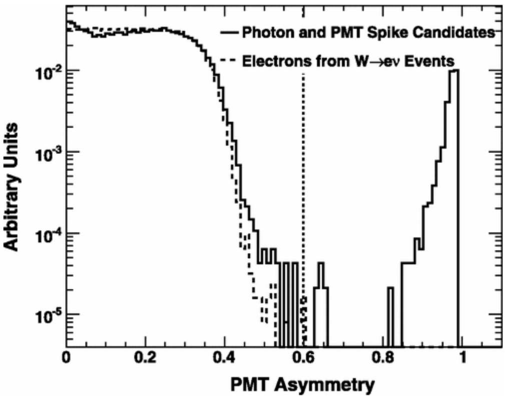


FIGURE 8.14: A comparison of the PMT asymmetry, $A_P < 0.6$, for a photon $+ MET$ sample that contains both PMT spikes and real photons, and a sample of electrons from $W \to e\nu$ events. PMT spikes can be effectively removed by requiring the asymmetry to be less than 0.6 [142].

8.3.1 Noncollision fake photons

The most common sources of noncollision backgrounds include cosmic ray muons, as they traverse the detector, and beam interactions with the beam pipe, the so-called "beam halo". Each of these sources produce an energy deposit in the calorimeter that is reconstructed as a photon and has a peculiar time distribution. The time distributions for cosmic rays and beam halo are very distinct from the prompt photons one, and allow to get an estimate of noncollision photons. Cosmic rays (CR) and beam halo (BH) events have similar signatures in the detector: $\gamma + MET$. In fact, they tend to produce a photon candidate with an energy imbalance (MET) in the opposite direction of photon, while photons produced by real collisions have one or more jets to balance its momentum and can produce a little MET; therefore, a combined subsample of CR and BH events is considered. The *noncollision sample* consists of events with a photon candidate selected with the photon preselection criteria in Table 7.3, with no reconstructed track, with the spikes killer cut and having high missing transverse energy MET/E_T^{γ} >0.8. The reason for this last requirement can be easily seen in Figure 8.15, where the observed MET/E_T^{γ} distribution in data is compared to the expected one (MC inclusive photons). The signal events region corresponds to MET/E_T^{γ} <0.8. while background events have MET/E_T^{γ} <0.8.

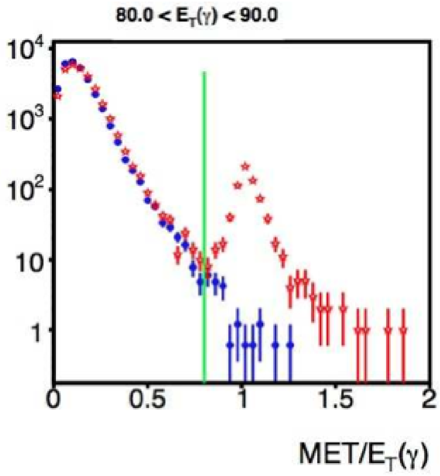


FIGURE 8.15: MET/E_T^{γ} distribution for the inclusive photon sample (red) and a Monte Carlo inclusive photon sample (blue) for 80<E_T^{γ}(GeV)<90 GeV. The vertical line at 0.8 corresponds to the separation between the signal region (MET/E_T^{γ}<0.8) and the background region (MET/E_T^{γ}>0.8).

Cosmic Rays

Cosmic rays are charged particles that originate in outer space and then interact with the earth's atmosphere producing secondary charged particles that then shower down to the earth's surface. If these particles have an energy of few GeV they can reach the surface of the earth [21]. Cosmic rays may interact with the detector and produce an electromagnetic cluster via a bremsstrahlung or a catastrophic showering within the EM calorimeter. If a cosmic ray reach the detector this can lead to both an incorrectly assigned photon to a vertex that had nothing to do with its production, as well as leaving an imbalance of energy in the detector which is misidentified as missing energy. The CR subsample selection relies on the typical geometry of these events; in order to separate cosmics from BH, the following variables are used:

seedWedge : the number of CEM containing towers in the same wedge as the seed tower of the cluster with energy deposition above 100 MeV;

nHadPlugTowers : the number of Plug HAD towers with energy deposition above 100 MeV in the same wedge as γ. Cosmics typically do not deposit any energy in the Plug hadronic calorimeter, while BH normally deposits some energy at the entry or exit points (typically about 6 towers).

phiWedge : correspond to identity the calorimeter wedge (they are a total of 24). You will see in the next paragraph that Beam Halo photon candidates are concentrated in wedges 0 and 23, which corresponds to small (ϕ) angle with respect the beam line (see Chapter 4).

muonStubs is the number of the hit muon stubs.. The muon detection system is contained on the outer radius of the CDF detector; a photon from a true collision has not deposit in the muon detector, while a cosmic ray may show activity in this outer detector within a close angle to the electromagnetic cluster giving an indication that a particle may have passed from the outside of the detector inwards.

Cosmic Rays subsample is obtained from the noncollision sample by adding the requirements listed in Table 8.2: A correlation plot of seedWedge vs. nHadTowers in Figure 8.16 shows two distinct regions. The cut on *seedWedge* and *nHadTowers* corresponds to the CR region. Typically BH produces much more activity: seedWedge greater than 8 and nHadPlugTow greater than 3.

Table 8.2: Set of requirements used to identify the photon candidates as originating from cosmic rays.

$seedWedge<8$ and $nHadTowers=0$
$phiWedge>0$ and $phiWeedge<23$
$muonStubs>1$

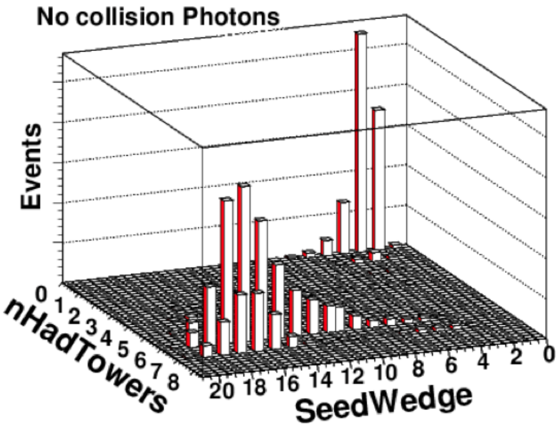


Figure 8.16: A correlation plot of seedWedge vs. nHadTowers. In order to separate BH from Cosmics, the number of EM towers and HAD towers with EM energy > 0.1 GeV in the same photon wedge is used. Beam halo events usually produce more activity

The time distribution of CR (grey histogram) is shown in Figure 8.17 together with the time distribution of all photon candidates (cyan histogram). Photon candidates from cosmic rays are not correlated in time with collisions, and therefore their t^γ distribution is roughly flat within the time energy integration window start and end times[4]. The CR distribution is scaled in order to match the collision distribution in the flat region, for display reasons. The collision events lie within the "signal timing window" defined as $|t_\gamma|$ <6.6 ns. The events falling outside this region correspond to the noncollision ones. Thus, the latter events are strongly rejected by imposing the cut $|t_\gamma|$ <6.6 ns, but a component still remains. To predict the number of these events in the signal timing window, a cosmic rays-dominated time window, called from now on "side band region", is defined in the interval of $\{30, 90\}$ ns, (i.e. well away from the signal timing window). Figure 8.17 highlights the two timing regions. The estimate of the cosmic events in the signal region, a is obtained as follows:

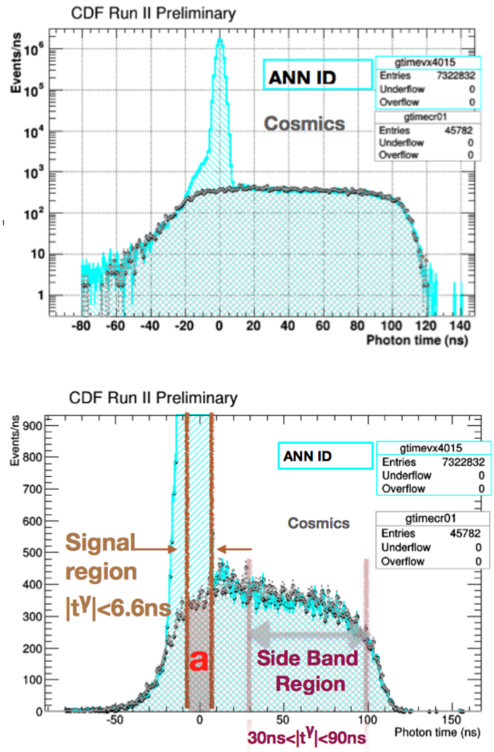


FIGURE 8.17: Time distribution of all photon candidates (cyan) and cosmic rays (grey): the collision events correspond to the peak around zero. The CR distribution is scaled in order to match the first distribution in the flat region, for display reasons. Photon candidates from cosmic rays (CR) are not correlated in time with collisions, and therefore their t^γ distribution is roughly flat allowing to estimate the rate of cosmics in the signal region from data. Fraction of cosmic photons remaining in the signal region can be estimated by studing their distributions in two regions: the time signal region ($|t_\gamma|$ <6.6 ns) and the side band region 30< $|t_\gamma|$ <6.6 ns.

$$a = b \times \frac{c}{d} \tag{8.3.3}$$

---

[4] the fall at -30 and 100 ns is due to the 132 ns energy integration window for the ADMEM

where b (d) is the number of all photon candidates (*cosmics*) in the side band region and c is the number of cosmic rays events in the signal region. The contribution from CR is found to be less than 1%, so it is considered negligible.

Beam Halo

Beam halo events are caused by beam particles (mostly from the more intense proton beam) that hit the beam pipe upstream of the detector and produce muons. These muons travel almost parallel to the proton beam direction and shower into the EM calorimeter to create a photon candidate. They feature either a long, narrow trail of energy in the central calorimeter, or else a high-energy spike in one or both plug calorimeter. These events usually do not have tracks or significant activity in the calorimeter outside the wedge with the energy deposition.

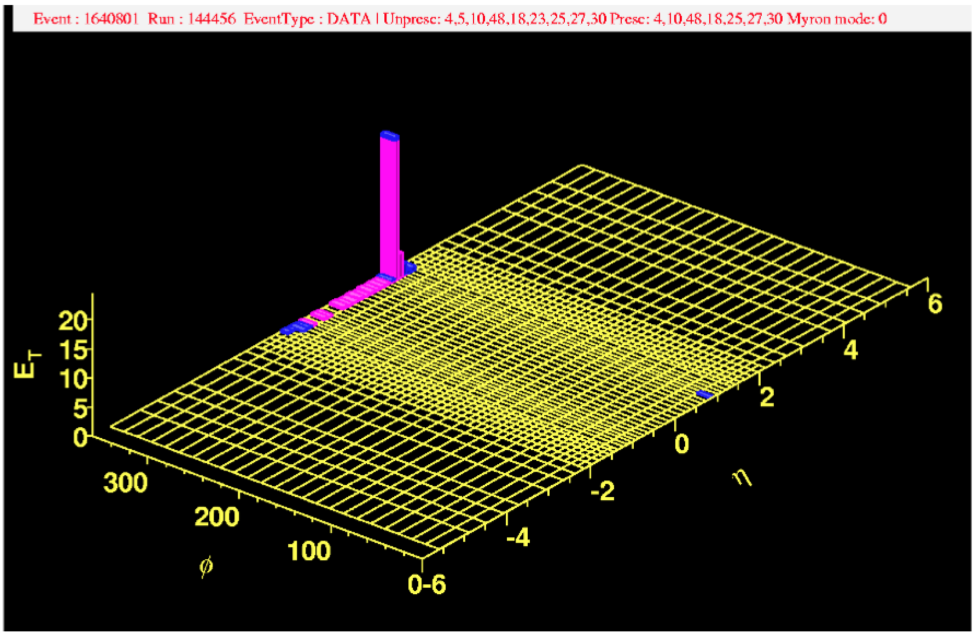


FIGURE 8.18: Energy deposited in the calorimeters for a beam halo event. Photons from the beam halo travel in the plane of the beam orbits and tend to have a large spread in η. Protons enter from the $-\eta$ side.

Figure 8.18 shows an *event display*[5] of an off-axis muon producing a hard bremsstrahlung in the central calorimeter. The muon "track" is identified by the set of contiguous towers in η at constant azimuthal angle ϕ. The vertical scale is the energy in the calorimeter transverse to the beam. These events form backgrounds to physics signals involving photons and missing transverse energy.

---

[5] *Event display* details have been discussed in Chapter 4

Photons from the beam halo travel parallel to the beam (corresponding to $phiWeedge=0$ and $phiWeedge=23$).

The BH subsample selection relies on the typical geometry of these events. The requirements listed in Table 8.3 are applied The time distribution of BH is shown in Figure 8.19.

TABLE 8.3: Set of requirements used to identify the photon candidates as originating from beam halo sources.

$seedWedge>8$ and $nHadTowers>3$
$phiWeedge=0$ and $phiWeedge=23$
$muonStubs=0$

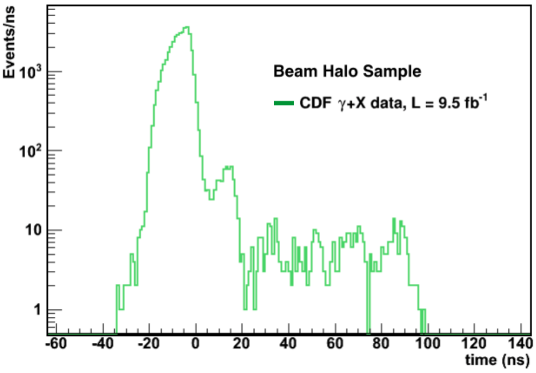


FIGURE 8.19: Time distribution of beam halo events selected from photon data by applying to the noncollision presample, the set of cuts in Table 8.3. These beam halo *photons* typically arrive a few nanoseconds earlier than prompt photons from collisions owing to the nature of the beam structure.

The main feature in the timing distribution is the peak around -10 ns. Beam halo *photons* typically arrive a few ns earlier than prompt photons for geometric reasons as shown in Figure 8.20. Beam Halo flies parallel to the z axis of the detector and interacts with a tower, while the collision event occurs in the middle of the detector and produces a photon that then travels to the tower. Quantitatively, it takes T_C time for the particles traveling parallel to the z axis to reach the center of the detector, then there is a T_T extra time to reach a specific tower.

However in this case, while the rate is lower, the photon candidate can also have a t_{corr} of 19 ns (and multiples later and earlier) if the muon was created in one of the beam interactions that can occur every 19 ns in the accelerator.

The BH events are less than 1% with respect the photon candidates, so this background source is considered negligible.

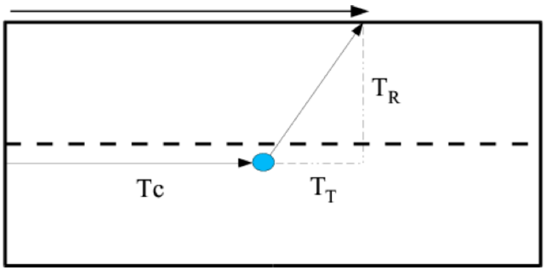


FIGURE 8.20: Sketch illustrating time difference between photons from $p\bar{p}$ collisions and beam halo. The path of beam halo (black arrow) is shorter than the collision data path (grey arrows), thus BH reaches a specific tower earlier.

8.4 Conclusion

In summary, backgrounds for the $\gamma + X$ signature come from noncollision processes and SM processes, including QCD production of hadronic jets. The rate at which noncollision backgrounds populate the signal region is estimated from collision data using events with no identified collision. The contribution from these photons is reduced to less than 1%, so it is considered negligible. The main challenge for this measurement is the suppression of the huge background coming from hadronic jets misidentified as prompt photons. The signal fraction (purity) estimate has been performed by developing a fit technique which exploits the outcomes of an Artificial Neural Network (ANN), dedicated to the discrimination of prompt (signal) photons from "fake" photons, coming from light neutral mesons decays (mostly π^0). The fitting algorithm has been validated. The obtained results will be used in next chapter.

Chapter 9

Cross Section Measurement

This chapter presents the measured differential cross section for the production of prompt isolated photons. Results are compared to three theoretical predictions.

Contents

9.1 Definition of the Cross Section

The differential cross section for the production of isolated prompt photons is measured in 15 E_T^γ bins, from 30 GeV up to 500 GeV. In a given E_T^γ bin it is calculated as:

$$\frac{d^2\sigma}{dE_T^\gamma d\eta^\gamma} = \frac{N f^\gamma}{\mathcal{L}\Delta E_T^\gamma \Delta\eta^\gamma A \times \epsilon} \tag{9.1.1}$$

where N is the number of data events after applying the full selection (Chapter 7), f^{γ} is the signal fraction evaluated in Chapter 8; the cross section is measured for $|\eta^{\gamma}| < 1.0$, so $\Delta\eta^{\gamma}$ is 2.0. ΔE_T^{γ} is the width of the E_T^{γ} bin, $\mathcal{L}$ is the integrated luminosity corresponding to 9.5 fb^{-1}; $A \times \epsilon$ is a correction factor discussed in the next Section.

9.2 Acceptance times efficiency

The cross section is unfolded to particle level for comparison with theory. The unfolding procedure, which relies on Monte Carlo simulations, corrects for acceptance and detector effects, photon reconstruction efficiency and selection efficiency. The factor *acceptance times efficiency*, $A \times \epsilon$, takes into account all of these contributions. $A \times \epsilon$ values are evaluated from the bin-by-bin ratio of the number of reconstructed prompt photons (N^{γ}_{rec}) to the number of particle-level prompt photons (N^{γ}_{gen}), in the signal events simulated with Pythia:

$$A \times \epsilon = \frac{N^{\gamma}_{rec}(photon\ ID\ cuts,\ trigger\ cuts, |\eta^{\gamma}| < 1.0, E_T^{\gamma} > 30GeV)}{N^{\gamma}_{gen}(iso_{cal} < 2GeV, |\eta^{\gamma}| < 1.0, E_T^{\gamma} > 30GeV)} \tag{9.2.1}$$

- The numerator is calculated by applying the same requirements as those applied to data: a photon candidate in the central calorimeter passing trigger and loose photon ID cuts with reconstructed transverse energy in the range 30< $E_T^{\gamma}(GeV)$ <500. The numerator includes a tuning factor for the calibration of the photon reconstruction efficiency, which accounts for the differences between real and simulated detector performance; the value quoted by previous published measurements [38, 147] is used[1].

- The denominator is obtained by applying the same kinematic and isolation selection on the generated-particles quantities. Generated particles are at "particle level" which implies a collection of stable particles from parton shower generators (therefore including QCD+QED effects), without any simulation of the interaction of these particles with the detector components or any additional proton-antiproton interactions. The isolation energy is calculated by summing transverse momentum of all the generated

---

[1] This tuning factor was obtained by comparing the selection efficiencies for $Z \to e^+e^-$ events in data and in simulation

stable particles in the cone of 0.4 around the photon. In this case, sum goes over stable particles instead of towers. Underlying events are included in the isolation energy calculation

Figure 9.1 shows the obtained correction factors as a function of E_T^γ. Values range from 65% to 72% and do not have a strong E_T dependence.

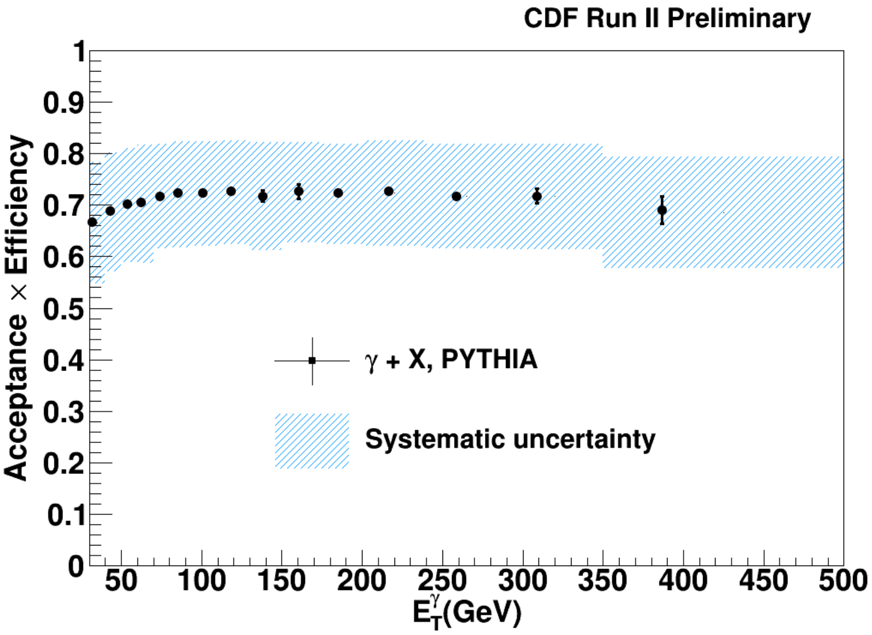


FIGURE 9.1: $A \times \epsilon$ corrections factors as a function of the photon transverse energy, E_T^γ. The correction factors are evaluated from the bin-by-bin ratio, using the PYTHIA signal simulation, of reconstructed prompt photons to particle-level prompt photons. The error bars represent the MC statistical errors and the azure bands represent the systematic errors. Values range from 65% to 72% and do not have a strong E_T^γ dependence.

9.2.1 Systematic uncertainty on $A \times \epsilon$ factors

Figure 9.2 shows the total systematic error on the unfolding factors, with the single contributions. The following systematic effects on the $A \times \epsilon$ factors are taken into account:

Photon energy scale A value of ± 1.5% systematics in the photon energy scale is considered according to the studies in [108]. This uncertainty takes into account both geometrical and energy dependence differences between data and MC. This effect is studied by varying the E_T^γ by ± 1.5% in MC. This

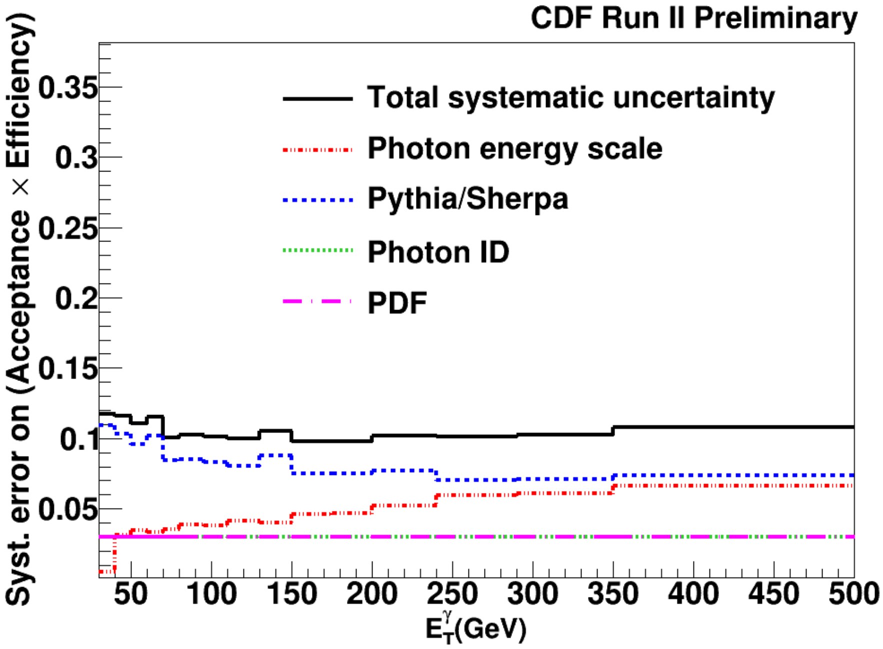


FIGURE 9.2: Total systematic uncertainty and single contributions on the factors ($A \times \epsilon$) as a function of photon E_T. The total systematic errors are approximately 10% and are dominated by the uncertainties in the photon-energy scale at high E_T ($\sim$ 6%) and by the choice of generator.

will make events migrate between E_T^γ bins. This systematic effect is of $\sim$ 6% at high E_T.

Photon ID the assigned systematic error of 3% is taken from [147].

PDF uncertainties a 3% error is assigned, based on the previous measurement [38, 147].

Pythia vs Sherpa : the stability on the $A \times \epsilon$ factors due to the choice of MC generator is computed by considering SHERPA for the bin-by-bin ratio instead of PYTHIA. The stability is related to photon reconstruction and identification. It also probes the uncertainty on the signal reconstruction due to an alternative fragmentation mechanism. The observed shift (of $\sim$8%), is taken into account for systematic uncertainty estimation in the correction factors.

The total systematic errors on the correction factors are approximately 11%. They are dominated by the choice of generator and, at high E_T^γ ($\sim 6\%$), by the uncertainties in the photon-energy scale.

9.3 Results

The measured differential cross section for the production of prompt isolated photons within the pseudorapidity region $|\eta^\gamma| < 1.0$ is shown in Figure 9.3. The measured values, also listed in Table 9.1.

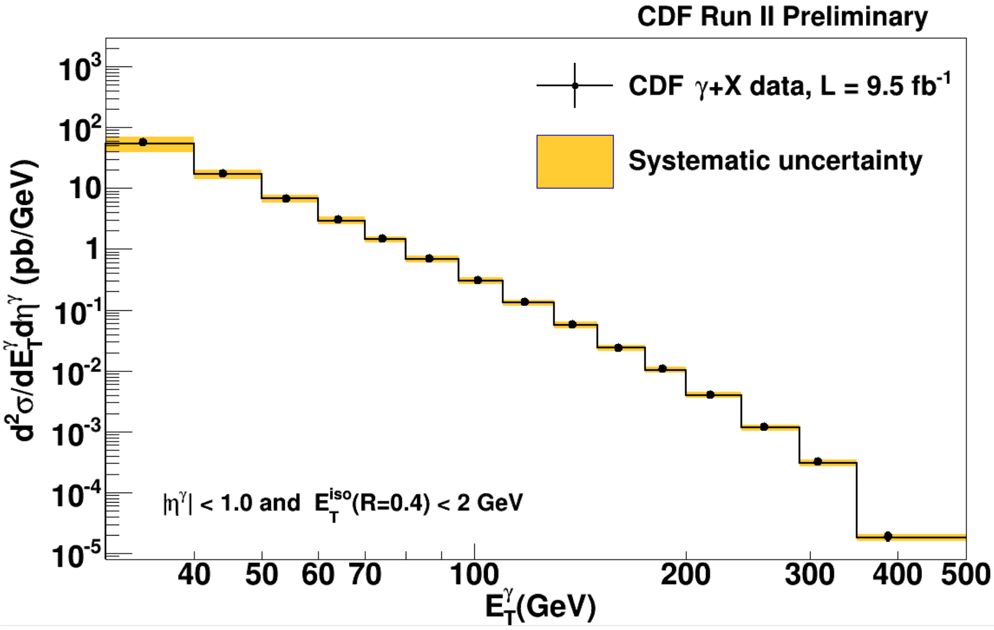


FIGURE 9.3: The measured $\gamma + X$ cross section compared. The vertical error bars show the statistical uncertainties, while the shaded areas show the systematic uncertainties. The 6% luminosity uncertainty on the data is not included.

9.3.1 Systematic Uncertainties

Systematic uncertainties for f^γ and $A \times \epsilon$, discussed in Section 8.2.5 and in Section 9.2.1, are propagated in the calculation of the cross section according to Equation 9.1.1. Table 9.2 summarizes all the considered systematic sources. When affecting both f^γ and $A \times \epsilon$, systematic uncertainties are considered as 100% correlated. Figure 9.4 shows the fractional systematic errors on the the measured cross sections as a function of E_T^γ. The largest contribution to the total uncertainty at low E_T^γ is caused by the modeling of ANN templates for the f^γ estimation. For E_T^γ =30 GeV, the calorimeter isolation energy requirement gives a contribution of 16%, which decreases to ~3% for higher E_T^γ. The choice of the generator (PYTHIA/SHERPA) strongly affects photon reconstruction and

Table 9.1: The measured differential cross section for the production of prompt isolated photons within the pseudorapidity region $|\eta^\gamma| < 1.0$, in bins of E_T^γ. $< E_T^\gamma >$ is the average E_T^γ within each bin. The uncertainties are statistical. The column $\delta\sigma_{syst}$ represents the systematic uncertainty. The additional 6% luminosity uncertainty is not included in the table.

E_T^γ (GeV)	$< E_T^\gamma >$ (GeV)	$d^2\sigma/dE_T^\gamma d\eta^\gamma$ (pb/GeV)	$\delta\sigma_{syst}$ (%)
30-40	34.1	$(5.49 \pm 0.41)\times10^{1}$	28.6
40-50	44.3	$(1.72 \pm 0.23)\times10^{1}$	17.9
50-60	54.3	$(6.72 \pm 0.11)\times10^{0}$	14.3
60-70	64.4	$(2.95 \pm 0.04)\times10^{0}$	13.9
70-80	74.5	$(1.45 \pm 0.02)\times10^{0}$	12.9
80-90	86.5	$(6.87 \pm 0.10)\times10^{-1}$	12.3
90-110	101.7	$(3.03 \pm 0.05)\times10^{-1}$	11.8
110-130	118.7	$(1.32 \pm 0.03)\times10^{-1}$	11.8
130-150	138.8	$(5.65 \pm 0.15)\times10^{-2}$	12.1
150-175	160.9	$(2.37 \pm 0.08)\times10^{-2}$	11.6
175-200	185.9	$(1.03 \pm 0.03)\times10^{-2}$	11.5
200-240	216.8	$(4.01 \pm 0.12)\times10^{-3}$	11.9
240-290	259.2	$(1.16 \pm 0.05)\times10^{-3}$	12.2
290-350	309.4	$(3.08 \pm 0.23)\times10^{-4}$	13.0
350-500	387.6	$(1.83 \pm 0.29)\times10^{-5}$	13.6

Table 9.2: Summary of the systematic uncertainties considered for cross section measurement. Uncertainties marked with $\surd$ are treated as 100% correlated.

Systematic	f^γ	$A \times \epsilon$
EIso4	$\surd$	$\surd$
Had/Em	$\surd$	$\surd$
CES/CEM	$\surd$	$\surd$
Photon energy scale	$\surd$	$\surd$
PDF	-	*
PYTHIA/SHERPA	-	*
ANN binning	*	-
Fitting Method	*	-

identification. It also probes the uncertainty on the signal reconstruction due to an alternative fragmentation mechanism. The uncertainty on the cross section due to this contribution is about 10% over the whole E_T^γ range. Finally, there is an additional 6% uncertainty on the integrated luminosity.

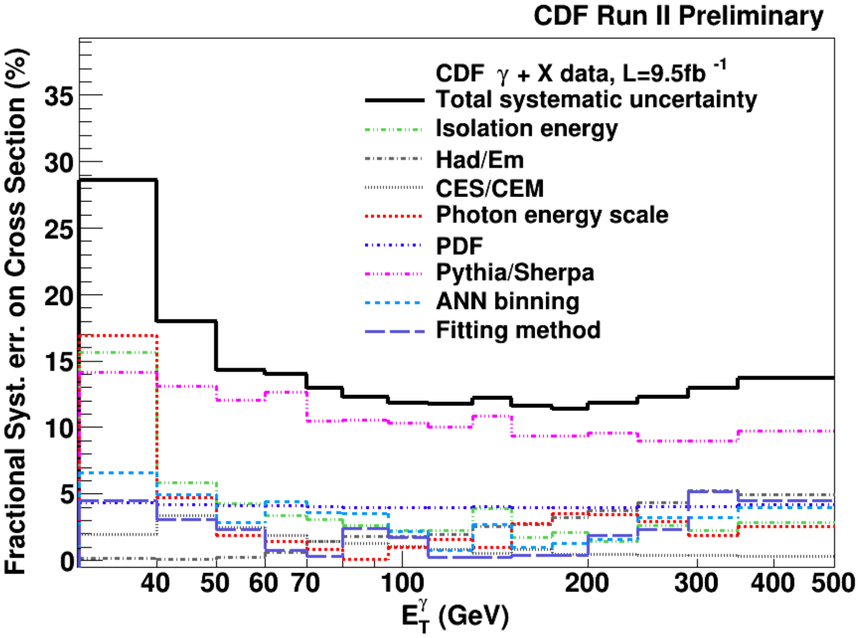


FIGURE 9.4: The fractional systematic errors on the measured $\gamma + X$ cross section. The continuous line is total systematic uncertainty while the dashed lines correspond to the single contribution. .

9.4 Theoretical Predictions

The expected prompt-photon production cross section is calculated using the fixed-order NLO program MCFM 6.8 (see Section 2.4.2).

The MCFM predictions are at parton-level and need to be corrected to the particle level. The correction factor C_{UE} accounts for the presence of contributions from the underlying event and parton-to-hadron fragmentation, which tend to increase the energy in the isolation cone. Such a correction reduces the predicted cross section, since the presence of extra activity results in some photons failing the isolation requirements. C_{UE} is determined as the ratio between the isolated fraction of the total prompt photon cross section at the particle level and the same fraction obtained after turning off both multiple-parton interactions and hadronization. The applied correction factor $C_{UE} = 0.91 \pm 0.03$, was estimated averaging the results in PYTHIA MC generated with the Tune A or Tune DW [38]; the corresponding plots are shown in Figure 9.5.

The parton-level isolation, defined as the total E_T from the partons produced with the photon inside a cone of radius $R = 0.4$ around the photon direction, is required

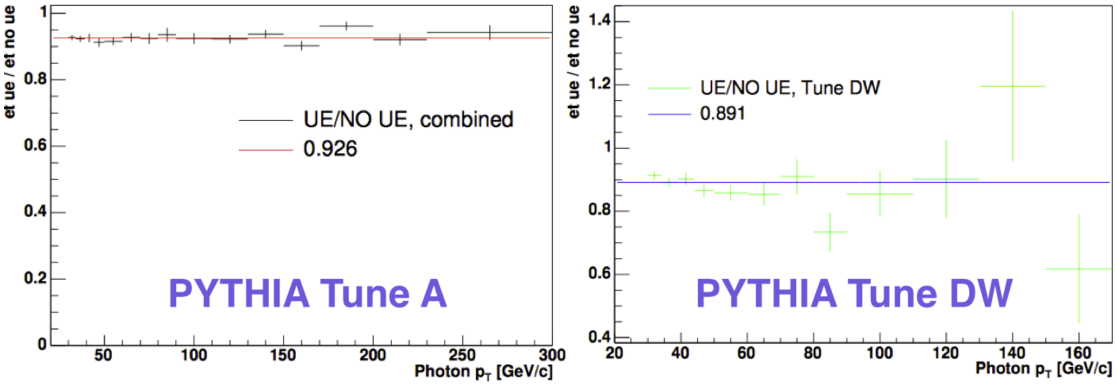


FIGURE 9.5: For comparisons of parton level MC to data, a correction (C_{UE}) for hadronization and underlying events is applied to parton-level MC. C_{UE} is estimated averaging the results in PYTHIA MC generated with the Tune A or DW [38].

to be smaller than 2 GeV. The nominal renormalization (μ_R), factorization (μ_F) and fragmentation (μ_f) scales were set to the photon transverse energy ($\mu_R = \mu_F = \mu_f = E_T^\gamma$). The scale uncertainty is evaluated by varying the three scales following the constraints $\mu_R = \mu_F = \mu_f = E_T^\gamma/2$ and $\mu_R = \mu_F = \mu_f = 2E_T^\gamma$. Variation of the scales by these factors of two induces changes between 10% and 15% in the theoretical predictions.

The measured cross section is also compared to those from the LO parton-shower generators PYTHIA and SHERPA. These models were described in Chapter 8. Both are calculated at the particle level, meaning that the photon isolation energy is estimated using generated stable particles to which cuts are applied; therefore, they both are directly comparable to the measurement.

9.4.1 Comparison to the theoretical predictions

The results are compared to the theoretical predictions (Figures 9.6, 9.7 and 9.8). The ratio of measured cross section over the predicted ones can be seen in Figure 9.8 and in Figure 9.7. In Figure 9.7 the full error bars on the data points represent statistical and systematic uncertainties summed in quadrature. The inner error bars show statistical uncertainties. The MCFM predictions are shown with their theoretical uncertainties (red curves), due to the choice of factorization and renormalization scales as well as the fragmentation scale, as discussed in Section 9.4.

The NLO calculations agree with the data up to the highest E_T^γ considered. The

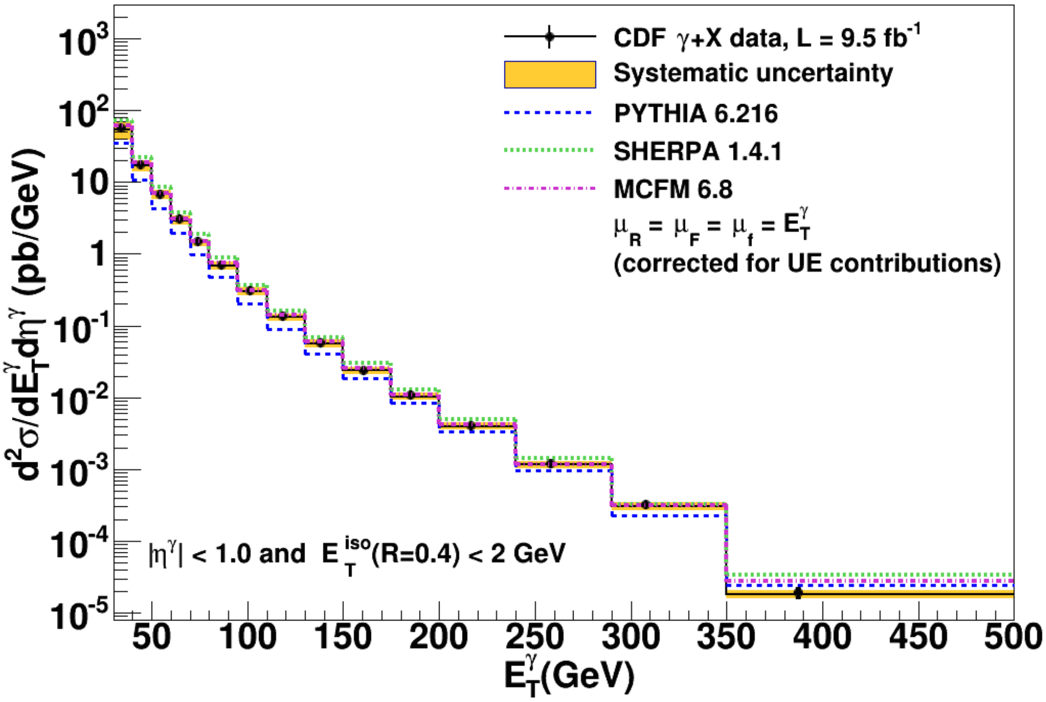


FIGURE 9.6: The measured $\gamma + X$ cross section compared with three theoretical predictions: PYTHIA, SHERPA and MCFM. The vertical error bars show the statistical uncertainties, while the shaded areas show the systematic uncertainties. The 6% luminosity uncertainty on the data is not included. A correction to account for extra activity (C_{UE}) is applied to the MCFM theoretical predictions, as explained in the text.

data are somewhat higher than the central NLO calculation for low E_T^γ , but agree within the theoretical uncertainty of the NLO calculation.

The predictions of PYTHIA and SHERPA are also shown in Figures 9.3 and 9.7. The shape of the cross section is well described by both models. PYTHIA prediction undershoots the data nearly uniformly across the E_T^γ range. Very likely this deficit in cross section strength comes from the lack of higher then LO terms in the PYTHIA photon+jet matrix elements. The SHERPA calculation overshoots the data nearly uniformly across the E_T^γ range. This is a NNLL calculation, including up to three jet emissions associated with the observed photon, but it is missing virtual corrections in the subprocess matrix elements. The lack of virtual terms could possibly be a reason for the excessive strength of the SHERPA cross section. Other possible reasons could be related with soft QCD mechanisms, such as mistuned fragmentation subprocesses leading to excessive rates of photon production through fragmentation.

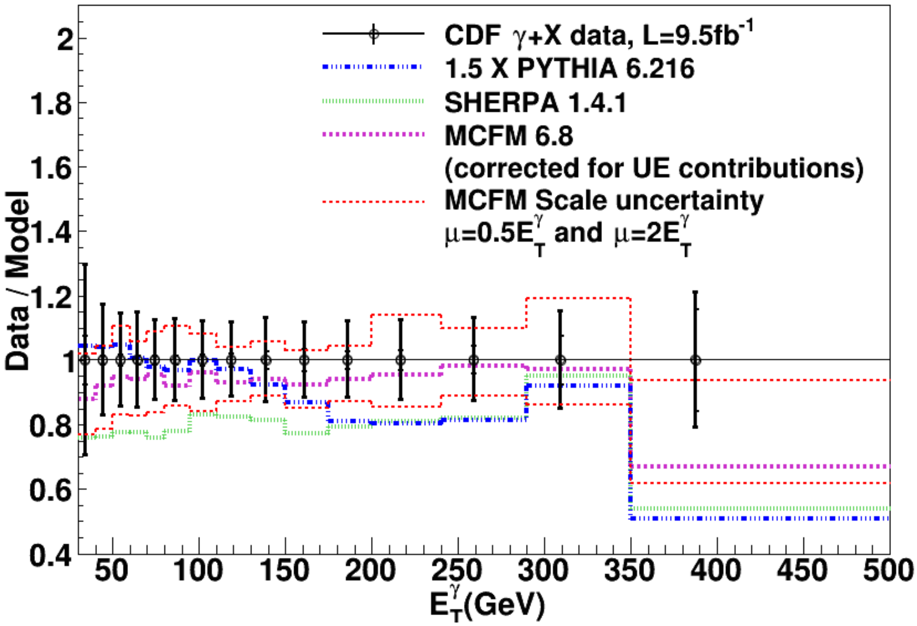


FIGURE 9.7: Data points centered at 1.0 and Data/Theory ratio (dashed lines) of the inclusive prompt photon cross section as a function of the photon transverse energy, E_T^γ , in the central η^γ region. The inner error bars on the data points show statistical uncertainties. The full error bars show statistical and systematic uncertainties added in quadrature. The 6% luminosity uncertainty on the data is not included. The PYTHIA prediction has been multiplied by a factor 1.5 for display reasons. The two sets of curves show the uncertainties on the MCFM theoretical predictions due to their dependency on the renormalization, factorization, and fragmentation scales.

.

9.5 Conclusions

A measurement of the differential cross section for the inclusive production of isolated prompt photons in $p\bar{p}$ collisions at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV is presented using the full dataset collected with the CDF II detector at the Tevatron. The cross section is measured as a function of photon transverse energy E_T^γ in the central pseudorapidity region ($|\eta^\gamma| < 1.0$). The E_T^γ kinematic range of this measurement spans from 30 GeV to 500 GeV.

Comparisons of the measurement to three theoretical predictions are discussed. Both PYTHIA and SHERPA describe the shape of the differential cross section. The PYTHIA generator predicts a smaller cross section compared to SHERPA and the data because of missing higher order corrections. The data have an overall good agreement with the fixed-order NLO MCFM calculation.

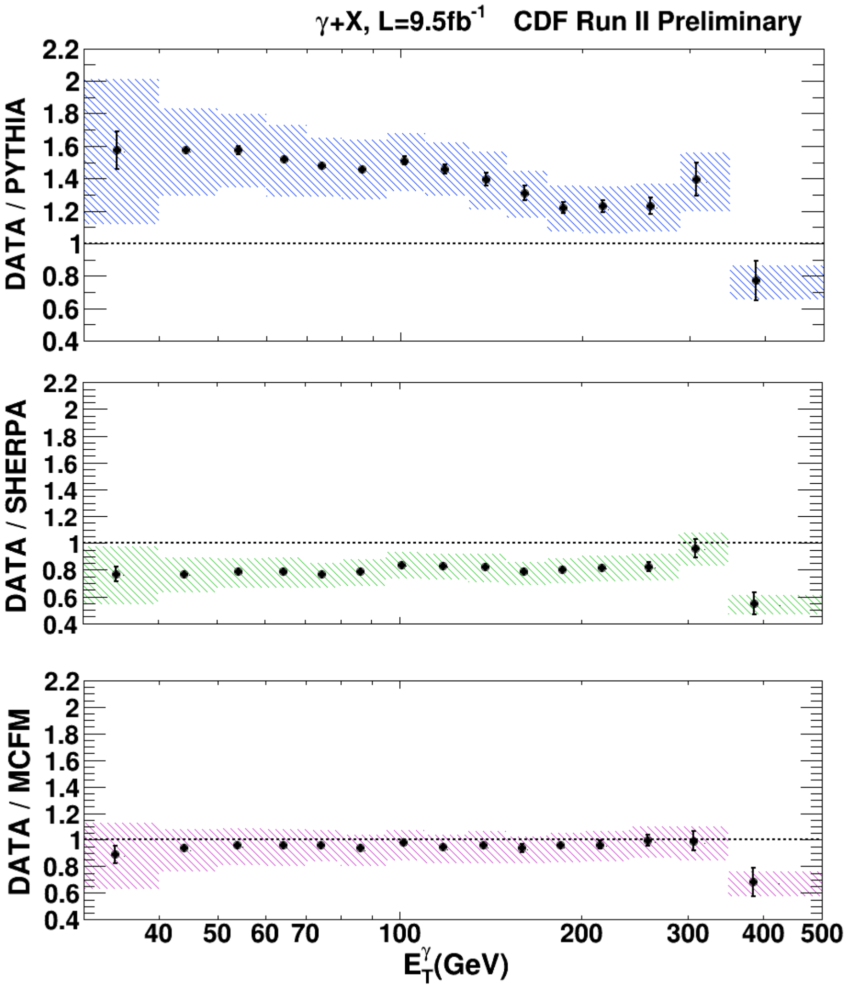


FIGURE 9.8: Ratio of the measured $\gamma + X$ cross section to three theoretical predictions: PYTHIA (upper part), SHERPA (central part) and MCFM (bottom part). The vertical error bars show the statistical uncertainties, while the shaded areas show the systematic uncertainties. The 6% luminosity uncertainty on the data is not included. A correction to account for extra activity (C_{UE}) is applied to the MCFM theoretical predictions, as explained in the text.

.

Chapter 10

Conclusions

A measurement of the differential cross section for the inclusive production of isolated prompt photons in $p\bar{p}$ collisions at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV is presented using the full data set collected with the CDF II detector at the Tevatron. The cross section is measured as a function of photon transverse energy E_T^γ in the central pseudorapidity region ($|\eta^\gamma| < 1.0$). The E_T^γ kinematic range of this measurement spans from 30 GeV to 500 GeV, thus extending the measured kinematic range previously published by CDF [38].

The prompt photon cross section measurement offers a unique opportunity to test the photon tools over a large energy range. In this thesis a new technique to suppress the irreducible isolated photons from meson decays is presented. The method is based on a likelihood fit of the Artificial Neural Network output distribution in the data to pure signal and background templates for every bin in the photon E_T. The ANN templates modeling is based only on the calorimeter information. The background subtraction technique developed for this measurement is currently being used for the measurement of the differential cross section of prompt photons associated with jets.

Data are unfolded back to particle level to correct for efficiencies and detector acceptance and resolution effects; the procedure is implemented in an array derived from a PYTHIA Monte Carlo. The unfolding factors do not present strong dependence on the photon transverse energy.

Results are compared to leading-order (LO) and next-to-leading order (NLO) pQCD calculations. The PYTHIA calculation predicts a smaller cross section compared to SHERPA and the data; both PYTHIA and SHERPA describe the shape of the differential cross section. The data have an overall good agreement with the prediction given by the fixed-order NLO MCFM calculation based on the MRST2008 NLO PDFs and GdRG LO FFs and renormalization, fragmentation and factorization scales set equal to the transverse momentum of the photon. The MCFM theoretical predictions are corrected for non perturbative QCD effects.

Finally, the measurement has been approved by the CDF Collaboration and it is in process for a publication in the Journal *Physical Review D*.

Appendix A

Lagrangian of the Standard Model

In Chapter 2, the SM Lagrangian densities for the QCD and the Electroweak Theory have been discussed. When gathering together all the ingredients described in the corresponding sections (Sections 2.2.1, 2.2.2, 2.2.3, 2.2.4), the SM Lagrangian density [15] is written as follows:

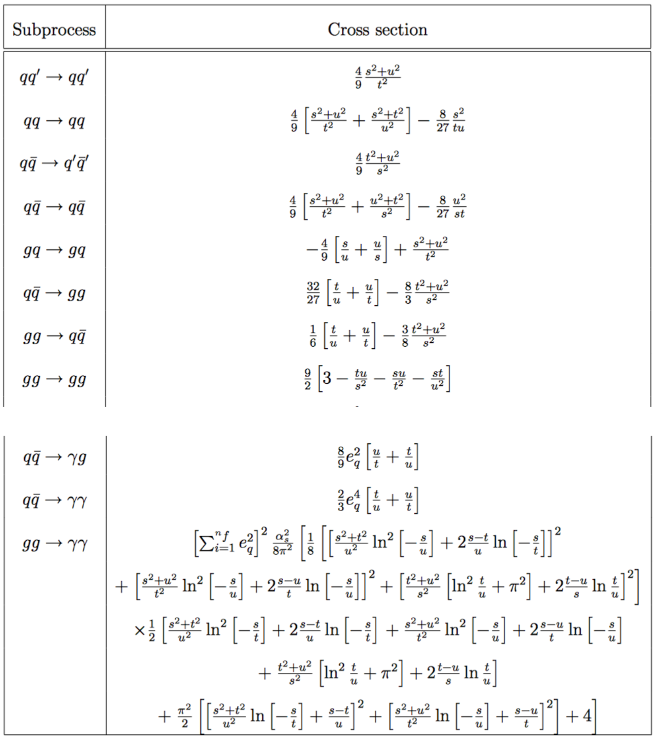
$$\begin{aligned}
\mathcal{L}_{SM} \;=\; & i\overline{\psi}_L \not{\partial}\psi_L + i\overline{\psi}_R \not{\partial}\psi_R + \\
& -\frac{1}{2}W^{+\mu\nu}W^{-}_{\mu\nu} - \frac{1}{4}Z^{\mu\nu}Z_{\mu\nu} - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \\
& -ig\cos(\theta_W)[(W^-_\mu W^+_\nu - W^-_\nu W^+_\mu)\partial^\mu Z^\nu + W^+_{\mu\nu}W^{-\mu}Z^\nu - W^-_{\mu\nu}W^{+\mu}Z^\nu] \\
& +ie[(W^-_\mu W^+_\nu - W^-_\nu W^+_\mu)\partial^\mu A^\nu + W^+_{\mu\nu}W^{-\mu}A^\nu - W^-_{\mu\nu}W^{+\mu}A^\nu] + \\
& +g^2\cos^2(\theta_W)(W^+_\mu W^-_\nu Z^\mu Z^\nu - W^-_\mu W^{+\mu}Z_\nu Z^\nu) + \\
& +g^2(W^+_\mu W^-_\nu A^\mu A^\nu - W^-_\mu W^{+\mu}A_\nu A^\nu) + \\
& +\mathrm{eg}\cos(\theta_W)[(W^+_\mu W^-_\nu(Z^\mu A^\nu + Z^\nu A^\mu) - 2W^+_\mu W^{-\mu}Z_\nu A^\nu] + \qquad \text{(A.0.1)} \\
& +\frac{\mathrm{g}^2}{2}[(W^+_\mu W^-_\nu(W^{+\mu}W^{-\nu} - W^{+\nu}W^{-\mu})] + \\
& +\frac{\mathrm{g}}{\sqrt{2}}[W^+_\mu(\overline{\nu}_L\gamma^\mu l_L + V_{CKM}\overline{u}_L\gamma^\mu d_L) + \mathrm{h.c.}] + \\
& +\frac{\mathrm{g}}{\cos\theta_W}Z_\mu(T^3 - Qsin^2\theta_W)\overline{\psi}_{L,R}\gamma^\mu\psi_{L,R} - eQA_\mu\overline{\psi}\gamma^\mu\psi + \\
& -\sqrt{2}(M\overline{\psi}_L\psi_R + \mathrm{h.c.}) + M^2_W W^{+\mu}W^-_\mu + \frac{M^2_{Z^0}}{2}Z^\mu Z_\mu + \\
& +\frac{1}{2}\partial_\mu H\partial^\mu H + \frac{\mathrm{g}^2}{4}(2v + H)H(W^{+\mu}W^-_\mu + \frac{M^2_{Z^0}}{2}Z^\mu Z_\mu) - \sqrt{2}\frac{M}{v}H\overline{\psi}\psi + \\
& +\lambda(\frac{3}{4}v^4 + 2v^3H + 2v^2H^2 + vH^3 + \frac{1}{4}H^4) + \\
& -\frac{1}{4}G^{\mu\nu}_a G^a_{\mu\nu} - g_S\overline{\psi}^a\gamma_\mu T^c_S\psi_b A^\mu_c
\end{aligned}$$

where $l = e, \mu, \tau$, $\nu = \nu_e, \nu_\mu, \nu_\tau$.

Appendix **B**

Parton-parton two-body scattering differential cross sections

The parton-parton two-body scattering differential cross sections are listed below.



Subprocess	Cross section
$qq' \to qq'$	$\frac{4}{9}\frac{s^2+u^2}{t^2}$
$qq \to qq$	$\frac{4}{9}\left[\frac{s^2+u^2}{t^2}+\frac{s^2+t^2}{u^2}\right]-\frac{8}{27}\frac{s^2}{tu}$
$q\bar{q} \to q'\bar{q}'$	$\frac{4}{9}\frac{t^2+u^2}{s^2}$
$q\bar{q} \to q\bar{q}$	$\frac{4}{9}\left[\frac{s^2+u^2}{t^2}+\frac{u^2+t^2}{s^2}\right]-\frac{8}{27}\frac{u^2}{st}$
$gq \to gq$	$-\frac{4}{9}\left[\frac{s}{u}+\frac{u}{s}\right]+\frac{s^2+u^2}{t^2}$
$q\bar{q} \to gg$	$\frac{32}{27}\left[\frac{t}{u}+\frac{u}{t}\right]-\frac{8}{3}\frac{t^2+u^2}{s^2}$
$gg \to q\bar{q}$	$\frac{1}{6}\left[\frac{t}{u}+\frac{u}{t}\right]-\frac{3}{8}\frac{t^2+u^2}{s^2}$
$gg \to gg$	$\frac{9}{2}\left[3-\frac{tu}{s^2}-\frac{su}{t^2}-\frac{st}{u^2}\right]$
$q\bar{q} \to \gamma g$	$\frac{8}{9}e_q^2\left[\frac{u}{t}+\frac{t}{u}\right]$
$q\bar{q} \to \gamma\gamma$	$\frac{2}{3}e_q^4\left[\frac{t}{u}+\frac{u}{t}\right]$
$gg \to \gamma\gamma$	$\left[\sum_{i=1}^{n_f} e_q^2\right]^2 \frac{\alpha_s^2}{8\pi^2}\left[\frac{1}{8}\left[\left[\frac{s^2+t^2}{u^2}\ln^2\left[-\frac{s}{u}\right]+2\frac{s-t}{u}\ln\left[-\frac{s}{t}\right]\right]^2\right.\right.$ $\left.+\left[\frac{s^2+u^2}{t^2}\ln^2\left[-\frac{s}{u}\right]+2\frac{s-u}{t}\ln\left[-\frac{s}{u}\right]\right]^2+\left[\frac{t^2+u^2}{s^2}\left[\ln^2\frac{t}{u}+\pi^2\right]+2\frac{t-u}{s}\ln\frac{t}{u}\right]^2\right]$ $\times\frac{1}{2}\left[\frac{s^2+t^2}{u^2}\ln^2\left[-\frac{s}{t}\right]+2\frac{s-t}{u}\ln\left[-\frac{s}{t}\right]+\frac{s^2+u^2}{t^2}\ln^2\left[-\frac{s}{u}\right]+2\frac{s-u}{t}\ln\left[-\frac{s}{u}\right]\right.$ $\left.+\frac{t^2+u^2}{s^2}\left[\ln^2\frac{t}{u}+\pi^2\right]+2\frac{t-u}{s}\ln\frac{t}{u}\right]$ $\left.+\frac{\pi^2}{2}\left[\left[\frac{s^2+t^2}{u^2}\ln\left[-\frac{s}{t}\right]+\frac{s-t}{u}\right]^2+\left[\frac{s^2+u^2}{t^2}\ln\left[-\frac{s}{u}\right]+\frac{s-u}{t}\right]^2\right]+4\right]$

FIGURE B.1: The parton-parton two body scattering differential cross sections. Factors of $\pi\alpha_S^2/s^2$, $\pi\alpha\alpha_S/s^2$ and $\pi\alpha_{2/s^2}$ have been factored out of the purely strong interaction, the single photon production and the double photon production processes, respectively. The e_q is the electric charge of quark. The s, t and u are the Mandelstam variables [1].

Appendix C

Dataset

Data sample

At L3, online data are divided into data streams based on similar triggers. The CDF's "c-stream" corresponds to events selected by high$-p_T$ triggers. In particular the following data samples are used: **cph1ad**, **cph1ah**, **cph1ai**, **cph1aj**, **cph1ak**, **cph1am**, and **cph1ap**.

Monte Carlo samples

The following Pythia photon MC samples are used to estimate the photon fractions and to unfold the measurement back to the hadron level.

The generation has been done with different cuts of $\hat{p}_T$ to guarantee enough statistics along the p_T range considered for the measurement of the cross section. Here, $\hat{p}_T$ is defined as the p_T of the outgoing partons in a $2 \rightarrow 2$ process in the center-of-mass frame relative to the axis defined along the trajectory of the incoming partons.

gq0sqd: $\hat{p}_T > 13$ GeV/c, $p_T^{\gamma\ GEN} > 22$ GeV/c, 323 pb^{-1};
pq0sj0: $\hat{p}_T > 20$ GeV/c, $p_T^{\gamma\ GEN} > 40$ GeV/c, 323 pb^{-1};
gq0s07: $\hat{p}_T > 30$ GeV/c, $p_T^{\gamma\ GEN} > 70$ GeV/c, 6692 pb^{-1};
gq0s15: $\hat{p}_T >$70 GeV/c, $p_T^{\gamma\ GEN} > 150$ GeV/c, 206 fb^{-1};

The following Sherpa photon MC samples are used for comparison to theory predictions
gq0s02: $p_T^{\gamma GEN} > 20$ GeV/c, 747.9 pb^{-1};
gq0s06: $p_T^{\gamma GEN} > 60$ GeV/c, 56.4 fb^{-1};

All available Pythia di-jet MC samples listed on the CDF II QCD MC webpage [1] are used to estimate the background. Only events with reconstructed and corrected $E_T > 1.1 \times \hat{p}_T$ are used.
$\hat{p}_T > 18$ GeV/c: **bt0sqb**, **bq0sqc**
$\hat{p}_T > 40$ GeV/c: **q8is01**, **bt0srb**, **bq0src**
$\hat{p}_T > 60$ GeV/c: **jq0s06**, **bt0ssb**, **bq0ssc**, **bq0ssd**
$\hat{p}_T > 70$ GeV/c: **jq0s07**
$\hat{p}_T > 80$ GeV/c: **jq0s08**
$\hat{p}_T > 90$ GeV/c: **jq0s09**, **bt0stb**, **bq0stc**, **bq0snd**
$\hat{p}_T > 110$ GeV/c: **jq0s11**
$\hat{p}_T > 120$ GeV/c: **bt0sub**, **bq0suc**, **bq0sud**
$\hat{p}_T > 130$ GeV/c: **jq0s13**
$\hat{p}_T > 150$ GeV/c: **jq0s15**, **bt0svb**, **bq0svc**
$\hat{p}_T > 170$ GeV/c: **jq0s17**
$\hat{p}_T > 200$ GeV/c: **jq0s20**, **bq0swd**
$\hat{p}_T > 230$ GeV/c: **jq0s23**
$\hat{p}_T > 300$ GeV/c: **bq0sxd**
$\hat{p}_T > 400$ GeV/c: **bq0syd**
$\hat{p}_T > 500$ GeV/c: **bq0szd**
$\hat{p}_T > 600$ GeV/c: **bq0s0d**

---

[1] CDF II QCD MC webpage: `http://www-cdf.fnal.gov/internal/physics/qcd/qcd.html`.

Appendix D

Correction to MC templates

The isolation correction is an E_T- dependent function derived in [38]. Its term are:

f_{shift}: **Align the data and template peaks**

f_{offset}: **Center the template at zero isolation**

f_{weight}:**Weight the Monte Carlo peak width**

Center the signal template back on its position

The correction formula is

$$E_T^{iso}(E_T) = (E_{Tuncorr}^{iso} - f_{shift}(E_T) - f_{offset}(E_T)) \times f_{weight}(E_T) + f_{offset}(E_T) \quad \text{(D.0.1)}$$

where $E_{Tuncorr}^{iso}$ is the isolation as it comes out from the photon Monte Carlo, f_{shift} aligns the data and template peaks, f_{offset} centers the template at 0 before being reweighted, $f_{weights}$ weights the peak width in the Monte Carlo template.

Appendix E

Fit validation - additional plots

This Appendix contains additional plots related to the fit validation procedure described in Chapter 8. Residual and pull distributions are shown for each bin of photon transverse energy. One can easily see that the mean values of the residual distributions are around zero and the pull widths are either ∼1 or close to 1.

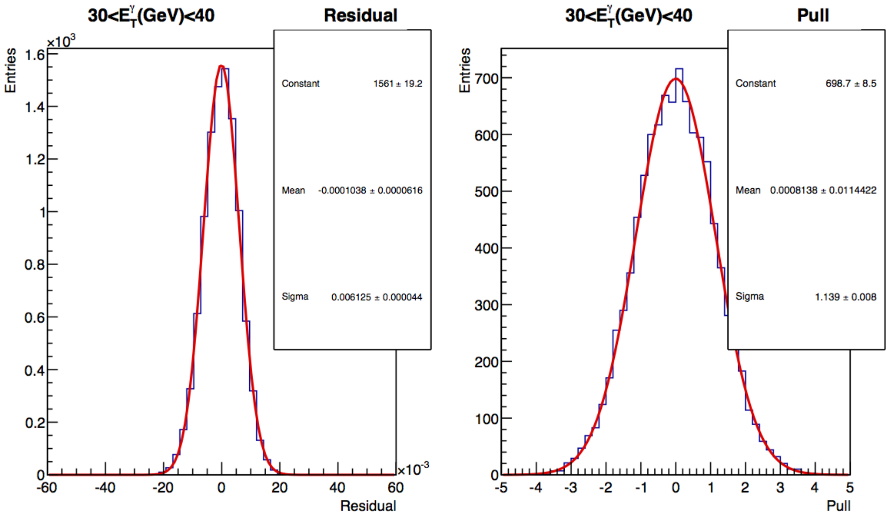


FIGURE E.1: Residual and pull distributions for pseudo-experiments with a number of entries N = 10000; distributions are shown for one bin of photon transverse energy. The mean of the residual distribution represent the difference between the expected signal fraction (from Data fit to MC templates) and the “trial” signal fraction (from pseudo-experiments). See text for more details (Section 8.2.2.2).

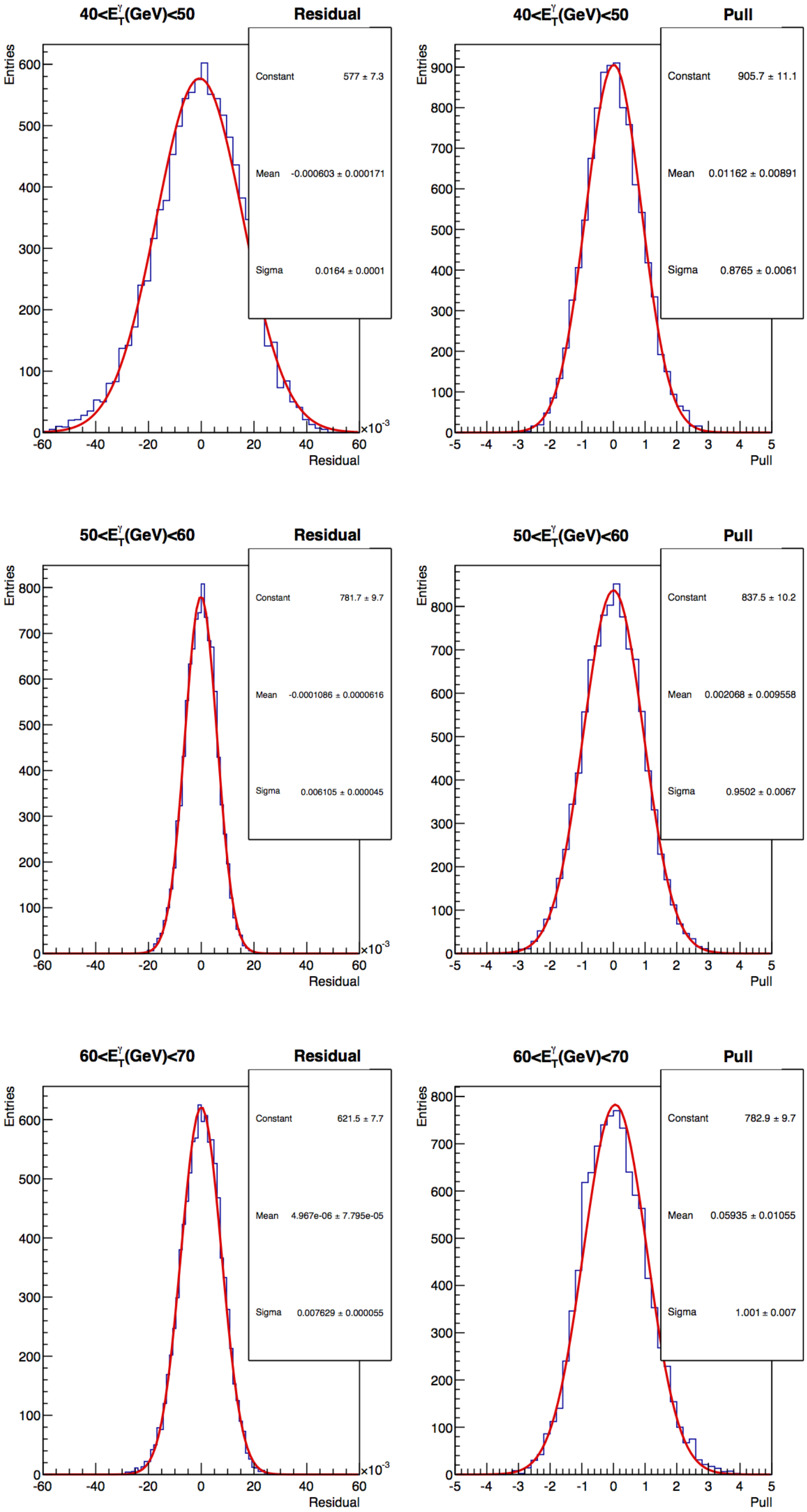


FIGURE E.2: Residual and pull distributions for pseudo-experiments with a number of entries N = 10000; distributions are shown in different bins of photon transverse energy. The mean of the residual distribution represent the difference between the expected signal fraction (from Data fit to MC templates) and the “trial” signal fraction (from pseudo-experiments). See text for more details (Section 8.2.2.2).

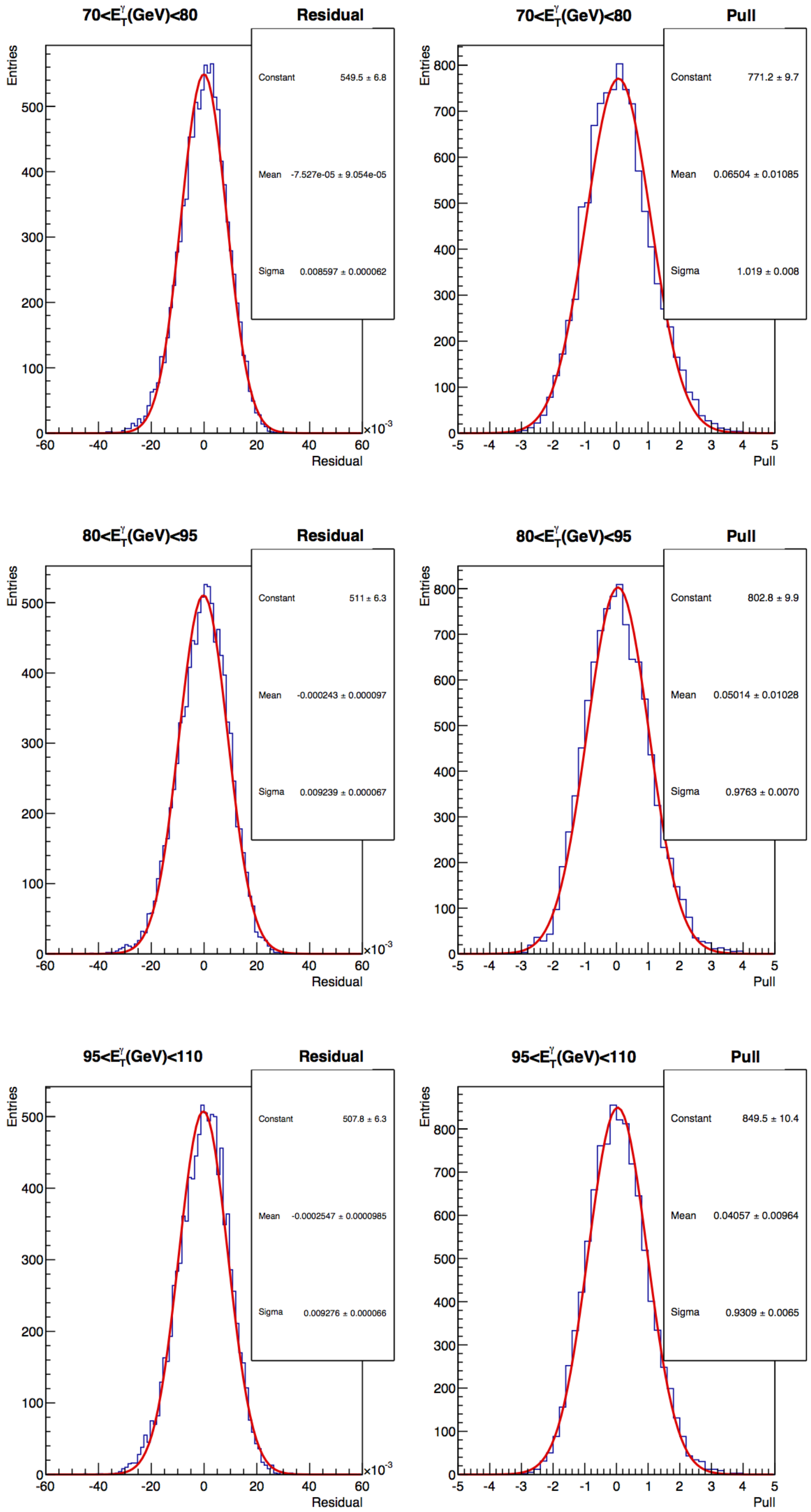


Figure E.3: Residual and pull distributions for pseudo-experiments with a number of entries N = 10000; distributions are shown in different bins of photon transverse energy. The mean of the residual distribution represent the difference between the expected signal fraction (from Data fit to MC templates) and the “trial” signal fraction (from pseudo-experiments). See text for more details (Section 8.2.2.2).

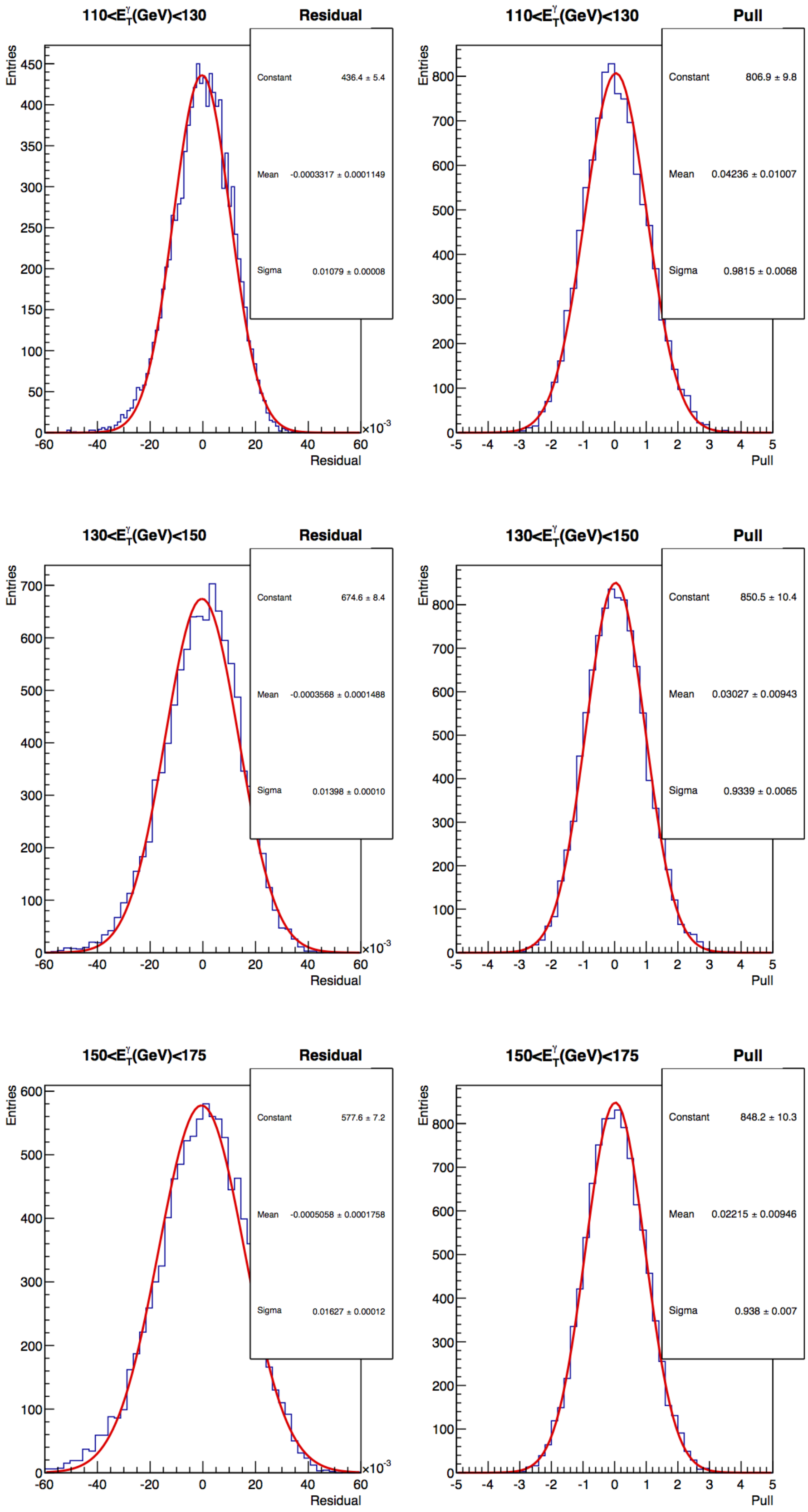


FIGURE E.4: Residual and pull distributions for pseudo-experiments with a number of entries N = 10000; distributions are shown in different bins of photon transverse energy. The mean of the residual distribution represent the difference between the expected signal fraction (from Data fit to MC templates) and the "trial" signal fraction (from pseudo-experiments). See text for more details (Section 8.2.2.2).

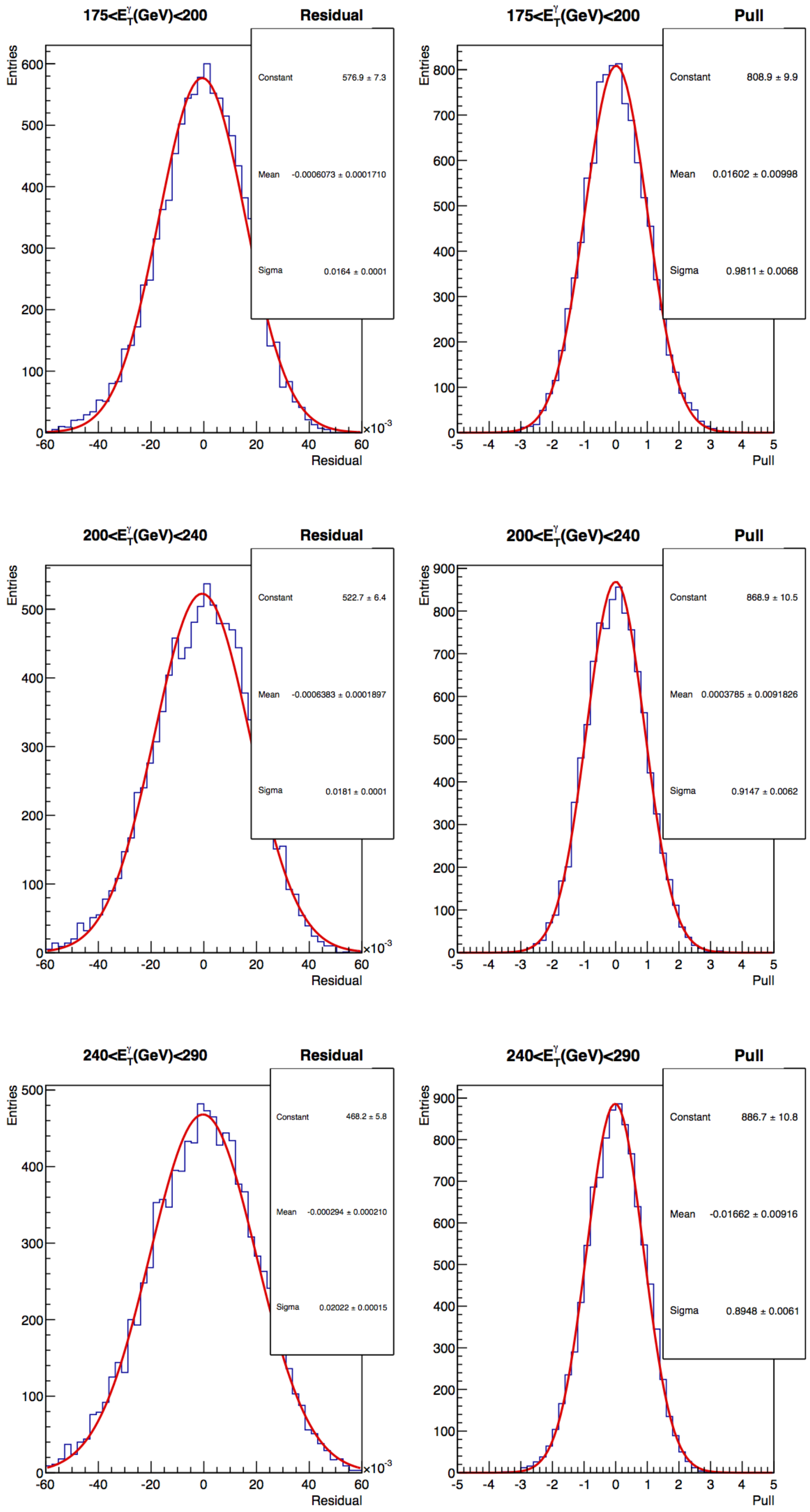


FIGURE E.5: Residual and pull distributions for pseudo-experiments with a number of entries N = 10000; distributions are shown in different bins of photon transverse energy. The mean of the residual distribution represent the difference between the expected signal fraction (from Data fit to MC templates) and the "trial" signal fraction (from pseudo-experiments). See text for more details (Section 8.2.2.2).

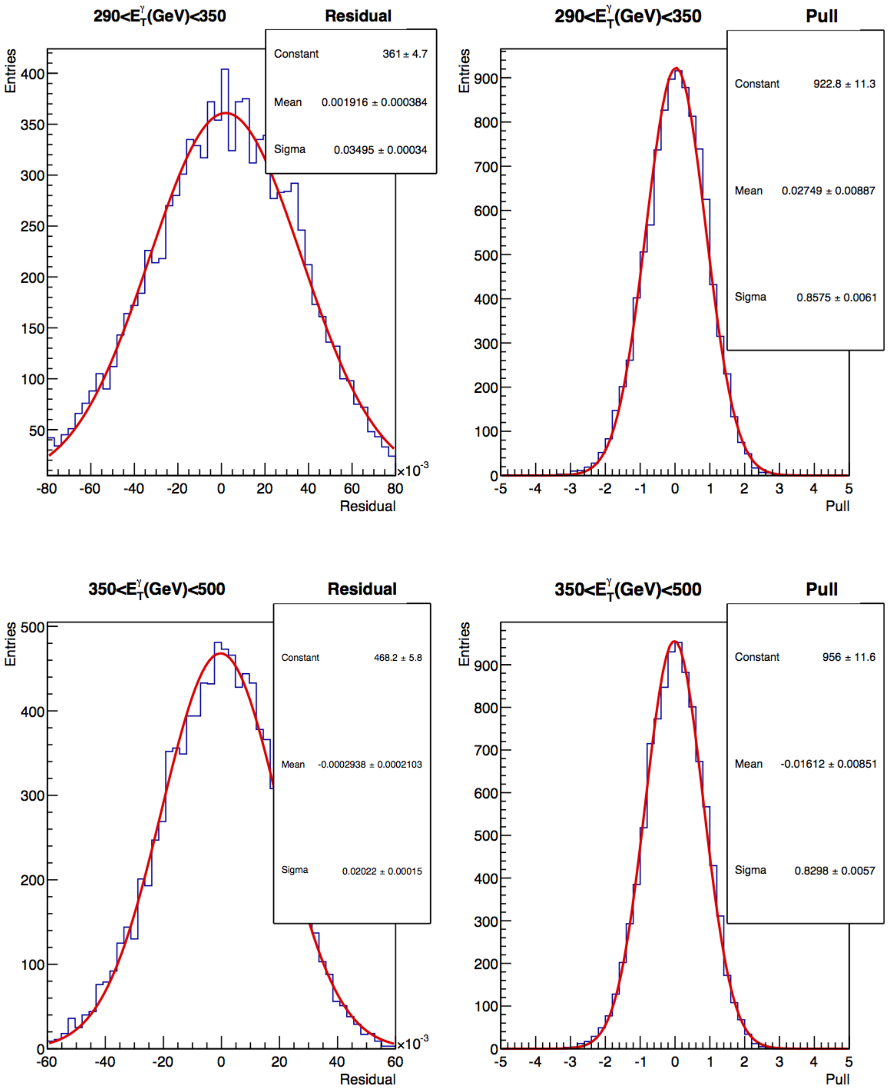


FIGURE E.6: residual and pull distributions for pseudo-experiments with a number of entries N = 10000; distributions are shown in different bins of photon transverse energy. The mean of the Residual distribution represent the difference between the expected signal fraction (from Data fit to MC templates) and the "trial" signal fraction (from pseudo-experiments). See text for more details (Section 8.2.2.2).

Appendix F

Two-Bins Fit

This fitting method, relies on the use of two bins distributions: for all the corresponding E_T^{γ} intervals, the ANN output distributions are divided into two bins. This means that instead of having a smooth and continuous distribution of the probability to have a true photons, there are only two possible values: 0 (corresponding to background) and 1 (corresponding to signal). This removes all details of the shapes. The 2-bins fit formula is:

$$N_i^D = f^S N_i^S + (1 - f^S) N_i^B \tag{F.0.1}$$

where N_i^D, N_i^S, N_i^B is the $i - th$bin content for Data (D), Signal MC (S) and Background MC (B) respectively; f^S is the Signal fraction. By defining

$$R_1^D = \frac{N_1^D}{N_1^D + N_2^D} \tag{F.0.2}$$

$$R_2^D = \frac{N_2^D}{N_1^D + N_2^D} \tag{F.0.3}$$

and similarly

$$R_1^S = \frac{N_1^S}{N_1^S + N_2^S} \tag{F.0.4}$$

$$R_2^S = \frac{N_2^S}{N_1^S + N_2^S} \tag{F.0.5}$$

$$R_1^B = \frac{N_1^B}{N_1^B + N_2^B} \tag{F.0.6}$$

$$R_2^B = \frac{N_2^B}{N_1^B + N_2^B} \tag{F.0.7}$$

Equation F.0.1 becomes:

$$f^S = \frac{R^D - R^B}{R^S - R^B} \tag{F.0.8}$$

The same result holds for the 1^{st} or 2^{nd} bin.

Appendix G

Residual Background

Diphoton background

Diphoton production is a SM process where a pair of prompt photon is produced, as shown in Figure G.1. In such a process, there is an high probability to lose one of the photon in a crack [1], in this case the event will have an associated high MET. As described in Chapter 7, among the selection requirements the high MET events are rejected. Furthermore, even if one of the photon in the final state will pass the analysis selection criteria and the second one will be identified as a jet, making difficult to distinguish from the single prompt photon production process, this is not a concern because the diphoton production cross section is smaller by orders of magnitude.
For these reasons, this source of background is completely negligible.

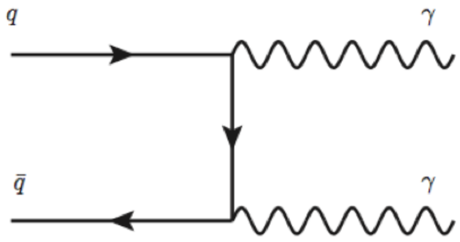


FIGURE G.1: Example of Feynman diagram for diphoton production in the SM.

---

[1] The rate of having one of the photons in an unintrumented region of the detector is twice as large as in $\gamma + X$ event

Electroweak background

A possible residual background could arise from electrons from Z and W decays, whose showers in the calorimeter can be misidentified as those of the photons. These electrons are rejected by retaining no more than 1 track pointing to the EM cluster. One extra track is allowed to take into account underlying event and pile-up energy around the cluster. If there is one track pointing to the photon cluster, this track has to be soft with respect to the measured energy of the photon. The cut on the missing transverse energy (MET) further remove electrons coming from W decays. An estimate of a residual 1% from other measurements [38] of electrons in the first two bins, makes this a minor background.

.

Acknowledgements

I would like to express my special thanks to my advisor from University of Roma "Tor Vergata" Prof. Giovanni Carboni, for his continuous availability and precious suggestions on improvements to this project.
I am pleased to acknowledge my advisor at Fermilab Dr. Ray Culbertson, for introducing me into the high energy collider analysis world. His advice on both research as well as on my career have been inestimable.

I also would like to thank my referees Dr. Giorgio Chiarelli and Dr. Matteo Palutan for their valuable feedback on this thesis and Prof. Massimo Bianchi, Prof. Rita Bernabei and Prof. Annalisa D'Angelo, for serving on my dissertation committee.

I express my gratitude to the CDF spokepeople Prof. Giorgio Bellettini and Prof. David Toback, for their strong support to my research activity at Fermilab, and to Dr. Stefano Miscetti, Dr. Fabio Happacher, Prof. Matteo Martini and Dr. Tommaso Spadaro from Laboratori Nazionali di Frascati of INFN.

During this work I have also benefited of fruitful collaborations and discussions with many CDF colleagues at Fermilab (alphabetically oredered): Dr. Marco Trovato, Dr. Costas Vellidis and Dr. Tingjun Yang.
Many thanks to Dr. Mike Albrow, Dr. Bo Jayatilaka, Dr. Christina Mesoropian, Prof. Pekka Sinervo, Dr. Diego Tonelli and Dr. Jon Wilson for their helpful contribution in reviewing the analysis results and the paper preparation.
I also would like to thank Dr. Willis Sakumoto, Dr. Stephan Lammel, Dr. Pasha

Murat and Dr. Andrei Gaponenko.
Thanks to Anna Driutti (now Dr.), the best PhD-mate I could ever had.
Thanks to Barbara Hehner, Cindy Kennedy and Barbara Kristen for being always nice to me.

Sono grata a Paolo, Lisa e Caterina.

Grazie infinite a mamma, papà, Pino, Gabri, Manu, Adri, Livi, Sasha, Silvanello, Paja, Dora, Bab e naturalmente Asse.
Dulcis in fundo, *nullius boni sine socio iucunda possessio est.*
Ago gratiam tibi, Secco.

Bibliography

[1] Owens, J. F *Large-momentum-transfer production of direct photons, jets, and particles*, Rev. Mod. Phys.**59**, 465–503 (1987)

[2] Bertfried Fauser, Jurgen Tolksdorf, Eberhard Zeidler Editors, *Quantum field theory: competitive models*, - Birkhauser Verlag AG, Basel Boston Berlin - Springer Science BusinessMedia

[3] D. J. Gross and F. Wilczek, *Asymptotically free gauge theories* Phys. Rev. D **8**, 3633 (1973);

H. D. Politzer, *Asymptotic freedom: An approach to strong interactions* Phys. Rev. Lett. **14**, 129 (1974).

[4] S. L. Glashow, *Partial symmetries of weak interactions* Nucl. Phys. **22**, 579-588 (1961);

A. Salam and J. C. Ward, *Electromagnetic and weak interactions* Phys. Lett. **13**, 168-171 (1964);

S. Weinberg, *A model of leptons* Phys. Rev. Lett. **19**, 1264 (1967).

[5] Povh, *Particles and Nuclei - Introduction to the Physical Concepts*, 5e Springer (2006)

[6] Particle Physics Resources for PHYS 4040/5040 http://www.hep.yorku.ca/menary/oldhep/

[7] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13**, 508 (1964).

[8] Aad G. *et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, arXiv:1207.7214.

Chatrchyan S. *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, arXiv:1207.7235.

[9] C. O'Luanaigh, *New results indicate that new particle is a Higgs boson*, CERN (14 March 2013). Retrieved 2013-10-09.

[10] I. J. R. Aitchison & A.J.G. Hey, *Gauge Theories in Particle Physics (second edition)*, Douglas F. Brewer, M.A., D.Phil. (1989).

[11] Beringer, J. and others *Review of Particle Physics (RPP)*, Phys. Rev. D **86**, 010001 (2012).

[12] D.J. Griffiths, *Introduction to Elementary Particles*, John Wiley & Sons (1987).

[13] Francis Halzen, and Alan D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics,* John Wiley & Sons, (1984).

[14] R.K. Ellis, W.J. Stirling and B.R. Webber, *QCD and Collider Physics*, Cambridge University Press (1996).

[15] Michael E. Peskin, Daniel V. Schroeder, *An Introduction to Quantum Field Theory*, Westview Press (1995).

[16] K. Hagiwara *et al. Measuring The WWZ Coupling At The Tevatron* Phys. Rev. D **41**, 2113-2119, (1990).

[17] Wesley Robert Ketchum, *Search for $ZW/ZZ \to l^+l^- +$ Jets Production in $p\bar{p}$ Collisions at CDF*, University of Chicago, Chicago U., FERMILAB-THESIS-2012-36 (2012).

[18] Yu. L. Dokshitzer, V.A Khoze, A.H. Mueller and S.I. Troyan, *Basic of Perturbative QCD*. Ed. J. Tran Thanh Van, Edditions Frontires, Gif-sur-Yvette, 1991.

[19] C.G. Callan Jr., *Broken scale invariance in scalar field theory*. Phys. Rev. **D 2** (1970) 1541; K. Symanzik, *Small distance behavior in field theory and power counting*. Comm. Math. Phys. 18 (1970) 227.

[20] R. P. Feynman, Proceedings of the 3rd Topical Conference on High Energy Collision of Hadrons, Stony Brook, N. Y. (1969)

[21] K. A. Olive *et al.* [Particle Data Group Collaboration], Chin. Phys. C **38**, 090001 (2014). doi:10.1088/1674-1137/38/9/090001

[22] G. Altarelli and G. Parisi, Nucl. Phys. **B126** (1977) 298;

[23] R. K. Ellis, W. J. Stirling and B. R. Webber, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **8**, 1 (1996).

[24] MRST Collaboration, `http://www.phys.psu.edu/cteq`

[25] MSTW Collaboration, `http://mstwpdf.hepforge.org/`

[26] A. D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C **63**, 189, (2009), arXiv:0901.0002 [hep-ph]

[27] Sherstnev, A. and Thorne, R.S.", *Parton distributions for LO generators*, Eur. Phys. J. C, **55**, 553 (2008), arXiv:0711.2473 [hep-ph]

[28] David d'Enterria and Juan Rojo, *Quantitative constraints on the gluon distribution function in the proton from collider isolated-photon data*, arXiv:1202.1762 [hep-ph]

[29] T. Binoth, *Two photon background for Higgs boson searches at the LHC.* QCD and high energy hadronic interactions. Proceedings, 35th Rencontres de Moriond, Les Arcs, France, March 18-25, 2000 (2000) pp. 151–154, arXiv:hep-ph/0005194 [hep-ph].

[30] L. Randall and R. Sundrum, *A large mass hierarchy from a small extra dimension*, Phys. Rev. Lett. 83,594 33703373 (1999).

[31] M. Dine, W. Fischler, and M. Srednicki, *Supersymmetric Technicolor*, Nucl. Phys. B596 189, 575593 (1981).

[32] E. Anassontzis *et al.*, *High p_T Direct Photon Production in pp Collisions*, Z. Phys. C **13** (1982) 277–289.

[33] A.L.S. Angelis *et al.*,*Direct photon production at the CERN ISR*, Nuclear Physics B, Volume **327**, Issue 3, 4 December 1989, Pages 541-568, ISSN 0550-3213,

[34] C. Albajar *et al.* (UA1 Collaboration), *Direct photon production at the CERN proton-antiproton Collider*, Phys. Lett. B **209** (1988) 385.

[35] M. Werlen *et al.* (UA6 Collaboration), *A New determination of* α_S *using direct photon production cross-sections in pp and* $p\bar{p}$ *collisions at* $\sqrt{s} = 24.3$*-GeV*, Phys. Lett. B **452** (1999) 201–206.

[36] V. M. Abazov et al. (D0 Collaboration), “Measurement of the isolated cross section in $p\bar{p}$ collisions at $\sqrt{s}$ =1.96 TeV”, Phys. Lett. **B** 639 (2006) 151

[37] D. Acosta et al. (CDF Collaboration), *Comparison of the isolated direct photon cross sections in* $p\bar{p}$ *collisions at* $\sqrt{s} = 1.8$TeV *and* $\sqrt{s} = 0.63$TeV, Phys. Rev. D 65, 112003 (2002).

[38] T. Aaltonen *et al.* (CDF Collaboration), *Measurement of the inclusive isolated prompt photon cross section in* $p\bar{p}$ *collisions at* $\sqrt{s}$ *=1.96 TeV using the CDF detector*, Phys. Rev. D **80** (2009) 111106(R).

[39] G. Aad et al. (ATLAS Collaboration), *Measurement of the inclusive isolated prompt photon cross section in pp collisions at* $\sqrt{s} = 7$ TeV *with the ATLAS detector*, Phys. Rev. D 83, 052005 (2011).

[40] G. Aad et al. (ATLAS Collaboration), Phys. Lett. B 706, 150 (2011).

[41] G. Aad et al. (ATLAS Collaboration), *Measurement of the inclusive isolated prompt photons cross section in pp collisions at* $\sqrt{s} = 7$ TeV *with the ATLAS detector using* 4.6 fb^{-1}, Phys. Rev. D 89, 052004 (2014).

[42] V. Khachatryan et al. (CMS Collaboration), *Measurement of the Isolated Prompt Photon Production Cross Section in pp Collisions at* $\sqrt{s} = 7$ TeV, Phys. Rev. Lett. 106, 082001 (2011).

[43] S. Chatrchyan et al. (CMS Collaboration), *Measurement of the differential cross section for isolated prompt photon production in pp collisions at 7 TeV*, Phys. Rev. D 84, 052011 (2011).

[44] M. Cacciari, G. P. Salam, and S. Sapeta, *On the characterisation of the underlying event*, JHEP **04** (2010) 065.

[45] T. Aaltonen *et al.* (CDF Collaboration), *Measurement of the* $WW + WZ$ *Production Cross Section Using the lepton + jets Final State at CDF II*, Phys. Rev. Lett. **104**, 101801 (2009).

[46] F. Abe *et al.*, *Observation of Top Quark Production in ppbar Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74**, 2626-2631 (1995).

[47] T. Aaltonen *et al.* (CDF Collaboration), *Combination of searches for the Higgs boson using the full CDF data set*, [arXiv:1301.6668 (2013)].

[48] P. Aurenche, R. Baier, M. Fontannaz, J. F. Owens and M. Werlen, *Gluon content of the nucleon probed with real and virtual photons*, Phys. Rev. D39 (1989), 3275.

[49] Aurenche, Catani, Fontannaz, Binoth, Guillet, Pilon, Werlen,

[50] J. M. Campbell and R. K. Ellis, MCFM home page, http://mcfm.fnal.gov

[51] L. Bourhis, M. Fontannaz and J. P. Guillet, Eur. Phys. J. C **2**, 529 (1998) [arXiv:hep-ph/9704447].

[52] A. Gehrmann-De Ridder, E. W. N. Glover, Eur. Phys. J. **C7**, 29-48 (1999). [hep-ph/9806316].

[53] Link to Fermilab Beam Division (Linac)

[54] Link to Fermilab Beam Division (Main Injector)

[55] Link to Fermilab Beam Division

[56] D. Acosta *et al.*. (CDF Collaboration), *The CDFII Detector Technical Design Report*, Fermilab Report No. Fermilab-Pub-96/390-E
Phys. Rev. D **71**, 032001 (2005).

[57] F. Abe *et al.* (CDF Collaboration), *Evidence for top quark production in* $\bar{p}p$ *collisions at* $\sqrt{s} = 1.8 \; TeV$, Fermilab-Pub-94/097-E.

[58] C.S. Hill (CDF Collaboration), *Initial Experience with the CDF Layer 00 Silicon Detector* Nucl. Instrum. Methods A **511**, 118 (2004).

[59] Diego Tonelli, *First observation of the* $B_s \rightarrow K^+K^-$ *decay mode, and measurement of the* B^0 *and* B_s *mesons decay-rates into two-body charmless final states at CDF.*, Scuola Normale Superiore, Pisa, FERMILAB-THESIS-2006-23 (2006).

[60] D. Amidei *et al.* (CDF Collaboration), *The Silicon Vertex Detector of the Collider Detector at Fermilab*, Nucl. Instr. and Meth. A **350** 73-130 (1994); S. Cihangir *et al.*, *The SVX-II silicon vertex detector upgrade at CDF*, Nucl. Instr. and Meth. A **360**, 118-124 (1995).

[61] A. Affolder *et al.*, *Intermediate silicon layers detector for the CDF experiment*, Nucl. Instrum. Methods, A **453**, 84-88 (2000).

[62] C. S. Hill (CDF Collaboration), *Operational experience and performance of the CDFII silicon detector*, Nucl. Instrum A **530**, 1-6 (2004).

[63] A. Abulencia *et al.* (CDF Collaboration), *Measurements of Inclusive W and Z Cross Sections in $p\overline{p}$ Collisions at $ps = 1.96$ TeV*, J. Phys. G: Nucl. Part. Phys. **34**, 2457-544 (2007).

[64] T. Affolder *et al.* (The CDF Collaboration), *CDF central outer tracker*, Nucl. Instrum. Meth. A **526**, 249-299 (2004).

[65] T. Aaltonen *et. al* (CDF Collaboration), *Operational Experience, Improvements, and Performance of the CDF Run II Silicon Vertex Detector*, FERMILAB-PUB-13-015-E.

[66] L. Balka *et al.*, *The CDF Central Electromagnetic Calorimeter*, Nucl. Instrum. Methods A **267**, 272 (1988).

[67] M. Gallinaro *et al.* (CDF Collaboration), *New Scintillator TileFiber Preshower Detector for the CDF Central Calorimeter*, IEEE Trans. Nucl. Sci. **52**, 879 (2005) [arXiv:physics/0411056];
M. Albrow *et al.* (CDF Collaboration), *A preshower detector for the CDF plug upgrade: Test beam results*, Nucl. Instrum. Meth. A **431**, 104-111 (1999);
A. Artikov *et al.*, *CDF Central Preshower and Crack Detector Upgrade.*, FNAL-PUB-07-023-E (2007).

[68] S. Bertolucci *et al.*, *The CDF Central and Endwall Hadron Calorimeter*, Nucl. Instrum. Methods A **267**, 301-314 (1988).

[69] M. Albrow *et al.*, *The CDF plug upgrade electromagnetic calorimeter: test beam results.*, Nucl. Instr. Meth. A 48Nucl. Instrum. Methods A **480**, 524-546 (2002).

[70] G. Apollinari *et al.*, *Shower maximum detector for the CDF plug upgrade calorimeter* Nucl. Instrum. Methods A **412**, 515-526 (1998).

[71] C. M. Ginsburg for the CDF COllaboration, *CDF Run 2 muon system*, Eur. Phys. J. C **33**, S1002-1004 (2004).

T. Dorigo (CDF Collaboration), *The muon system upgrade for the CDF II experiment*, Nucl. Instrum. Meth. A **461**, 560-2 (2001).

[72] D. Acosta *et al.* (CDF Collaboration), *The CDF Luminosity Monitor*, Nucl. Instrum. Meth. A **461**, 540-44 (2001);
D. Acosta *et al.* (CDF Collaboration), *The performance of the CDF luminosity monitor*, Nucl.Instrum. Meth. **A**, 57-62 (2002).

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

[74] E.J. Thomson *et al.*, IEEE Trans. on Nucl. Science. **49**, 1063-1070 (2002).

[75] B. Ashmanskas *et al.* (CDF Collaboration), *The CDF Silicon Vertex Trigger*, Nucl. Instrum. Meth. A **518**, 532-36 (2004);
I. Vila for the CDF Collaboration, *Performance and First Physics Results of the SVT Trigger at CDF II*, ECONF C0304052:FO008 (2003).

[76] A. Adelman, *et al.*, *The Silicon Vertex Trigger upgrade at CDF*, Nucl. Instrum. Methods A **572**, 361-64 (2007).

[77] A. Canepa *et al.*, *Level-3 Calorimetric Resolution available for the Level-1 and Level-2 CDF Triggers*, [arXiv:0810.3738] (2008).

[78] J. Mariner, *Secondary Vertex Fit With Mass and Pointing Constraints (CTVMFT)*, CDF Internal Note 1996 (1993).

[79] C. Hays *et al.*, *The COT Pattern Recognition Algorithm and Offline Code* CDF Internal Note 6992.

[80] B.R. Ko *et al. SVXII Stand-alone Tracking* CDF internal note 6440.

[81] Y. Huang *et al. Inside-Out Tracking* CDF internal note 6707.

[82] K. Bloom *et al.* Outside-In *Silicon Tracking at CDF* CDF internal note 5991.

[83] J. Goldstein *et al. Silicon Tracking for Plug Electrons* CDF internal note 5970.

[84] D. Acosta *et al.* (CDF Collaboration), *Measurement of the $t\bar{t}$ production cross section using lepton+jets events with secondary vertex b-tagging*, Phys. Rev. D **71**, 052003 (2005).

[85] Abe F *et al*, *Evidence for top quark production in $p\bar{p}$ collisions at $\sqrt{s}$ =1.8 TeV*, Phys. Rev. D **50**, 2966-3026 (1994).

[86] D. Acosta *et al.* (CDF Collaboration), *Direct Photon Cross Section with Con- versions at CDF*, Phys. Rev. D **70**, 074008 (2004);
Alon Attal, Anadi Canepa, *Photon Conversion Removal Efficiency*, CDF Internal Note 8073.

[87] Spreitzer *et al*, *Electron ID Efficiency and Scale Factor with TopNtuple* 6.1.4 *for Summer 2006* CDF Internal Note 8274.

[88] A. Taffard, *Run II Cosmic Ray Tagger Performances*, CDF Internal Note 6255.

[89] Grundler *et al*, *High-P_T muons recommended cuts and efficiencies for Summer 2006*, CDF Internal Note 8262.

[90] Ulysses, Grundler, Lubomir, Lovas, Anyes, Taffard, *High-Pt muons recommended cuts and efficiencies for Winter 2007* CDF Internal Note 8618.

[91] F. Abe *et al.* (The CDF Collaboration), *Measurement of the W Boson Mass* Phys. Rev. D **52**, 4784-4827 (1995).

[92] S. Eidelman *et al.*, *Review of particle physics. Particle Data Group* , Phys. Lett. B **592**, 1 (2004).

[93] S.R. Hahn *et al.*, Nucl. Instrum. Meth. A **267**, 351-366 (1988);
S. Bertolucci *et al.*, *The CDF Central and Endwall Hadron Calorimeter*, Nucl. Instrum. Meth. A **267**, 301 (1988).

[94] A. Bhatti *et al.*, *Determination of the jet energy scale at the collider detector at Fermilab*, Nucl. Instrum. Meth. A **566**, 375-412 (2006).

[95] S.V. Chekanov, *Jet algorithms: a minireview*, [arxiv:0211298v1] (2002).

[96] G. C. Blazey *et al.*, *Run II jet physics* [arXiv:0005012] (2000).

[97] Maria Spiropopulu, *A Blind search for supersymmetry in* $p\overline{p}$ *collisions at* $\sqrt{s}$ *= 1.8TeV using the missing energy plus multijet channel*, Harvard U. FERMILAB-THESIS-2000-16 (2000).

[98] M. Goncharov *et al.*,*The Timing System for CDF Electromagnetic Calorimeters*, FERMILAB-PUB-05-543-E

[99] CDF Note 10607

[100] G. Pauletta *et al.* (CDF Collaboration), *Operation and performance of the CDF calorimeters*,J. Phys.: Conference Series 160, 012007 (2009)

[101] Jean-Francois Arguin and Beate Heinemann, *Underlying Event Corrections for Run II*, CDF Internal Note 6239, 2003 (unpublished).

[102] R. Wagner, *Electron Identification for Run II: Algorithms*, CDF Internal Note 5456, 2003 (unpublished).

[103] M. Riveline, *CES Clustering in Run II*, CDF Internal Note 5863, 2002 (unpublished).

[104] K. Yasuoka, S. Mikamo, T. Kamon, and A. Yamashita (CDF Collaboration), R*esponse maps of the CDF central electromagnetic calorimeter with electrons*, Nucl. Instrum. Meth. A 267, 315 (1988).

[105] R. Wagner et al. (CDF Collaboration), *Cosmic ray test of the CDF central calorimeters*, Nucl. Instrum. Meth. A 267, 330 (1988).

[106] J. Marriner, *Secondary vertex fit with mass and pointing constrains*. CDF Note (1996).

[107] F. D. Snider *et al.*, *Tracking at CDF: Algorithms and Experience from Run I and Run II*. Nucl. Instrum. Methods. A 566 (2006), p. 133.

[108] S.-S. E. Yu and R. Culbertson, *Studies of Electron and Photon Energy Scales*, CDF Internal Note 9591, 2008 (unpublished).

[109] A. Hoecker *et al.* PoS ACAT **040** (2007).

[110] T. Sjöstrand *et al.*, J. High Energy Phys. **05**, 26 [arXiv:hep-ph/0603175] (2006).

[111] M. Goncharov *et al.*, Nucl. Instrum. Methods A **565**, 5643 (2006).

[112] H.L. Lai *et al.* (CTEQ Collaboration), *Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions*, Eur. Phys. J. C **12**, 375-392 (2000).

[113] T. Stelzer and W.F. Long, *Automatic generation of tree level helicity amplitudes*, Comput. Phys. Commun. **81**, 357–371 (1994).

[114] Fabio Maltoni and Tim Stelzer, *MadEvent: Automatic event generation with MadGraph*, JHEP **0302**, 027 (2003);
Johan Alwall *et al.*, *MadGraph/MadEvent v4: The New Web Generation*, JHEP **09** 028 (2007).

[115] V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. **15**, 438-450 (1972);
Yu. L. Dokshitzer, Sov. J. Phys. JETP **46**, 641 (1977);
G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. B **126**, 298 (1977).

[116] Link to P. Skands' talk (1/2);
Link to P. Skands' talk (2/2)

[117] S. Catani, F. Krauss, R. Kuhn, B.R. Webber, *QCD Matrix Elements + Parton Showers*, CERN-TH/2000-367 [hep-ph/0109231] (2001).

[118] X. Artru and G. Mennessier, *String model and multiproduction*, Nucl. Phys. B **70**, 93-115 (1974);
M.G. Bowler, *e^+e^- Production of Heavy Quarks in the String Model*, Zeit. Phys. C **11**, 169 (1981);
B. Andersson, G. Gustafson and B Söderberg, *A General Model for Jet Fragmentation*, Zeit. Phys. C **20**, 317 (1983); Nucl. Phys. B **264**, 29 (1986).

[119] R.D. Field and R.P. Feynman, *Quark Elastic Scattering as a Source of High Transverse Momentum Mesons* , Phys Rev. D **15**, 2590-2616 (1977);

[120] R.D. Field and S. Wolfram, *A QCD Model for e^+e^- Annihilation*, Nucl. Phys B **213**, 65 (1983);
B.R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, Nucl. Phys. B **238**, 492 (1984).

[121] Rene Brun, Federico Carminati, and Simone Giani, *GEANT Detector Description and Simulation Tool. 1994. CERN Program Library Long Writeup.*

[122] CDF Simulation, http://www-cdf.fnal.gov/cdfsim/cdfsim main.html.

[123] Guenter Grindhammer, M. Rudowicz, and S. Peters, *The Fast Simulation Of Electromagnetic And Hadronic Showers*, Nucl. Instrum. Meth. A **290**, 469 (1990).

[124] Aaltonen T. *et al.* (CDF Collaboration), *Observation of Single Top Quark Production and Measurement of $|V_{tb}|$ with CDF*, Phys. Rev. D **82**, 112005 (2010).

[125] Aaltonen T. *et al.*, *Measurement of cross sections for b jet production in events with a Z boson in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV*, Phys. Rev. D **79**, 052008 (2009).

[126] U. Langenfeld, S. Moch, and P. Uwer. *Measuring the running top-quark mass.*, Phys. Rev. D **80**, 054009 (2009);
A. Abulencia *et al.* (CDF Collaboration), *Measurement of the top quark mass using template methods on dilepton events in proton antiproton collisions at $\sqrt{s}$ = 1.96 TeV* Phys. Rev. D **73**, 112006 (2006).

[127] Aaltonen T. *et al.* (CDF Collaboration), *Measurement of the cross section for W^- boson production in association with jets in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV*, Phys. Rev. D **77**, 011108 (2008).

[128] N. Kidonakis, *Single top quark production cross section at hadron colliders*, arXiv:1005.3330 (2010);
N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. D **81**, 054028 (2010).

[129] N. Kidonakis, *Single top production at the Tevatron: Threshold resummation and finite-order soft gluon corrections*, Phys. Rev. D **74**, 114012 (2006).

[130] Particle Data Group Collaboration, J. Beringer et al., *Review of Particle Physics (RPP)*, Phys.Rev. D**86** (2012) 010001.

[131] MCnet Collaboration, A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. H o che, et al., *General-purpose event generators for LHC physics*, Phys.Rept. **504** (2011) 145, 1101.2599.

[132] T. Sj o strand, *Monte Carlo Generators*, `hep-ph/0611247`.

[133] Michael H. Seymour and Marilyn Marx, *Monte Carlo Event Generators*, arXiv:1304.6677v1 [hep-ph]

[134] https://mstwpdf.hepforge.org

[135] T. Aaltonen *et al.* (CDF Collaboration), *Studying the underlying event in Drell-Yan and high transverse momentum jet production at the Tevatron*, Phys. Rev. D **82**, 034001 (2010).

[136] R. Field and R. C. Group (CDF Collaboration), *PYTHIA Tune A, HERWIG, and JIMMY in Run 2 at CDF*, arxiv:0510198, 2005.

[137] F. Bedeschi *et al.*, *Design and Construction of the CDF Central Tracking Chamber*, Nucl. Instr. Meth. A **268**, 50-74 (1988).

[138] G. Apollinari *et al.* (CDF Collaboration), *Shower maximum detector for the CDF plug upgrade calorimeter*, Nucl. Instrum. Meth. A **412**, 515 (1998).

[139] [arXiv:hep-ph/0602031].

[140] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, J. High Energy Phys. **02**, 007 (2009).
S. Hoeche, S. Schumann, and F. Siegert, Phys. Rev. **D** 034026, 007 (2010).

[141] R. Culbertson *et al.*, *Search for Anomalous Production of di-photon+MET Events in 2/fb of Data.*, CDF Internal Note 9184.

[142] T. Aaltonen et al. (CDF Collaboration), *Search for Heavy, Long-Lived Neutralinos that Decay to Photons at CDF II Using Photon Timing*, Phys. Rev. D 78 (2008) 032015, [arXiv:0804.1043 [hep-ex]];

[143] J. Ray, R. C. Group, and R. Culbertson, *Training a Multivariate Photon ID Against Jets*, CDF Internal Note 10282, 2010 (unpublished)

[144] A. Hoecker et al. ,"TMVA 4 - Toolkit for Multivariate Data Analysis with ROOT", Users Guide, arXiv:physics/0703039 [Data Analysis, Statistics and Probability] CERN-OPEN-2007-007

[145] R. Brun and F. Rademakers, "ROOT - An Object Oriented Data Analysis Framework", Nucl. Inst. Meth. in Phys. Res. **A** 389, 81 (1997)

[146] R. Barlow and C. Beeston, Comp. Phys. Comm. **77** (1993) 219-228, and http://www.hep.man.ac.uk/ roger/hfrac.f

[147] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. 111, 042003 (2013).