

October 1994

INFN PI/AE 94/10

University of Pisa - Faculty of Science

Ph.D. THESIS IN PHYSICS

VI cycle of Doctorate School (November 1, 1990 - October 31, 1993)

**Top-Quark Search with Secondary Vertexing  
Tags in  $W +$  Multijet Events in  $p\bar{p}$  Collisions  
at  $\sqrt{s} = 1.8$  TeV**

Simone Dell'Agnello  
INFN - Pisa

Thesis Advisors: Prof. G. Bellettini, Dr. F. Bedeschi  
Chairman of the Dissertation Committee: Prof. P. Menotti



UNIVERSITÀ DEGLI STUDI DI PISA

DIPARTIMENTO DI FISICA



**Istituto Nazionale di Fisica Nucleare**

Sezione di Pisa

October 1994

INFN PI/AE 94/10

University of Pisa - Faculty of Science

Ph.D. THESIS IN PHYSICS

VI cycle of Doctorate School (November 1, 1990 - October 31, 1993)

**Top-Quark Search with Secondary Vertexing  
Tags in  $W +$  Multijet Events in  $p\bar{p}$  Collisions  
at  $\sqrt{s} = 1.8$  TeV**

Simone Dell'Agnello  
INFN - Pisa

Thesis Advisors: Prof. G. Bellettini, Dr. F. Bedeschi  
Chairman of the Dissertation Committee: Prof. P. Menotti

To my parents

# Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>   | <b>10</b> |
| <b>2</b> | <b>The Top Quark in the Standard Model</b>  | <b>13</b> |
| 2.1      | The Standard Model . . . . .  | 13        |
| 2.1.1    | Electroweak Interactions . . . . .  | 13        |
| 2.1.2    | Strong Interactions . . . . .   | 15        |
| 2.2      | Limits on the Top-Quark Mass . . . . .  | 17        |
| 2.3      | Top-Quark Production and Decay in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV . . . . . | 18        |
| 2.4      | CDF Top-Quark Search Program Within the Minimal Standard Model . . . . .                  | 19        |
| 2.5      | CDF Top-Quark Searches Beyond the Minimal Standard Model . . . . .                        | 22        |
| <b>3</b> | <b>The CDF Experiment</b>   | <b>24</b> |
| 3.1      | The Tevatron Collider of Fermilab . . . . .   | 24        |
| 3.2      | CDF, the Collider Detector at Fermilab . . . . .  | 25        |
| 3.2.1    | Tracking System . . . . .   | 27        |
| 3.2.2    | Calorimetry . . . . .   | 29        |
| 3.2.3    | Muon Detection . . . . .  | 30        |
| 3.2.4    | Trigger System . . . . .  | 32        |
| 3.2.5    | High $P_t$ Electron and Muon Triggers . . . . .   | 32        |
| 3.2.6    | Jet Triggers . . . . .  | 33        |
| <b>4</b> | <b>The Silicon VerteX Detector (SVX)</b>  | <b>36</b> |
| 4.1      | Mechanical Characteristics . . . . .  | 37        |
| 4.2      | Electrical Characteristics . . . . .  | 40        |
| 4.3      | Calibrations . . . . .  | 41        |
| 4.4      | Tracking with the SVX . . . . .   | 41        |
| 4.4.1    | Track Parametrization . . . . .   | 41        |
| 4.4.2    | Track Reconstruction Algorithm . . . . .  | 42        |
| 4.4.3    | Track Reconstruction Efficiency . . . . .   | 44        |
| 4.4.4    | Detector Alignment . . . . .  | 45        |
| 4.4.5    | Track Impact Parameter Resolution . . . . .   | 45        |
| <b>5</b> | <b>Secondary Vertex Tagging with the SVX</b>  | <b>52</b> |
| 5.1      | Finding Primary $p\bar{p}$ Interaction Vertices . . . . .                                 | 52        |
| 5.1.1    | Average $p\bar{p}$ Beam Line Position . . . . .   | 53        |

|          |   |            |
|----------|---|------------|
| 5.1.2    | Multiple Primary Interactions . . . . .   | 53         |
| 5.1.3    | The “VXPRIM” Primary Vertex Finding Algorithm . . . . .   | 55         |
| 5.1.4    | Results on the High $P_t$ Electron and Muon Sample . . . . .  | 56         |
| 5.2      | Tagging Secondary Decay Vertices . . . . .  | 65         |
| 5.2.1    | Impact Parameter-Azimuth ( $D - \phi$ ) Space . . . . .   | 65         |
| 5.2.2    | Track Selection . . . . .   | 65         |
| 5.2.3    | The “ $D - \phi$ ” Secondary Vertex Tagging Algorithm . . . . .                                       | 66         |
| 5.2.4    | Jets and Leptons Associated with Secondary Vertices . . . . .   | 68         |
| 5.2.5    | Backward Decays . . . . .   | 69         |
| <b>6</b> | <b>Characterization of the <math>D - \phi</math> Algorithm As a <math>b</math> Tagger</b>             | <b>76</b>  |
| 6.1      | The $b$ -Enriched Sample: Inclusive Muons . . . . .   | 77         |
| 6.1.1    | Identification of Jets Associated with Muons . . . . .  | 79         |
| 6.1.2    | Identification of Jets Recoiling Against Muon Jets . . . . .  | 79         |
| 6.1.3    | Event Selection . . . . .   | 79         |
| 6.1.4    | Measurement of the $b$ Fraction ( $F_b$ ) for Muon Jets . . . . .                                     | 80         |
| 6.1.5    | Measurement of the $b$ Fraction ( $F_b$ ) for Recoil Jets . . . . .                                   | 85         |
| 6.1.6    | $b$ -Tagging Efficiency ( $\epsilon_b$ ) . . . . .  | 88         |
| 6.1.7    | Measurement of $\epsilon_b$ for Muon Jets . . . . .   | 89         |
| 6.1.8    | Measurement of $\epsilon_b$ for Recoil Jets . . . . .   | 90         |
| 6.1.9    | Measurement of $\epsilon_b$ for Muon Jets Using Double $b$ Tags . . . . .                             | 90         |
| 6.1.10   | Montecarlo $b\bar{b} \rightarrow \mu X$ Sample . . . . .  | 93         |
| 6.1.11   | Decay Length of $b$ Tags . . . . .  | 94         |
| 6.2      | The Low $b$ -Content Sample: Inclusive Jets . . . . .   | 97         |
| 6.2.1    | The “Jet_50” Sample . . . . .   | 98         |
| 6.2.2    | The “Jet_100” and “ $\Sigma E_t$ ” Samples . . . . .  | 101        |
| 6.3      | Summary of Results . . . . .  | 105        |
| 6.4      | Appendix to 6.1.9: Calculation of $\epsilon_b$ for Muon Jets Using Double $b$ Tags . . . . .          | 106        |
| <b>7</b> | <b>Inclusive <math>W +</math> Multijet Selection</b>  | <b>111</b> |
| 7.1      | $W \rightarrow e + \bar{\nu}_e, \mu + \bar{\nu}_\mu$ Selection . . . . .                              | 111        |
| 7.1.1    | Identification of High $E_t$ Electrons . . . . .  | 111        |
| 7.1.2    | Identification of High $P_t$ Muons . . . . .  | 112        |
| 7.1.3    | Measurement of the Missing Transverse Energy ( $E_t$ ) . . . . .                                      | 113        |
| 7.2      | Final Event Selection . . . . .   | 115        |
| <b>8</b> | <b>Top-Quark Search with the <math>D - \phi</math> <math>b</math> Tagger</b>                          | <b>119</b> |
| 8.1      | The Counting Experiment . . . . .   | 119        |
| 8.2      | Observed $D - \phi$ Tags in the $W +$ Multijet Sample . . . . .                                       | 120        |
| 8.3      | Background Sources . . . . .  | 127        |
| 8.3.1    | Estimate of Background from $Wb\bar{b}, Wc\bar{c}$ and Fake Secondary Vertices:<br>Method 1 . . . . . | 127        |
| 8.3.2    | Estimate of Background from $Wb\bar{b}, Wc\bar{c}$ and Fake Secondary Vertices:<br>Method 2 . . . . . | 128        |
| 8.3.3    | Other Background Sources . . . . .  | 128        |

|       |  |     |
|-------|--|-----|
| 8.3.4 | Summary of Background Estimate . . . . .   | 131 |
| 8.4   | Statistical Significance of the Result . . . . .   | 131 |
| 8.5   | Estimate of the $b$ -Tagging Efficiency for $t\bar{t} \rightarrow W + \text{Multijet}$ Events . . . . .    | 132 |
| 8.6   | Estimate of the Expected Number of $b$ -Tagged $t\bar{t} \rightarrow W + \text{Multijet}$ Events . . . . . | 134 |
| 8.7   | Study of the $t\bar{t}$ Production Hypothesis . . . . .  | 135 |

# List of Figures

|     |   |    |
|-----|---|----|
| 2.1 | Tree level diagram for $t\bar{t}$ pair production via $q\bar{q}$ annihilation, followed by the Standard Model decay chain. . . . .  | 19 |
| 2.2 | An example of $W$ + multijet production. . . . .  | 21 |
| 2.3 | The Standard Model tree level prediction of the cross section of $t\bar{t}$ , $W$ + 3 jet and $W$ + 4 jet production. . . . .   | 23 |
| 3.1 | Schematic view of several components of the CDF detector. The figure shows a quarter side-view of CDF. The detector is forward-backward symmetric about the interaction region (which is at the lower-right corner of the figure), and symmetric around the line of the beams. The CDF cartesian and polar coordinate systems are indicated at the upper-left corner of the figure. . . . . | 26 |
| 3.2 | $R - \phi$ view of the wire hits of an event with reconstructed tracks in the central tracking chamber (CTC); the inset on the top left corner shows the cell geometry of outer superlayers 4 to 8, where sense wires are indicated by crosses. . . . .   | 28 |
| 3.3 | Distribution of the energy lost in the central hadron calorimeter by $P_t > 15$ GeV/c muon candidates detected in central muon chambers (CMU), with and without the requirement of simultaneous detection in the central muon upgrade detector (CMP). . . . .   | 34 |
| 3.4 | Geometry of the CMU. The chambers are located outside each azimuthal slice (wedge) of the hadron calorimeter. Between chambers of adjacent wedges there are uninstrumented regions of $\sim 2.4^\circ$ , which give an overall loss of geometrical acceptance of $\sim 16\%$ . . . . .  | 35 |
| 4.1 | Isometric view of an SVX barrel. . . . .  | 38 |
| 4.2 | Schematic view of the components of a ladder. . . . .   | 39 |
| 4.3 | Track of a particle with positive charge $Q$ . . . . .  | 43 |
| 4.4 | Distribution of SVX track residuals before (dashed histogram, $4 \mu\text{m}/\text{bin}$ ) and after (squared points and gaussian fit) internal alignment performed with SVX tracks. . . . .  | 47 |
| 4.5 | SVX impact parameter distribution measured with respect to the average beam line as a function of track $P_t$ in the data. . . . .  | 48 |
| 4.6 | SVX impact parameter resolution as a function of track $P_t$ in the data. . . . .   | 49 |
| 4.7 | SVX impact parameter resolution as a function of track $P_t$ in the Montecarlo. . . . .   | 50 |
| 4.8 | $P_t$ resolution of CTC only and CTC + SVX tracks in the data. . . . .  | 51 |

|      |  |    |
|------|--|----|
| 5.1  | Distribution of the maximum value of the variable $Q$ (see text) in events with electron and muons of $P_t > 20$ GeV/ $c$ (fig. a) and in minimum bias events (fig. b). . . . .  | 54 |
| 5.2  | Schematic description of the VXPRIM primary vertex finding algorithm. . .  | 58 |
| 5.3  | Resolution in the $x$ and $y$ coordinates for the primary vertex found by VXPRIM (top plot) and for the seed primary vertex (bottom plot). Each event enters the plot twice, once for the $x$ and once for the $y$ coordinate. . . . .   | 59 |
| 5.4  | Difference (improvement) between the VXPRIM and the seed primary vertex resolution in $x$ and $y$ (top plot); improvement in $x$ and $y$ resolution as a function of the number of tracks used in the primary vertex fit (bottom plot). Each event enters the plot twice. . . . .  | 60 |
| 5.5  | Difference between the VXPRIM and the seed primary vertex $x$ and $y$ coordinates (top plot); this difference normalized by its error (“pulls” distribution, bottom plot). Each event enters the plot twice. . . . .   | 61 |
| 5.6  | (1) $x$ - $y$ profile of the colliding beams as reconstructed by VXPRIM (see text); (2) projection of the beam profile on the $x$ axis; (3) beam profile in the $x$ - $z$ plane. . . . .   | 62 |
| 5.7  | $Z$ -distance between the primary vertex found by VXPRIM and the high $P_t$ lepton ( $Z_{diff} = \Delta Z$ ); events with $ \Delta Z  > 5$ cm are rejected as potential multiple primary vertex events; $ \Delta Z  < 5$ cm events are then accepted only if the lepton passes the high $P_t$ trigger. . . . .   | 63 |
| 5.8  | $Z$ -distance between the primary vertex found by VXPRIM and the high $P_t$ lepton ( $Z_{diff} = \Delta Z$ ), when the maximum- $Q$ primary interaction is not used as a seed for the VXPRIM fit. . . . .  | 64 |
| 5.9  | Graphic illustration of a $D - \phi$ cluster with three tracks and three links; also shown is the road of association of electrons and muons with $D - \phi$ clusters. . . . .   | 71 |
| 5.10 | $D - \phi$ plot for an event with two three-track $b$ tags. A $P_t = 13.6$ GeV/ $c$ electron is associated with the cluster at $\phi \sim 2.2$ radians (triangular symbol labelled “notag, in road”); this electron does not belong to the links (“notag”), because it has a low value of the impact parameter, but it lies in the road of cluster-to-lepton association (“inroad”). . . . .   | 72 |
| 5.11 | $R - \phi$ view of an event with a five-track $b$ tag, candidate for the semileptonic decay $b \rightarrow e\nu X$ . . . . .   | 73 |
| 5.12 | Magnification of the candidate $b \rightarrow e\nu X$ event in a region of radius = 2 cm around the primary vertex. . . . .  | 74 |
| 5.13 | Magnification of $B^0 \rightarrow J/\Psi K^{*0} \rightarrow \mu^+ \mu^- K^+ \pi^-$ candidate event in a region of radius = 2 cm around the PV. . . . .   | 75 |
| 6.1  | Inclusive electron and muon sample: reconstruction of $D^0 \rightarrow K\pi$ decays associated with electrons and muons. To reduce the combinatorial background, $K\pi$ candidate pairs are selected by requiring their transverse decays length, $L_{xy}$ , be larger than the relative uncertainty, $\sigma_{L_{xy}}$ . The detection of $c$ hadrons associated with electrons and muons strengthens the hypothesis that the inclusive electron and muon sample is enriched of $b \rightarrow cl\bar{\nu}$ decays. . . . . | 78 |
| 6.2  | Definition of signed impact parameter, $\bar{D}$ . . . . .   | 81 |

|      |   |     |
|------|---|-----|
| 6.3  | Signed impact parameter ( $\overline{D}(\mu)$ ) distributions for $b \rightarrow \mu X$ decays in the Montecarlo before (a) and after (b) full CDF simulation; $\overline{D}(\mu)$ for $c \rightarrow \mu X$ decays in the Montecarlo after simulation (c). . . . .   | 83  |
| 6.4  | Various signed impact parameter distributions used to model decays in flight. . . . .   | 84  |
| 6.5  | Muon signed impact parameter distribution in the data compared to those expected for $b$ decays, $c$ decays, and decays in flight (DIF Decay Model, dashed histogram). . . . .  | 86  |
| 6.6  | Muon signed impact parameter distribution in the data compared to the sum of those expected for $b$ decays, $c$ decays, and decays in flight . . . . .  | 87  |
| 6.7  | $E_t$ spectrum of muon ( <i>away</i> ) jets contained in the SVX fiducial volume before $b$ tagging in figure a (figure c) and after $b$ tagging in figure b (figure d). The background subtraction and normalization by $F_b(\mu)$ ( $F_b(\textit{away})$ ) has not been performed. . . . .  | 91  |
| 6.8  | $E_t$ spectrum of muon ( <i>away</i> ) jets with $\geq 2$ well measured SVX tracks before $b$ tagging in figure a (figure c) and after $b$ tagging in figure b (figure d). The background subtraction and normalization by $F_b(\mu)$ ( $F_b(\textit{away})$ ) has not been performed. . . . .  | 92  |
| 6.9  | Proper lifetime distribution of $b$ tags associated with SVX-fiducial muon jets. The exponential fit gives $c\tau_{btag} = 444 \mu\text{m}$ , in agreement with the value expected for $b$ hadrons. . . . .   | 95  |
| 6.10 | Proper lifetime distribution of $b$ tags associated with SVX-fiducial recoil jets. The exponential fit gives $c\tau_{btag} = 461 \mu\text{m}$ , in agreement with the value expected for $b$ hadrons. . . . .   | 96  |
| 6.11 | $E_t$ spectrum of inclusive jets (“Jet_50” sample) before $b$ tagging (a), $E_t$ spectrum of positive $b$ -tagged jets (b), $E_t$ spectrum of negative $b$ -tagged jets (c). . . . .  | 99  |
| 6.12 | Proper lifetime distribution of negative $b$ tags ( $c\tau < 0$ ) and positive $b$ tags ( $c\tau > 0$ ). The dashed histogram is obtained by summing: (1) the $c\tau$ distribution of negative $b$ tags (which models fake SV’s) and (2) an exponential distribution with slope equal to $c\tau_{btag}(\textit{away}) = 461\mu\text{m}$ , as measured for recoil jets in the inclusive muon sample, normalized by the difference of the total number of positive $b$ tags and the total number of negative $b$ tags (which models the HF component). The good agreement between data points with $c\tau > 0$ and the histogram demonstrates that the actual composition of $b$ -tagged inclusive jets in terms of fake SV’s and HF’s is understood correctly. . . . . | 100 |
| 6.13 | Fraction of inclusive fiducial jets associated with positive $b$ tags as a function of the number of SVX tracks in the jet, in three different jet $E_t$ bins. . . . .  | 102 |
| 6.14 | Fraction of inclusive fiducial jets associated with negative $b$ tags as a function of the number of SVX tracks in the jet, in three different jet $E_t$ bins. . . . .  | 103 |
| 6.15 | Inclusive muon sample: the top (bottom) plot shows the $E_t(\mu - jet)$ ( $E_t(\mu - jet) - E_t(\textit{away jet})$ ) spectrum for events with double $b$ tags (data, fiducial jets). . . . .   | 109 |
| 6.16 | Inclusive muon sample: the top (bottom) plot shows the $E_t(\mu - jet)$ ( $E_t(\mu - jet) - E_t(\textit{away jet})$ ) spectrum for events with double $b$ tags (ISAJET, fiducial jets). . . . .   | 110 |

|     |   |     |
|-----|---|-----|
| 7.1 | $E_t$ distribution of $t\bar{t}$ events of $M_{top} = 160 \text{ GeV}/c^2$ (ISAJET Montecarlo) selected to have an electron of $E_t(e) > 20 \text{ GeV}$ ; the efficiency of the $E_t > 20 \text{ GeV}$ cut is $\sim 90\%$ , almost insensitive to $M_{top}$ in the range 120-180 $\text{GeV}/c^2$ . . . . .  | 114 |
| 7.2 | Evidence for $W \rightarrow e\bar{\nu}_e, \mu\bar{\nu}_\mu$ decays confirmed by the transverse mass distribution of $(e, \cancel{E}_t)$ and $(\mu, \cancel{E}_t)$ pairs. . . . .  | 116 |
| 7.3 | Transverse mass distribution of $(e, \cancel{E}_t)$ and $(\mu, \cancel{E}_t)$ pairs in the final $W + \geq 3$ jets sample, compared to the predictions obtained using: (1) the VECBOS Montecarlo program (generator of $p\bar{p} \rightarrow W + \text{multijet events}$ ) + CDF simulation; (2) the ISAJET Montecarlo program (generator of $p\bar{p} \rightarrow t\bar{t}$ events) + CDF simulation. . . . .                          | 117 |
| 8.1 | Distribution of the number of jets in $t\bar{t}$ ISAJET Montecarlo events of mass $M_{top} = 140 \text{ GeV}/c^2$ (solid histogram) and $M_{top} = 160 \text{ GeV}/c^2$ (dashed histogram). . . . .   | 121 |
| 8.2 | Comparison of the $W \rightarrow e/\mu + \text{multijet}$ cross sections observed in the data, compared to the VECBOS prediction ( $W + \text{multijet background}$ ) and to the ISAJET prediction ( $t\bar{t} \rightarrow W + \text{multijet signal}$ ), as a function of the number of jets and of the top-quark mass. . . . .  | 122 |
| 8.3 | Display of the <i>electron + multijet</i> candidate event in the impact parameter-azimuth ( $D - \phi$ ) plane, in which each track is a point. The open circles indicate tracks which do not pass the quality cuts. $D$ is measured in cm, $\phi$ is measured in radians. The $D - \phi$ algorithm tags the 4 tracks with large value of the impact parameter around $\phi = 1$ as a potential $b$ -hadron decay. . . . .              | 123 |
| 8.4 | <i>Electron + multijet</i> candidate event: $R - \phi$ display, with the tracks crossing the four SVX layers; the energy deposited in the central calorimeter is shown outside the SVX; the primary vertex coordinates w.r.t. the CDF reference system are shown on the top right corner. The azimuth is counted counter-clockwise, $\phi = 0$ being the horizontal semi-axis $x > 0$ at the right of the event primary vertex. . . . . | 124 |
| 8.5 | <i>Electron + multijet</i> candidate event: magnification of the $R - \phi$ region of radius = 2 cm around the primary vertex. The coordinates of the secondary vertex associated to the $b$ tag are shown in the top right corner. . . . .   | 125 |
| 8.6 | Proper lifetime distribution, $c\tau$ , of positive $b$ tags in the $W + \geq 1$ jets sample; the dashed histogram represents the $b$ tags in $W + \geq 3$ jets. The data correspond to a statistics of $21.4 \text{ pb}^{-1}$ . . . . .  | 126 |
| 8.7 | Number of events with positive - negative $b$ tags (+DL excess after tagging) as a function of the number of jets for the inclusive, generic jet sample ("Jet.50" sample) and for a HERWIG Montecarlo sample of jets containing heavy flavors. . . . .  | 129 |
| 8.8 | Number of jets in the electron/muon + multijet event sample ( $21.4 \text{ pb}^{-1}$ of data) before and after $b$ tagging. Also shown is the background calculated with the matrix $M^+$ (background from +DL) and the total expected background. . . . .  | 133 |
| 8.9 | $t\bar{t}$ production cross section measured in the data using method 1 (the most conservative) as background estimate (points) compared to the Standard Model prediction (curves). . . . .   | 137 |

# List of Tables

|     |   |    |
|-----|---|----|
| 2.1 | Fermions of the GSW model and their quantum numbers: third component of the weak isospin, $T_3$ , weak hypercharge, $Y$ , and electric charge, $Q$ , related by $Q = T_3 + \frac{1}{2}Y$ . . . . .  | 14 |
| 2.2 | Standard Model gauge bosons and their properties. . . . .   | 16 |
| 2.3 | Decay modes for a $t\bar{t}$ pair and their approximate branching ratios (to lowest order) assuming charged-current decays. The symbol $q$ stands for a light quark: $u, d, c, s$ . . . . .   | 20 |
| 3.1 | Mechanical parameters of the central tracking chamber (CTC) . . . . .   | 29 |
| 3.2 | Summary of CDF calorimeter properties. CEM (CHA), PEM (PHA) and FEM (FHA) stand respectively for central, plug, forward EM (HAD) calorimeter. Calorimeters cover $2\pi$ in azimuth and $ \eta  < 4.2$ . The symbol $\oplus$ indicates that the constant term is added in quadrature in the resolution. The energy resolutions of EM calorimeters are for incident electrons and photons, the energy resolutions of HAD calorimeters are for incident isolated pions. Energy ( $E$ ) is given in GeV; the transverse energy is $E_t = E \sin \theta$ . Thicknesses are given in radiation lengths ( $X_0$ ) for the electromagnetic (EM) calorimeters and in interaction lengths ( $\lambda_I$ ) for the hadron (HAD) calorimeters. Spatial resolution is given in $\text{cm}^2$ at 50 GeV energy. The CEM spatial resolution is measured with proportional chambers with wire and orthogonal strip readout (CES), embedded in the EM calorimeter towers at the depth of $6X_0$ . The CES chambers are useful to separate single EM showers ( $e^\pm$ and $\gamma$ ) from neutral-pion double EM showers ( $\pi^0 \rightarrow \gamma\gamma$ ). . . . . | 30 |
| 3.3 | Summary of characteristics of central muon detectors. . . . .   | 31 |
| 4.1 | Positions and dimensions of the silicon detectors. Silicon crystals are $280 \mu\text{m}$ thick. . . . .  | 39 |
| 6.1 | Inclusive muon sample: $b$ -tagging efficiencies for SVX-fiducial jets in the data and in the Montecarlo. . . . .   | 93 |
| 6.2 | Inclusive muon sample: $b$ -tagging efficiencies for jets with $\geq 2$ well measured SVX tracks in the data and in the Montecarlo. . . . .   | 94 |
| 6.3 | Fractions (in %) of fiducial jets with positive and negative $b$ tags in the inclusive jet sample ("Jet_50" sample). . . . .  | 98 |

|     |  |     |
|-----|--|-----|
| 6.4 | Comparison between the number of positive and negative $b$ tags observed in the “Jet_100”, “ $\Sigma E_t$ ” samples and the number of tags predicted using the $M^+$ and $M^-$ matrices built from the “Jet_50” sample. . . . .  | 103 |
| 6.5 | Events with a muon and an away jet: composition (and corresponding number of events, $N_i$ , with $i=1,2,3,4$ ) in terms of jets originated by $b$ quarks (“ $b$ ”) and not originated by $b$ quarks (fake or “ $f$ ”); $N_1 + N_2 + N_3 + N_4 = N$ , where $N$ is the total number of events with fiducial muon and away jets or with muon and away jets both with $\geq 2$ well measured SVX tracks. . . . . | 105 |
| 7.1 | Summary of the number of inclusive $W \rightarrow e\bar{\nu}_e, \mu\bar{\nu}_\mu$ events as a function of the number of jets in the event, $N_{jets}$ . . . . .  | 117 |
| 8.1 | Summary of the number of $W +$ multijet events before and after $b$ tagging as a function of the number of jets in the event, $N_{jets}$ . . . . .   | 119 |
| 8.2 | Summary of the number of expected background and of observed candidate $W +$ multijet events after $b$ tagging. . . . .  | 129 |
| 8.3 | Summary of scale factors between data and Montecarlo needed to estimate the $b$ -tagging efficiency for $t\bar{t} \rightarrow e/\mu + \geq 3$ jet events. . . . .  | 132 |
| 8.4 | Summary of the acceptance, of the $b$ -tagging efficiency, of the expected number of $b$ -tagged $t\bar{t} \rightarrow e/\mu + \geq 3$ jets in $21.4 \text{ pb}^{-1}$ of data; also shown is the $t\bar{t}$ production cross section predicted by the Standard Model and the observed cross section obtained by using both method 1 and method 2 as background estimates. . . . .                              | 135 |

# Chapter 1

## Introduction

This thesis completes the work done by the candidate for the 6th cycle of the Italian Doctorate School in physics at the University Pisa (academic years from November 1, 1990 to October 31, 1993). It is an experimental thesis in the field of high-energy physics and it is one of the searches for the top ( $t$ ) quark performed by the CDF experiment (“Collider Detector at Fermilab”), which studies proton-antiproton ( $p\bar{p}$ ) collisions at a center-of-mass energy ( $\sqrt{s}$ ) of 1.8 TeV. The  $p\bar{p}$  accelerator used by CDF is the TEVATRON collider of FERMILAB, located in Batavia, Illinois (USA). The thesis is based on  $(21.4 \pm 1.5) pb^{-1}$  of data collected by CDF during the 1992-93 collider run. The candidate is a member of the CDF Collaboration since the beginning of 1988.

The theoretical frame of this top-quark search is the *minimal* Standard Model (SM). Possible decays of top-antitop ( $t\bar{t}$ ) pairs in which one of the  $W$  bosons decays into an electron (or muon) and a neutrino are studied:

$$t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow l\nu + 4 \text{ jets} \quad (\text{lepton} + \text{jets or } W + \text{multijet channel}); \quad (1.1)$$

two of the final state jets are originated by bottom ( $b$ ) quarks, and contain, therefore, long-lived  $b$  hadrons ( $c\tau_B \sim 450\mu\text{m}$ ). The dominant background to this signal is given by the associated production of  $W$  bosons and hadronic jets, which, in this case, do not mainly contain  $b$  hadrons. The signal-to-background ratio (S/N) is  $< 1$  for the top-quark mass range favoured by the SM (see sections 2.1 and 8.1). To reduce the  $W + \text{multijet}$  background we try to signal  $b$  quarks ( $b$  tagging) by looking for decay vertices of  $b$  hadrons in jets. The  $b$ -tagging algorithm developed by the candidate for this top-quark search in the lepton + jets channel fully exploits the high resolution tracking provided by the Silicon Vertex detector (SVX) of CDF. The SVX, which measures the track azimuthal angle and impact parameter in the plane transverse to the  $p\bar{p}$  beams, was installed in CDF at the beginning of 1992.

The candidate has been a member of the SVX construction group since the end of 1989, when the SVX project was officially approved. He worked on tests of the SVX prototypes which used diode-emitted infrared light and pion beams of 220 GeV/ $c$  momentum, on the analysis of the results of these tests, and wrote a paper describing the prototypes' performance. Since the summer 1991, he has also been an active member of the newly formed BTAG group, studying the use of SVX information for  $b$  tagging in the lepton + jets top-quark search. His original contributions to the BTAG group analysis program were

the development of a high resolution primary vertex finder and a high S/N  $b$ -tagging algorithm, both using SVX tracks. These two algorithms are directly incorporated in two papers recently published by the CDF Collaboration reporting evidence for top-quark production in the 1992-93 CDF data [1]: (1) the results obtained with his  $b$ -tagging algorithm provide further supporting evidence for a possible  $t\bar{t}$  signal in the lepton + jets channel when an SVX-based  $b$ -tagging analysis is applied; (2) his primary vertex finder is used by top-quark searches in *all* decay channels (the lepton+jets channel, where one  $W$  boson decays to leptons, the dilepton channel, where both  $W$  bosons decay to leptons, the hadronic channel, where both  $W$  bosons decay to quarks); his high resolution primary vertex finder is essential to the lepton + jets + SVX  $b$  tagging: in fact, if the average position of the  $p\bar{p}$  beams is used instead of the primary vertex finder, the  $b$ -tagging efficiency for  $t\bar{t}$  events in this channel drops by  $\sim 20\%$ ; finally, the candidate also contributed to the development of a dedicated event display program and data storage bank which could suitably accommodate the  $b$ -tagging information provided by any of the CDF  $b$ -tagging algorithms (whether the algorithm is based on the detection of displaced  $b$  decay vertices or  $b$  semileptonic decays). During the last year of the Doctorate School program, in addition to working for the top-quark analysis, he was heavily involved in the data taking, with 4 months of on-line shifts as data ACquisition Expert (ACE). This activity, together with his heavy involvement in the SVX prototype tests, were his original contributions to the commissioning and operation of the CDF detector for the 1992-93 data taking.

The operation of the SVX and the analysis of its data have been successful, despite of the complexity of the collider environment: short time interval between  $p\bar{p}$  bunch crossings ( $3.5 \mu\text{sec}$ ), multiple  $p\bar{p}$  interactions in the same event, high multiplicity of charged particles and high radiation dose emitted close to the interaction region. In addition, the integration of the SVX in the data acquisition system was difficult due to its high number of channels ( $\sim 46,000$  SVX channels, compared to a CDF total of  $\sim 100,000$  channels without the SVX). These difficulties were splendidly solved by the SVX and BTAG groups, after several years of intensive work. Similar problems of tracking at high luminosities will be faced, on a larger scale, by future experiments at the CERN Large Hadron Collider (LHC).

The rest of this document is organized in the following way:

- chapter 2 describes the theoretical frame of this top-quark search; present limits on the top-quark mass are discussed, as well as the mechanism of  $t\bar{t}$  production and decay in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.8 \text{ TeV}$  and the previous searches undertaken by CDF;
- chapter 3 describes the upgraded CDF detector components and trigger system;
- chapter 4 describes in detail the mechanical and electrical characteristics of the SVX, the SVX tracking algorithm and some selected results showing the high spatial resolution achieved by such an algorithm;
- chapter 5 describes the original contribution of the candidate to the lepton + jets analysis: the primary vertex finding and the  $b$ -tagging algorithms;
- chapter 6 contains a full characterization of the  $b$ -tagging algorithm with Montecarlo and CDF control data samples (efficiency, fake rate, etc.);

- chapter 7 describes the selection of the  $W + \text{multijet}$  sample;
- chapter 8 describes the use of the SVX  $b$  tagging to extract a potential  $t\bar{t}$  signal from the  $W + \text{multijet}$  sample; this implies a thorough evaluation of the residual background after  $b$  tagging and the comparison of the number of observed  $b$ -tagged candidate events with the number of  $b$ -tagged events expected from SM  $t\bar{t}$  pairs in the  $W + \text{multijet}$  channel.

This work has been supported by the University of Pisa, the CDF collaboration and, above all, by the Italian *Istituto Nazionale di Fisica Nucleare* (INFN).

# Chapter 2

## The Top Quark in the Standard Model

This chapter briefly describes the theoretical frame of this top-quark search. In particular, present limits on the top-quark mass, the characteristics of top-quark production at the Tevatron and its decay modes are discussed.

### 2.1 The Standard Model

The Standard Model (SM) describes the electroweak and strong interactions of elementary particles. The gauge theory of Glashow, Salam and Weinberg (GSW) [2], based on the  $SU(2)\times U(1)$  group, is the part of the SM which describes the electroweak interactions. Quantum Chromodynamics (QCD), a gauge theory based on the  $SU(3)$  group, is the SM part which describes the strong interactions.

#### 2.1.1 Electroweak Interactions

The elementary particles of the GSW theory are:

- the gauge vector bosons:  $Z^0$ ,  $W^\pm$  and the photon,  $\gamma$ .  $Z^0$  and  $W$  carry the weak interaction and have non-zero mass, while the photon carries the electromagnetic interaction and has zero mass. Two of the most important results obtained in recent years are the discovery and mass measurement of the  $W$  and  $Z$  bosons.  $M_Z$  is measured with high precision by experiments at the LEP collider of CERN [4], while  $M_W$  is measured by experiments at the  $p\bar{p}$  colliders of CERN [5] and of FERMILAB [6].
- the spin  $\frac{1}{2}$  fundamental fermions, divided into two categories: the *quarks*, which undergo strong and electroweak interactions, and the *leptons*, involved only in electroweak interactions. Fermions are coupled to the  $SU(2)$  gauge bosons by their weak isospin,  $T$ , and to the  $U(1)$  boson, by their electric hypercharge,  $Y$ , where  $Q = T_3 + \frac{Y}{2}$  ( $Q$  is the electric charge);
- the scalar Higgs boson, which in the SM provides the gauge boson masses by means of spontaneous symmetry breaking. The Higgs boson has not been observed yet.

In the *minimal* electroweak model there are three *families* (or generations) of fermion pairs and a single physical Higgs boson, with zero spin and zero electric charge. The left-handed fermions of each family transform like doublets under SU(2) ( $T = \frac{1}{2}$ ), while the right-handed fermions transform like SU(2) singlets ( $T = 0$ ). Of the neutrinos, only left-handed states have been observed so far. Renormalizability of the theory requires an equal number of quark and lepton families. Thus, possible anomalies due to fermion loops are avoided, once the strong color charge (see below) of each quark type (“flavor”) is taken into account. Table 2.1 shows the three fermion families known today, along with their quantum numbers. Despite the SM’s outstanding success, there is only indirect evidence for the top quark,  $t$ , which is required in the SM as the weak-isospin partner of the bottom quark,  $b$ . One of the main goals of the CDF and D0 experiments at the Tevatron is the direct observation of top quark production.

| Family:  | 1  | 2  | 3  | $T_3$                           | Y                              | Q                               |
|----------|--|--|--|---------------------------------|--------------------------------|---------------------------------|
| Leptons: | $\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$ | $\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$ | $\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$ | $\frac{1}{2}$<br>$-\frac{1}{2}$ | -1<br>-1                       | 0<br>-1                         |
| Quarks:  | $\begin{pmatrix} u \\ d \end{pmatrix}_L$     | $\begin{pmatrix} c \\ s \end{pmatrix}_L$         | $\begin{pmatrix} t? \\ b \end{pmatrix}_L$          | $\frac{1}{2}$<br>$-\frac{1}{2}$ | $\frac{1}{3}$<br>$\frac{1}{3}$ | $\frac{2}{3}$<br>$-\frac{1}{3}$ |

Table 2.1: Fermions of the GSW model and their quantum numbers: third component of the weak isospin,  $T_3$ , weak hypercharge,  $Y$ , and electric charge,  $Q$ , related by  $Q = T_3 + \frac{1}{2}Y$ .

The existence of three fundamental fermion families is experimentally confirmed by the precision measurements of LEP experiments: the shape of the  $e^+e^-$  annihilation cross section around  $\sqrt{s} = M_Z$  agrees with the GSW prediction for three light neutrino flavors [7]. Indirect experimental evidence for the top-quark existence, as bottom-quark weak-isospin partner, comes from:

- 1) the suppression of flavor changing neutral currents (FCNC) in the decays of neutral  $B$  mesons. According to the GSW model,  $B^0 \rightarrow l^+l^-$  decays are allowed only by radiative corrections, and have, therefore, very low branching ratios (BR):  $\text{BR}(B^0 \rightarrow l^+l^-) < 10^{-10}$  [8, 9]. The experimental 90 % CL upper limits [9],  $\text{BR}(B^0 \rightarrow \mu^+\mu^-) < 5.9 \times 10^{-6}$ ,  $\text{BR}(B^0 \rightarrow e^+e^-) < 5.9 \times 10^{-6}$ , exceed the SM predictions. Without a top quark, i.e. if the  $b$  were member of an SU(2) left-handed singlet, it has been shown [10] that the BR should be  $\sim 1.3 \times 10^{-2}$ : the above experimental limits rule out the five-quark model.
- 2) the measured value of  $e^+e^- \rightarrow b\bar{b}$  [11] forward-backward asymmetry. This asymmetry is sensitive to the  $b$ -quark coupling to the  $Z$  boson and, therefore, to the  $b$ -quark weak isospin,  $T(b)$ . Measurements at the LEP, PETRA/PEP and TRISTAN  $e^+e^-$  colliders give [12]  $T_3(b_L) = 0.490^{+0.015}_{-0.012}$  for the left-handed state and  $T_3(b_R) = 0.028 \pm -0.056$  for the right-handed state. Thus, the  $b$ -quark is a member of a weak isospin doublet;

- 3) (i) the measured value of neutral  $B$  meson mixing and (ii) the results of LEP global fits to precision electroweak measurements. These fits are sensitive to the existence of the top quark and to the value of its mass ( $M_{top}$ ), through electroweak radiative corrections having the top quark as fermion propagator; limits on  $M_{top}$  from these two classes of measurements are discussed in section 2.2.

The quark states shown in table 2.1 are weak interaction eigenstates, and they are different from quark mass eigenstates of definite flavor. The mixing matrix relating the two bases of quark states has been introduced for three fermion families by Kobayashi and Maskawa [13] in 1973. This  $3 \times 3$  matrix generalizes the case of two fermion families, in which the matrix is function of a single parameter, the Cabibbo angle [14]. By convention, charge  $+\frac{2}{3}$  quarks remain unchanged and the whole mixing is given by a  $3 \times 3$  unitary matrix,  $V$ , operating on charge  $-\frac{1}{3}$  quarks:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \quad (2.1)$$

where the weak interaction eigenstates are indicated with primed symbols. This mixing is not observed in the leptonic sector, consistently with the hypothesis that neutrinos have zero mass. The values of individual elements of the Cabibbo - Kobayashi - Maskawa (CKM) matrix can in principle be determined from weak decays of hadrons containing the relevant quarks or from neutrino deep inelastic scattering (DIS) experiments. From these measurements, from the constraints given by the unitarity of  $V$ , and assuming only three quark families, one finds the following 90% CL limits on the matrix elements involving the top quark (see [7], III.65):

$$0.03 < V_{td} < 0.018, \quad 0.030 < V_{ts} < 0.054, \quad 0.9985 < V_{tb} < 0.9995. \quad (2.2)$$

These values imply that, if the top quark exists, it decays via weak interaction to a  $W$  boson and a  $b$  quark ( $t \rightarrow Wb$ ), with  $\text{BR}(t \rightarrow Wb) \simeq 1$ . The  $W$  boson is virtual or real depending on the value of  $M_{top}$ . This prediction can be radically different if more than three families are allowed. In this case, not excluded by present data, Ref. [7] indicates that at 90% CL  $0 < V_{tb} < 0.9995$  (see [7], III.66), which might have an effect on  $\text{BR}(t \rightarrow Wb)$ .

### 2.1.2 Strong Interactions

The fundamental particles of QCD are *gluons* and *quarks*. The gluons are the eight  $SU(3)$  gauge bosons, with spin = 1, zero mass and zero electric charge, carrying the strong interaction. Quarks of any flavor are coupled to gluons by their three strong charges ("colors"), corresponding to the three-state base of the  $SU(3)$  fundamental representation. Since  $SU(3)$  is a non-abelian group, gluons are colored and can interact among themselves. The gluon-gluon interactions determine the behavior the renormalized strong coupling constant,  $\alpha_s(Q^2)$ , as a function of the  $Q^2$  scale.  $\alpha_s$  is conventionally expressed at the leading order in  $1/\ln(Q^2)$  as:

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 22n_f) \ln(Q^2/\Lambda_{n_f}^2)}, \quad (2.3)$$

where  $n_f$  is the number of quark flavors with mass  $< Q^2$  and  $\Lambda_{n_f}$  is a fundamental dimensional parameter to be determined from experiment. This expression of  $\alpha_s$  illustrates the principle of *asymptotic freedom*: the QCD coupling constant is small only for  $Q \gg \Lambda_{n_f}$ , and it is only in this domain that perturbative QCD can provide reliable predictions (for example cross sections) or precision tests like those performed in Quantum Electrodynamics (QED). In the  $\overline{MS}$  [15] renormalization scheme the average of LEP and DIS measurements gives  $\alpha_s(M_z) = 0.1134 \pm 0.0035$ , which corresponds to  $\Lambda_4 = 260_{-46}^{+54}$  MeV and  $\Lambda_5 = 175_{-34}^{+41}$  MeV [7].

Perturbative QCD can be tested in many ways in high-energy hadron colliders and over many orders of magnitude in cross section [16]. Given the value of  $\alpha_s$ , quantitative tests are useful only if the process under study has been calculated in perturbative QCD beyond leading order. The production of hadrons with large transverse momentum in hadron-hadron collisions provides a direct probe of the scattering of quarks and gluons:  $qq \rightarrow qq$ ,  $qg \rightarrow qg$ ,  $gg \rightarrow gg$ , etc. For example, the Fermilab Tevatron collider (and the CERN  $SppS$  before) provides a  $p\bar{p}$  center-of-mass energy sufficiently high that these processes can be unambiguously identified in two-jet production at large transverse momentum. Recent higher-order QCD calculations of inclusive jet production rates and shapes around pseudorapidity  $\eta = 0$  as a function of jet  $P_t$  [17] are in impressive agreement with experimental data [18]. These predictions combine the QCD *parton* scattering amplitudes with the proton *structure functions*; the partons are the proton constituents, i.e., quarks and gluons, while the structure functions are distributions which define the probability that partons inside the proton carry a fraction  $x$  of the proton momentum (“Parton Distribution Functions”, PDF). Other typical examples are the measurements of the two-jet production cross section as a function of the two-jet invariant mass ( $M_{jj}$ ) [19], of the jet angular distributions [20], of the  $b$ -quark production cross section [21] and of color coherence effects [23]. The study of the  $M_{jj}$  spectrum is particularly interesting; in fact, at CDF it allowed to impose limits on the production of possible new heavy particles decaying to two jets and on the energy scale of a possible quark compositeness [22]. These are not high-precision measurement yet, but they clearly indicate that QCD provides a reliable prediction of quark and gluon interactions in  $p\bar{p}$  colliders, in a wide range of kinematic variables ( $\sqrt{s}$ ,  $P_t(\text{jet})$ ,  $M_{jj}$ ,  $\eta(\text{jet})$ , etc.). The good agreement between theoretical predictions and experimental data justify the use of QCD to study top-quark production at the Tevatron.

Finally, table 2.2 shows the properties of the  $SU(3) \times SU(2) \times U(1)$  gauge bosons of the Standard Model.

| Gauge boson                 | Spin | Electric Charge | Color | Mass ( $GeV/c^2$ )     |
|-----------------------------|------|-----------------|-------|------------------------|
| $\gamma$                    | 1    | 0               | no    | 0                      |
| $Z^0$                       | 1    | 0               | no    | $91.187 \pm 0.007$ [7] |
| $W^\pm$                     | 1    | $\pm 1$         | no    | $80.22 \pm 0.26$ [7]   |
| $g_i$ ( $i = 1, \dots, 8$ ) | 1    | 0               | yes   | 0                      |

Table 2.2: Standard Model gauge bosons and their properties.

## 2.2 Limits on the Top-Quark Mass

In recent years, many experiments have performed direct searches for the top quark. Searches at the CERN  $p\bar{p}$  collider ( $\sqrt{s} = 0.63$  TeV) by the UA1 and UA2 Collaborations have provided, respectively, the 95%CL lower limits  $M_{top} > 60$  GeV/ $c^2$  [24] and  $M_{top} > 69$  GeV/ $c^2$  [25]. These limits assume a SM top-quark decay via weak charged current to a virtual  $W$  boson and a  $b$  quark. Direct searches in  $e^+e^-$  collisions have given the limit  $M_{top} > 46$  GeV/ $c^2$ , independent of top decay modes [26]. Another decay-mode-independent limit,  $M_{top} > 44$  GeV/ $c^2$  at 95% CL, has been obtained by CDF from the value of the  $W$ -boson width ( $\Gamma_W$ ) extracted from the ratio of the  $W$  and  $Z$  production cross sections [27]. Combining the estimates of  $\Gamma_W$  by CDF [27], UA1 [29] and UA2 [28], one gets the limit  $M_{top} > 55$  GeV/ $c^2$  at 95 % CL [30]. Recently, CDF raised this limit to  $M_{top} > 62$  GeV/ $c^2$  [31]. These limits are independent of top decay modes, because they rely on the fact that  $W \rightarrow t\bar{b}$  decays, if possible, increase the  $W$  width. In particular, they are valid regardless of the structure of the Higgs sector, i.e., even in the presence of top decays not predicted by the minimal SM, like  $t \rightarrow H^+b$  (see section 2.5).

Since in the SM the top quark is the weak-isospin partner of the  $b$  quark, its existence (and the value of its mass) influences  $b$ -physics processes involving loops having top quarks as fermion propagators. Typical examples are the weak decays of  $b$  hadrons via “penguin” diagrams, and  $B^0\bar{B}^0$  mixing. A lower limit on  $M_{top}$  of about 50 GeV/ $c^2$  is derived from a fit of SM parameters to the measured level of  $B^0\bar{B}^0$  [32].

The top quark is also relevant to radiative corrections of several SM parameters, which are measured with high precision by the LEP experiments [4]. The leading  $M_{top}$  dependence is quadratic and global fits of these parameters to data from LEP, DIS( $\nu$ -nuclei), UA2 and CDF yield a favoured mass of [4]:

$$M_{top} = 164_{-17-21}^{+16+18} \text{ GeV}/c^2, \quad (2.4)$$

where the central value and the first uncertainty are for a Higgs boson mass,  $M_H$ , of 300 GeV/ $c^2$ , and the second uncertainty corresponds to a variation of the central value in the range  $60 < M_H < 1000$  GeV/ $c^2$ . The experimental data used in [4] to estimate  $M_{top}$  do not allow a meaningful determination of  $M_H$ , because the leading  $M_H$  dependence is logarithmic. However, with a direct measurement of  $M_{top}$  and by further reducing the uncertainty on  $M_W$ , one can impose constraints on  $M_H$ , since in the SM  $M_W$ ,  $M_{top}$  and  $M_H$  are quantities related by electroweak radiative corrections (see for example [33]). Direct searches of a neutral Higgs scalar give the limit  $M_H > 58.4$  GeV/ $c^2$  at 95% CL [34].

Direct top-quark searches have been performed by CDF, using an integrated luminosity of  $4.1$  pb $^{-1}$  of  $p\bar{p}$  collisions, collected in 1988-1989, yielding  $M_{top} > 91$  GeV/ $c^2$  at 95% CL [35, 36]. Recently, this limit has been raised to  $M_{top} > 131$  GeV/ $c^2$  by the D0 Collaboration [37]. The CDF and D0 limits exclude, in the context of the minimal SM, the existence of decays  $W \rightarrow tb$ , with a real  $W$  boson. In the following two sections, we discuss: (1) the production and decay of top quarks at Tevatron energies; (2) the CDF searches for a minimal SM top quark and new strategies developed to analyze the five-fold increased statistics integrated in the 1992-1993 data taking, with the upgraded CDF detector (the upgraded detector is described in chapter 3 and 4). Among the new strategies the topic of

this thesis is also indicated. Finally, present results of CDF searches for a non-minimal SM top quark are reported for completeness in the last section of this chapter (section 2.5).

## 2.3 Top-Quark Production and Decay in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV

CDF is in a privileged position to search for the top quark, because the Tevatron  $p\bar{p}$  collider provides the highest center-of-mass energy available to date ( $\sqrt{s} = 1.8$  TeV) and, therefore, the highest available top-quark production cross section. A  $160 \text{ GeV}/c^2$  mass top quark should yield  $2.6 \pm 1.0$  events in the decay mode exploited by this thesis (see section 8.6).

The SM predicts that at Tevatron energies, for  $M_{top} < 250 \text{ GeV}/c^2$ , top quarks are mainly pair-produced, via strong interactions:

$$p\bar{p} \rightarrow t\bar{t}. \quad (2.5)$$

The lowest order ( $\alpha_s^2$ ) pair production is given at parton level by quark-antiquark annihilation and gluon-gluon fusion processes:

$$q\bar{q} \rightarrow t\bar{t}, \quad \text{and} \quad gg \rightarrow t\bar{t}, \quad (2.6)$$

which are calculable in QCD perturbation theory because  $M_{top} \gg \Lambda$ . For  $M_{top} \sim 100 \text{ GeV}/c^2$  the rates of the two processes are comparable, while for  $M_{top} > 100 \text{ GeV}/c^2$   $q\bar{q} \rightarrow t\bar{t}$  dominates, since the quark PDF is harder than for gluons.  $qg$  and  $\bar{q}g$  processes contribute at  $\alpha_s^3$ , representing  $\approx 2\%$  of the total  $t\bar{t}$  production for any value of  $M_{top}$ . For  $M_{top} \gtrsim 250 \text{ GeV}/c^2$ , predictions indicate that the dominant mechanism is the semi-weak  $W$  gluon fusion (for example  $ug \rightarrow dt\bar{b}$ ); this higher values of  $M_{top}$  are unlikely, given the experimental constraints from the LEP experiments described in the previous section, and they are outside CDF top mass reach with the current integrated luminosity of  $21.4 \text{ pb}^{-1}$  (note that  $W$  gluon fusion does not lead to an observable signal in  $100 \text{ pb}^{-1}$  samples at  $\sqrt{s} = 1.8$  TeV [38]). Thus the phenomenological parameters required to predict the top-quark production cross section,  $\sigma_{t\bar{t}}$ , are the quark PDF for  $x \sim 0.1$  and the value of  $\alpha_s$ . The  $\alpha_s^3$  cross section of partonic processes has been available for some time [39, 40, 41, 42]; recently, also higher order contributions have been included [43]. The minimum value of the predicted  $\sigma_{t\bar{t}}$  accepted by theory is required for setting lower limits on  $M_{top}$  together with experimental upper limits on the cross section; on the other hand, if a top signal is found, the central value of the predicted  $\sigma_{t\bar{t}}$  can be used, together with the central value of the experimental cross section, to estimate  $M_{top}$ . Thus, it is extremely important that a reliable and precise theoretical prediction be available. The uncertainty on the prediction, that is, the difference between its upper and lower limits allowed by theory, is usually determined by varying the renormalization scale (typically from  $M/2$  to  $2M$ , where  $M = M_{top}$  or  $M = \sqrt{M_{top}^2 + P_t^2}$ ) and using different PDF parametrizations. Including the  $\alpha_s^3$  corrections leads to a reduction in theoretical uncertainty, when compared to tree graph level calculations. A partial resummation of higher order terms indicates a further 10% increase in  $\sigma_{t\bar{t}}$ . The most recent theoretical prediction [43] quotes a  $\sigma_{t\bar{t}}$  fractional uncertainty varying from  ${}_{-13\%}^{+28\%}$  for  $M_{top} = 120 \text{ GeV}/c^2$  to  ${}_{-10\%}^{+14\%}$  for  $M_{top} = 180$

GeV/ $c^2$ ; note that this uncertainty is due only to variation of a parameter related to the soft gluon emission, which is the higher order contribution included in this calculation; it does not include the variation of the renormalization scale and of the PDF parametrizations. For this reason, other references indicate that including all known effects, a reasonable estimate of the theoretical uncertainty on  $\sigma_{t\bar{t}}$  for high values of  $M_{top}$  is  $\pm 20\%$  (see for example [38]).

The minimal SM predicts that  $t\bar{t}$  pairs decay as:

$$t\bar{t} \rightarrow W^+bW^-\bar{b}, \quad (2.7)$$

where the two  $W$  bosons are real, given the D0 limit  $M_{top} > 131$  GeV/ $c^2$ , and decay into a charged lepton and its neutrino or a pair of light quark-antiquark (see figure 2.1).

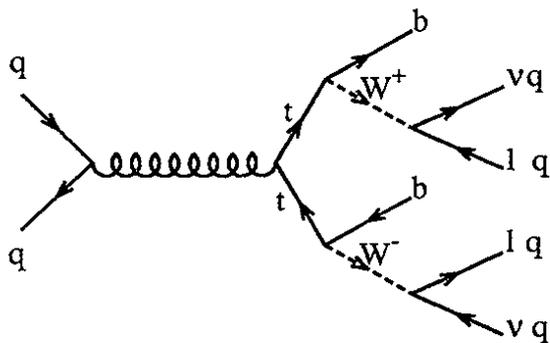


Figure 2.1: Tree level diagram for  $t\bar{t}$  pair production via  $q\bar{q}$  annihilation, followed by the Standard Model decay chain.

Different top-quark decays, namely  $t \rightarrow H^+b$ , predicted by plausible theories beyond the minimal SM are discussed in section 2.5.

## 2.4 CDF Top-Quark Search Program Within the Minimal Standard Model

CDF searches for the top quark naturally divide into three categories, depending on the decay modes of the  $W^+W^-$  boson pair. If only one  $W$  decays leptonically, the event contains one high  $P_t$  lepton, one high  $P_t$  neutrino, and four jets (*lepton+jets* or  $W + multijet$  mode, BR  $\simeq 15\%$ /lepton). If both  $W$  bosons decay leptonically, the event contains two high  $P_t$  leptons, two high  $P_t$  neutrinos and two jets (*dilepton* or  $WW$  mode, BR  $\simeq 5\%$ ). If none of the  $W$ 's decays leptonically, the event contains six jets (*all hadronic* mode, BR  $\simeq 44\%$ ). Table 2.3 shows the list of possible  $t\bar{t}$  decays and the corresponding BR's.

Note that, since  $t\bar{t}$  events result from high energy processes, they may contain additional jets from initial and final state gluon radiation, besides those from top decays:  $\geq 6$  jets in the all hadronic mode,  $\geq 4$  jets in the lepton+jets mode, and  $\geq 2$  jets in the dilepton mode.

| Decay mode                                      | Branching Ratio |
|---|-----------------|
| $t\bar{t} \rightarrow q\bar{q}bq\bar{q}\bar{b}$ | 36/81           |
| $t\bar{t} \rightarrow q\bar{q}be\nu\bar{b}$     | 12/81           |
| $t\bar{t} \rightarrow q\bar{q}b\mu\nu\bar{b}$   | 12/81           |
| $t\bar{t} \rightarrow q\bar{q}b\tau\nu\bar{b}$  | 12/81           |
| $t\bar{t} \rightarrow e\nu b\mu\nu\bar{b}$      | 2/81            |
| $t\bar{t} \rightarrow e\nu b\tau\nu\bar{b}$     | 2/81            |
| $t\bar{t} \rightarrow \mu\nu b\tau\nu\bar{b}$   | 2/81            |
| $t\bar{t} \rightarrow e\nu be\nu\bar{b}$        | 1/81            |
| $t\bar{t} \rightarrow \mu\nu b\mu\nu\bar{b}$    | 1/81            |
| $t\bar{t} \rightarrow \tau\nu b\tau\nu\bar{b}$  | 1/81            |

Table 2.3: Decay modes for a  $t\bar{t}$  pair and their approximate branching ratios (to lowest order) assuming charged-current decays. The symbol  $q$  stands for a light quark:  $u, d, c, s$ .

**The all hadronic mode:** this mode has the highest BR ( $\simeq 44\%$ ), but it suffers from a huge background due to QCD processes producing multijet final states. This background can be reduced (and this mode be reasonably used) only by tagging  $b$  quarks that should be present in top events. For  $M_{top} = 160 \text{ GeV}/c^2$ , by using just kinematic cuts and no  $b$  tagging  $S/N \sim 1/100$  can be obtained, while  $b$  tagging allows for  $S/N \sim 1/10$  to be achieved. The advantage is that in this mode there are no neutrinos in the final state and that by measuring the 4-vectors of the six jets one can in principle determine  $M_{top}$ . However, the jet energy resolution may not be good enough to isolate an invariant mass peak around  $M_{top}$ . In the case of kinematic fits to the  $t\bar{t}$  hypothesis, tagging of  $b$  quarks reduces the number of possible combinations of three-jet pairs coming from a top (or anti-top) decay. A top search in the all hadronic mode has been undertaken for the first time by CDF using the 1992-1993 data [44].

**Lepton+jets mode:** a much better signal-to-noise ratio can be achieved in this mode, compared to the all hadronic mode, thanks to CDF good lepton identification capabilities; this is even more true for the dilepton channel. The presence of large missing transverse energy ( $E_t$ , see chapter 7) can be used to signal high  $P_t$  neutrinos. Considering only electrons and muons ( $\tau$  detection is difficult), this mode has approximately a 30% BR, six times higher than the dilepton channel. However, the associated QCD production of a  $W$  boson and jets produces a significant background. The  $W +$  multijet production has already been studied by CDF in [45]. Tree level calculations of this process [48] yield cross sections comparable or higher to  $t\bar{t}$  production rates. Figures 2.1 and 2.2 show, respectively, two diagrams for signal and background processes. Figure 2.3 shows the tree level cross sections for  $t\bar{t}$ ,  $W + 3$  jets and  $W + 4$  jets [38].

The  $b$ -tagging requirement provides a strong background rejection, since all top events have a  $b\bar{b}$  pair, while  $W$ +jet events are depleted of  $b$ -quark jets [49]. The usefulness of muon identification, as a tag of the semileptonic decay of  $b$  quarks in  $t\bar{t} \rightarrow W^+bW^-\bar{b}$  decays,

has been demonstrated by CDF in [35, 36], using the 1988-1989 data. In those references, muons with  $2 < P_t < 15$  GeV/c were searched for outside an  $\eta - \phi$  cone of radius=0.5 around the two highest  $E_t$  jets, in order to minimize the background from pion and kaon decays in flight and from hadron punchthrough from the calorimeter into the muon chambers. Despite its low efficiency (4.5%), the low- $P_t$  muon search, combined with the dilepton channel results, allowed a more stringent lower limit on  $M_{top}$  (from 85 GeV/c<sup>2</sup> to 91 GeV/c<sup>2</sup>). This technique has been furtherly developed and it is adopted in the analysis of the present data. The following improvements have been introduced: (i) the identification of electrons (besides muons) of  $P_t$  down to 2 GeV/c, using, in addition, the information of a pre-radiator which samples the early development of electromagnetic showers and a  $dE/dx$  measurement, which were not previously available (see sections 3.2.1 and 3.2.2); (ii) better S/N for muon identification, due to the substantially upgraded muon detector (see section 3.2.3); (iii) the search for low- $P_t$  electrons and muons also inside jets.

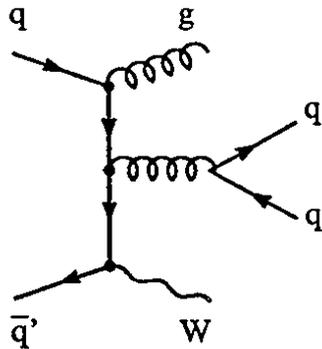


Figure 2.2: An example of  $W +$  multijet production.

Other, new  $b$ -tagging techniques take advantage of the high resolution tracking provided by the new Silicon microstrip Vertex detector (SVX, see chapter 4) of CDF, to search inside jets for secondary decay vertices, whose distance from the  $p\bar{p}$  interaction vertex is consistent with the average  $b$ -hadron lifetime ( $\tau_B \sim 1.46 \pm 0.08$  psec,  $c\tau_B \sim 446 \pm 26$   $\mu\text{m}$ , see [50]). These two classes of  $b$ -tagging algorithms, the low  $P_t$  lepton identification (“soft lepton tagging” or SLT [52]) and the secondary vertex identification (“SVX tagging” or simply SVX, see chapter 5 and [53]), are the standard techniques adopted by CDF in the top-quark search in the lepton+jets channel. The  $\tau$ -lepton identification in the single lepton and dilepton modes can profit from both the SLT and SVX tagging algorithms, since  $\text{BR}(\tau \rightarrow e/\mu\nu\nu) \simeq 18\%$  and since the  $\tau$  lifetime is  $0.305 \pm 0.006$  psec or  $c\tau_\tau \sim 100\mu\text{m}$  ([7], VI.21). Top-quark searches in decay channels involving  $\tau$ 's have been initiated by CDF using the 1988-1989 data, thus covering all SM top decay modes listed in table 2.3. Finally,  $b$  tagging is extremely useful when performing kinematic fits to  $t\bar{t}$  candidate events with  $\geq 4$  jets which allow a direct measurement of  $M_{top}$ , since  $b$  tagging reduces the number of possible triplet combinations of jets, leptons and  $E_t$  compatible with  $t \rightarrow Wb$  decays.

**The dilepton mode:** this mode has a background level lower than the previous one;

its main limitation is the total detection efficiency (including BR) versus the value of  $M_{top}$ , that is, the total number of detectable  $t\bar{t}$  events.  $b$  tagging in this decay channel is very critical, because it may reduce the number of candidate events (efficiency  $\epsilon_{tag} \sim 20\%$ , see chapter 6). On the other hand, if the number of candidates is sufficiently high ( $M_{top}$  not too large),  $b$  tagging can improve S/N. In addition, if a significant  $t\bar{t}$  signal is observed, from the measurement of the transverse energy of  $b$ -tagged jets,  $E_t(b-jet)$ , one could determine the mass of the parent top quarks, because the average value of  $E_t(b-jet)$  is related to  $M_{top}/2$  [54]. Exploiting the kinematics of dilepton candidate events to measure  $M_{top}$ , is made difficult by the presence of two high  $P_t$  neutrinos. However, the possibility is actively pursued by CDF and it looks promising.

In summary,  $b$  tagging is crucial in every top-quark search in order to: (1) reduce the background; (2) measure directly  $M_{top}$ , using the final state kinematics.

The work described in this thesis is a minimal SM top-quark search in the  $e/\mu +$  multijet + SVX  $b$ -tagging mode.

## 2.5 CDF Top-Quark Searches Beyond the Minimal Standard Model

Some extensions of the minimal SM, like the two-Higgs doublet (THD) model and the minimal supersymmetric extension of the SM (MSSM), predict the existence of a charged scalar Higgs boson,  $H^+$ , with mass  $M_{H^+}$  such that the decay  $t \rightarrow H^+b$  is kinematically allowed. In this case, if  $M_{H^+} + M_b < M_{top} < M_W + M_b$ ,  $t \rightarrow H^+b$  could dominate over  $t \rightarrow Wb$  [55], thus reducing the sensitivity of the CDF search in the latter channel. The eventual  $H^+$  boson should decay to the heaviest available leptons and quarks ( $c\bar{s}$  and  $\tau\nu_\tau$ ) with relative BR unconstrained by the theory. A lower limit on  $M_{H^+}$  independent of the explicit structure of the Higgs sector of the SM, comes from direct searches performed by the LEP experiments [58]:  $M_{H^+} > 45 \text{ GeV}/c^2$  at 95 % CL. A lower limit on  $M_{top}$  independent of the Higgs structure comes from the estimate of  $\Gamma_W$  recently obtained by CDF:  $M_{top} > 62 \text{ GeV}/c^2$  at 95 % CL [31]. Therefore, current limits on  $M_{H^+}$  and  $M_{top}$  do not exclude the existence of  $t \rightarrow H^+b$  decays. CDF has published a search for the decay chain  $t \rightarrow H^+b$ ,  $H^+ \rightarrow \tau^+\nu_\tau$ ,  $\tau \rightarrow hadrons$ , using the  $4.1 \text{ pb}^{-1}$  collected in 1988-1989 [56], based on the hypothesis that  $\text{BR}(H^+ \rightarrow \tau^+\nu_\tau) \geq 0.5$ . This search finds no evidence for top decays to a charged Higgs boson and leads to the exclusion of most of the region of the  $(M_{top}, M_{H^+})$  plane where the  $t \rightarrow Wb$  decays could be suppressed ( $M_{top} > 62 \text{ GeV}/c^2$ ,  $M_{H^+} > 45 \text{ GeV}/c^2$ ,  $M_{H^+} + M_b < M_{top} < M_W + M_b$ ). Recently, CDF completed (and submitted to Phys. Rev. Lett.) also a search for  $t \rightarrow H^+b$ , followed by  $H^+ \rightarrow \tau^+\nu_\tau$ ,  $\tau \rightarrow e/\mu X$ , using the  $21.4 \text{ pb}^{-1}$  collected in 1992-1993 [57]. This search also finds no evidence for  $t \rightarrow H^+b$ , and extends the regions of the  $(M_{top}, M_{H^+})$  plane excluded by ref.[56], as a function of  $\text{BR}(H^+ \rightarrow \tau^+\nu_\tau) \geq 0.5$ . The outcome of both searches is also interpreted in the context of the THD model (see [56], [57]).

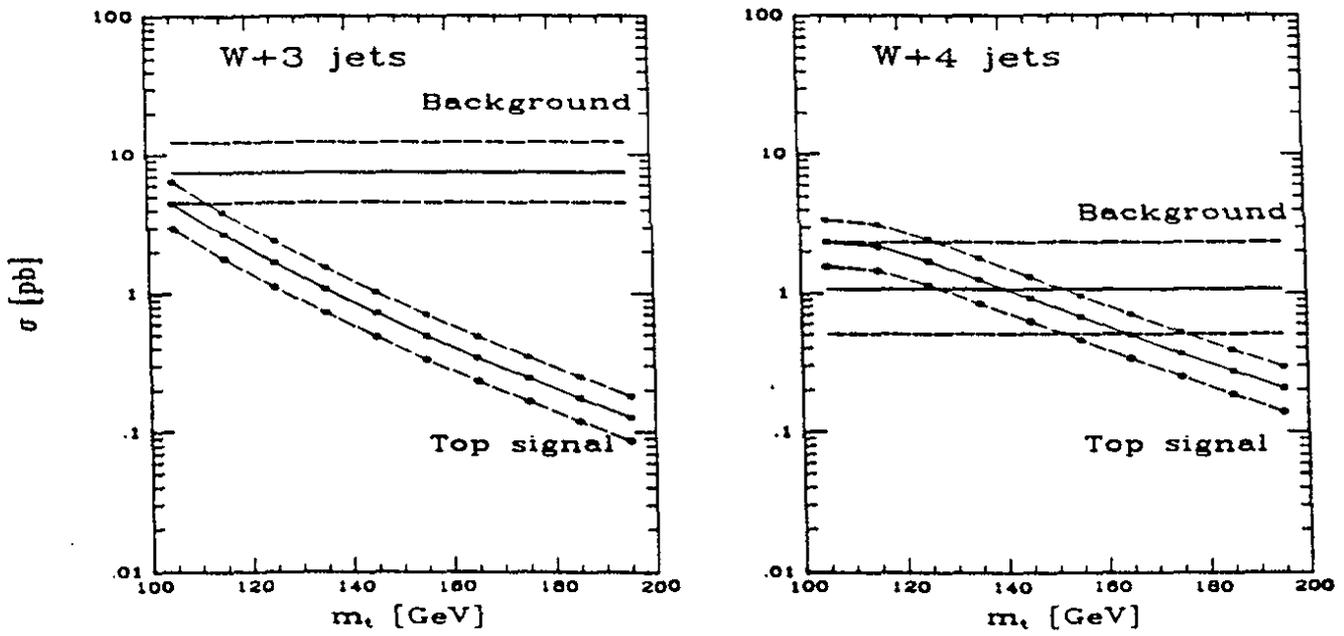


Figure 2.3: The Standard Model tree level prediction of the cross section of  $t\bar{t}$ ,  $W + 3$  jet and  $W + 4$  jet production.

# Chapter 3

## The CDF Experiment

This chapter briefly describes the Fermilab collider and the CDF experiment, while the new silicon vertex detector that was installed in CDF for the 1992-93 data taking is described in greater detail in the next chapter.

### 3.1 The Tevatron Collider of Fermilab

The Fermilab laboratory located in Batavia, Illinois (USA), has the proton-antiproton collider with the world-highest center of mass energy ( $\sqrt{s} = 1.8$  TeV), the TEVATRON. During the years 1992-93, the Tevatron was operated with 6 proton and 6 antiproton bunches, which collided in the two interaction regions where the CDF and D0 experiments are located. Each proton (antiproton) bunch contained  $N_p \sim 12 \times 10^{10}$  ( $N_{\bar{p}} \sim 4 \times 10^{10}$ ) particles and circulated with a frequency  $f_0 \sim 50$  KHz. The time interval between two consecutive bunch crossings at the interaction regions was  $\sim 3.5\mu\text{sec}$ .

An important parameter of every accelerator is the instantaneous Luminosity,  $\mathcal{L}$ , defined by the relationship:

$$N_e = \mathcal{L}\sigma, \quad (3.1)$$

where  $N_e$  is the number of events per second for a process whose cross section is  $\sigma$  (in  $\text{cm}^2$ ). Using the accelerator characteristics  $\mathcal{L}$  can also be determined by the equation:

$$\mathcal{L} = \frac{N_p N_{\bar{p}} B f_0}{A_{eff}}, \quad (3.2)$$

where  $B = 6$  is the number of bunches in each of the two beams and  $A_{eff}$  is the effective area of overlap of the beams in the plane transverse to their direction of motion. During the last data taking the average and maximum values of  $\mathcal{L}$  were approximately  $3 \times 10^{30} \text{cm}^{-2} \text{sec}^{-1}$  and  $9 \times 10^{30} \text{cm}^{-2} \text{sec}^{-1}$ . The integrated luminosity delivered by the Tevatron was  $\sim 30.2 \text{pb}^{-1}$  and CDF collected  $21.4 \text{pb}^{-1}$ , which corresponds to a data acquisition efficiency of about 70%. A total of about 15.5 million events were recorded to tape. Finally, we define two terms of common use throughout this thesis are defined here: "store" and "run". The word store indicates the time interval between the beginning of the collisions and the ejection of the beams from the Tevatron, which happens when the instantaneous luminosity drops to levels well below the average - and acceptable - value; during the period of the store the

data taking is divided into several parts, called runs, for ease of detector operation and data storage.

## 3.2 CDF, the Collider Detector at Fermilab

CDF is an experiment built to study high energy  $p\bar{p}$  collisions [59]. Its main goals are the identification of leptons and jets and the measurement of their energy and momentum on the largest possible solid angle. To do this, tracking chambers, sampling calorimeters and muon detectors are installed around the interaction region. A schematic view of several CDF components is shown in figure 3.1. To describe these components and the analysis based on the information which they provide, we define the coordinate systems adopted by the experiment (see figure 3.1).

**Cartesian coordinates:** the  $z$  (*longitudinal*) axis is taken along the  $p\bar{p}$  beams, with  $z > 0$  in the direction of motion of the protons, which circulate from the CDF interaction point B0, towards east; the  $z = 0$  or  $x, y$  (*transverse*) plane is the CDF median plane which goes through the nominal center of the interaction region.

**Polar coordinates:**  $\phi$  is the (*azimuthal*) angle about the  $z$  axis, starting from the  $x$  axis;  $\theta$  is the (*polar*) angle w.r.t. the  $z$  axis, starting from the  $z > 0$  semi-axis, and  $R$  is the (*radial*) distance from the  $z$  axis. The pseudorapidity, a quantity of frequent use in high energy physics, is defined as  $\eta = -\ln(\tan(\theta/2))$ .



### 3.2.1 Tracking System

CDF has four separate tracking subsystems, immersed in a 1.4 tesla solenoidal magnetic field generated by a superconducting coil whose axis is parallel to the beams. In the following we describe these subsystems, going from the interaction region to the outside.

1. The precision silicon microstrip tracking detector (SVX). Installed in 1992, this detector, located immediately outside the vacuum tube containing the beams, represents the most important CDF upgrade. It provides a single-hit resolution of  $13\ \mu\text{m}$  and an impact parameter resolution (in the transverse plane) for  $P_t > 20\ \text{GeV}/c$  of about  $17\ \mu\text{m}$ . A detailed description of the SVX is given in the following chapter.
2. The vertex time projection chamber (VTX) [60], which tracks in the region  $R < 22\ \text{cm}$  and  $|\eta| < 3.25$ . The VTX is used to measure the  $z$  coordinate of primary  $p\bar{p}$  interaction vertices and to help identifying photon conversions into  $e^+e^-$  pairs.
3. The central tracking chamber (CTC) [61]. This is a cylindrical drift chamber which fills almost completely the solenoid internal volume. It provides good spatial resolution,  $\sim 200\ \mu\text{m}$ , and good transverse momentum ( $P_t$ ) resolution,  $\sigma_{P_t}/P_t \sim 0.002 \times P_t(\text{GeV}/c)$ , in the  $|\eta| < 1.2$  region.  $P_t$  is related to the total track momentum ( $P$ ) by  $P_t = P \sin \theta$ . By requiring the tracks to emanate from the average beam line position, one finds  $\sigma_{P_t}/P_t \sim 0.0011 \times P_t(\text{GeV}/c)$ . The CTC is used also to reconstruct high  $P_t$  tracks at “trigger” level (see section 3.2.4). Note that because of the great redundancy of the measurements of its 84 sense-wire layers, the CTC is the most important CDF detector component for track pattern recognition and track separation in dense environments, like events with high energy collimated multijets. On the other hand, the SVX low redundancy (4 layers of space-point measurements) and uses only tracks pre-identified by the CTC. However, the excellent SVX spatial resolution allows for a large improvement of the impact parameter resolution and, to a lesser extent, of the  $P_t$  resolution (see section 4.4.5) for CTC tracks contained in its geometrical acceptance ( $\sim 60\%$ ). The  $P_t$  resolution of the CTC + SVX system is  $(\sigma_{P_t}/P_t)^2 \sim (0.0009 \times P_t(\text{GeV}/c))^2 + (0.0066)^2$ . During the last data collection, the 54 outermost CTC layers have been instrumented in order to measure the track energy release by ionization ( $dE/dx$ ). This information has been used for a better separation of electrons from charged pions for  $P_t < 4\ \text{GeV}/c$ , and to help identifying  $K^\pm$  for  $P_t < 700\ \text{MeV}/c$ ; the  $dE/dx$  resolution is  $\sim 15\%$ . Finally, some of the mechanical characteristics of the CTC are summarized in table 3.1, while figure 3.2 shows the  $R - \phi$  view of an event with reconstructed tracks.
4. The drift tubes (CDT) [62]. Located between the CTC and the solenoid, they provide a correlated  $R - \phi - z$  space point measurement in the  $|\eta| < 1.0$  region.

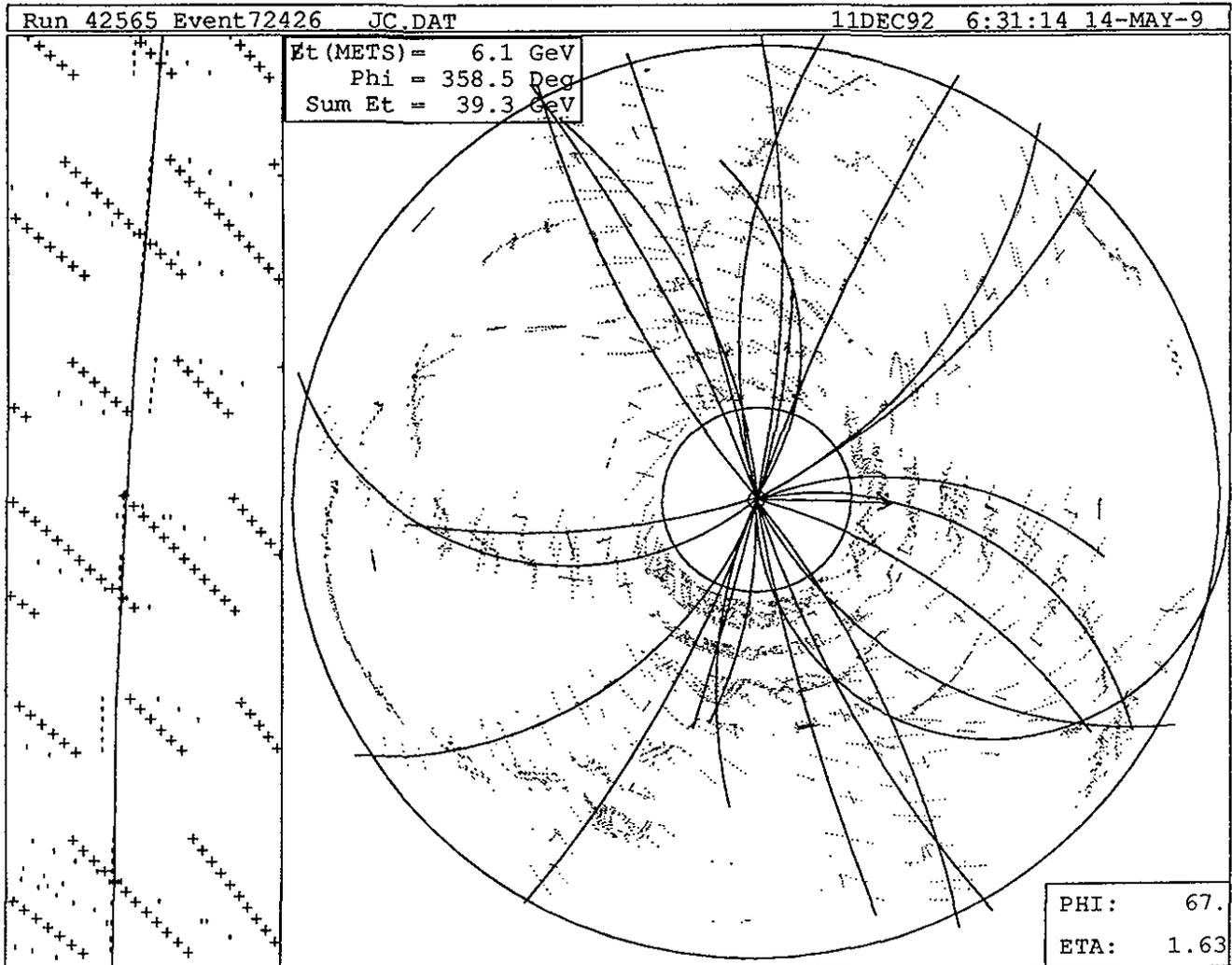


Figure 3.2:  $R - \phi$  view of the wire hits of an event with reconstructed tracks in the central tracking chamber (CTC); the inset on the top left corner shows the cell geometry of outer superlayers 4 to 8, where sense wires are indicated by crosses.

| Mechanical parameters          |   |
|--------------------------------|---|
| Number of layers               | 84  |
| Number of superlayers          | 9   |
| Stereo angles                  | $0^\circ +3^\circ 0^\circ -3^\circ 0^\circ +3^\circ 0^\circ +3^\circ 0^\circ$ |
| Number of supercells/layer     | 30, 42, 48, 60, 72, 84, 96, 108, 120  |
| Number of sense wires/cell     | 12, 6, 12, 6, 12, 6, 12, 6, 12  |
| Sense wire spacing             | 10 mm in plane of wires   |
| “Tilt” angle of cells          | $45^\circ$  |
| Radius of innermost sense wire | 309 mm  |
| Radius of outermost sense wire | 1320 mm   |
| Wire length                    | 3214 mm   |
| Sense wire diameter/tension    | 40 $\mu\text{m}$ gold-plated tungsten/135g                                    |
| Gas                            | argon-ethane-alcohol (49.6%:49.6%:0.8%)                                       |
| Drift field $E_0$              | 1350 V/cm   |
| Drift field uniformity         | $dE_0/E_0$ 1.5%(rms)  |
| Maximum drift time             | 800 nsec  |

Table 3.1: Mechanical parameters of the central tracking chamber (CTC)

### 3.2.2 Calorimetry

The electromagnetic (EM) and hadronic (HAD) calorimeters are divided into two categories: central calorimeters ( $30^\circ \leq \theta \leq 150^\circ$ ), where the active elements are scintillators, and forward calorimeters, “Plug” and “Forward” ( $2^\circ \leq \theta \leq 30^\circ$ ,  $150^\circ \leq \theta \leq 178^\circ$ ), which use proportional chambers. Their configuration is  $\phi$ -symmetric around the beams and forward-backward symmetric w.r.t. the median plane of CDF. An important feature of the calorimetry is its fine segmentation. As a whole, calorimeters are divided into approximately 5,000 projective *towers*, that is, detector elements covering some portion of solid angle and pointing to the nominal center of the interactions (and of the whole detector). The  $\Delta\eta \times \Delta\phi$  sizes of the towers are:  $0.1 \times 15^\circ$  for the central calorimeters and  $0.1 \times 5^\circ$  for the forward ones. Table 3.2 shows some features of the various CDF calorimeters.

The electron and photon identification has been improved with the installation (in the year 1992) of a set of proportional wire chambers (CPR or “Central Preradiator” [69]) between the magnet (acting as a preradiator, with its  $\sim 1X_0$  thickness) and the electromagnetic central calorimeter (CEM). The CPR increases the electron- $\pi^\pm$  and photon- $\pi^\pm$  separation by a factor 2-3.

| Calorimeter | $\eta$ Coverage      | Energy Resolution              | Thickness      | Spatial Resolution |
|-------------|----------------------|--------------------------------|----------------|--------------------|
| CEM [63]    | $ \eta  < 1.1$       | $13.5\%/\sqrt{E_t} \oplus 2\%$ | $18X_0$        | $0.2 \times 0.2$   |
| PEM [64]    | $1.1 <  \eta  < 2.4$ | $28\%/\sqrt{E} \oplus 2\%$     | $18 - 21X_0$   | $0.2 \times 0.2$   |
| FEM [65]    | $2.4 <  \eta  < 4.2$ | $25\%/\sqrt{E} \oplus 2\%$     | $25X_0$        | $0.2 \times 0.2$   |
| CHA [66]    | $ \eta  < 1.3$       | $75\%/\sqrt{E_t} \oplus 3\%$   | $4.5\lambda_I$ | $10 \times 5$      |
| PHA [67]    | $1.3 <  \eta  < 2.4$ | $90\%/\sqrt{E} \oplus 4\%$     | $5.7\lambda_I$ | $2 \times 2$       |
| FHA [68]    | $2.4 <  \eta  < 4.2$ | $130\%/\sqrt{E} \oplus 4\%$    | $7.7\lambda_I$ | $3 \times 3$       |

Table 3.2: Summary of CDF calorimeter properties. CEM (CHA), PEM (PHA) and FEM (FHA) stand respectively for central, plug, forward EM (HAD) calorimeter. Calorimeters cover  $2\pi$  in azimuth and  $|\eta| < 4.2$ . The symbol  $\oplus$  indicates that the constant term is added in quadrature in the resolution. The energy resolutions of EM calorimeters are for incident electrons and photons, the energy resolutions of HAD calorimeters are for incident isolated pions. Energy ( $E$ ) is given in GeV; the transverse energy is  $E_t = E \sin \theta$ . Thicknesses are given in radiation lengths ( $X_0$ ) for the electromagnetic (EM) calorimeters and in interaction lengths ( $\lambda_I$ ) for the hadron (HAD) calorimeters. Spatial resolution is given in  $\text{cm}^2$  at 50 GeV energy. The CEM spatial resolution is measured with proportional chambers with wire and orthogonal strip readout (CES), embedded in the EM calorimeter towers at the depth of  $6X_0$ . The CES chambers are useful to separate single EM showers ( $e^\pm$  and  $\gamma$ ) from neutral-pion double EM showers ( $\pi^0 \rightarrow \gamma\gamma$ ).

### 3.2.3 Muon Detection

In the central region ( $|\eta| < 1.0$ ) muons which penetrate the calorimeters are detected by means of drift chambers located outside the CHA. In the forward region ( $3^\circ < \theta < 16^\circ$  with respect to the two beams) muons are detected by means of a spectrometer (FMU, [70]) made up by a toroidal magnet, planes of drift chambers and scintillation counters for the trigger. In the intermediate region muon candidates can be found as isolated CTC tracks (if available) which are minimum ionizing in the calorimeters.

The central muon (CM) detector is particularly relevant for the top search in the leptonic channels. The CM detector configuration of the 1988-89 data taking described in the previous paragraph showed the following problems:

- the  $\lambda_I$  thickness of the material in front of the CM detectors was insufficient to obtain the desired level of containment of hadronic particles and showers;
- the angular coverage was too much limited, especially around  $\theta = 90^\circ$  ( $\eta = 0$ ).

These problems have been solved for the 1992-1993 data taking, thanks to the following upgrades [71]:

- behind the existing four layers of chambers (Central MUon chambers or CMU [72]; see figures 3.4 and 3.1) 0.6 m of steel was added, as well as four more layers of drift chambers (Central Muon uPgrage or CMP [73]); the CMP allows for a sizeable reduction of the background due to (interacting and non-interacting) penetration of hadrons through the CHA, as shown by figure 3.3;

- the CMU+CMF coverage has been extended, with the installation of another  $\mu$  detector (Central Muon EXtension or CMEX) formed by drift chambers (CMX [74]) sandwiched between and two layers of scintillation counters for the trigger (CSX [75]); an azimuthal region of  $\sim 90^\circ$  corresponding to the floor of the CDF collision hall is not instrumented.

In all CM detectors, muon  $P_t$  is measured with charged tracking and has a tracking resolution discussed in section 3.2.1. Table 3.3 summarizes some features of the central muon detectors.

| Characteristics          | CMU                         | CMF                         | CMEX                        |
|--------------------------|-----------------------------|-----------------------------|-----------------------------|
| $ \eta $ coverage        | [0,0.63]                    | [0,0.60]                    | [0.62,1.0]                  |
| $\phi$ coverage          | 84 %                        | 63 %                        | 71 %                        |
| Minimum detectable $P_t$ | 1.4 GeV/c                   | 3.0 GeV/c                   | 1.7 GeV/c                   |
| Thickness in front       | [5.1,6.2] $\cdot \lambda_I$ | [7.8,9.2] $\cdot \lambda_I$ | [6.2,7.2] $\cdot \lambda_I$ |

Table 3.3: Summary of characteristics of central muon detectors.

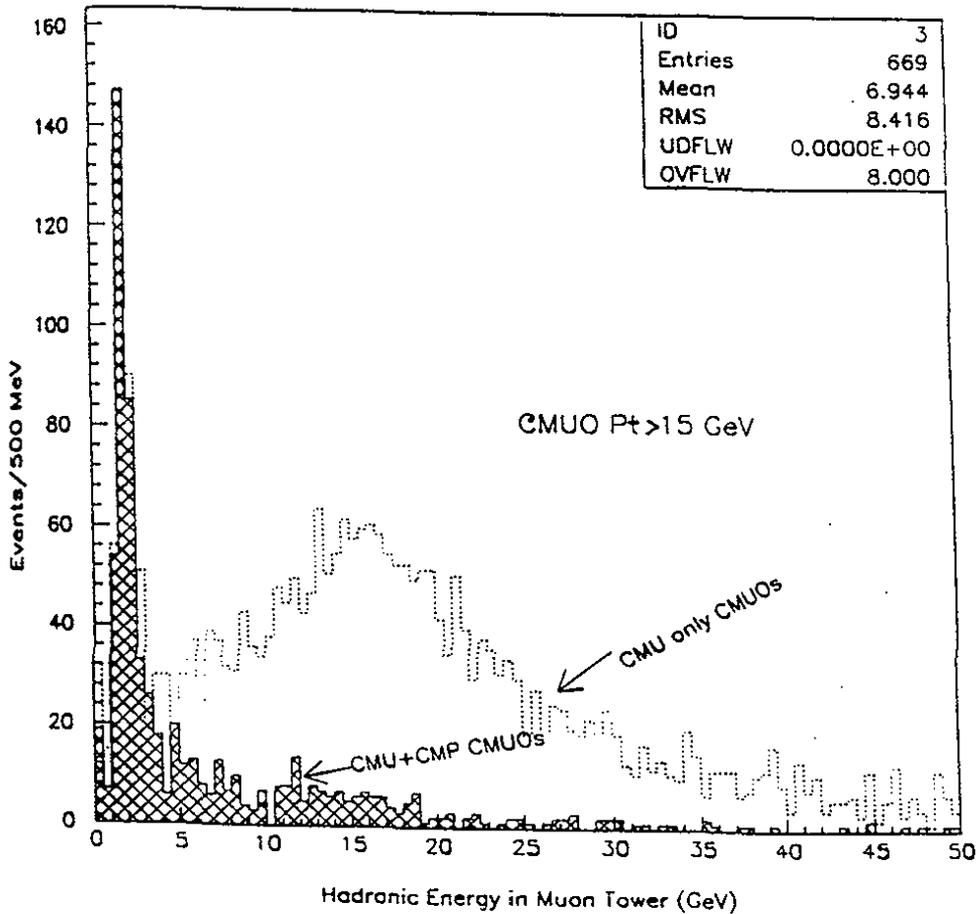


Figure 3.3: Distribution of the energy lost in the central hadron calorimeter by  $P_t > 15$  GeV/c muon candidates detected in central muon chambers (CMU), with and without the requirement of simultaneous detection in the central muon upgrade detector (CMF).

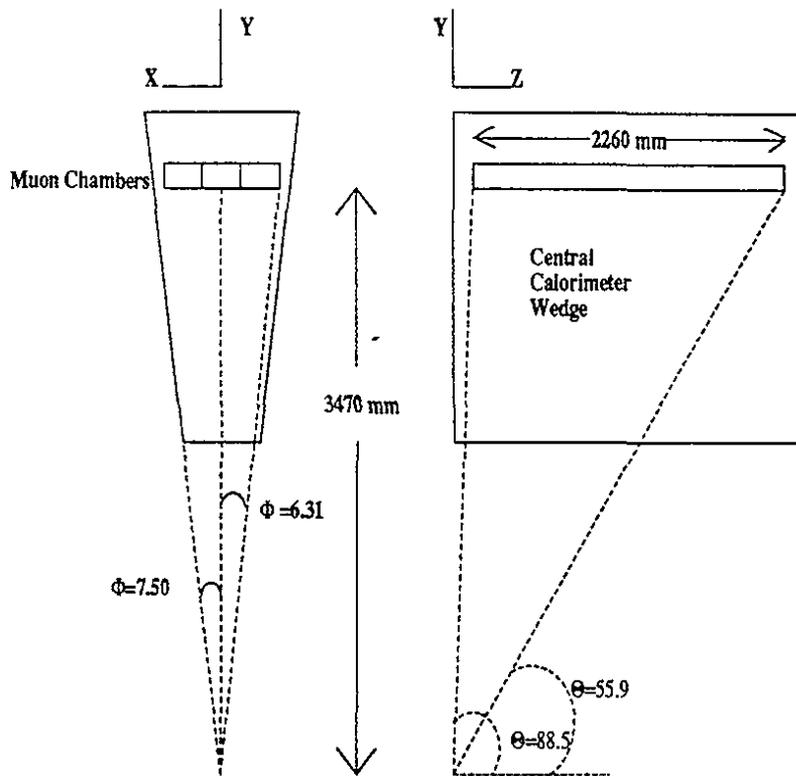


Figure 3.4: Geometry of the CMU: The chambers are located outside each azimuthal slice (wedge) of the hadron calorimeter. Between chambers of adjacent wedges there are uninstrumented regions of  $\sim 2.4^\circ$ , which give an overall loss of geometrical acceptance of  $\sim 16\%$ .

### 3.2.4 Trigger System

The  $3 \times 10^{30} \text{cm}^{-2} \text{sec}^{-1}$  average luminosity of the 1992-1993 data taking corresponds to a Tevatron frequency of interactions of approximately 130 KHz, while, on the other hand, events can be recorded on magnetic tapes only at a frequency of a few Hz. The necessary reduction factor is obtained by means of a three-level trigger system [76]. Each level is a logical OR of a certain number of triggers designed to select events containing electrons, muons or jets. The lowest level trigger ("Level 1") uses fast outputs from the CMU chambers to trigger on muon events and fast outputs from all calorimeters to trigger on electron and jet events. The calorimeter information is summed in *trigger* towers of size  $(\Delta\phi = 15^\circ) \times (\Delta\eta = 0.2)$  both for the EM and the HAD calorimeter. At a typical luminosity of  $5 \times 10^{30} \text{cm}^{-2} \text{sec}^{-1}$ , the rate of level 1 triggers is approximately 1 KHz.

The second trigger level ("Level 2") uses trigger tower information in a more sophisticated manner. A hardware processor is used to form a list of contiguous groups of towers with  $E_t$  greater than a certain threshold (*calorimetric cluster*), and for each of them the transverse energy,  $E_t$ , average  $\eta$  and average  $\phi$  are determined. These data are combined with a list of two-dimensional tracks (in the  $R - \phi$  plane), provided by another hardware

processor (the central fast tracker or CFT), which uses fast timing information from the CTC as input. The momentum resolution of the CFT is  $\delta P_t/P_t^2 \simeq 0.035 \text{ (GeV/c)}^{-1}$  and its efficiency is  $(93.5 \pm 0.3)\%$  for tracks of  $P_t > 10 \text{ GeV/c}$ . At the second trigger level the CMU, CMP and CMX information is also available. The CFT tracks can be combined with calorimetric clusters with high electromagnetic content to find electron candidates, or they can be combined with the track segments from the muon chambers to find muon candidates. At a typical luminosity of  $5 \times 10^{30} \text{ cm}^{-2} \text{ sec}^{-1}$ , the level 2 output rate is approximately 12 Hz.

The third trigger level ("Level 3") is a software trigger, which uses a system of Silicon Graphics computers, and has the power of about  $3 \times 10^9$  instructions per second. A VME interface allows for an input frequency up to 20 Hz. Events from level 2 are passed through an appropriate sequence of Fortran offline reconstruction algorithms. Among these, the three-dimensional CTC track reconstruction takes most of the execution time. The level 3 data output frequency is  $\sim 5 \text{ Hz}$ , and events are then written to magnetic tapes for offline processing.

### 3.2.5 High $P_t$ Electron and Muon Triggers

The top search in the leptonic channels is based on events collected with high  $P_t$  electron and muon triggers. This section describes how the information from the several CDF detector components is used to collect events with electrons or muons of  $P_t > 20 \text{ GeV/c}$  in the  $|\eta| < 1.0$  region.

The first level of central electron trigger (TE) uses only calorimeter information and requires at least an EM trigger tower of  $E_t > 6 \text{ GeV}$ . The first level of central muon trigger (TM) [77] is based only on muon chamber information. The track direction is measured using the arrival times of drift electrons at the sense wires. At least two hits on radially aligned wires in the CMU and CMX chambers are required. Assuming the track to be originated from the average beam line, its  $P_t$  is determined from its deflection in the magnetic field. The TM requires a CMU track segment of  $P_t > 6 \text{ GeV/c}$  in coincidence with hits in the CMP, or CMX track segment of  $P_t > 10 \text{ GeV/c}$  in coincidence with hits in the CSX scintillators located on both sides of the chambers.

The level 2 TE and TM use the  $R - \phi$  projections of the CFT tracks. The TE requires a CFT track of  $P_t > 9.2 \text{ GeV/c}$ , pointing to an EM cluster of  $E_t > 9 \text{ GeV}$ . An EM cluster is a contiguous group of towers with  $E_t > 7 \text{ GeV}$  each, with at least one tower (the *seed*) of  $E_t > 9 \text{ GeV}$  and with a ratio of hadronic to electromagnetic energy less than 0.125. The TM requires a match between a CFT track of  $P_t > 9.2 \text{ GeV/c}$  and a muon track segment which passed the level 1 TM.

The third level TE requires the reconstructed EM cluster to have  $E_t > 18 \text{ GeV}$ , and a  $P_t > 13 \text{ GeV/c}$  track pointing to it. The TM requires a reconstructed track of  $P_t > 18 \text{ GeV/c}$  distant in  $R - \phi$  by less than 10 cm from a muon track segment which passed the level 2 TM. In addition the deposited energy in the corresponding CHA tower must not exceed 6 GeV.

The trigger efficiency has been measured as a function of electron  $E_t$  and muon  $P_t$  using event samples collected with independent triggers. From a study of  $W$  and  $Z$  events, one finds that the TE efficiency for isolated electrons of  $20 < E_t < 150 \text{ GeV}$  is  $92.8\% \pm 0.2\%$ .

The TM efficiency for CMU muons (CMX) of  $P_t > 20$  GeV/ $c$  is  $86.8\% \pm 1.9\%$  ( $54.4\% \pm 5.5\%$ ), within the chamber geometrical acceptance.

The definition of high  $E_t$  ( $P_t$ ) electron (muon) from the point of view of the offline analysis is given in section 7.1.1 (7.1.2).

### 3.2.6 Jet Triggers

The event samples collected with inclusive jet triggers are used in chapter 6 to characterize the  $b$ -tagging algorithm by studying the tagging background in  $W +$  multijet events.

The level 1 jet trigger requires a single trigger tower with  $E_t$  above a threshold varying with the kind of calorimeter (6 GeV for the CEM, 8 GeV for the CHA/PEM/FEM, 25 GeV for the PHA/FHA). The level 2 requires a localized deposit of transverse EM or HAD energy. In the trigger system, a *jet* is a cluster initiated by a seed tower of  $E_t > 3$  GeV and consists of all contiguous towers of  $E_t > 1$  GeV contiguous in  $\eta$  and  $\phi$ . Four separate thresholds are imposed on  $E_t$  of such clusters, 20, 50, 70 and 100 GeV, defining the so-called “Jet\_20, Jet\_50, Jet\_70 and Jet\_100” triggers.

The definition of jet from the point of view of the offline analysis is given in sections 6.1.1, 6.1.2.

## Chapter 4

# The Silicon Vertex Detector (SVX)

The expression “vertex detector” indicates a tracking system capable of providing a precise measurement of charged particle trajectories in a region as close as possible to the interaction vertex. In particular, it measures the track direction and impact parameter, resolving those coming from secondary decay vertices from the ones originating from primary interactions.

In order to achieve high resolution tracking, silicon ( $S_i$ ) detectors are used. These consist of large numbers of  $n - p$  junctions arranged as closely spaced strips. The diodes are inversely polarized in order to increase the thickness of the silicon depletion region up to the actual thickness of the silicon crystal itself. The passage of minimum ionizing charged particles (MIPs) releases a charge of about 20,000 electron-hole pairs (corresponding to 3.2 fC) in a silicon thickness of about 280  $\mu$ . Due to the electric field present in the silicon crystal, the liberated charge migrates to the  $p^+$  strips (collecting holes) and to the  $n^+$  area (collecting electrons). The charge reaching the strips is then collected by the microelectrodes located on their surface and is readout by the “front-end” electronics. The typical distance between electrodes allows for a spatial resolution of the order of 10  $\mu$ m.

The VLSI (Very Large Scale Integrated circuit) technology permits integration of a large number of preamplifier channels on circuits of a few millimeters, with a spacing comparable to the silicon strip pitch. This, in turn, allows to assemble the tracking detectors as telescopes located close to the interaction region, to place the front-end electronics on the detectors themselves and to microbond the single electronic channels to the strip electrodes. In addition, the VLSI circuits are capable of integrating the charge in parallel from all the connected electrodes and of communicating the information serially (multiplexing). In this way the number and mass of the output cables from the detector can be very much reduced.

With its impact parameter resolution of about 17  $\mu$ m for  $P_t > 20$  GeV/ $c$ , the CDF Silicon Vertex detector is well suited to the identification of  $b$ -hadron (and  $c$ -hadron) decays, characterized by a relatively long lifetime,  $c\tau_B = 446 \pm 26 \mu$ m [50]. The SVX is therefore essential to search for the top quark and to study the properties of the  $b$  quark. Among these properties, CDF published the observation and mass measurement of the  $B_s$  meson [78], the measurement of the inclusive  $b$  lifetime averaged over all  $b$ -hadron states ( $c\tau_B$ , see [50]) and of the separate lifetimes of  $B_u^\pm$  and  $B_d^\pm$  mesons [51].

The SVX has been described in detail in several technical documents, concerning its mechanical and electronical characteristics [79, 80, 81, 82, 83], the performance of its prototypes [84, 85, 86, 87], the progress in its construction [88, 95, 90] and its data acquisition

system [91, 92]. A complete description of the SVX and its operation in  $p\bar{p}$  collisions can be found in [93]. Since the information provided by the SVX are extremely relevant for the top-quark search [94, 95] and for this thesis work, the SVX is described in the following in greater detail than other CDF detectors.

The SVX is located around the interaction region, inside the VTX and CTC, covering the region  $|\eta| < 2.0$ . It provides high resolution  $R - \phi$  tracking, using pre-fit CTC tracks (see section 3.2.1). The following sections illustrate the main mechanical and electrical features of the detector, as well as some of the calibrations needed to characterize its performance.

## 4.1 Mechanical Characteristics

The SVX is divided into two longitudinal sections (“barrels”) located right and left of the center of CDF. Its total length (51 cm) has been chosen to cover as much as possible the Tevatron interaction region, which has a length (RMS) of  $\sim 30$  cm. The SVX geometrical acceptance for single tracks is  $\sim 60\%$ . The amount of SVX material has been kept to minimum, since: (1) secondary vertices from particle interactions in the detector and photon pair conversions represent a source of background in the offline analysis and in the trigger; (2) the multiple scattering worsens the position resolution for low  $P_t$  tracks. For example, each silicon crystal is mounted on a very low mass Rohacell and carbon fiber support, and all these supports are then tied to a Beryllium structure (“bulkhead”). In addition, the section of the beam pipe around CDF is made of Beryllium, in order to reduce multiple scattering before the SVX position measurements. The average thickness of material traversed by a track of  $|\eta| < 1.2$  is  $3\% \times X_0$ .

Each barrel is divided into 12 “wedges” of  $30^\circ$  azimuthal aperture and each wedge consists of 4 layers of silicon detectors (“ladders”) with microstrips parallel to the beam direction (tracking occurs, therefore, only in  $R - \phi$ ). The ladder surfaces are tilted by  $3^\circ$  with respect to the normal to the straight line linking the ladder center to the SVX nominal center. To limit the total number of channels (46080), 3 silicon crystals of 8.5 cm length are assembled on each ladder and the electrodes of the strips bonded together by microsoldering. The average SVX temperature is kept around  $20^\circ$ - $25^\circ$  with a good safety margin, by means of a cooling system with water circulating in aluminum tubes in thermal contact with the bulkhead in the vicinity of the ladder end where the readout electronics is located (“ear card”). The SVX has been built with great mechanical accuracy ( $\sim 25\mu\text{m}$ ). Misalignments w.r.t. nominal positions have been measured with optical methods during assembly and they have been used for the first track reconstruction with real particles. The positions and dimensions of the silicon detectors are summarized in table 4.1. A general view of the SVX is shown in figure 4.1, while figure 4.2 shows the various components of a ladder.

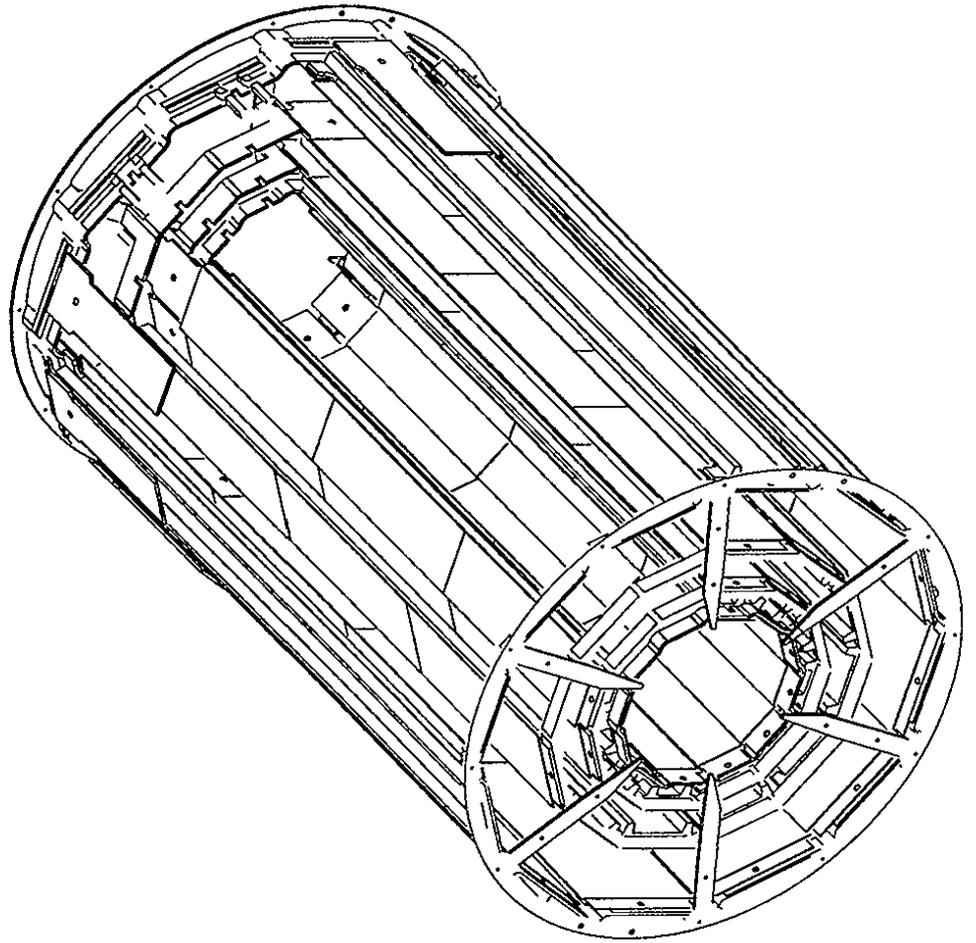


Figure 4.1: Isometric view of an SVX barrel.

| Layer | Side with $S_i$ | Radius (cm) | Total width (cm) | Active width (cm) | Pitch ( $\mu\text{m}$ ) | Nr. of Strip | Nr. of Chip |
|-------|-----------------|-------------|------------------|-------------------|-------------------------|--------------|-------------|
| 0     | IN              | 3.005       | 1.6040           | 1.536             | 60                      | 256          | 2           |
| 1     | OUT             | 4.256       | 2.3720           | 2.304             | 60                      | 384          | 3           |
| 2     | OUT             | 5.687       | 3.1400           | 3.072             | 60                      | 512          | 4           |
| 3     | IN              | 7.866       | 4.2930           | 4.224             | 55                      | 768          | 6           |

Table 4.1: Positions and dimensions of the silicon detectors. Silicon crystals are  $280 \mu\text{m}$  thick.

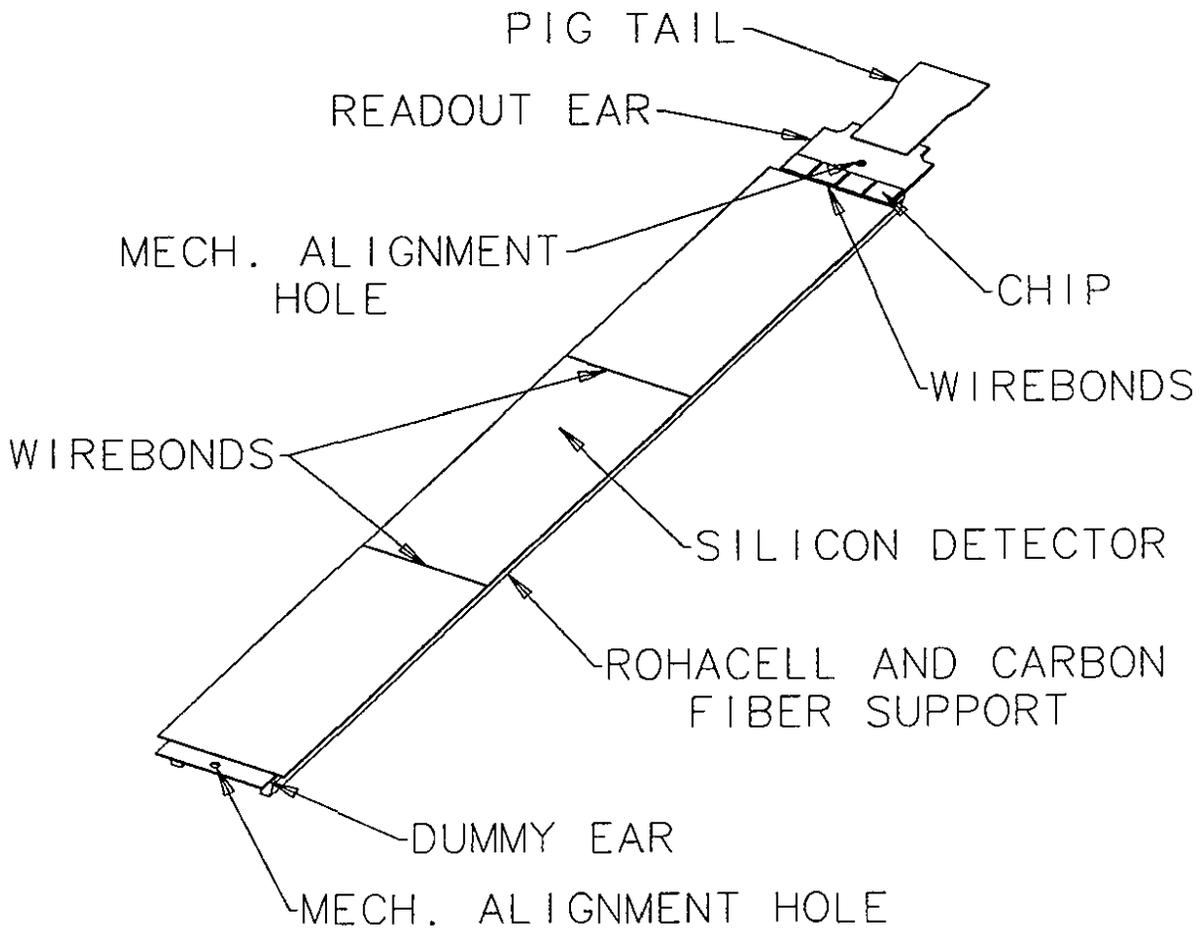


Figure 4.2: Schematic view of the components of a ladder.

## 4.2 Electrical Characteristics

The SVX detectors are  $n$ -type,  $280\ \mu\text{m}$  thick silicon crystals. The junction side (“J side”) of each detector has  $p$ -type strips, obtained from Boron implantation. The electrodes used for charge collection and as ohmic contacts (“ $\Omega$  side”) to provide inverse bias voltage are made of aluminum. All detectors are  $8.5\ \text{cm}$  long, while their width and strip pitch depend on the layer they belong to (see table 4.1).

For each strip, or channel, one can read out its address (that is, the  $R-\phi$  strip position) and the electric charge released in it by charged particles crossing it. This charge is collected directly by the electrodes running along the strips ( $DC$  coupling). Typically, the total charge released by MIP’s (charge “cluster”) is divided among 1, 2 or 3 strips. The position of the point in which the MIP crossed the crystal is taken as the cluster position, calculated as the center-of-gravity of the charges of the hit strips. The position resolution of one-strip clusters is  $\sim 15\ \mu\text{m}$ , in agreement with the strip pitch ( $60\ \mu\text{m}/\sqrt{12}$ ); the two-strip cluster resolution is  $\sim 13\ \mu\text{m}$ , since the use of the charge center-of-gravity improves the resolution; for three-strip clusters the resolution is  $25\ \mu\text{m}$ . The hybrid circuit with the strip readout electronics (“SVX chip”) is located on one of the two ladder ends (“ear card”) and it is bonded to the strips by microsoldering. Each chip reads out 128 strips. The channel input capacitance is  $\sim 30\ \text{pF}$  (for a total strip length of  $25.5\ \text{cm}$ ) and  $S/N \sim 9$ .

Since the ear card electronics is not radiation resistant and since the charge collection occurs with  $DC$  coupling, the radiation absorbed by the SVX during the data taking caused the following effects: (i) a partial damage of the ear card electronics; (ii) an increase of the strip leakage currents (due to minority charge carriers). The leakage current increase can change the working point of the preamplifiers, eventually saturating them, thus lowering their gain and compromising the detection of the signal. The absorbed radiation caused therefore slightly worsened the SVX performance towards the end of the data taking period; for example, the  $S/N$  of the inner layer decreased to  $6/1$ .

During  $p\bar{p}$  collisions the time available for SVX data acquisition is  $\sim 2\ \text{msec}$ . This makes it impossible to readout all the SVX channels. In order to reduce the readout time and the data size, only strips whose charge content (signal) exceeds a variable threshold are readout (*sparse readout*). To do this, the readout chip is used in “quadruple sample and hold” mode: two subsequent charge integrations are performed, one with and another one without the bunch crossing, and the strip signal is taken to be the difference of the two charges. This allows for a “hardware” subtraction of the inverse current contribution to the pedestal. During one of the integrations a charge is injected through the calibration line, in such a way as to compare the signal with a variable threshold (that is, the injected charge). The sparse readout signifies the acquisition of data from the strips whose signal exceeds the chosen threshold plus the two immediately adjacent strips (“neighbor” logic). In this way, the readout time and data size are determined by the effective number of channels hit by particles (the effective occupancy) and not by the total number of channels. The average SVX occupancy is  $< 10\%$ , which corresponds to the readout time of  $\sim 0.77\ \text{msec}$ . In order to reduce the volume and mass of the cables, the channels are readout serially, with a single line for all strips in a wedge. To drive these lines from the SVX to the FASTBUS digitizers located outside CDF, specially designed “port cards” mounted on the bulkheads are used.

The sparse readout worked well and reliably, both in the test carried out with proto-

types before the installation in CDF [85], and during the operation of the whole SVX in  $p\bar{p}$  collisions. The hardware threshold has been optimized to achieve a cluster reconstruction efficiency (see section 4.4.3) of about 90 % for layer 1, 95 %, for layers 2, 3, 4, and the quoted occupancy of < 10%. However, at the end of the data taking period the layer 1 efficiency decreased down to 63 % due to the absorbed radiation dose.

### 4.3 Calibrations

Special calibrations are performed periodically between  $p\bar{p}$  collision runs, in order to monitor the functioning of the detector and to calculate the constants to be used in the offline analysis: pedestals, noise, gain and thresholds for the sparse readout. Coherent pedestal oscillations have not been observed. The measured noise level was of the order of  $\sim 2200$  electrons, uniform at the 20 %. Gains were uniform at 5 %, while less uniform were the wedge thresholds. Since only two charge injection lines are available per wedge (one for layer 1 and another one for the three outer layers), in order to keep a good efficiency for all the layers, one has to accept occupancy variations from 5 to 15 % (10 % on the average). This, however, had no consequence on the quality of the collected data. The fraction of malfunctioning channels due to high inverse currents, low gains and high noise values was  $\sim 1.5\%$ . Most of them had been already found during the ladder test prior to installation.

### 4.4 Tracking with the SVX

This section describes how the information provided by the SVX is used for high precision tracking of charged particle trajectories; some important equations used in chapter 5 are also defined.

#### 4.4.1 Track Parametrization

In a homogeneous magnetic field charged particles move along a helix whose axis is parallel to the field (see figure 4.3). This trajectory can be described by five parameters:

$$\vec{\alpha} = (\cot \theta, C, z_0, D, \phi_0), \quad (4.1)$$

where:

- $\cot \theta$  = cotangent of the polar angle at the point of closest approach to the  $z$  axis,
- $C$  = half curvature (same sign of the electric charge),
- $z_0$  =  $z$  coordinate of the point of closest approach to the  $z$  axis,
- $D$  = impact parameter or distance of closest approach,
- $\phi_0$  = azimuthal direction at the point of closest approach.

The CDF magnetic field is directed towards  $z < 0$ . If  $(x_0, y_0)$  is the center of the trajectory projected onto the transverse plane, the impact parameter is:

$$D = Q \cdot (\sqrt{x_0^2 + y_0^2} - \rho), \quad (4.2)$$

where  $\rho = \frac{1}{|2C|} = \frac{1}{2QC}$  is the radius of the circle. Using the parametrizations of equation 4.1 one can obtain the following equations (see [99]), describing the particle trajectory as a function of its distance from the  $z$  axis:

$$\phi(R, \vec{\alpha}) = \phi_0 + \sin^{-1} \left( \frac{CR + (1 + CD)D/R}{1 + 2CD} \right) \quad (4.3)$$

$$Z(R, \vec{\alpha}) = z_0 + \frac{\cot \theta}{C} \sin^{-1} \left( C \sqrt{\frac{R^2 - D^2}{1 + 2CD}} \right), \quad (4.4)$$

where  $\vec{\alpha} = (\cot \theta, C, z_0, D, \phi_0)$  is the vector of parameters.

#### 4.4.2 Track Reconstruction Algorithm

The SVX tracking algorithm used by CDF is derived and compared to other methods in [98]. In this section the main ideas concerning the SVX tracking algorithm are described. A more detailed discussion can be found in [99].

The algorithm used is the so-called *progressive* method: it combines track fits from the different detectors, including the effect of coulomb multiple scattering. The starting point is the track fit in the CTC, located outside the SVX. A *road* whose size is based on its covariance matrix is formed around the CTC track, and the corresponding SVX charge cluster (*hit*) is looked for within this road, on the outermost SVX layer. When one such hit is found, the track is re-fitted including the new position measurement and the new candidate track (and road) is propagated inside the SVX to the next layer.

For each new hit found, new track parameters are calculated in the following way:

$$\vec{\alpha} = \vec{\alpha}_0 + [(V_0^*)^{-1} + A]^{-1} \cdot \frac{\partial x}{\partial \vec{\alpha}} \cdot \frac{x^{track} - x^{hit}(\vec{\alpha}_0)}{\sigma_x^2} \quad (4.5)$$

where:

$$\begin{aligned} \alpha_0 (\alpha) &= \text{parameters before (after) the new fit,} \\ x^{track} &= \text{hit position measured in the } R - \phi \text{ plane,} \\ x^{hit}(\vec{\alpha}_0) &= \text{hit position predicted by the previous fit,} \\ \sigma_x &= \text{resolution on the measurement of } x^{track}, \\ \frac{\partial x}{\partial \vec{\alpha}} &= \text{derivatives w.r.t. track parameters (see [99]).} \end{aligned}$$

$$A_{ij} = \frac{1}{\sigma_x^2} \cdot \frac{\partial x}{\partial \alpha_i} \frac{\partial x}{\partial \alpha_j}, \quad (4.6)$$

$$V_0^* = V_0 + V_{MS}, \quad (4.7)$$

where  $V_0$  is the covariance matrix before the fit and  $V_{MS}$  is the multiple scattering contribution (see [99]).

The new covariance matrix is calculated using the following formula:

$$V' = [(V_0^*)^{-1} + A]^{-1}. \quad (4.8)$$

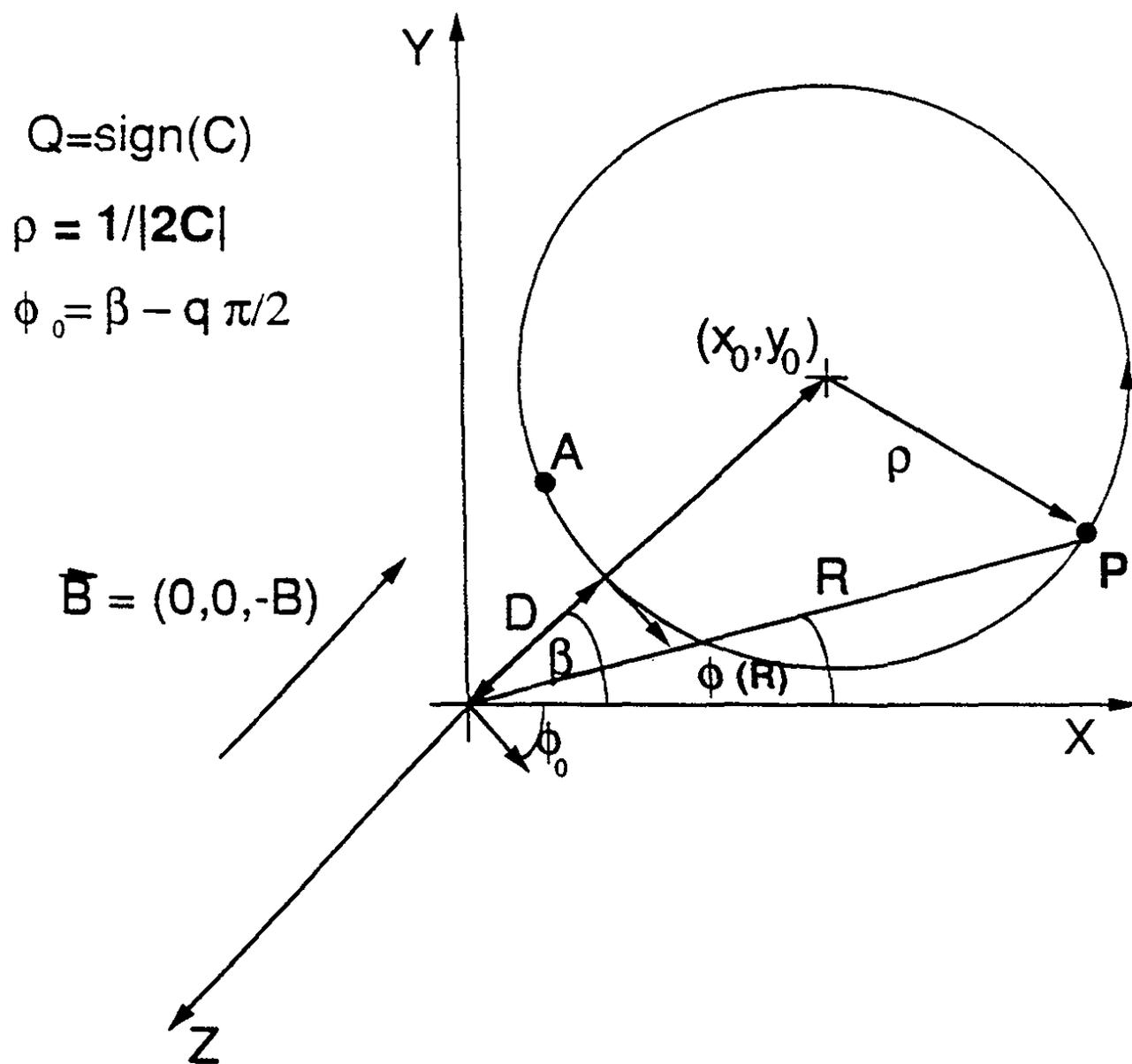


Figure 4.3: Track of a particle with positive charge  $Q$ .

For each new hit progressively added to the track, the difference between the new and old  $\chi^2$  is:

$$\delta\chi^2 = \chi^2 - \chi_0^2 = (\Delta\vec{\alpha})^t (V_0^*)^{-1} \Delta\vec{\alpha} + \frac{(x^{track} - x^{hit}(\vec{\alpha}_0))^2}{\sigma_x^2}, \quad (4.9)$$

where  $\Delta\vec{\alpha} = \vec{\alpha} - \vec{\alpha}_0$  is the difference between the new and old parameters and  $\chi_0^2$  refers to the previous fit. The first term of  $\delta\chi^2$  expresses the error in the track extrapolation due to the covariance matrix of the previous fit ( $V_0$ ) and to multiple scattering ( $V_{MS}$ ). The second term contains the SVX position resolution. The  $\chi^2$  value is strictly related to the CTC resolution, because when the first SVX hit is found,  $V_0$  is just the CTC covariance matrix. On the contrary the dependence of  $\delta\chi^2$  on the CTC error scale (that is, on the CTC track fit chi-squared,  $\chi_{CTC}^2$ ) is much weaker to first order. Often a variable number of different hits can be associated to the same CTC track. In the presence of different SVX track candidates, the lowest  $\delta\chi^2$  candidate is chosen [100], in order to minimize the effect of the uncertainty in the CTC error scale. In addition, the  $\delta\chi^2$  cut is based on the assumption that the SVX hits are consistent among themselves and, as such, the cut is less dependent on the global detector alignments than would be a cut on the value of the total  $\chi^2$ .

Once all possible hits in all SVX layers have been assigned to the track, the final SVX track chi-squared,  $\chi_{SVX}^2$ , is formed. Since  $\chi_{SVX}^2 - \chi_{CTC}^2$  has a much weaker dependence on the CTC error scale than  $\chi_{SVX}^2$  (like  $\delta\chi^2$  of equation 4.9), this difference is used to impose quality cuts on SVX tracks.

### 4.4.3 Track Reconstruction Efficiency

The hit (or cluster) reconstruction inefficiency of a given SVX layer, that is,  $1 - \epsilon_{hit}(N_{layer})$  is measured by selecting tracks with  $P_t > 1$  GeV/c, and  $\geq 3$  hit in the same wedge, contained in the wedge acceptance, and computing the ratio of the number of tracks which do not leave hits on that particular layer to the total number of tracks. The measured values are:

$$\epsilon_{hit}(N_{layer} = 1) = 90\%,$$

$$\epsilon_{hit}(N_{layer} = 2) = 95\%,$$

$$\epsilon_{hit}(N_{layer} = 3) = 95\%,$$

$$\epsilon_{hit}(N_{layer} = 4) = 95\%.$$

The SVX track reconstruction efficiency,  $\epsilon_{track}(N_{hit})$ , as a function of the number of its hits,  $N_{hit}$ , is obtained by selecting good quality CTC tracks with  $P_t > 1$  GeV/c, contained in the SVX acceptance, not crossing regions with malfunctioning strips, and computing the ratio of the number of tracks with  $N_{hit}$  hits found by the SVX to the total number of CTC tracks. The measured values are:

$$\epsilon_{track}(N_{hit} = 4) = 73.7\%,$$

$$\epsilon_{track}(N_{hit} = 3) = 19.8\%,$$

$$\epsilon_{track}(N_{hit} \geq 3) = 93.5\%.$$

#### 4.4.4 Detector Alignment

Despite the fact that the SVX has been built with great care, trying to keep mechanical misalignment to the lowest possible level [88, 90], after its installation an accurate alignment of the whole SVX and its parts is necessary in order to reach the position resolution quoted in the previous sections.

Two kinds of alignments have been carried out [101]: the alignment of each barrel with respect to the CTC (*global*) and the alignment of the ladders within each barrel (*internal*). The two procedures are then iterated to stabilize remaining correlations between them. An effort has been made to concentrate the effect of possible CTC systematics on the global alignment, for which a lower accuracy than is required for the internal alignment.

Treating the barrels as rigid bodies, and keeping in mind that the SVX does not measure the  $z$  coordinate, the global alignment requires knowledge of 5 parameters: three rotation angles and two transverse plane displacements  $x$  and  $y$ . For a given data acquisition run, a measurement is made of the average beam line with CTC tracks and with SVX track belonging separately to each barrel, thus determining the  $x$  and  $y$  displacements and two rotation angles between the SVX and the CTC, used as a reference frame. The remaining parameter is a relative rotation around the  $z$  axis, which is determined by optimizing the distance between the extrapolation of CTC tracks to the SVX strips and the position of the hits actually reconstructed by the SVX. An estimate of the systematic error on the relative barrel alignment is given by the residual distance of the average beam line positions reconstructed independently with the two barrels, after correcting for the global alignments (as observed in an event sample different from the one used to measure the alignment parameters):  $< 10\mu\text{m}$ . The systematic error on the alignment of both barrels to the CTC is  $\sim 40\mu\text{m}$ .

The internal alignment procedure uses the optical measurements performed during detector assembly as a starting point for the first reconstruction of SVX tracks from real particles. An effort was made to identify the degrees of freedom which most influence the resolution, keeping to a minimum the number of parameters to be determined. For each ladder the most relevant parameters are: one translation parallel to the silicon plane and perpendicular to the direction of the strips; one rotation around the axis perpendicular to the silicon plane; the radial distance of the ladder from the beam line. For the four ladders of a given wedge these three parameters have been estimated keeping the position of the outermost ladder fixed and using the average beam line as an additional constraint (for a given data taking run). The systematic uncertainty on ladder radial positions is  $\sim 40\mu\text{m}$ ; different wedges show relative positions which differ by about  $20\mu\text{m}$ . Figure 4.4 shows the distribution of SVX track residuals before and after the application of the internal alignments: the width of residuals decreases from  $17.6\mu\text{m}$  (before) to  $10.6\mu\text{m}$  (after). This result implies a position resolution (averaged over single hits with any number of strips) of  $13\mu\text{m}$ , to be compared to the  $12\mu\text{m}$  specified in the initial SVX project proposal dated 1985.

#### 4.4.5 Track Impact Parameter Resolution

The measurement of the SVX impact parameter (IP, or  $D$ ) resolution,  $\sigma_D(P_t)$  is of the greatest importance. The transverse plane profile of the colliding beams is well described

by a circular gaussian with a sigma ( $\sigma_{beam}$ ) of  $36 \mu\text{m}$  (see section 5.1.1), which is negligible compared to the IP resolution of low  $P_t$  tracks ( $P_t < 1 \text{ GeV}/c$ ). Figure 4.5 shows  $D$  calculated with respect to the average beam line vs.  $P_t$ , for good quality tracks (4 hits, well measured in  $z$ ,  $(\chi_{SVX}^2 - \chi_{CTC}^2) < 20$ ). Assuming most of the tracks to originate from primary interactions of the beams, the width of the distribution is the sum in quadrature of  $\sigma_D(P_t)$  and of  $\sigma_{beam}$ . One observes that the distribution broadens at low  $P_t$ , where it is dominated by multiple scattering, and gets narrow at high  $P_t$ , where it is dominated by  $\sigma_{beam}$ .  $\sigma_D(P_t)$  as a function of  $P_t$  can be measured by slicing the distribution of figure 4.5 in vertical regions of  $100 \text{ MeV}/c$  amplitude, projecting the measured points onto the ordinate axis and fitting the distribution thus obtained to a gaussian.  $\sigma_D(P_t)$  is related to  $\sigma_{beam}$  and to the  $\sigma$  of such a fit by the following relationship:

$$\sigma_D(P_t) = \sqrt{\sigma^2 - \sigma_{beam}^2}. \quad (4.10)$$

The result is shown in figure 4.6, where the dashed curve represents a fit of the form:

$$\sigma_D(P_t) = A + B \frac{1}{P_t}, \quad (4.11)$$

where  $A \approx 10 \mu\text{m}$  is the asymptotic ( $P_t \rightarrow \infty$ ) SVX resolution, and  $B \approx 41 \mu\text{m GeV}/c$  is a function of the amount of material which causes multiple scattering. The value of  $A$  depends on the distance of the first silicon plane from the beam line and on the intrinsic position resolution of the silicon detectors. The fit to the data shown in figure 4.6 does not provide a reliable value of the asymptotic SVX resolution (for  $P_t \rightarrow +\infty$ ), because such a fit is limited to the low  $P_t$  region  $\in [0.4, 2.0] \text{ GeV}/c$ ; a similar fit performed for high  $P_t$  muon tracks from  $Z \rightarrow \mu\mu$  decays gives  $\sigma_D(\text{asymptotic}) \sim 17 \mu\text{m}$ . For  $P_t > 5 \text{ GeV}/c$ , the SVX IP resolution is  $\sim 25 \mu\text{m}$ , while the corresponding CTC resolution is  $\sim 250 \mu\text{m}$ . Note that the result of the fit for the multiple scattering contribution to  $\sigma_D(P_t)$  is in good agreement with the prediction of the CDF simulation program shown in figure 4.7.

The SVX also provides an improvement of the CTC  $P_t$  resolution of the CTC, by extending the lever arm of tracking of about 30 cm with the addition of four high-precision measurements, as shown by 4.8.

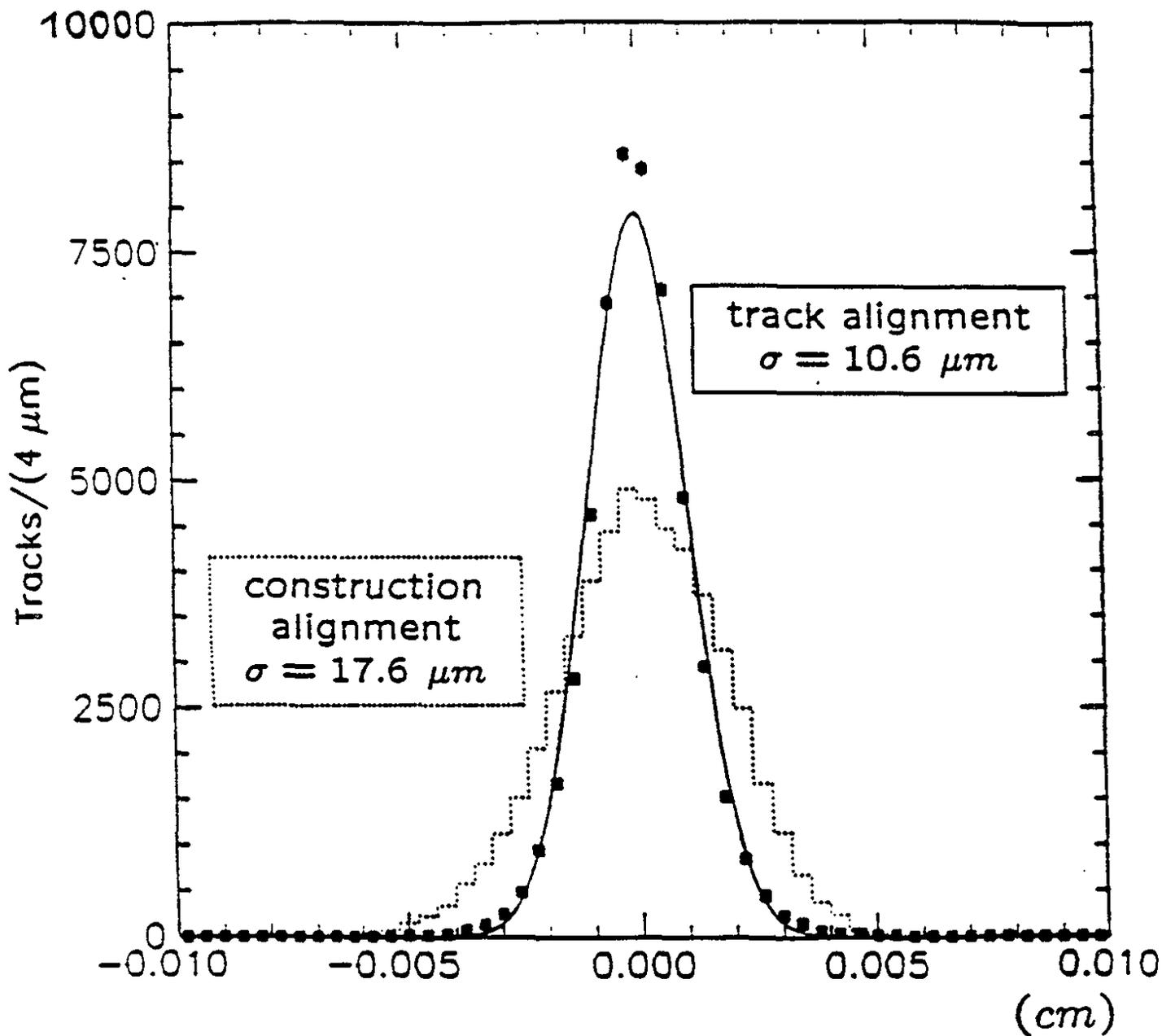


Figure 4.4: Distribution of SVX track residuals before (dashed histogram,  $4 \mu m/\text{bin}$ ) and after (squared points and gaussian fit) internal alignment performed with SVX tracks.

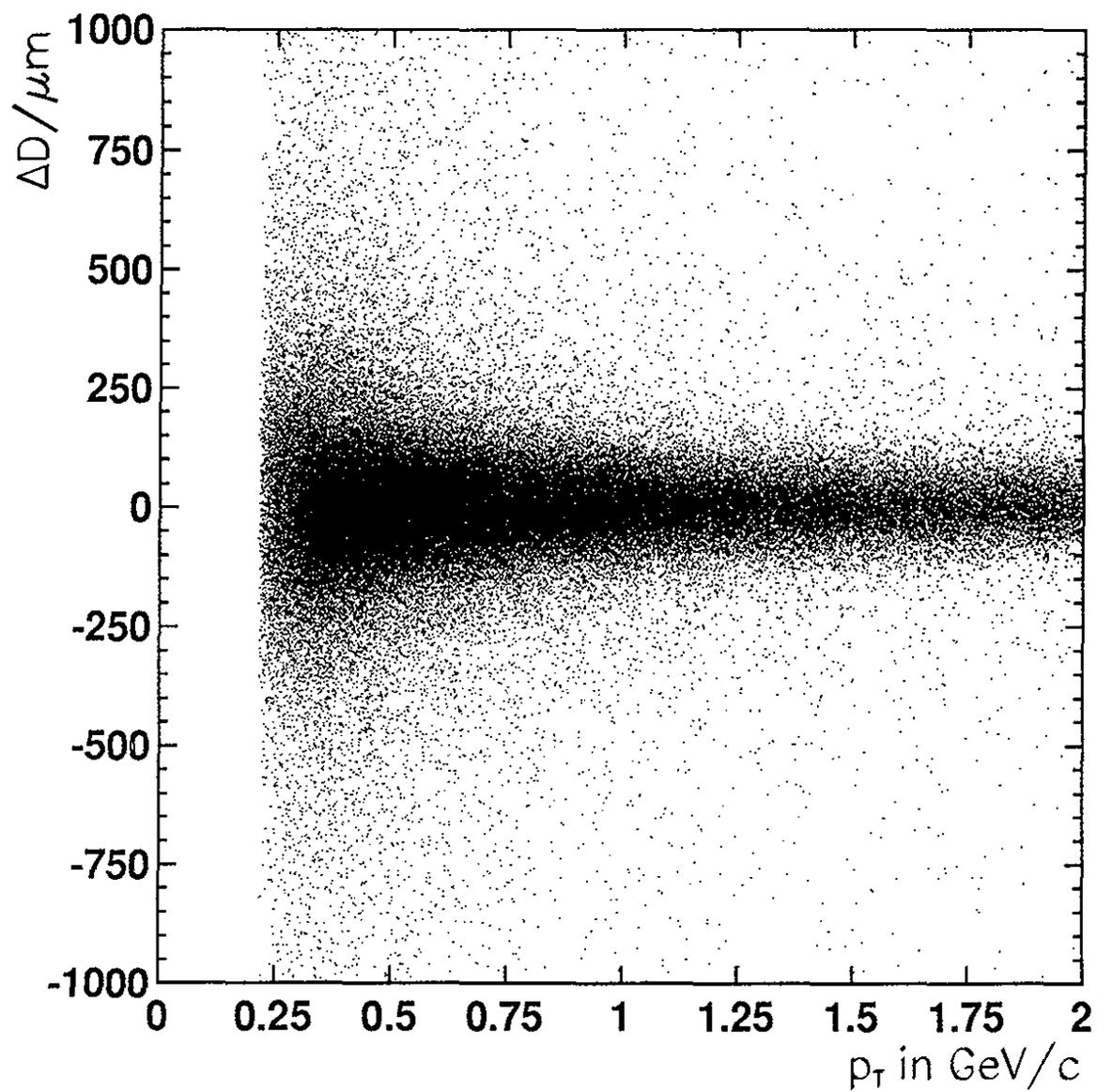


Figure 4.5: SVX impact parameter distribution measured with respect to the average beam line as a function of track  $P_t$  in the data.

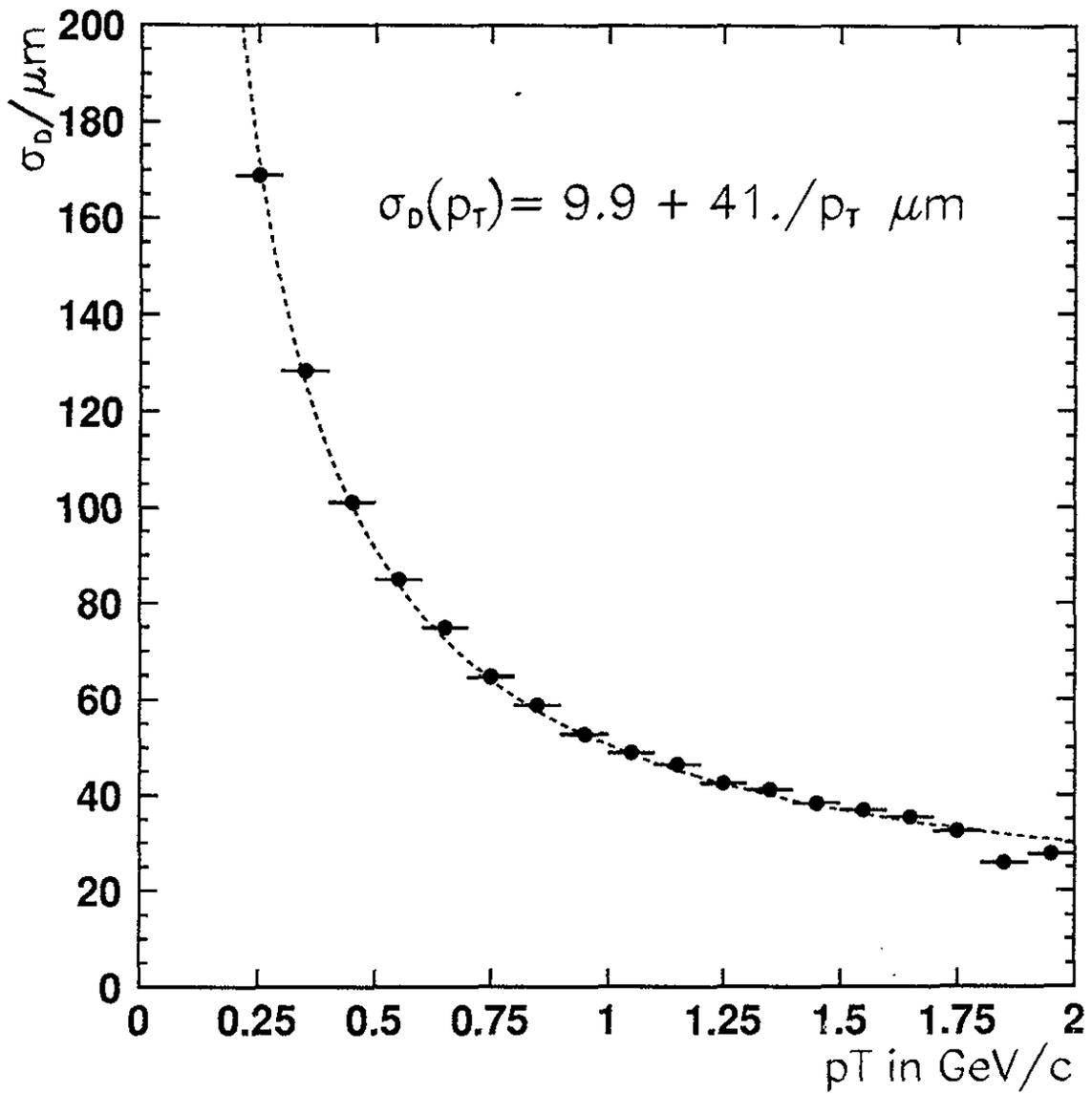


Figure 4.6: SVX impact parameter resolution as a function of track  $P_t$  in the data.

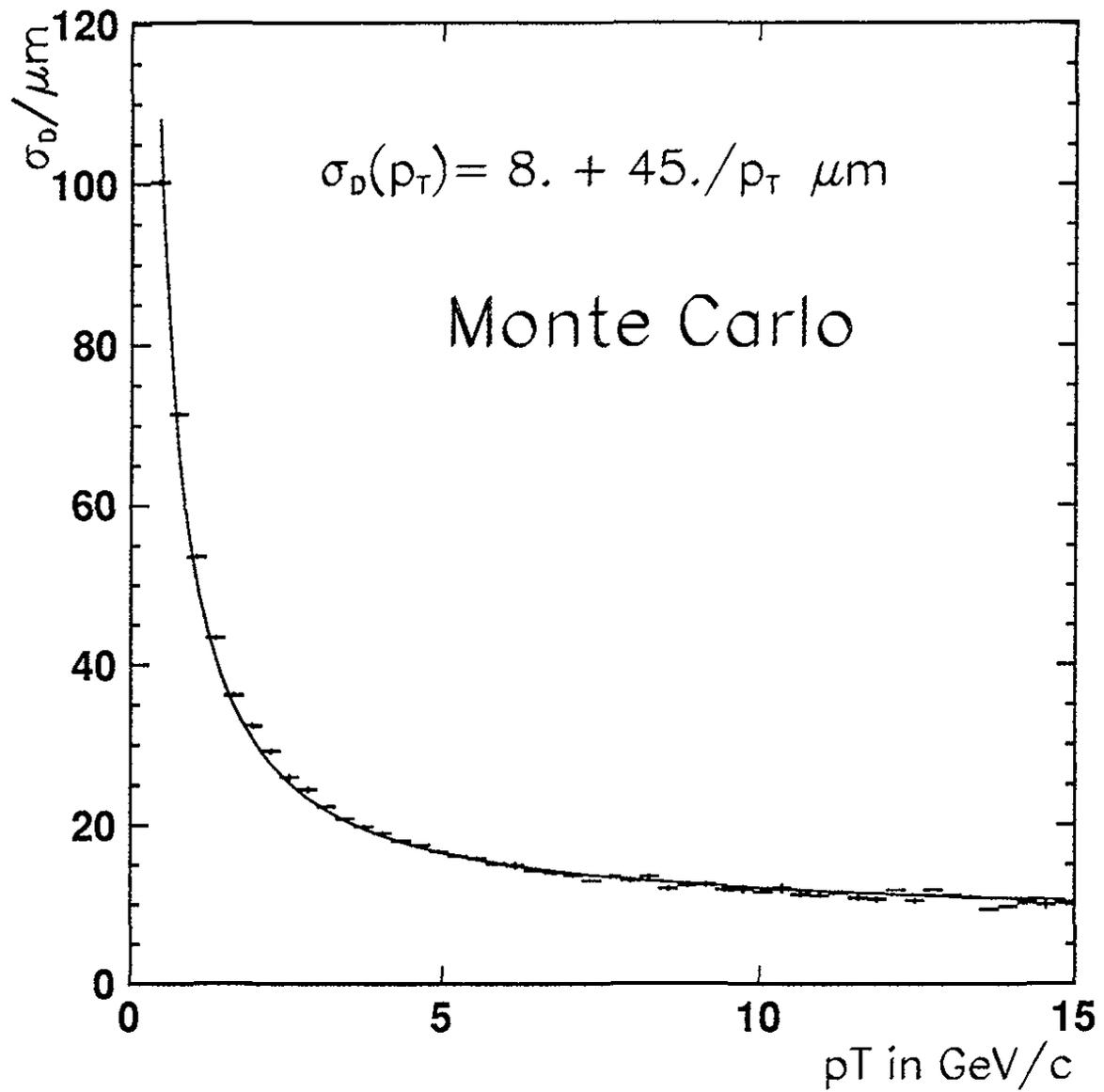


Figure 4.7: SVX impact parameter resolution as a function of track  $P_t$  in the Montecarlo.

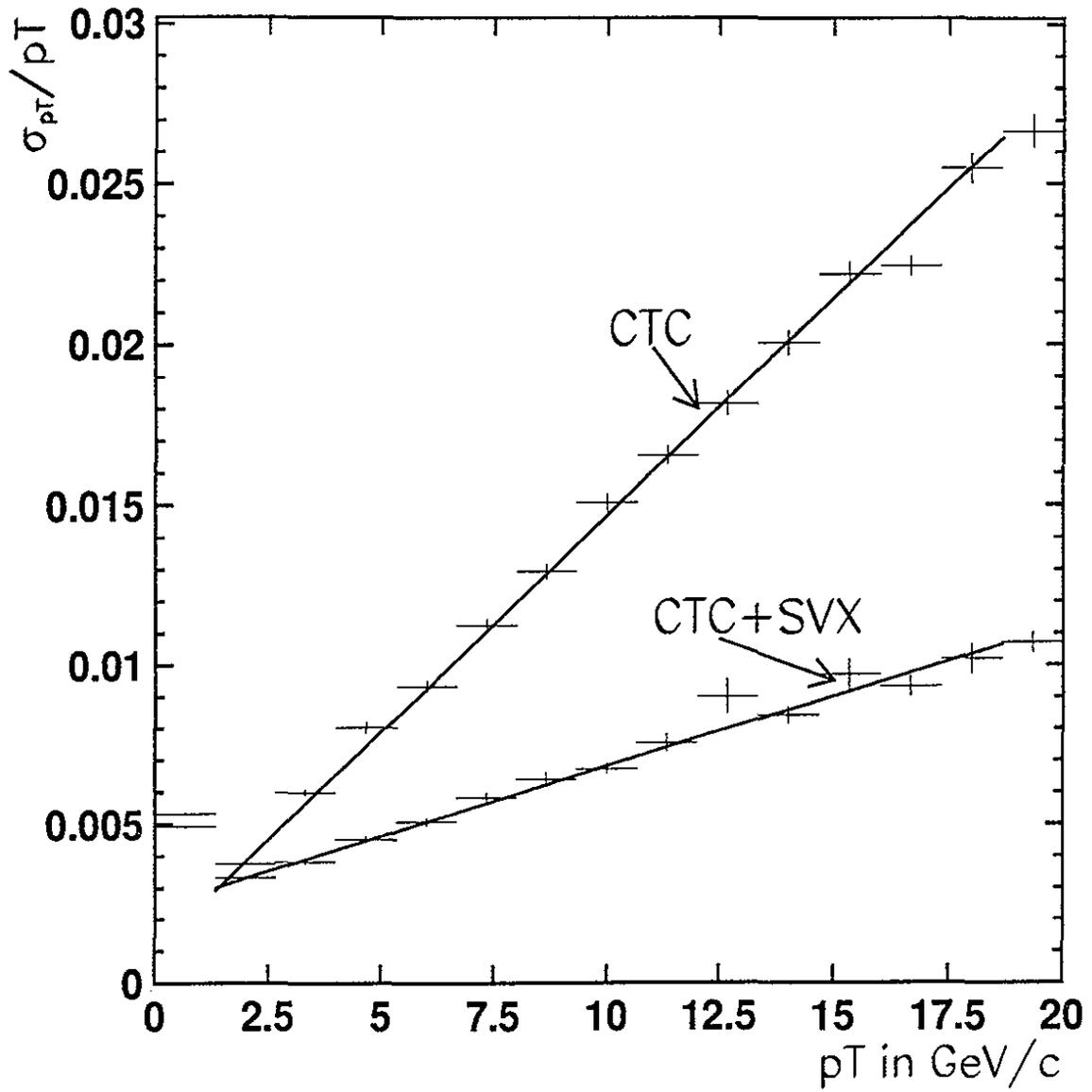


Figure 4.8:  $P_t$  resolution of CTC only and CTC + SVX tracks in the data.

# Chapter 5

## Secondary Vertex Tagging with the SVX

The recognition of  $b$ -hadron decays ( $c\tau_B = 446 \pm 26 \mu\text{m}$  [50]) requires high resolution tracking plus precise and reliable vertex reconstruction algorithms (both for primary interactions and for secondary decays). These prerequisites are essential in high-energy hadron collisions which produce in the final state many charged particles collimated in high-energy hadronic jets. In particular, in the  $W + \text{multijet} + b$  tag top-quark search we require  $\geq 3$  jets in  $|\eta| < 2.0$  (see section 7.2). Moreover, the high luminosity of the collisions may produce multiple primary interactions in the same event, making tracking and vertexing more difficult. The problem of multiple primary vertices (MPV's) is interesting especially in view of future experiments at the CERN "Large Hadron Collider" (LHC), where luminosity and  $\sqrt{s}$  will be even higher than at the Tevatron.

This chapter describes the algorithms for primary vertex (PV) and secondary vertex (SV) reconstruction. They represent the original contribution of the author of this thesis to the top-quark search program of the CDF experiment.

### 5.1 Finding Primary $p\bar{p}$ Interaction Vertices

The reconstruction of  $p\bar{p}$  primary interaction vertices is the first step to take in order to signal  $b$ -hadron decays with displaced secondary vertices. The primary vertex finder (PVF) tries to reach two basic goals:

- to provide the best possible resolution; this is needed because the top quark search is strongly limited by the number of produced  $t\bar{t}$  events; moreover, it is essential that the resolution scale is understood correctly, because it is used in the determination of IP resolution, together with the tracking resolution;
- to minimize systematic errors related to the event topology characteristics, like the number of tracks, of jets, of MPV's, and the presence of non-primary vertex tracks originated by SVs close enough to the PV to confuse the PV recognition; satisfying this requirement is essential in order not to spoil the effort of pushing the resolution to the limit.

The algorithm developed to reach these goal is an iterative fit procedure which combines the SVX tracking information available for each event with the average beam line position measured on a run-by-run basis. The use of SVX tracks on an event-by-event basis provides good resolution. The use of the average beam line on a run-by-run basis provides the fit with good stability, since the beam line information is not influenced by possible event pathologies. In the presence of MPVs, the fit is performed *around* the PV with the highest  $Q$ , where  $Q$  is defined as the scalar  $P_t$  sum of tracks attached to this vertex (see section 5.1.2). This choice is motivated by the fact that  $W +$  multijet events have a  $Q$  value much larger than soft  $p\bar{p}$  collisions with low track  $P_t$  and multiplicity (“minimum bias” events) which represent the majority of MPVs (see figures 5.1.a and 5.1.b). The fraction of events with  $\geq 2$  MPV’s is a function of the Tevatron instantaneous luminosity. For the average value of luminosity observed during the whole data taking period, this fraction was about 25 %. The average beam line and the calculation of the PV  $Q$  value are two important ingredients of the PVF algorithm; they are described in the two following sections.

### 5.1.1 Average $p\bar{p}$ Beam Line Position

The CDF interaction region has a length of about 30 cm. The average beam line is a straight line in the x-z, y-z planes:

$$x_{beam} = x_0 + a_x \cdot z_{beam}, \quad y_{beam} = y_0 + a_y \cdot z_{beam}, \quad (5.1)$$

where  $a_x, a_y \in [0, 5] \mu\text{m}/\text{cm}$  and  $x_0, y_0 \in [-1, 1] \text{ mm}$ .  $a_x, a_y, x_0, y_0$  are measured using SVX tracks from all events of a given run and exploiting the sinusoidal relationship between their  $D$ ’s and  $\phi$ ’s (see equation 5.6). From the fit to the impact parameter distribution vs. the azimuthal angle [106] the average beam line position can be extracted for any given run. The fit determines also the transverse beam profile, which turns out to be a circular gaussian with a sigma  $\sigma_{beam} = 36 \mu\text{m}$ . The average displacement of the beam line during a store is  $\Delta x_0 \sim 5 \mu\text{m}$  and  $\Delta y_0 \sim 11 \mu\text{m}$ . During a run  $\Delta x_0, \Delta y_0 < 4 \mu\text{m}$ , which is negligible with respect to the  $36 \mu\text{m}$  beam line width.

### 5.1.2 Multiple Primary Interactions

The longitudinal positions of MPV’s are measured by the VTX ( $Z_{VTX}, \sigma_{VTX}$ ). If there is no PV within the VTX acceptance ( $z \in [-143.5, 143.5] \text{ cm}$ ) the event is rejected. The  $i$ -th PV is characterized by the value of its  $Q(i)$ , defined as:

$$Q(i) = \sum_{j=1}^{T(i)} P_t(j), \quad (5.2)$$

where  $T(i)$  is the number of good quality CTC tracks satisfying the condition  $P_t(j) > 500 \text{ MeV}/c$  and  $\Delta z = |Z_{PV}(i) - z_0(j)| < 5 \text{ cm}$ . The MPV with the maximum  $Q$  value is used as a *seed* for the fit together with the reconstructed SVX tracks. The seed coordinates are:

$$\vec{X}_{seed} = (x_0 + a_x \cdot Z_{VTX}, y_0 + a_y \cdot Z_{VTX}, Z_{VTX}), \quad (5.3)$$

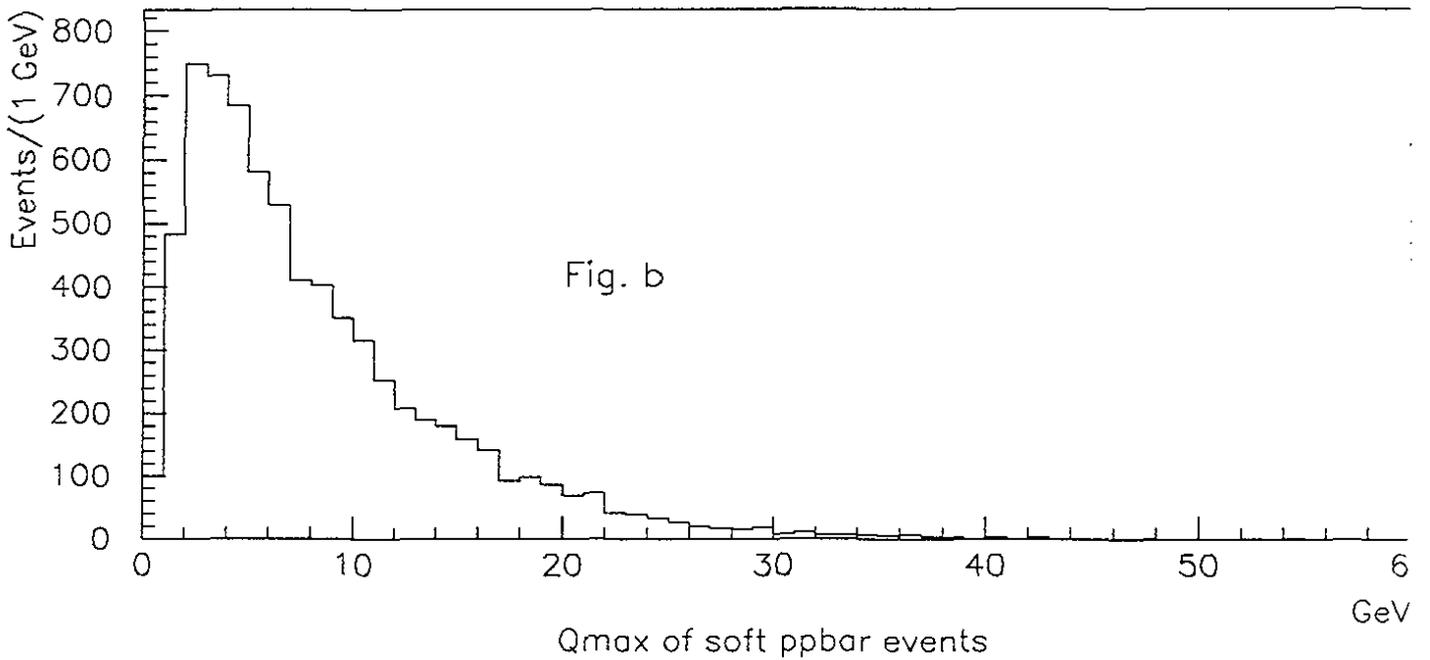
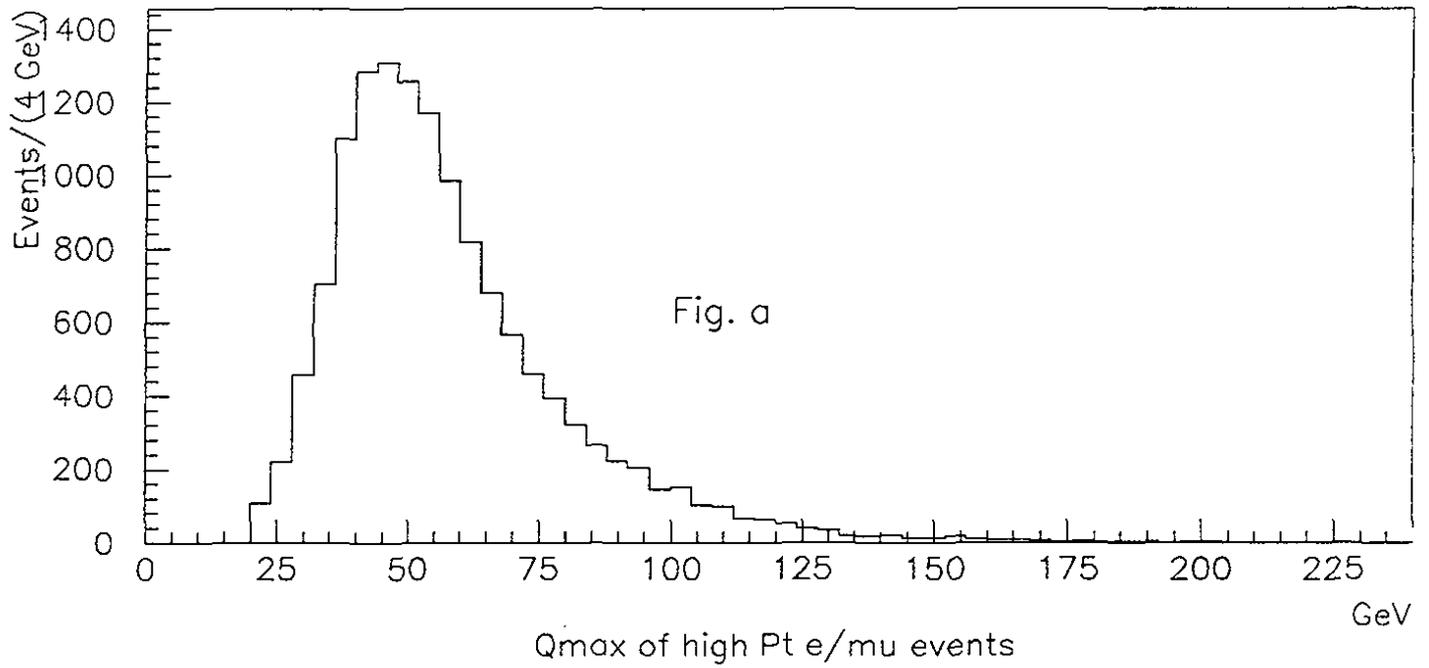


Figure 5.1: Distribution of the maximum value of the variable  $Q$  (see text) in events with electron and muons of  $P_t > 20$  GeV/ $c$  (fig. a) and in minimum bias events (fig. b).

while its covariance matrix is:

$$C_{seed} = \begin{pmatrix} \sigma_{beams}^2 & 0 & 0 \\ 0 & \sigma_{beams}^2 & 0 \\ 0 & 0 & \sigma_{VTX}^2 \end{pmatrix}. \quad (5.4)$$

Once the PV associated with the maximum- $Q$  primary interaction has been determined (with the algorithm described in the following section), the other primary interactions are considered in order to check that they do not contribute to the event topology under study. For example, in the  $W$  + multijet top-quark search one requires that the high  $P_t$  electron or muon and the tracks contained in the jets originated from the longitudinal region  $|\Delta z| < 5$  cm around the maximum- $Q$  PV. Events in which the  $W$  and the jets belong to distinct interactions are thus rejected.

The following sections describe the details of the PVF algorithm (which will be indicated from now on by the term VXPRI $M$ , from Ver $X$  PRI $M$ ary) and the results obtained with it.

### 5.1.3 The “VXPRI $M$ ” Primary Vertex Finding Algorithm

The VXPRI $M$  algorithm [107, 108, 109] is an iterative fit procedure, divided into three main steps.

**1. Initial Track Selection:** all three-dimensional SVX tracks with  $P_t > 400$  Mev/ $c$ ,  $\geq 3$  hit and  $|\Delta z| < 5$  cm are considered.

**2. Fit without variation (*steering*) of track parameters [102, 103, 104]:** tracks are approximated with straight lines tangent to their point of closest approach to  $\vec{X}_{seed}$  [105], and the point of minimum distance from the seed vertex and from the tracks,  $\vec{X}_0$ , is found by weighting the seed and the tracks with their respective covariance matrices. The residual of each track w.r.t.  $\vec{X}_0$ ,  $R_t$ , and the relative error,  $\sigma_{R_t}$ , are calculated and the track with the maximum value of  $R_t/\sigma_{R_t}$ , is rejected if  $R_t/\sigma_{R_t} > 3.5$ . This cut removes mismeasured PV tracks or non-PV tracks (like  $K_s \rightarrow \pi^+\pi^-$  and  $\Lambda \rightarrow \pi p$ ) which have high IP values. The fit is repeated (and  $\vec{X}_0$  updated) until all remaining tracks satisfy  $R_t/\sigma_{R_t} < 3.5$ . The value of the cut on  $R_t/\sigma_{R_t}$  has been established with studies on Montecarlo samples of  $p\bar{p} \rightarrow b\bar{b}$  events, in such a way to optimize the PV x and y resolution,  $\sigma_{PV}$ .

**3. Fit with steering of track parameters [102, 104]:** the tracks selected with step 2 are used for a more complex fit, where track parameters are varied in order to impose that all of them belong to a common vertex. This kind of fit allows for the determination of the coordinates of the common vertex and of the momenta of all tracks at that point. The initial value of the common vertex used by the fit is  $\vec{X}_0$  determined with step 2. The fit proceeds as previously by rejecting the track with the maximum value of  $R_t/\sigma_{R_t}$ , if  $R_t/\sigma_{R_t} > 3.5$ , and iterating the fit until  $R_t/\sigma_{R_t} < 3.5$  for all remaining tracks.  $\vec{X}_{PV}$  are the coordinates of the vertex found this way and  $\sigma_{PV}$  is the x-y coordinate resolution.

If in the event there are no SVX tracks satisfying the requisites of step 1, or if the iterative cut on track residual removes all tracks at steps 2 or 3, the above procedure is stopped and  $(\vec{X}_{seed}, C_{seed})$  are returned as a result.

### 5.1.4 Results on the High $P_t$ Electron and Muon Sample

The effectiveness of the PVF has been studied on all Montecarlo and CDF data samples relevant to the top-quark search, always giving consistent results. In this section we report results obtained on the inclusive  $P_t > 20$  GeV electron and muon sample (CDF data), from which  $W \rightarrow e\nu_e, \mu\nu_\mu$  events are selected. All figures shown and results quoted in the following refer to events where the number of tracks used in the PV fit,  $N_t$ , is  $\geq 1$ .

**Resolution.** Figure 5.3 shows the x and y resolution obtained with VXPRIM (top plot), and the distribution of  $\sigma_{beam}$  (bottom plot).  $\sigma_{PV}$  varies from 6 to 36  $\mu\text{m}$ ; the peak value of  $\sigma_{VP}$  is 13  $\mu\text{m}$ , the mean value is 20  $\mu\text{m}$ . The  $\sigma_{beam}$  distribution has a mean of 36  $\mu\text{m}$ ; events in which  $\sigma_{beam} = 50\mu\text{m}$  correspond to runs where the average beam line has been measured with low statistics. Figure 5.4 (top plot) shows that the gain in resolution ( $\sigma_{beam} - \sigma_{VP}$ ) is always positive: by combining even a single track with  $(\vec{X}_{seed}, C_{seed})$  one gets an improvement in resolution, that is the value of the xx, yy covariance matrix elements decreases after the fit. Figure 5.4 (bottom plot) shows that the gain in resolution increases with  $N_t$ . The average improvement in resolution obtained with VXPRIM w.r.t. the seed resolution (that is,  $\sigma_{beam}$ ) is about 44%.

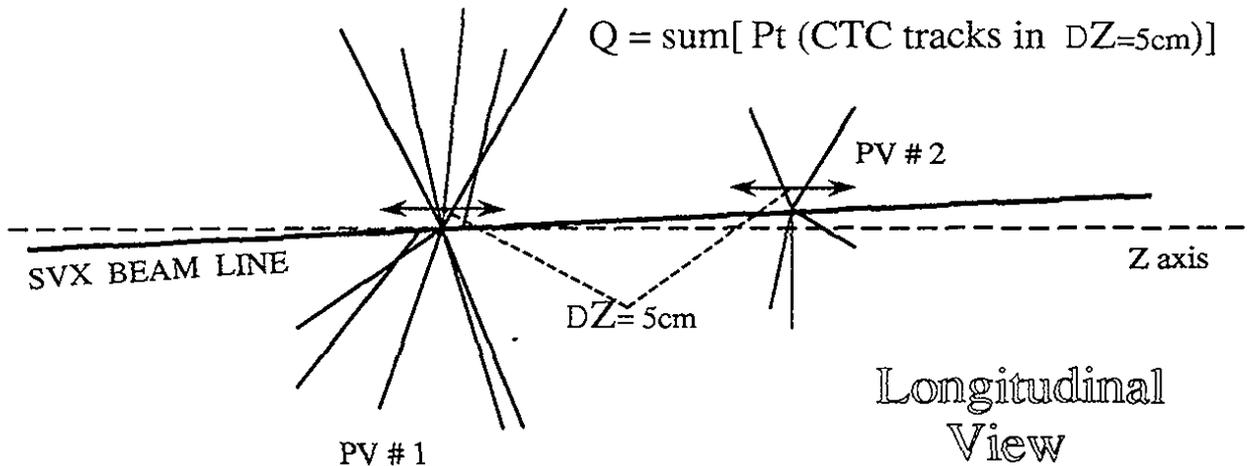
**Understanding of the Error Scale.** Figure 5.5 (top plot) shows the difference of the seed and the VXPRIM-found PV position in the x and y coordinates,  $\Delta X$ . The distribution is peaked around zero and has a width of about 27  $\mu\text{m}$ : the VXPRIM PV is basically located inside the beam envelope, but the high resolution SVX tracking allows to probe this envelope and to measure the position of a single primary interaction with an uncertainty lower than the size of the beam itself. Figure 5.5 (bottom plot) shows  $\Delta X$  divided by its error,  $\sigma(\Delta X) = \sqrt{\sigma_{beam}^2 - \sigma_{PV}^2}$ . If the error scale is understood correctly,  $\Delta X/\sigma(\Delta X)$  must be gaussianly distributed with a sigma = 1. This distribution is gaussian over about three orders of magnitude, and sigma = 0.984 (the statistical error on sigma is negligible): the PV error scale is understood correctly at the  $< 2\%$  level. Because of its relevance for  $b$  tagging, the effective understanding of the error scale has been further checked, with the following studies:

1. **FIT WITHOUT THE SEED VERTEX CONSTRAINT:** the VXPRIM algorithm has the flexibility to be run by using the seed as initial value of the fit but not as independent measurement which enters in the  $\chi^2$  minimized in the fit. While with the seed constraint,  $\sigma_{PV} < \sigma_{beam}$ , without this constraint  $\sigma_{VP}$  shows very long tails, to 100  $\mu\text{m}$  and above, due to events with few SVX tracks used for the fit. In addition, without the constraint,  $\sigma(\Delta X) = \sqrt{\sigma_{beam}^2 + \sigma_{PV}^2}$ , since the VXPRIM and the seed PV positions are uncorrelated. The results is that even in this case the  $\Delta X/\sigma(\Delta X)$  distribution indicates that the error scale given by the fit is correct at the 2 % level.
2. **STUDY OF THE BEAM PROFILE:** for a given run  $\sigma_{beam}$  can be measured using VXPRIM without the seed constraint (that is, the constraint of the average beam line), selecting events well contained in the SVX acceptance, selecting 4-hit tracks of  $P_t > 1$  GeV/c, and considering PV fits with  $\chi_{VXPRIM}^2 < 20$  (where  $\chi_{VXPRIM}^2$  is the  $\chi^2$  of the last iteration of the PV fit). Figure 5.6 shows: (1) the distribution of such PV's in the transverse  $x - y$  plane, (2) the projection of this distribution on the  $x$  axis, (3) the distribution of PV's in the  $x - z$  plane. Since well measured PV's have been

chosen ( $\sigma_{PV} \sim 15\mu\text{m}$ ) the width of the distribution projected on the  $x$  axis is basically dominated by  $\sigma_{beam}$ . A gaussian fit to this distribution gives  $\sigma = 40\mu\text{m}$ , which is consistent with  $\sigma_{beam} = 36\mu\text{m}$  and  $\sigma_{PV} = 15\mu\text{m}$ .

**Multiple Primary Vertices.** Figure 5.7 shows  $\Delta Z = Z_{VXPRIM} - z_0$  for electrons and muons of high  $P_t$ : the high  $P_t$  lepton is originated within  $|\Delta Z| < 5$  cm in  $> 99.7\%$  of the events. The events in the tails of the distribution ( $|\Delta Z| > 5$  cm) are cases in which there is an additional primary interaction with higher  $Q$ -value (perhaps in the form of jets) than the PV associated with the lepton. This fraction of events is  $< 0.3\%$ . In the  $W +$  multijet top-quark search, in order to avoid mixing jets and leptons produced by distinct MPV's, events with  $|\Delta Z| > 5$  cm are rejected. Events with  $|\Delta Z| < 5$  cm are then used in the analysis only if the high  $P_t$  lepton passes the trigger. If the maximum  $Q$  MPV is not used as seed of the VXPRIM fit, the fraction of events with  $|\Delta Z| > 5$  cm is about 10 % (see figure 5.8), to be compared with  $< 0.3\%$  when the maximum  $Q$  MPV is used as seed.

1. Use primary vertex with max Q



2. Compute (x,y) beam line position @ z of PV with max Q
3. Fit position of PV with max Q combining 1, 2 and SVX tracks in DZ

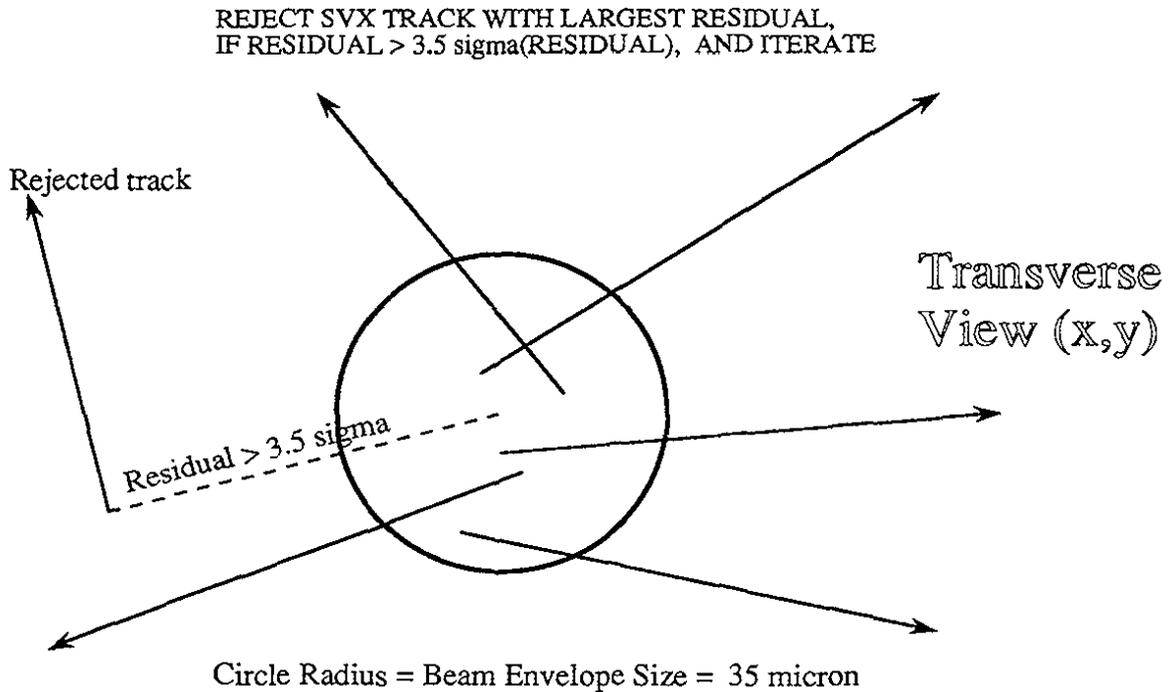


Figure 5.2: Schematic description of the VXPRIM primary vertex finding algorithm.

## *E/MU+JETS SAMPLE: X,Y RESOLUTION*

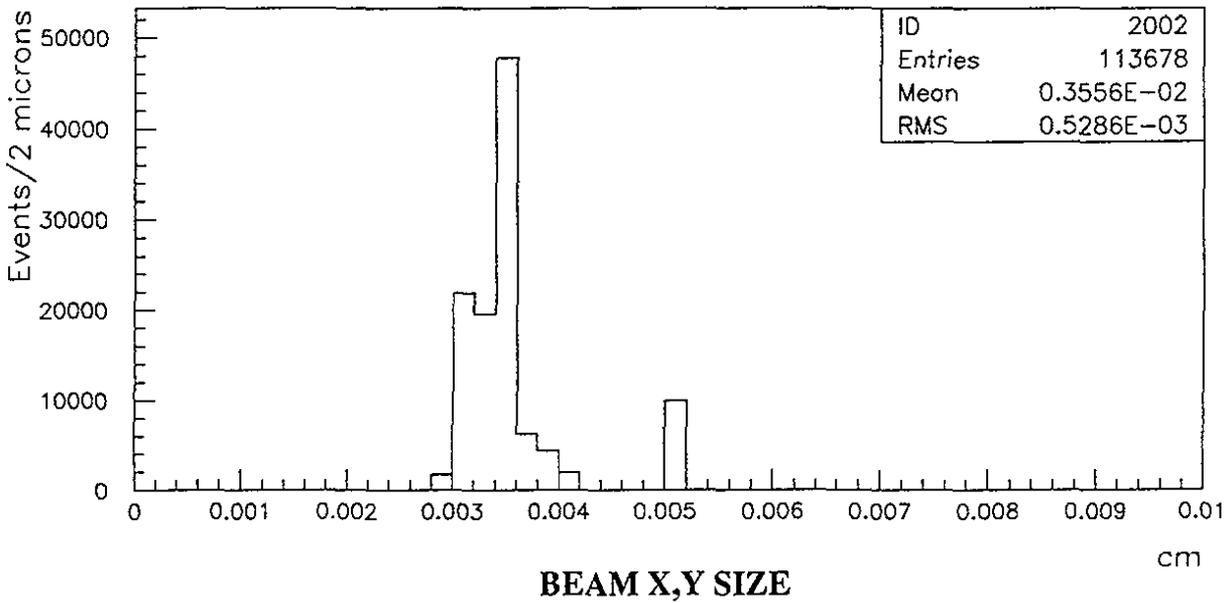
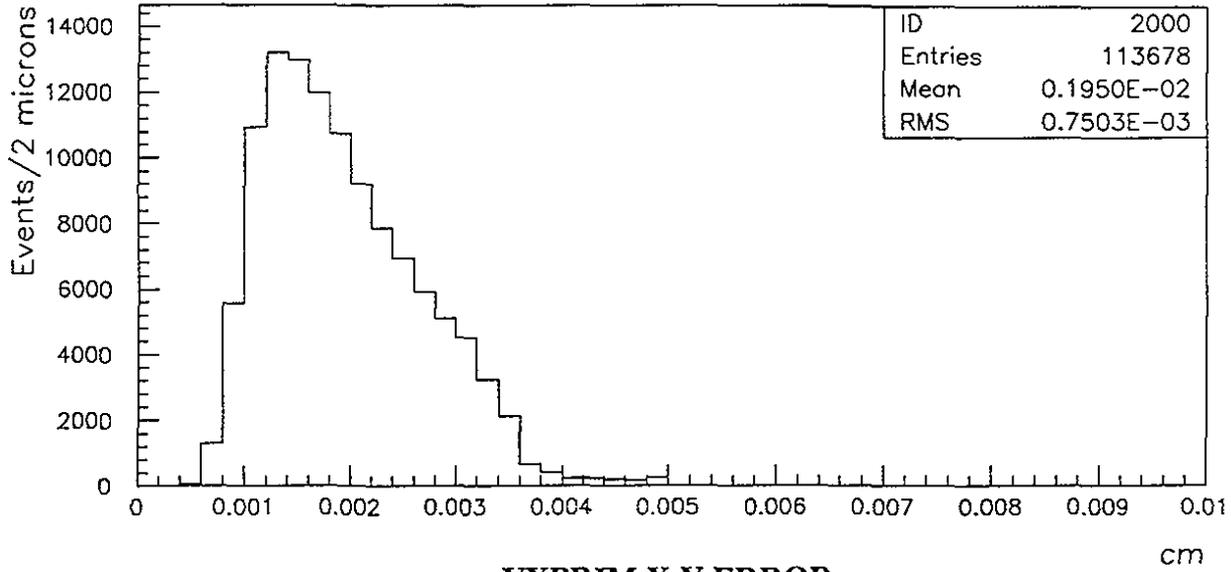
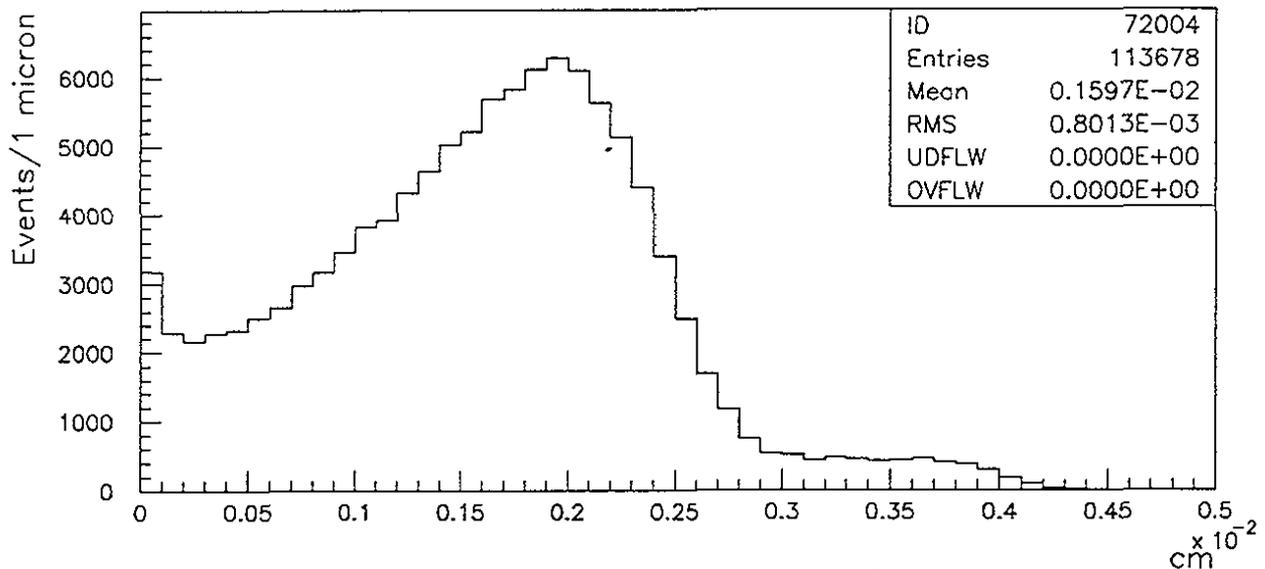
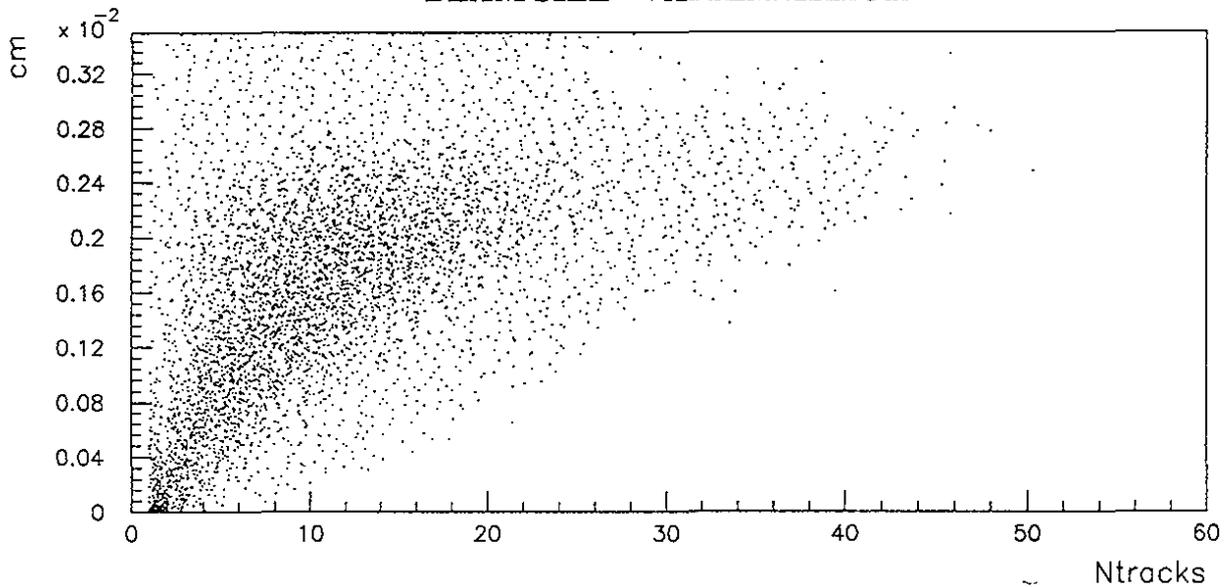


Figure 5.3: Resolution in the  $x$  and  $y$  coordinates for the primary vertex found by VXPRIM (top plot) and for the seed primary vertex (bottom plot). Each event enters the plot twice, once for the  $x$  and once for the  $y$  coordinate.

## ***E/MU+JETS: GAIN IN X,Y RESOLUTION***



**BEAM SIZE - VXPRIM ERROR**



**BEAM SIZE - VXPRIM ERROR VS NTRACKS**

Figure 5.4: Difference (improvement) between the VXPRIM and the seed primary vertex resolution in  $x$  and  $y$  (top plot); improvement in  $x$  and  $y$  resolution as a function of the number of tracks used in the primary vertex fit (bottom plot). Each event enters the plot twice.

## *E/MU+JETS: VXPRIM-BEAM X,Y PULLS*

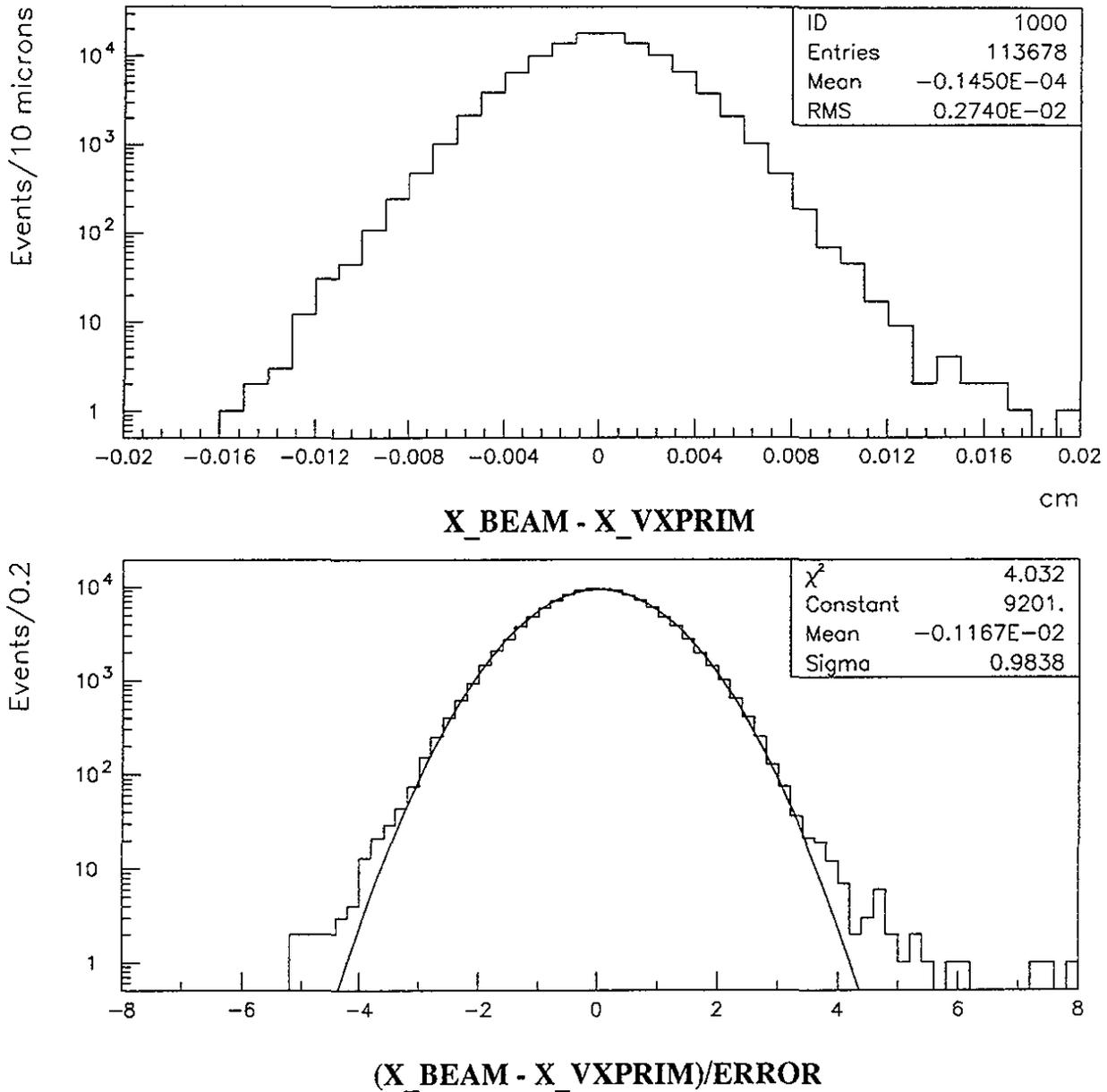


Figure 5.5: Difference between the VXPRIM and the seed primary vertex  $x$  and  $y$  coordinates (top plot); this difference normalized by its error (“pulls” distribution, bottom plot). Each event enters the plot twice.

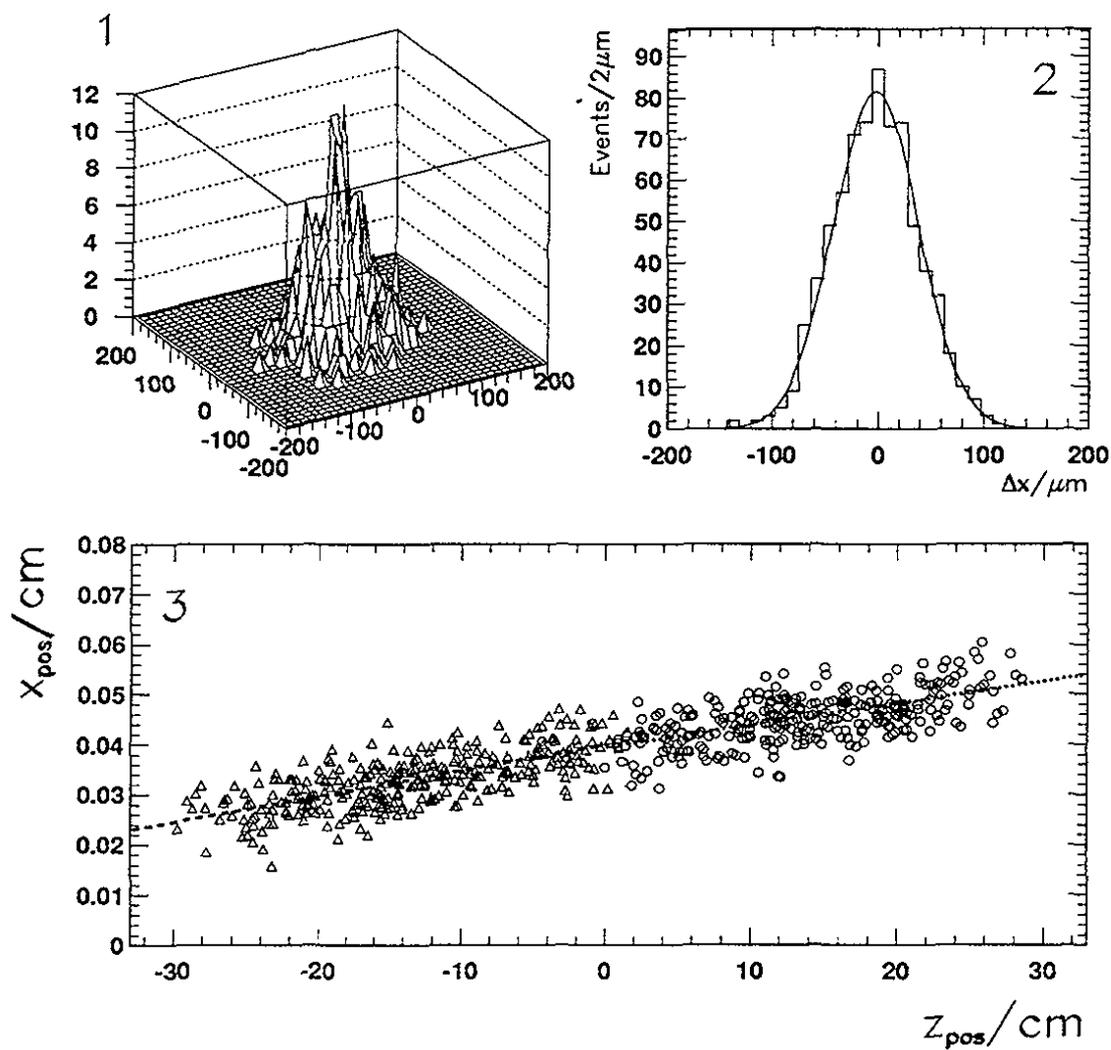


Figure 5.6: (1)  $x$ - $y$  profile of the colliding beams as reconstructed by VXPRIM (see text); (2) projection of the beam profile on the  $x$  axis; (3) beam profile in the  $x$ - $z$  plane.

## ***E/MU+JETS: HANDLING OF MULTIPLE PVs***

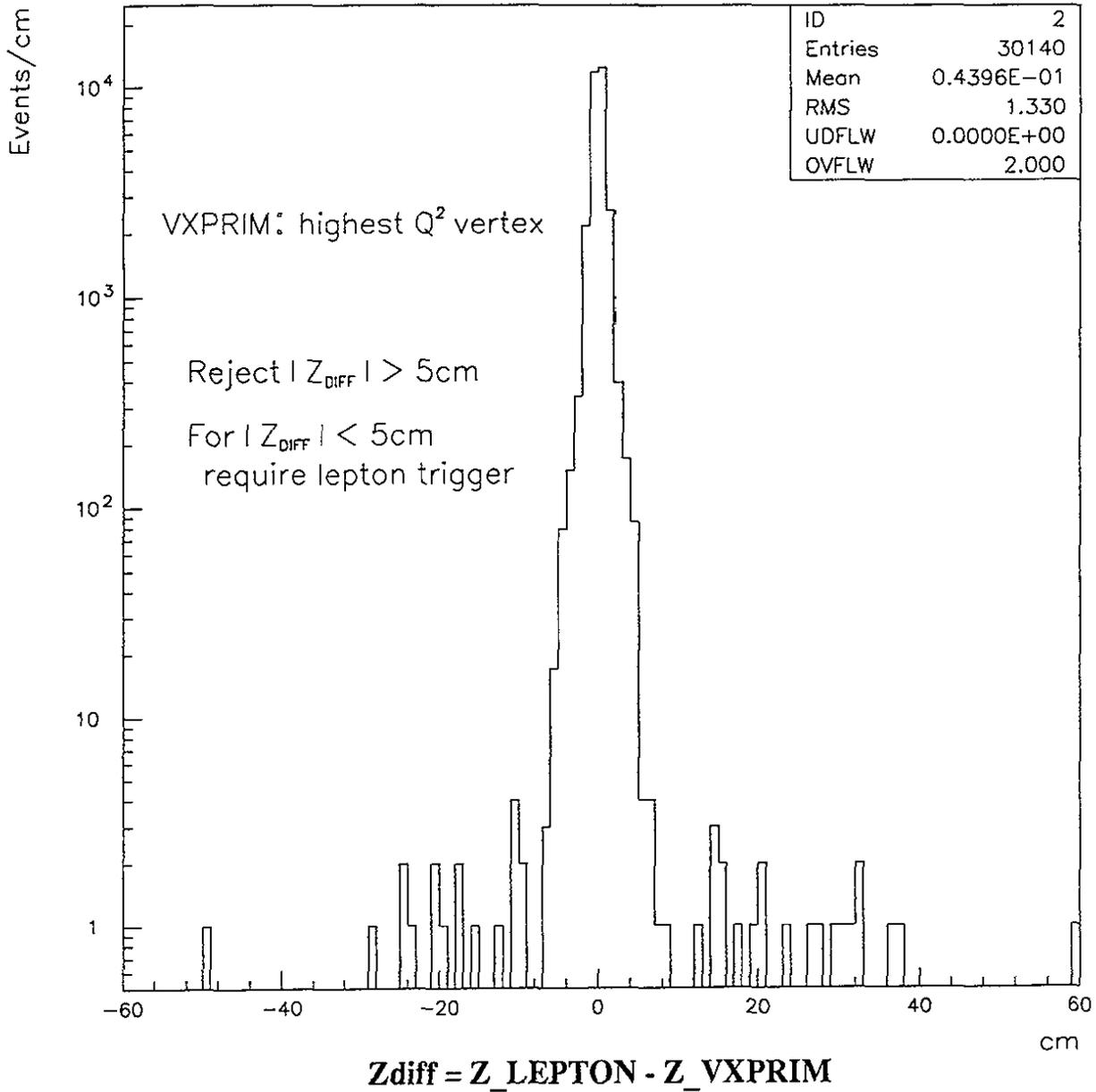


Figure 5.7:  $Z$ -distance between the primary vertex found by VXPRIM and the high  $P_t$  lepton ( $Z_{diff} = \Delta Z$ ); events with  $|\Delta Z| > 5$  cm are rejected as potential multiple primary vertex events;  $|\Delta Z| < 5$  cm events are then accepted only if the lepton passes the high  $P_t$  trigger.

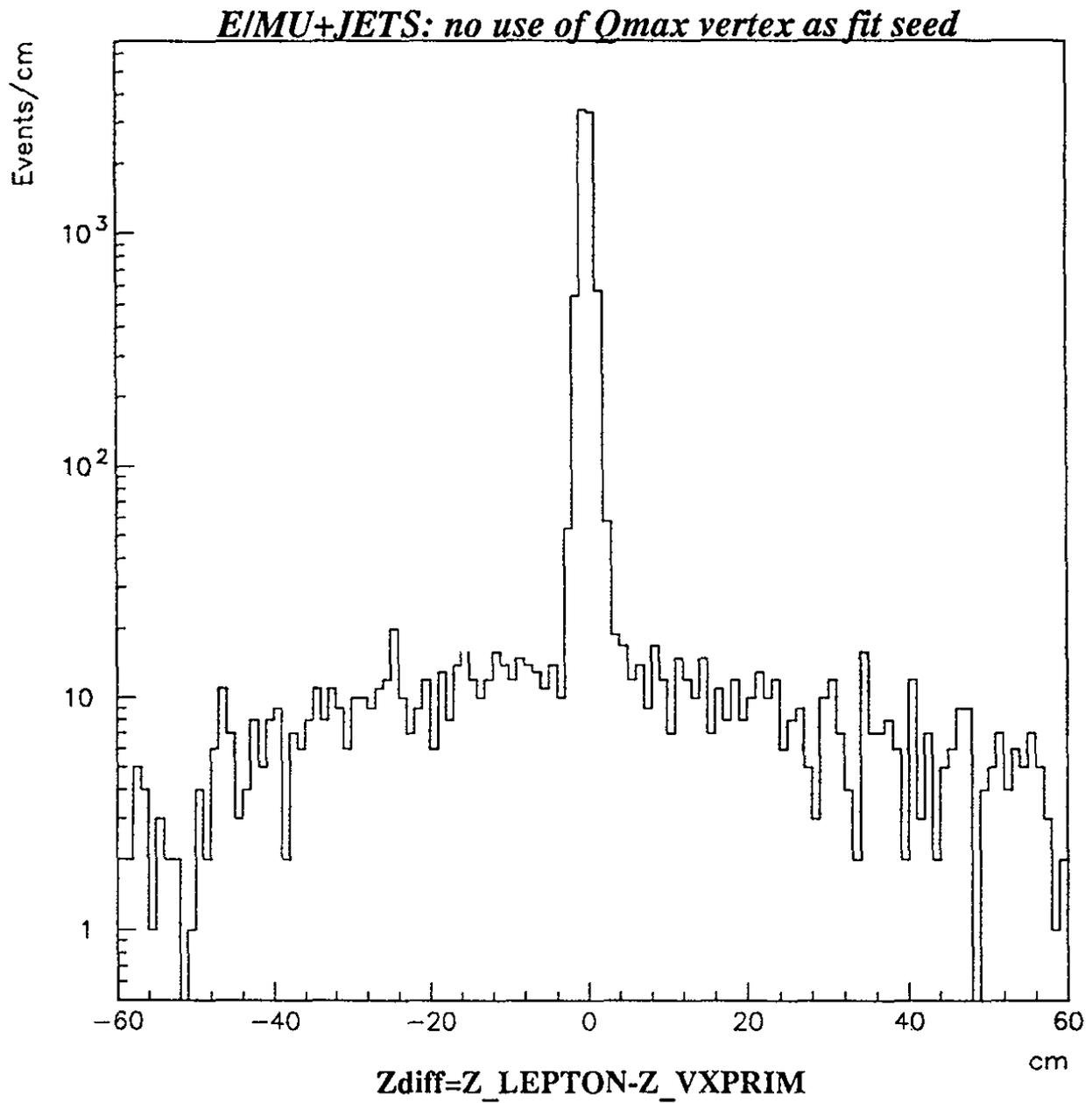


Figure 5.8:  $Z$ -distance between the primary vertex found by VXPRIM and the high  $P_t$  lepton ( $Z_{diff} = \Delta Z$ ), when the maximum- $Q$  primary interaction is not used as a seed for the VXPRIM fit.

## 5.2 Tagging Secondary Decay Vertices

The search for SV's exploits heavily the good SVX track IP resolution and the precise PV determination provided by VXPIM. The SV search algorithm described in this section has been developed with the goal of identifying  $b$ -hadron decays, whose lifetime is  $c\tau_B = 446 \pm 26 \mu\text{m}$  [50] (averaged over all  $b$ -hadron states). A complete characterization of the algorithm effectiveness as a tool for  $b$  tagging is given in the next chapter.

### 5.2.1 Impact Parameter-Azimuth ( $D - \phi$ ) Space

The first step in the SV search is taking the PV found by VXPIM as origin of the coordinates system and calculating the track parameters w.r.t. this vertex. The track equation 4.3, can be rewritten as:

$$\left( \frac{CR^2 + (1 + CD)D}{1 + 2CD} \right) = -R \sin(\phi_0 - \phi(R)), \quad (5.5)$$

where  $R$  is now the distance from the PV in the transverse plane.

Let us consider  $P_t > 400$  MeV/ $c$  tracks (which do not spiral in the CTC) originated from  $b$ -hadrons decays within the first SVX layer, that is, at a distance  $R < 2.8$  cm: one finds that  $|CD| \ll 1$  and  $|C|R^2 \ll |D|$ . For this kind of tracks/decays, the track equation becomes:

$$D = -R_B \sin(\phi - \phi_B), \quad (5.6)$$

where  $R_B$ ,  $\phi_B = \phi(R = R_B)$  are, respectively, the transverse decay distance and the azimuthal flight direction of the  $b$  hadron, and  $D$ ,  $\phi = \phi_0$  are the track parameters of the daughter particles (the  $b$  "prongs"). The SV  $x$  and  $y$  coordinates are:

$$(R_B \cdot \cos(\phi_B), R_B \cdot \sin(\phi_B)). \quad (5.7)$$

For  $b$  hadrons of sufficiently high  $P_t$ , the  $b$  prongs are collimated around  $\phi_B$ , and the relationship between  $D$  and  $\phi$  is linear.

The IP resolution,  $\sigma_D$ , is:

$$\sigma_D^2 = \sigma_{D_0}^2 + \sigma_{yy}^{(PV)} \cos^2 \phi + \sigma_{xx}^{(PV)} \sin^2 \phi - 2\sigma_{xy}^{(PV)} \cos \phi \sin \phi, \quad (5.8)$$

where  $\sigma_{D_0}^2$  is the track IP resolution w.r.t. the old reference system (the center of CDF) and  $\sigma_{xx}^{(PV)}$ ,  $\sigma_{xy}^{(PV)}$ ,  $\sigma_{yy}^{(PV)}$  ( $\text{cm}^2$ ) are the (1,1), (1,2), (2,2) elements of the PV covariance matrix.

### 5.2.2 Track Selection

The second step in the SV search is the selection of good quality tracks incompatible with the PV and inconsistent with  $K_s \rightarrow \pi\pi$  and  $\Lambda \rightarrow \pi p$  decays. This is achieved with the following cuts:

- A)  $P_t > 400$  MeV/ $c$  (the track crosses the whole CTC radius).

- B)  $(\chi_{SVX}^2 - \chi_{CTC}^2)/N_{hit} < 6$  (well measured track, see section 4.4.2).  $N_{hit} \geq 3$ ; among these, at least 2 hits must not be shared with other tracks, must not contain malfunctioning strips and must have a charge profile consistent with that of a single particle. Studies using inclusive samples of jet triggers show that 96 % of SVX tracks pass the chi-squared cut and 78 % pass the requirement of two such hits. Tracks passing the above selection with hits shared with nearby tracks are reconstructed again, by assigning to the shared hits the position of the geometric center of the hit (which, due to the track sharing, is usually made up by 3 or more strips), instead of the center of gravity of the strip charges, and a resolution equal to the geometrical width of the total number of strips contained in the hit (that is, the pitch of each strip multiplied by the total number of strips). This kind of reconstruction of tracks with shared hits represents a very conservative choice, which allows to keep the tagging rate for fake SV's to very low levels even in very dense tracking environments (see section 6.2.2).
- C)  $|Z_{PV} - z_0| < 5$  cm (track consistent with the primary  $p\bar{p}$  found by VXPTRIM).
- D)  $S_D = |D|/\sigma_D > 2.5$  (track incompatible with the PV in the transverse plane). If the track is recognized as an electron or muon of  $P_t(lepton) > 2$  GeV/c by the SLT algorithm, the cut on IP significance  $S_D$  is removed, since leptons represent by themselves a good tag of  $b$  hadrons. In addition, the requirement  $P_t(lepton) < 20$  GeV/c is applied, to avoid overlaps with the dilepton channel;
- E)  $K_s \rightarrow \pi\pi$  and  $\Lambda \rightarrow \pi p$  decays are removed in the following way:
- 1) track pairs consistent with  $M_{K_s}, M_\Lambda$  (within 3 sigma) are rejected if they originate at the PV (that is, if  $S_D(K_s) < 2.5$  or  $S_D(\Lambda) < 2.5$ );
  - 2)  $|D| < 0.2$  cm, to eliminate  $K_s$  and  $\Lambda$  decays not removed by the previous cut (for example when one of the two daughter particles is not reconstructed) and mismeasured tracks; the efficiency loss on the  $b$  prongs is negligibly small, given the value of  $c\tau_B$ .

### 5.2.3 The “ $D - \phi$ ” Secondary Vertex Tagging Algorithm

The third step of the SV search is the identification of  $\geq 3$  tracks consistent with a common vertex in the transverse plane and such that the direction of their total momentum is contained in the  $\eta - \phi$  cone of a jet. This identification exploits equation 5.6: tracks from a common vertex at a distance  $R_B$  from the PV, and azimuthal angle  $\phi_B$ , belong to the same sinusoid in the space of their parameters  $D - \phi$ . In this space, PV tracks are grouped around  $D = 0$  ( $\phi$  axis), with deviations due to the finite value of their IP resolution  $\sigma_D$ . If the event contains a SV, there can be tracks significantly displaced from the  $\phi$  axis, depending on the relative magnitude of  $\sigma_D$  and  $R_B$ . The significance of this deviation,  $S_D$ , depends of the track  $P_t$  since: (1) the average IP value of  $b$  prongs,  $\langle |D| \rangle$ , does not change with  $P_t$ , because  $R_B \propto P_t$  and  $|\phi - \phi_B| \propto 1/P_t$ , which makes the IP a suitable observable to study secondary decay vertices; (2)  $\sigma_D$  increases sizeably when  $P_t$  decreases in the range [0.4, 2.0] GeV/c, where the multiple scattering contribution is large (see figure 4.6). The situation is much more favorable when  $P_t(B)/M_B \gg 1$ :  $R_B \gg c\tau_B$ , the  $b$ -prong  $P_t$  and  $S_D$  increases

and they are more likely to be emitted along  $\phi_B$  (forward decays). In the  $D - \phi$  plane,  $b$  prongs tend to populate the sinusoid segment around  $\phi_B$  with negative slope, with a small dispersion angle. In the vicinity of  $\phi_B$ , equation 5.6 can be approximated as a straight line of angular coefficient  $-R_B$ , which intercepts the  $\phi$  axis in the point  $(\phi_B, 0)$ :

$$D \sim -R_B(\phi - \phi_B). \quad (5.9)$$

This approximation is good in  $t\bar{t}$  events where, for  $M_{top} > 91 \text{ GeV}/c^2$  [35, 36], one expects large  $P_t(B)$  [54]. In order to identify secondary vertices, tracks are considered in the  $D - \phi$  representation and the linear relationship of equation 5.9 between IP and azimuth is imposed. This is achieved with the following algorithm:

- 1) All possible segments (“links”) of pairs of tracks passing the cuts listed in the previous section are formed, and for each of them the following quantities are defined: the intercept of the link with the  $\phi$  axis,  $\phi_L$ , the angle between the link and the  $\phi$  axis,  $\theta_L$  ( $0 < \theta_L < \pi$ ), the pseudorapidity associated with the momentum vector equal to the sum of the momenta of the two particles of the link,  $\eta_L$ ;  $\phi_L$  and  $\tan(\theta_L)$  are estimators of  $\phi_B$  and  $R_B$ ; using these variables, the following kinematical cuts are applied:
  - $\pi/2 < \theta_L < \pi$ : links must have negative angular coefficient, to be consistent with forward decays;
  - $|\phi_L - \phi(i)| < \pi/2$ , where  $\phi(i)$  ( $i=1,2$ ) are the azimuths of the two tracks of the link; this cut is necessary to select forward decays, collimated around the direction  $\phi_B$  ( $\phi_B - \pi/2 < \phi(i) < \phi_B + \pi/2$ );
  - $|\eta(1) - \eta(2)| < 2 \times B_{cone}$ , where  $\eta(i)$  ( $i=1,2$ ) are the pseudorapidities of the two tracks of the link and  $B_{cone} = 0.8$  (see point 3 below); this cut ensures that the tracks from the SV are relatively close also in  $\eta$ .
- 2) All possible pairs of links are considered and at least one such pair is required whose links  $i$  and  $j$  are “aligned” in  $D - \phi$  and close in  $\eta$ , namely the two candidate links satisfy the following cuts:
  - $|\phi_L(i) - \phi_L(j)| < 0.15$ ,
  - $|\theta_L(i) - \theta_L(j)| < 0.15$ ,
  - $|\eta_L(i) - \eta_L(j)| < 2 \times B_{cone}$ ,

Groups of  $\geq 2$  links (that is,  $\geq 3$  tracks) passing these cuts represent a “ $D - \phi$  cluster” and a candidate SV. Figure 5.9 graphically illustrates a  $D - \phi$  cluster and all the variables used in its definition. Figure 5.10 shows two clusters found in a CDF events which fired the  $P_t > 12 \text{ GeV}/c$  electron trigger. Each point of the figure represents a track with the vertical error bar indicating the value of  $\sigma_D$ . The tracks used in the PV determination ( $S_D < 3$ ) are distributed along the  $\phi$  axis between 0 and  $2\pi$ ;  $S_D > 2.5$  tracks of the two clusters are connected by links; the azimuths of the two clusters (averaged over the  $\phi_L$  of their links) differ by  $\sim \pi$ ; in the calorimeter, two jets correspond to the direction of the two clusters; the electron track carries most of the  $P_t$  of the respective cluster/jet and has  $S_D \sim 0$ . This event is a good  $p\bar{p} \rightarrow b\bar{b}$  candidate, where one of the two  $b$ ’s decays semileptonically to the electron candidate observed.

- 3) The total momentum of each cluster is computed (sum of the momenta of its tracks) as well as the respective centroid in pseudorapidity and azimuth,  $(\eta_C, \phi_C)$ . The track with the maximum distance from the cluster centroid is searched for:

$$\Delta R(i) = \sqrt{(\eta_C - \eta(i))^2 + (\phi_C - \phi(i))^2}, \quad (5.10)$$

and such a track is removed from the cluster list if  $\Delta R(i) > B_{cone} = 0.8$ . Note that  $\Delta R(i)$  is a relativistically invariant quantity. This procedure is repeated until  $\Delta R(i) < 0.8$  for all tracks  $i$ . This cut is consistent with the requirement that all tracks originate from the same space point (the SV) and in addition it ensures that they are emitted in a cone of radius  $B_{cone} = 0.8$  ( $\sim 46^\circ$  in azimuth, around the direction of their total momentum).

- 4) For each cluster, points 3 and 4 are repeated until the list of its tracks becomes stable and the cluster is confirmed ( $\geq 2$  links) or rejected ( $< 2$  links). A confirmed cluster is now a  $D - \phi - \eta$  cluster, obtained using  $R_B$  and  $\phi_B$  estimators, without explicitly calculating them, as is customary in a standard vertex fit procedure. However, to make sure that the cluster is consistent with a decay occurring inside the SVX, one performs a fit to a vertex common to the tracks of the cluster, and the point thus found is required to lie inside the first SVX layer ( $R_B < 2.8$  cm). This is the *only* analysis cut based on the SVX fit results, and it is introduced to remove fake SV's (due to tracking errors) and interactions in the first silicon layer, located at 3.00 cm.

A  $D - \phi - \eta$  cluster which satisfies the above cuts is a “ $b$  tag” according to the  $D - \phi$  algorithm, that is, a set of tracks incompatible with the PV and consistent with a distinct SV. In the next chapter, it will be shown that the PV-SV distance of  $D - \phi$   $b$  tags is consistent with  $c\tau_B$  (that is, consistent with containing a very large fraction of  $b$  hadron decays).

High  $P_t$  tertiary vertices (like  $c$ -quark decays in the cascade  $t \rightarrow b \rightarrow c$ ) are often included in the clusters. However, the tagging efficiency for  $c$ -hadron decays from direct production,  $p\bar{p} \rightarrow c\bar{c}$ , is highly depressed w.r.t. the  $b$ -tagging efficiency, because: (i) the average charged multiplicity of  $c$ -hadron decays is  $\sim 3$ , compared to  $\sim 5$  for  $b$ -hadron decays; (ii)  $c$  prongs have on the average a lower  $P_t$  than  $b$  prongs; (iii)  $c\tau_D \sim 1/3 c\tau_B$ .

## 5.2.4 Jets and Leptons Associated with Secondary Vertices

The case in which jets and/or leptons (electrons or muons) are associated to the SVs is of particular interest. The identification of electrons and muons of high (low)  $P_t$  is described in detail in section 7.1 (in reference [52]). The identification of jets is described in section 6.1.2. The fourth and last step of  $b$  tagging is the search for jets and leptons associated with the  $b$  tag.

In  $t\bar{t}$  events one expects that  $b$  prongs deposit enough energy in the calorimeter to be reconstructed as a jet. Therefore, in the top-quark search one requires at least one jet of transverse energy  $E_t > 15$  GeV (see chapter 7.2) such that

$$\sqrt{(\eta_C - \eta_{jet})^2 + (\phi_C - \phi_{jet})^2} < R_{jet}, \quad (5.11)$$

where  $R_{jet}$  is the radius of the jet reconstruction cone in the  $\eta - \phi$  space.  $b$  tags which do not satisfy this requirement are rejected for the top-quark search.

Electrons and/or muons associated to SVs strengthen the evidence for  $b$ -hadron decays. An electron and/or muon is considered part of the  $b$  tag if:

- the corresponding SVX track is part of the cluster;
- the corresponding SVX track is not part of the cluster, but it lies within the cone of radius  $B_{cone}$  around  $(\eta_C, \phi_C)$ , and in the  $D - \phi$  plane it lies in a “road” of width  $\Delta\phi = 0.15$  and length = 0.2 cm around the cluster (see figure 5.9).

Figure 5.11 shows a CDF event with a  $D - \phi$  algorithm  $b$  tag having a jet and an electron associated with it. The figure shows a region in the transverse plane of radius  $R = 2$  cm around the PV. Tracks are represented with segments of length proportional to their  $P_t$  (tracks with  $P_t \geq 5$  GeV/ $c$  have, however, all the same length). Tracks belonging to the  $b$  tag are drawn with: (1) solid lines from their point of closest approach to the SV towards the SVX hits; (2) dashed lines from their point of closest approach to the SV towards the point of closest approach to the PV. The SV  $(x,y,z)$  coordinates are indicated in the top right corner of the figure. All tracks in the event (whether they have been used in the PV or SV fit, or they have been discarded by both fits) are represented by using the track parameters measured by the SVX (after proper translation of the origin to the PV); they are not represented by using the track parameters obtained from the fits (in which tracks are forced to belong to common primary or secondary vertices). In addition, tracks are shown in the  $D - \phi$  representation in the plot reported in the bottom left corner of the figure: tracks in the  $b$  tag are indicated with crosses, while other tracks are indicated with points. The  $P_t = 15.7$  GeV/ $c$  positron which fired the trigger is part of the  $b$  tag, and has  $|D| = 867 \mu\text{m}$ . The  $b$  tag invariant mass is  $\sim 4.6$  GeV/ $c$ , assuming the pion mass for all tracks except for the positron. The associated jet has  $E_t \sim 35$  GeV.

Figure 5.13 shows an analogous blow-up of the region (inside the circle) of radius  $R \leq 0.5$  cm for an event with two muons of opposite sign and with a  $b$  tag of  $R_B \sim 750 \mu\text{m}$  which contains also the two muons. This SV is a candidate for the exclusive decay:

$$B^0 \rightarrow J/\Psi K^{*0} \rightarrow \mu^+ \mu^- K^+ \pi^- . \quad (5.12)$$

The recoiling  $b$  tag is a candidate for the decay of the other possible  $b$  hadron of the event.

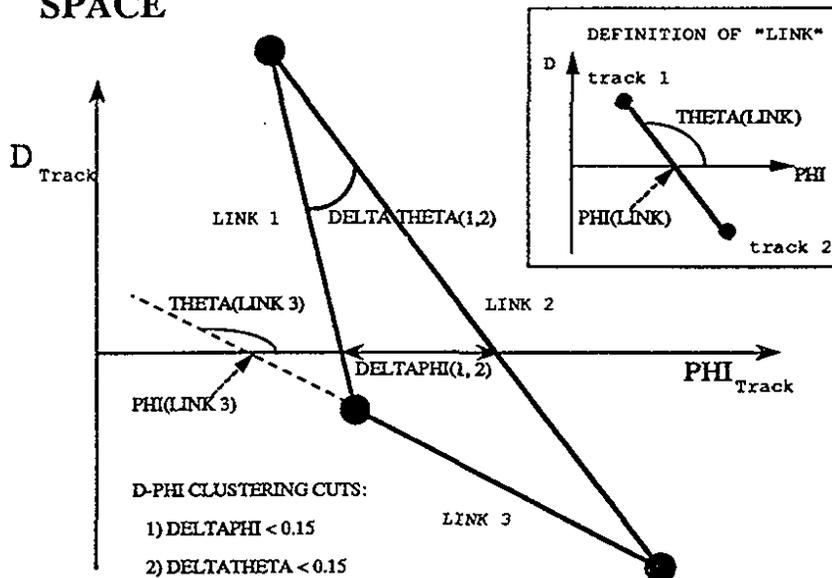
### 5.2.5 Backward Decays

The frequency of tagging errors of the  $D - \phi$  algorithm, that is, the fraction of  $b$  tags due to tracking errors and not to real decays of  $b$  or  $c$  hadrons (“heavy flavors”, HF), is one of its most important characteristics which will be studied quantitatively in section 6.2.1. From a technical point of view, in order to study the frequency of tagging errors, one defines the  $b$  tags associated to *backward* decays, in which the vector connecting the PV and SV in the transverse plane ( $\vec{L}_{xy} = \vec{X}_{SV} - \vec{X}_{PV}$ ) forms a  $> 90^\circ$  angle with the direction of the total transverse momentum of the tracks in the SV,  $P_{SV}$ . In the procedure for building  $D - \phi$  clusters described in the section 5.2.3, point 1, this is achieved by changing the cut

$\pi/2 < \theta_L < \pi$  (forward decays) to  $0 < \theta_L < \pi/2$  (backward decays), that is requiring that links have a positive angular coefficient in the  $D - \phi$  plane. Such  $b$  tags are called *negative  $b$  tags*, or negative decay length  $b$  tags.  $b$  tags discussed previously, associated to forward decays, which are the most likely configurations for HF decays of sufficiently high  $P_t$ , are called *positive  $b$  tags*, or positive decay length  $b$  tags.

In the following chapter the  $D - \phi$  algorithm will be fully characterized, by measuring its  $b$ -tagging efficiency and its frequency of tagging errors.

### 3 TRACK - 3 LINK CLUSTER IN D-PHI SPACE



### ASSOCIATION OF ELECTRONS AND MUONS TO CLUSTERS

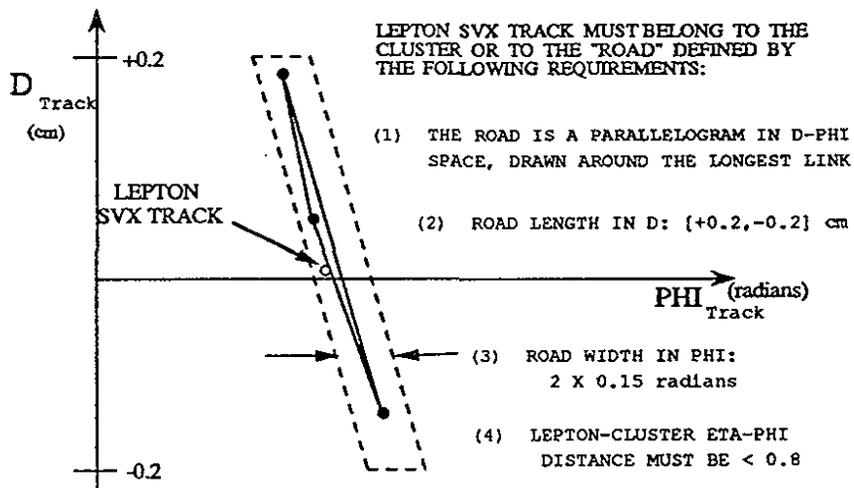


Figure 5.9: Graphic illustration of a  $D - \phi$  cluster with three tracks and three links; also shown is the road of association of electrons and muons with  $D - \phi$  clusters.

D-PHI CLUSTERING B TAGGING ALGO.

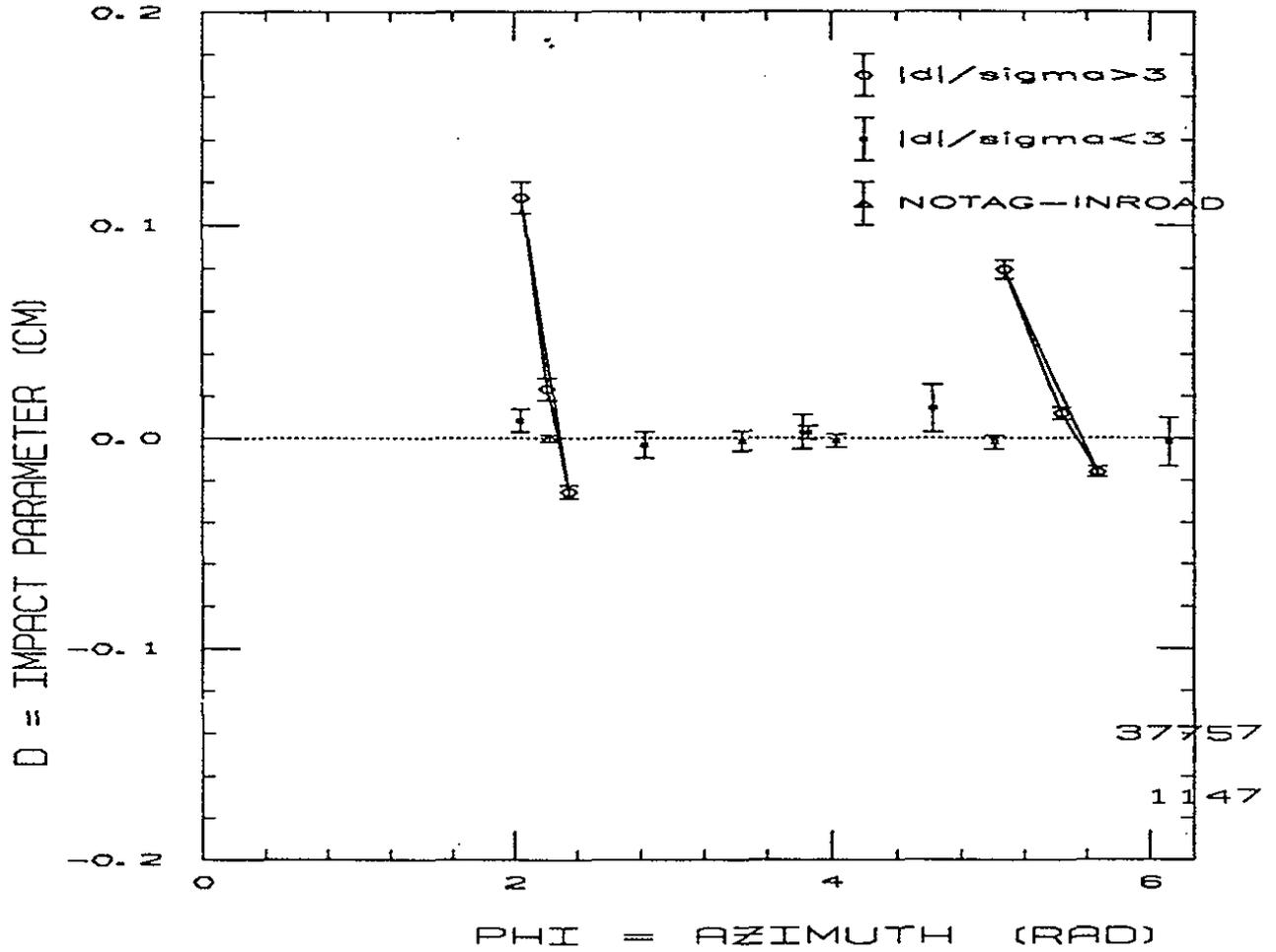


Figure 5.10:  $D - \phi$  plot for an event with two three-track  $b$  tags. A  $P_t = 13.6$  GeV/ $c$  electron is associated with the cluster at  $\phi \sim 2.2$  radians (triangular symbol labelled “notag, in road”); this electron does not belong to the links (“notag”), because it has a low value of the impact parameter, but it lies in the road of cluster-to-lepton association (“inroad”).

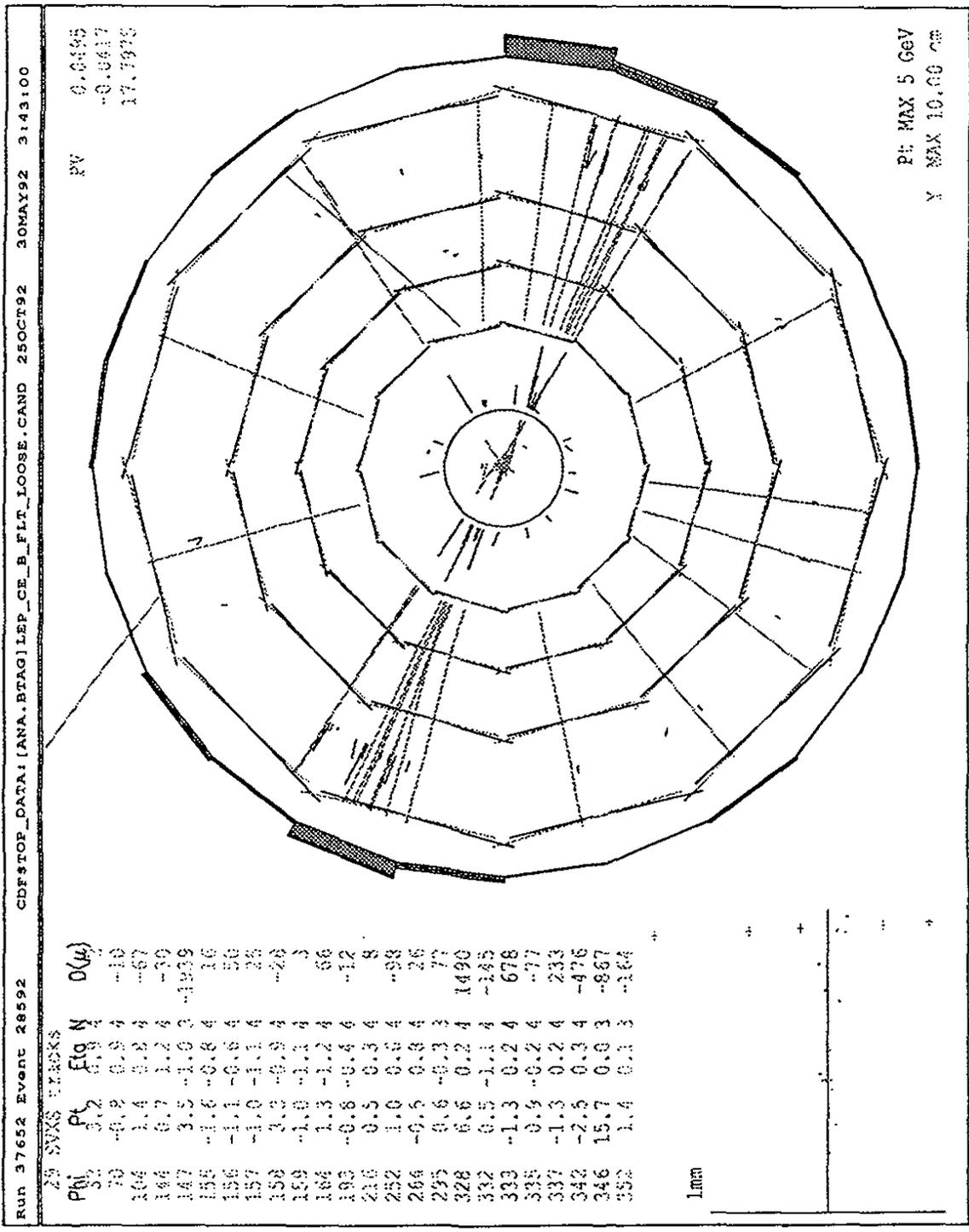


Figure 5.11:  $R - \phi$  view of an event with a five-track  $b$  tag, candidate for the semileptonic decay  $b \rightarrow e\nu X$ .

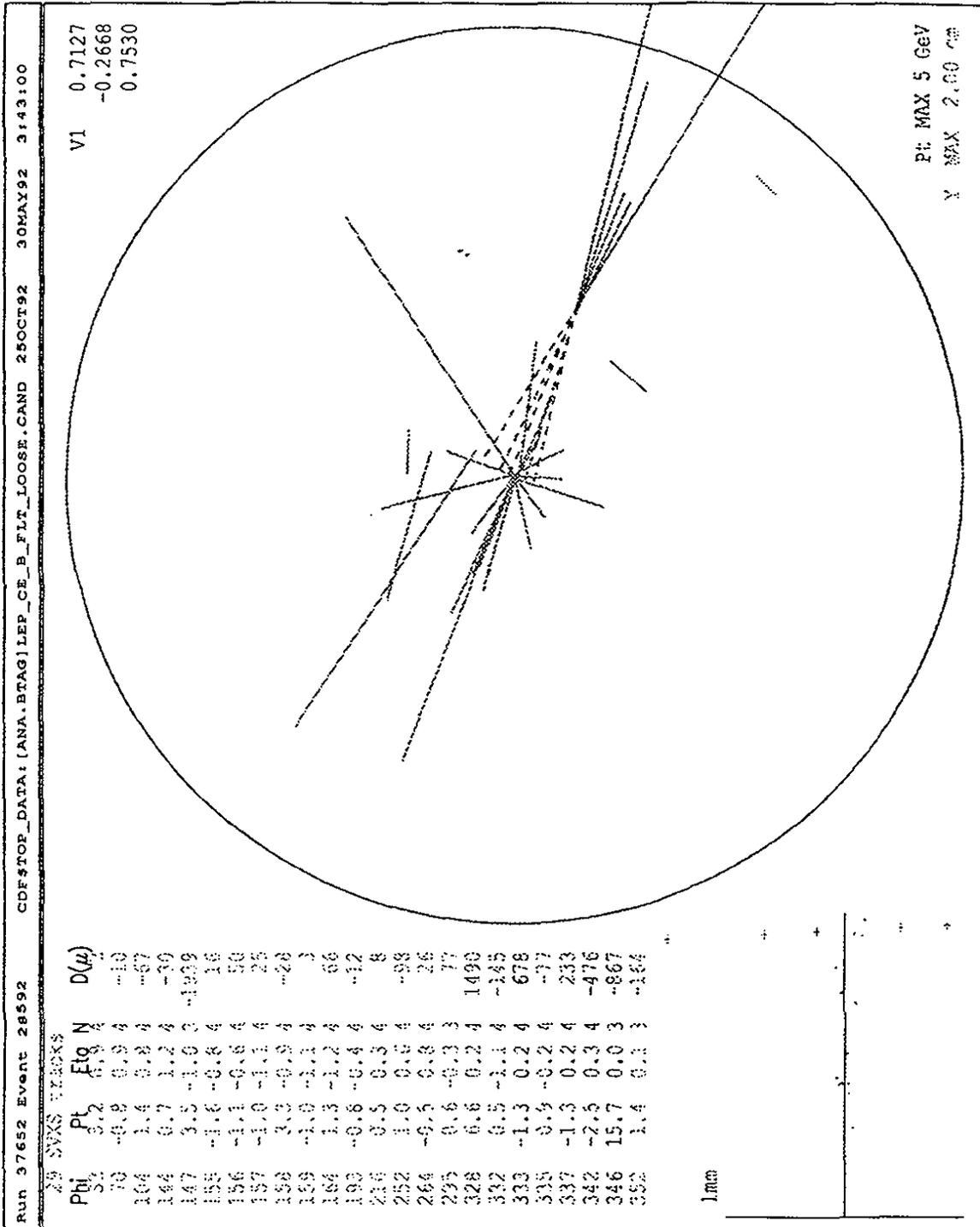


Figure 5.12: Magnification of the candidate  $b \rightarrow e\nu X$  event in a region of radius = 2 cm around the primary vertex.

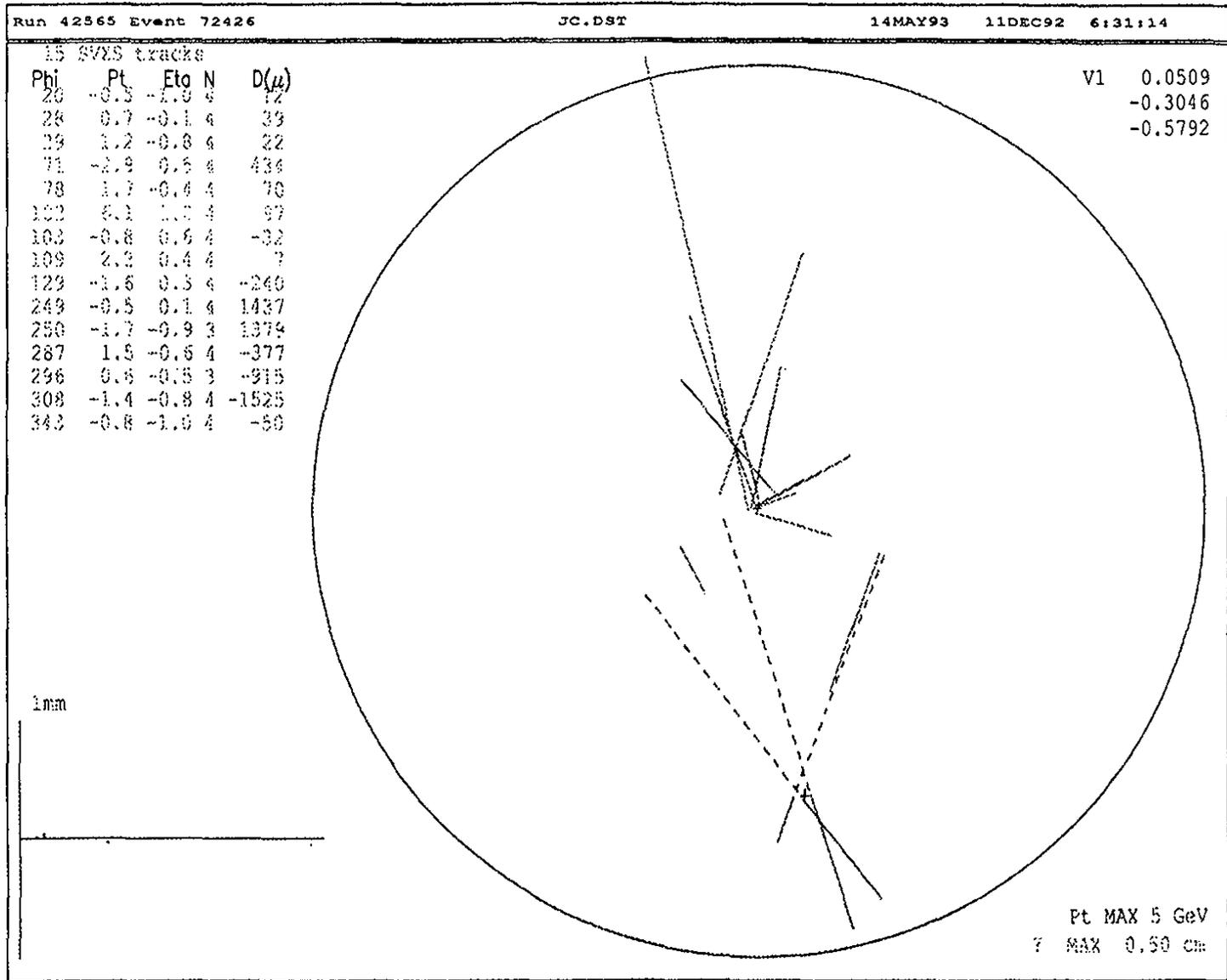


Figure 5.13: Magnification of  $B^0 \rightarrow J/\Psi K^{*0} \rightarrow \mu^+ \mu^- K^+ \pi^-$  candidate event in a region of radius = 2 cm around the PV.

# Chapter 6

## Characterization of the $D - \phi$ Algorithm As a $b$ Tagger

The  $D - \phi$  algorithm is the main tool used for  $b$  tagging in this top-quark search. In this chapter the  $D - \phi$  “ $b$  tagger” is characterized by using appropriate Montecarlo and CDF control data samples. For this purpose, the following quantities are measured:

- the  $b$ -tagging efficiency,  $\epsilon_b$ , using  $b$ -enriched samples, in which the fraction of events containing  $b$  hadrons,  $F_b$ , has been measured in an independent way (section 6.1); results are compared to results obtained with the corresponding Montecarlo sample (section 6.1.10); the  $b$ -tagging efficiency is measured per  $b$ -hadron jet contained in the SVX fiducial volume;
- the  $b$ -tag transverse decay length,  $L_{xy}$ ; the result is compared to what expected for  $b$  hadrons (section 6.1.11);
- the frequency of tagging errors (fake  $b$  tags, see section 5.2.5),  $\epsilon_f$ , using samples of low value of  $F_b$  (section 6.2); the rate of fake  $b$  tags is measured per jet contained in the SVX fiducial volume;
- systematic effects like the variation of  $\epsilon_b$  and  $\epsilon_f$  during the data-taking period (section 6.2.1) and the capability of predicting the number of  $b$  tags observed in data samples independent to those used to measure  $\epsilon_b$  and  $\epsilon_f$  (section 6.2.2).

The results of these measurements will then be used for the following purposes:

- to show that  $D - \phi$   $b$  tags are fully consistently with  $b$ -hadron decays;
- to determine quantitatively how accurately the ISAJET [112] + CLEO [113] Montecarlo programs and the CDF simulation program, CDFSIM [114], reproduce  $\epsilon_b$  of the data; this will allow to estimate the  $b$ -tagging efficiency for  $t\bar{t}$  events in section 8.5;
- to evaluate the background to the top signal after  $b$  tagging in  $W +$  multijet events, by extrapolating results obtained for inclusive jets (section 8.3).

## 6.1 The $b$ -Enriched Sample: Inclusive Muons

Events with muons identified both in the CMU and in the CMP  $\mu$  chambers are a suitable sample to study the  $b$ -tagging efficiency, since they are rich in  $b \rightarrow \mu\bar{\nu}c$  decays. Other sources of muons are  $c \rightarrow \mu\bar{\nu}X$  decays and decays in flight (DIF) of  $\pi$ 's and  $K$ 's; the background of fake muons due to penetration of hadrons through the calorimeter is negligible, due to the additional layer of steel between the CMU and the CMP.

The CMU-CMP muons are selected with the following cuts:

1.  $P_t > 9$  GeV/ $c$ ;
2. electromagnetic energy (in the muon tower)  $< 2.0$  GeV;
3. hadronic energy (in the muon tower)  $< 6.0$  GeV;
4.  $|D_{beam}| < 0.2$  cm;
5.  $|Z_{PV} - z_0| < 5$  cm;
6.  $|Z_{PV}| < 30$  cm;
7. distance  $\Delta x$  between the extrapolation of the CTC track up to the muon chambers and the position of the hit measured by the muon chambers:

$$\text{CMU: } |\Delta x| < 2.0 \text{ cm;}$$

$$\text{CMP: } |\Delta x| < 5.0 \text{ cm.}$$

A confirmation of the presence of  $b \rightarrow \mu\bar{\nu}c$  decays in this sample is given by the identification of  $c$  hadrons associated with the lepton. For example,  $D^0 \rightarrow K\pi$  decays are reconstructed by forming all possible combinations of  $K\pi$  track pairs in an  $\eta - \phi$  cone of radius  $r_\mu=0.5$  around the muon. Figure 6.1 shows a clear peak in the distribution of the  $M(K\pi)$  invariant mass around the  $D^0$  mass; this peak is not observed for  $K\pi$  track pairs in which the sign of the kaon charge and the sign of the associated lepton charge are not consistent with  $B$  meson decays. Because of the low efficiency for reconstructing these semi-exclusive ( $D^0 + \text{Lepton} + X$ ) decay channels, figure 6.1 has been obtained by using similar inclusive electron data samples.

In inclusive muon events, we consider the muon-associated jet (“ $\mu$  jet”) and the jet recoiling against the muon jet of highest transverse energy (“away” jet), but only if such jets are contained in the SVX fiducial volume (the definition of  $\mu$  jet, away jet, and SVX-fiducial jet are given in the following sections). The analysis goal is: (1) for the fraction of SVX-fiducial jets containing  $b$  hadrons (sections 6.1.4, 6.1.5) to measure the  $b$ -tagging efficiency (sections 6.1.6 to 6.1.9); (2) for SVX-fiducial jets  $b$ -tagged by the  $D - \phi$  algorithm, to show that the  $b$  tag decay length is consistent with the expectation for  $b$  hadrons.

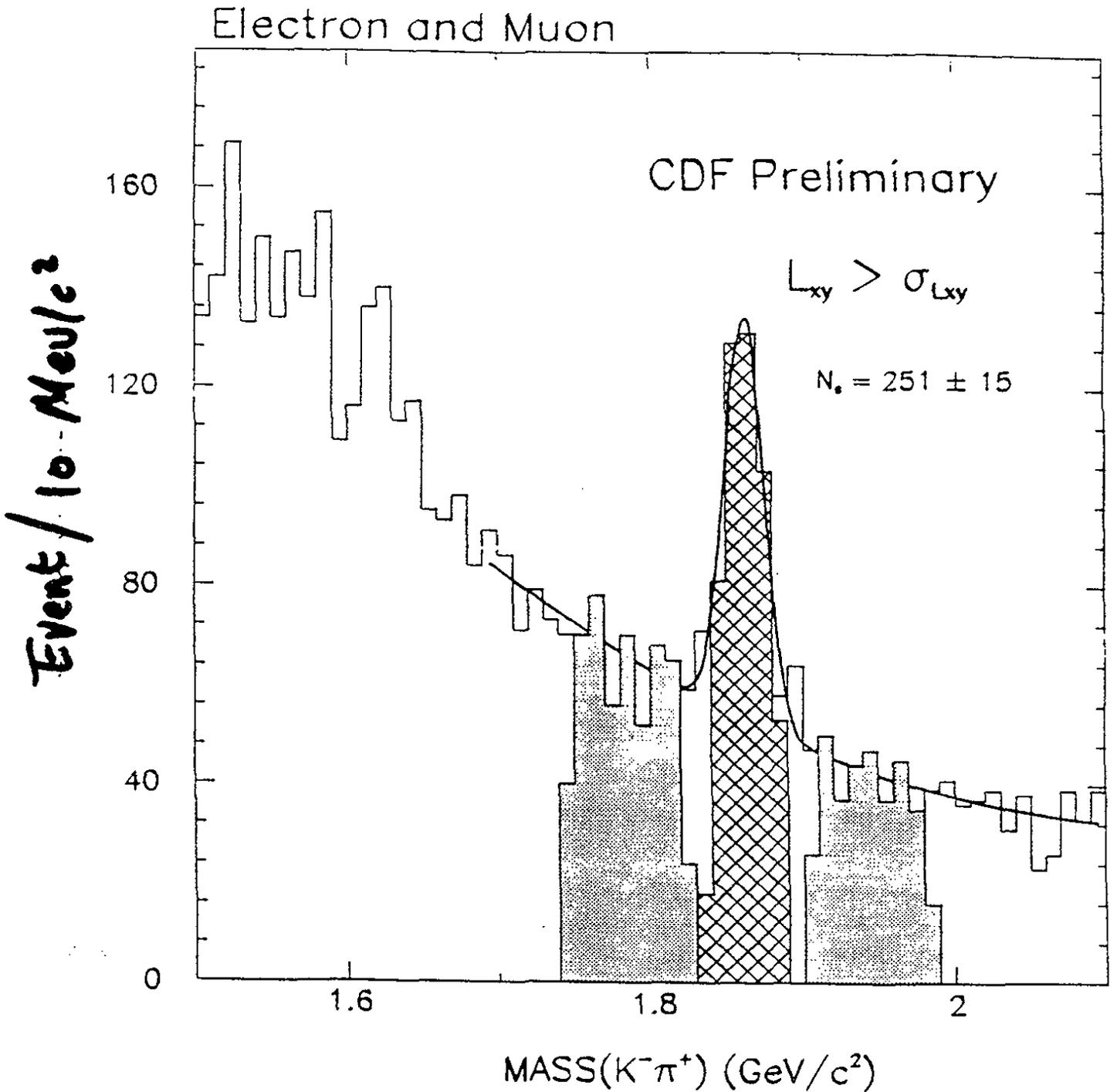


Figure 6.1: Inclusive electron and muon sample: reconstruction of  $D^0 \rightarrow K\pi$  decays associated with electrons and muons. To reduce the combinatorial background,  $K\pi$  candidate pairs are selected by requiring their transverse decays length,  $L_{xy}$ , be larger than the relative uncertainty,  $\sigma_{L_{xy}}$ . The detection of  $c$  hadrons associated with electrons and muons strengthens the hypothesis that the inclusive electron and muon sample is enriched of  $b \rightarrow c\bar{\nu}$  decays.

### 6.1.1 Identification of Jets Associated with Muons

Since the muon releases only a modest fraction of its energy in the calorimeter, we chose to define the jet *associated* with the muon using tracking information: we consider all good quality CTC tracks of  $P_t > 400$  MeV/c in an  $\eta - \phi$  cone of radius  $r_\mu = 0.5$  around the direction of the muon momentum (muon track included) and sum their momentum vectors; if there is no track satisfying these cuts the event is rejected; this set of tracks represents the muon jet, and its transverse energy,  $E_t(\mu - jet)$ , and direction  $(\eta_{\mu-jet}, \phi_{\mu-jet})$  are computed using the total momentum of the set of tracks. All the analysis described in this chapter is based only on events with a single muon; multiple muon events are rejected.

### 6.1.2 Identification of Jets Recoiling Against Muon Jets

Other jets in inclusive muon events (those not associated with muons) are defined using calorimetric information and the standard CDF jet reconstruction algorithm, which employs an  $\eta - \phi$  cone of fixed radius to define the jets. This algorithm begins by listing calorimeter towers with  $E_t > 1$  GeV (the seeds). Seed towers are grouped in clusters if their distance in the  $\eta - \phi$  space is  $< 0.4$ . The cluster  $\eta - \phi$  directions are calculated as sum of unit vectors applied to the centers of gravity of the energies of the towers in the cluster and pointing to the event primary vertex; each unit vector in the sum is weighted with the  $E_t$  of the corresponding tower. Towers with  $E_t > 100$  MeV, whose separation from the clusters is  $< 0.4$  in  $\eta - \phi$ , are added to the clusters, and the cluster directions are recomputed. This process is repeated until the list of towers in the clusters becomes stable. If two clusters share more than 50% of the energy, they are merged in a single cluster; otherwise the common towers are assigned to the nearest cluster. The jets are clusters which satisfy the cuts of this algorithm. The energy and momentum of each jet is computed in the following way:

$$(E^j, P_x^j, P_y^j, P_z^j) = (\sum_{i=1}^N E_i, \sum_{i=1}^N E_i \sin \theta_i \cos \phi_i, \sum_{i=1}^N E_i \sin \theta_i \sin \phi_i, \sum_{i=1}^N E_i \cos \theta_i), \quad (6.1)$$

where  $E_i$  is the energy of the  $i$ -th tower of the jet,  $\phi_i$  is the azimuthal angle of the energy center of gravity of the tower and  $\theta_i$  is the polar angle of the energy center of gravity of the tower w.r.t. the PV. The jet four-momentum thus defined is the momentum *observed* with the calorimeter; this vector is not corrected for calorimeter measurement errors, neither for energy losses outside the jet-cone definition nor for the energy flow of the so-called “underlying event”.

### 6.1.3 Event Selection

In the top-quark search described in this thesis (chapter 7 and 8) only jets with  $|\eta| < 2.0$  and observed transverse energy  $E_t > 15$  GeV are considered. For the characterization of the  $D - \phi b$  tagger with the inclusive muon control sample described in this chapter, the  $E_t$  cut is released to  $E_t > 10$  GeV in order to avoid a loss in statistics and only SVX-fiducial jets are considered. The definition of an SVX-fiducial jet is the following: the jet must contain  $\geq 2$  SVX tracks, satisfying the cuts:

1.  $P_t > 400$  MeV/c;

2.  $\sqrt{(\eta_{\mu-jet} - \eta_{track})^2 + (\phi_{\mu-jet} - \phi_{track})^2} < 0.4;$
3.  $|Z_{PV} - z_{track}| < 5 \text{ cm}.$

The two classes of jets studied in the following analysis are thus defined:

- (1)  $\mu$  jets with  $E_t(\mu-jet) > 10 \text{ GeV}$ ;  $|\eta_{\mu-jet}| < 2.0$  is always satisfied, since the CMU-CMP chambers cover  $|\eta| < 0.6$  and since the event sample has  $|Z_{PV}| < 30 \text{ cm}$ ;
- (2) recoil or away jets: the highest  $E_t$  jet with  $E_t(away) > 10 \text{ GeV}$ ,  $|\eta(away)| < 2.0$  and  $150^\circ < |\phi(away) - \phi(\mu-jet)| < 210^\circ$ .

The  $150^\circ < |\phi(away) - \phi(\mu-jet)| < 210^\circ$  cut is applied in order to be able to use the best available measurement of  $F_b$  for the away jet (see section 6.1.5). This azimuthal cut selects preferably two-jet configurations with the jets “back – to – back” in the transverse plane, which are typical of  $2 \rightarrow 2$  hard scattering processes, initiated by two parton in the initial and final state.

In the next two sections the methods used to estimate  $F_b(\mu)$  and  $F_b(away)$  are described.

#### 6.1.4 Measurement of the $b$ Fraction ( $F_b$ ) for Muon Jets

The value of  $F_b(\mu)$  can be measured by studying the *signed* IP distribution of SVX tracks corresponding to the muons,  $\bar{D}(\mu)$ , and taking into account the fraction of them coming from  $b$  decays,  $F_b(\mu)$ , from  $c$  decays,  $F_c(\mu)$ , and from DIF,  $F_{dif}$  [115]. The definition of signed IP is illustrated in figure 6.2; the unit vector along the  $\mu$  jet direction in the transverse plane,  $\vec{n}_{jet}(\mu)$ , is used to establish the sign:

$$\bar{D}(\mu) = |D| \text{sign}(\vec{n}_{jet}(\mu) \cdot \vec{D}), \quad (6.2)$$

where  $\vec{D}$  is the vector connecting the PV to the point of the track of closest approach to the PV in the transverse plane.

The characteristics of  $\bar{D}$  for  $\mu$  from  $b$  and  $c$  have been studied with Montecarlo samples:

- (i) the ISALEP program has been used to generate  $b$ -quark and  $c$ -quark events, by means of the following processes:

- Drell-Yan:  $q\bar{q} \rightarrow b\bar{b}$ ,
- gluon fusion:  $gg \rightarrow b\bar{b}$ ,
- flavor excitation:  $gb \rightarrow gb$ ,
- “gluon splitting”:  $gg \rightarrow gg, g \rightarrow b\bar{b}$ ;

- (ii) the CLEO Montecarlo has been used for  $b$ -hadron decays;
- (iii) the events thus obtained have been passed through a complete CDF simulation, which includes effects like uninstrumented regions between the calorimeter towers (“cracks”), malfunctioning channels of the tracking, photon conversions and the resolution of the various detectors [114].

TRACK 1:  $0 < \beta_1 < \pi/2 \Rightarrow \text{signed IP} > 0$

TRACK 2:  $\pi/2 < \beta_2 < \pi \Rightarrow \text{signed IP} < 0$

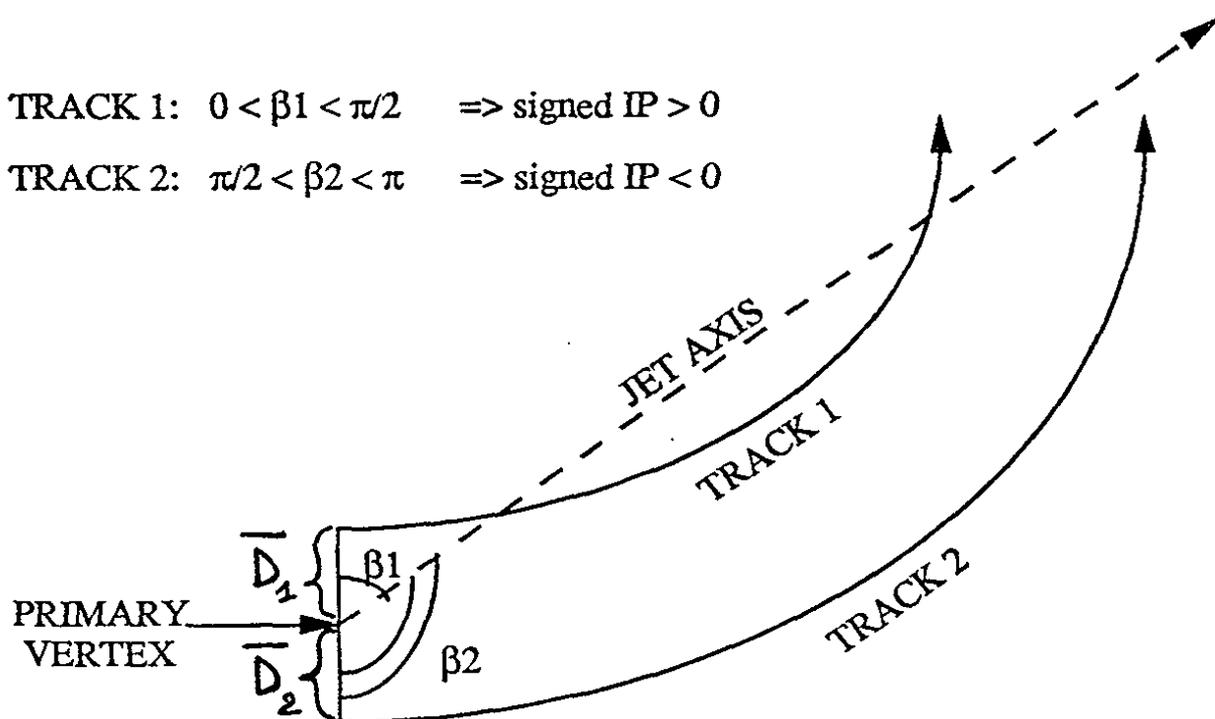


Figure 6.2: Definition of signed impact parameter,  $\bar{D}$ .

Figure 6.3.a shows  $\overline{D}(\mu)$  for  $b$  decays in the Montecarlo, before the detector simulation; in this case, in order to define  $\overline{D}(\mu)$  the true  $b$  direction is used; in figure 6.3.a the effect of the  $b$  lifetime is clearly visible. Figure 6.3.b shows  $\overline{D}(\mu)$  calculated using the jet direction and the SVX muon track, that is including the effect of the gluon splitting, of the fragmentation and of the detector resolution; while these effects produce negative tails, the width of the distribution is still dominated by the  $b$  lifetime.

The  $\overline{D}(\mu)$  of charm decays after the detector simulation shown in figure 6.3.c, is still asymmetric with positive mean, but the asymmetry is less than for the bottom quark.

For DIF's, one expects that the requirement of hits in the SVX produces a symmetric  $\overline{D}(\mu)$  distribution with zero mean, having a width related to the following effect: the  $\pi \rightarrow \mu$  decays occurring inside the CTC generate two different track segments (the  $\pi$  track and the  $\mu$  track) almost collinear ( $P_t(\mu) > 9$  GeV/c); the reconstruction of a single track from the two segments, together with the finite CTC resolution, determines a symmetric broadening of the distribution of the muon candidate around  $\overline{D}(\mu) = 0$ . The contribution of the non-zero DIF lifetime inside the SVX is negligible for muons of  $P_t > 9$  GeV/c. The DIF  $\overline{D}(\mu)$  distribution has been modeled with: (1)  $\pi \rightarrow \mu\bar{\nu}$  Montecarlo events, (2) generic PV track of  $P_t > 9$  GeV/c from the Montecarlo, (3) generic PV tracks of  $P_t > 9$  GeV/c from the data. Since the three distributions are very similar (see figure 6.4) the data distribution has been employed as a model (distribution number (3), dashed line of figure 6.4).

The distributions of figures 6.3.b, 6.3.c and 6.4 (dashed histogram) are used as models for a fit to the  $\overline{D}(\mu)$  distribution in the data. The average uncertainty on  $\overline{D}(\mu)$ , which contains CTC, SVX and PV contributions, is  $20 \mu\text{m}$ , very small compared to  $c\tau_B$ , and thus fully adequate for the separation of the different components. The event sample used for the fit contains about 5200 muon jets. The fractions determined by the fit are:

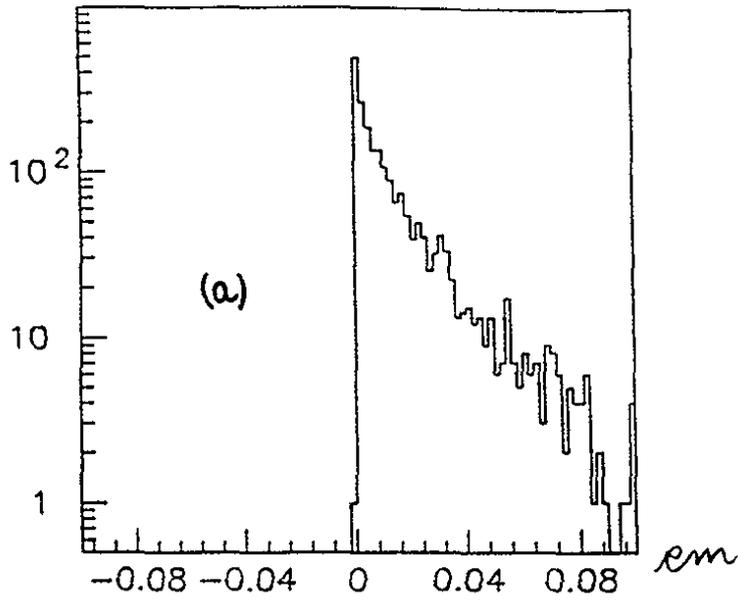
$$F_b(\mu) = 0.40 \pm 0.03, \quad (6.3)$$

$$F_c(\mu) = 0.14 \pm 0.04, \quad (6.4)$$

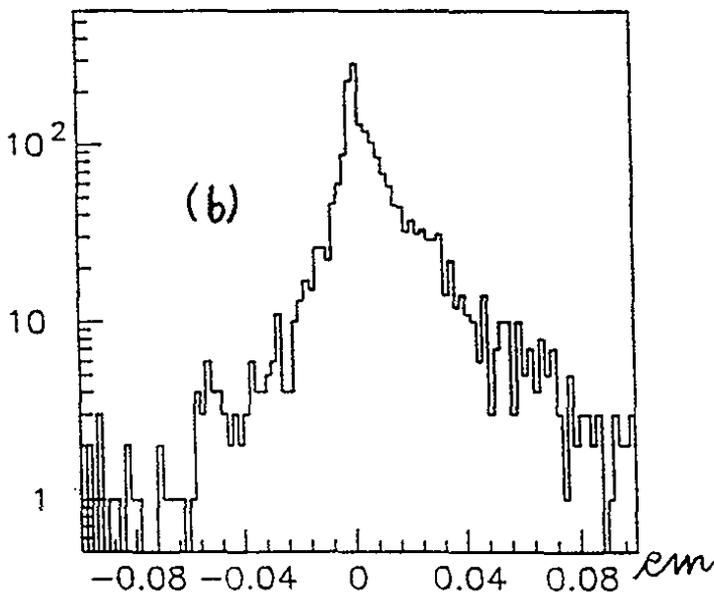
$$F_{DIF}(\mu) = 0.46 \pm 0.10. \quad (6.5)$$

The data  $\overline{D}$  distribution and the three components of the model, normalized by their respective fractions are shown in figure 6.5:  $F_b(\mu)$  is clearly dominated by the region of high IP. Figure 6.6 shows that the sum of the three components is a good description of the data. Moreover, the ratio  $F_b(\mu)/F_c(\mu) \sim 2.9$  is in reasonably good agreement with the Montecarlo prediction, which gives  $F_b(\mu)/F_c(\mu) \sim 2.3$ . The uncertainty on  $F_b(\mu)$  has been estimated by varying the width of the DIF distribution, the SVX resolution used to model the  $b$  and  $c$  decays, and by scaling the Montecarlo  $b$  distribution to take into account the different value of  $c\tau_B$  used in the Montecarlo and measured in CDF data [50]. Increasing the  $\overline{D}$  width of the DIF by 25%, causes a 2% increase of  $F_c(\mu)$ , while  $F_b(\mu)$  does not change; the other two effects give a 0.05 systematic error on  $F_b(\mu)$ . Note that potential correlations between  $F_{DIF}(\mu)$  and  $F_c(\mu)$ , if present, are not relevant to the following analysis, since it uses only  $F_b(\mu)$ , which is not sensitive to reasonable variations of the sum of  $F_{DIF}(\mu)$  and  $F_c(\mu)$  contributions. In conclusion, from the signed IP analysis of inclusive CMU-CMP of  $P_t > 9$  GeV/c one gets:

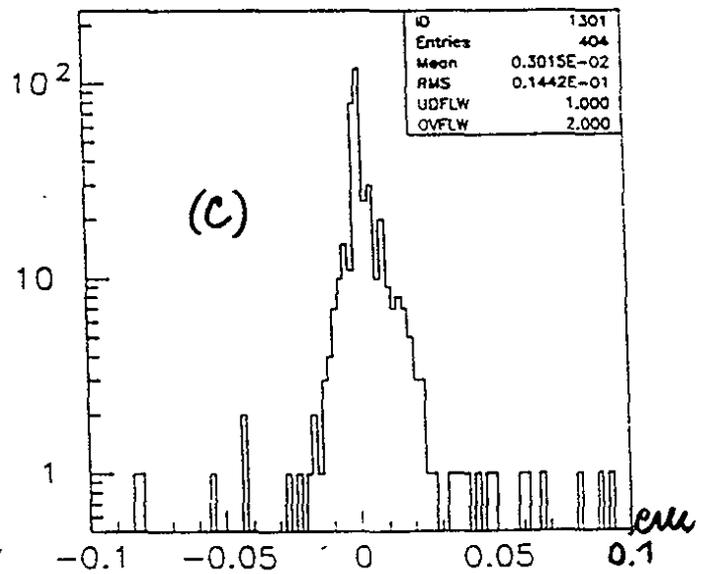
$$F_b(\mu) = 0.40 \pm 0.08 \quad (\text{method: } \overline{D}) \quad (6.6)$$



real B dir., OBS signed I.P.



JETS axis, SVX signed I.P.



c to mu signed impact parameter

Figure 6.3: Signed impact parameter ( $\overline{D}(\mu)$ ) distributions for  $b \rightarrow \mu X$  decays in the Monte-carlo before (a) and after (b) full CDF simulation;  $\overline{D}(\mu)$  for  $c \rightarrow \mu X$  decays in the Montecarlo after simulation (c).

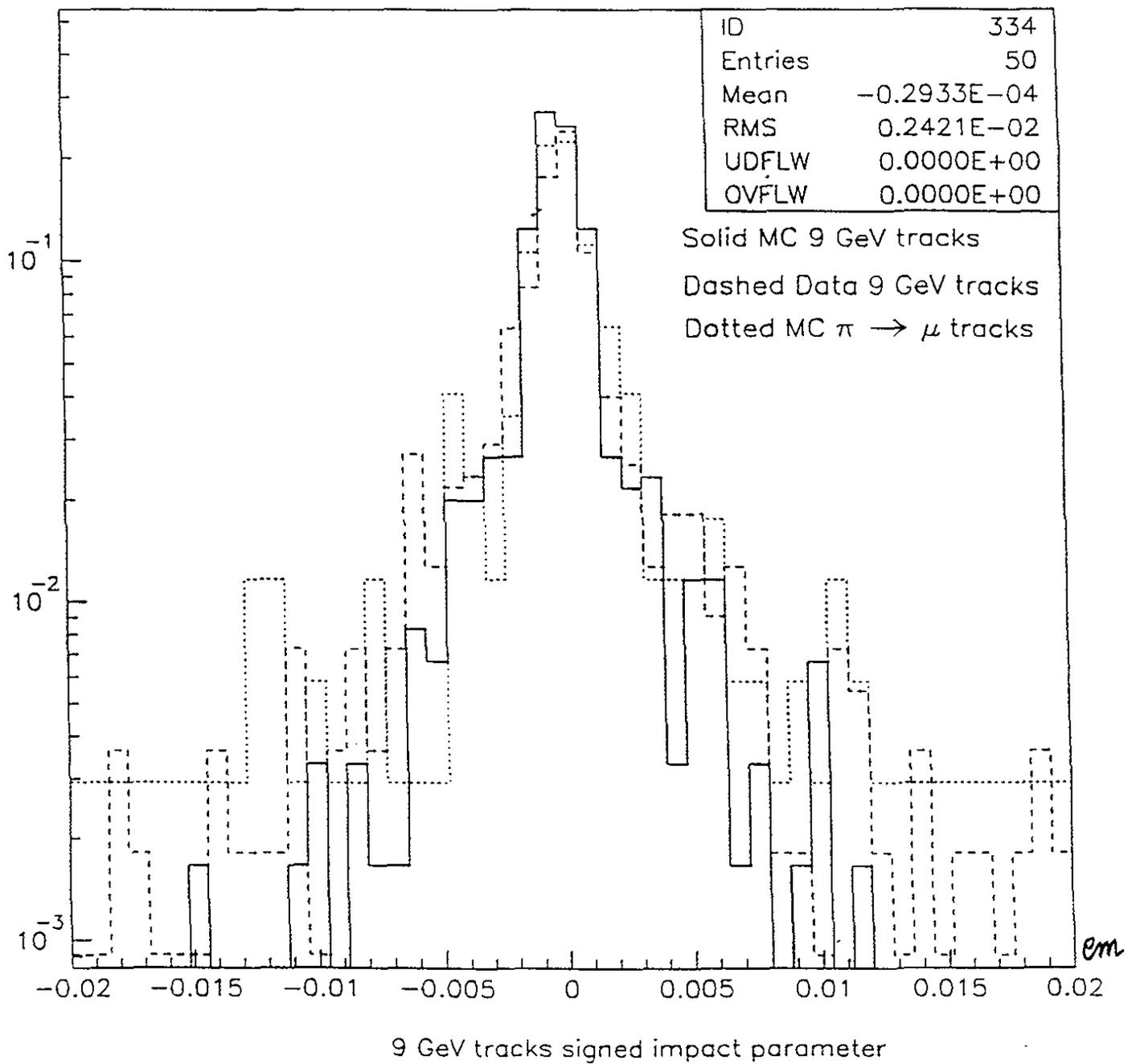


Figure 6.4: Various signed impact parameter distributions used to model decays in flight.

This results is confirmed by an independent measurement using a  $b$  tagging algorithm based on an appropriate jet probability function (“jet probability” or JETPRB). This algorithm was developed by the ALEPH collaboration [116] and further refined at CDF for the top-quark search [117]. The probability function is extracted directly from the data, by comparing  $\bar{D} > 0$  tracks in jets with the resolution function measured with the SVX. This probability distribution is uniform for light quark jets with zero lifetime, while showing a pronounced peak around zero for jets containing heavy flavors, that is, containing long lifetime  $b$  or  $c$  hadrons. From a physics point of view, the jet probability function represents the probability that the set of IP’s observed in the jet is consistent with the effect of the SVX resolution only;  $b$  jets correspond on the average to very low probabilities ( $< 0.01$ ), while light quarks are characterized by probabilities uniformly distributed between 0 and 1. From the shape of the probability function of  $\mu$  jets in our sample one extracts the following central values:  $F_b(\mu) = 0.38$ ,  $F_c(\mu) = 0.15$ , in good agreement with the results from the signed IP method (the uncertainty on these values is conservatively estimated to be around 15 % for  $F_b(\mu)$ , and around 30-40 % for  $F_c(\mu)$ ; further studies are in progress for the more accurate evaluation of the contribution of the systematic error).

### 6.1.5 Measurement of the $b$ Fraction ( $F_b$ ) for Recoil Jets

This section describes the two methods used to measure the  $b$  fraction of recoiling jets,  $F_b(away)$ .

(1) The first method uses the value of  $F_b(\mu)$  and the result of a theoretical calculation of heavy quark correlations, performed at parton level and complete to the third perturbative order in  $\alpha_s$ , (“heavy-quark correlations” Montecarlo, or HVQ Montecarlo, [118]). In  $b$ -quark events, when one of the two  $b$ ’s is fixed (for example the jet with  $b \rightarrow \mu X$ ), HVQ allows to estimate in which fraction of these events the highest  $E_t$  jet recoiling against the reference  $b$ -quark jet is also a  $b$  jet, rather than a gluon jet. For recoiling jets of  $E_t(away) > 10$  GeV (that is, for the highest  $E_t$  jets) with  $\eta - \phi$  distance from the  $\mu$  jet greater than 1.0, one finds:

$$F_b(away)/F_b(\mu) = 0.45 \pm 0.15, \quad (6.7)$$

where the quoted uncertainty comes from variation of the renormalization scale and the parton  $P_i$ . Assuming the measured value of  $F_b(\mu)$  of the previous section, HVQ predicts the central value  $F_b(away) = 0.18$ , with a large uncertainty due to the 33 % theoretical uncertainty of equation 6.7 and the 20 % systematic uncertainty on the experimental value of  $F_b(\mu)$ .

(2) The second method is independent of  $F_b(\mu)$ , because it uses the shape of the JETPRB function of *away* SVX-fiducial jets to make a fit to  $F_b(away)$  and  $F_c(away)$ , similarly to what has been done for the  $\mu$  jet [119]. In events in which the *away* jet is SVX-fiducial,  $E_t(away) > 10$  GeV,  $150^\circ < |\phi(away) - \phi(\mu - jet)| < 210^\circ$ , and  $|\eta(away)| < 2.0$ , one gets:

$$F_b(away) = 0.14 \pm 0.02 \quad (method : JETPRB), \quad (6.8)$$

$$F_c(away) = 0.055 \pm 0.016 \quad (method : JETPRB), \quad (6.9)$$

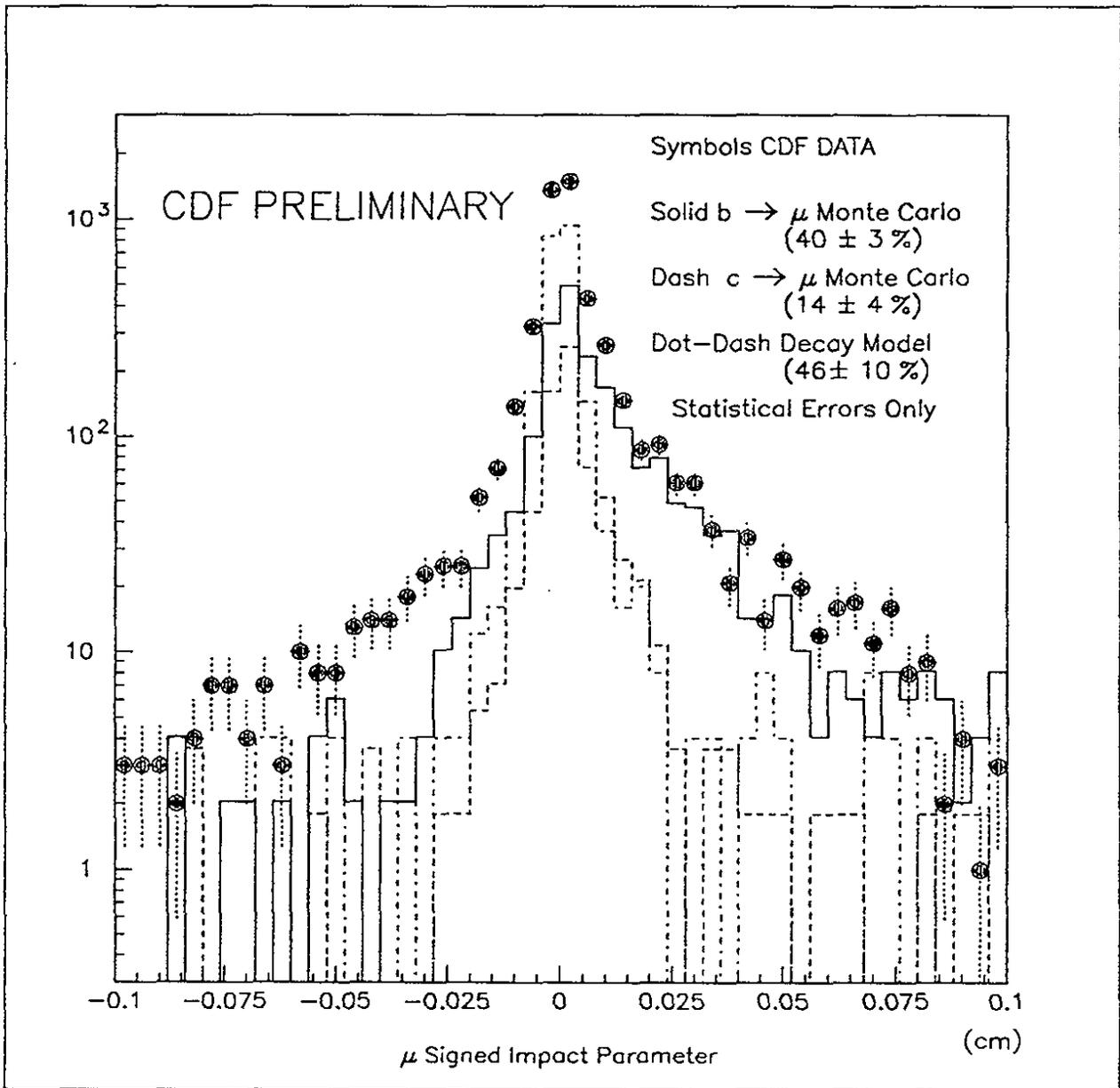


Figure 6.5: Muon signed impact parameter distribution in the data compared to those expected for  $b$  decays,  $c$  decays, and decays in flight (DIF Decay Model, dashed histogram).

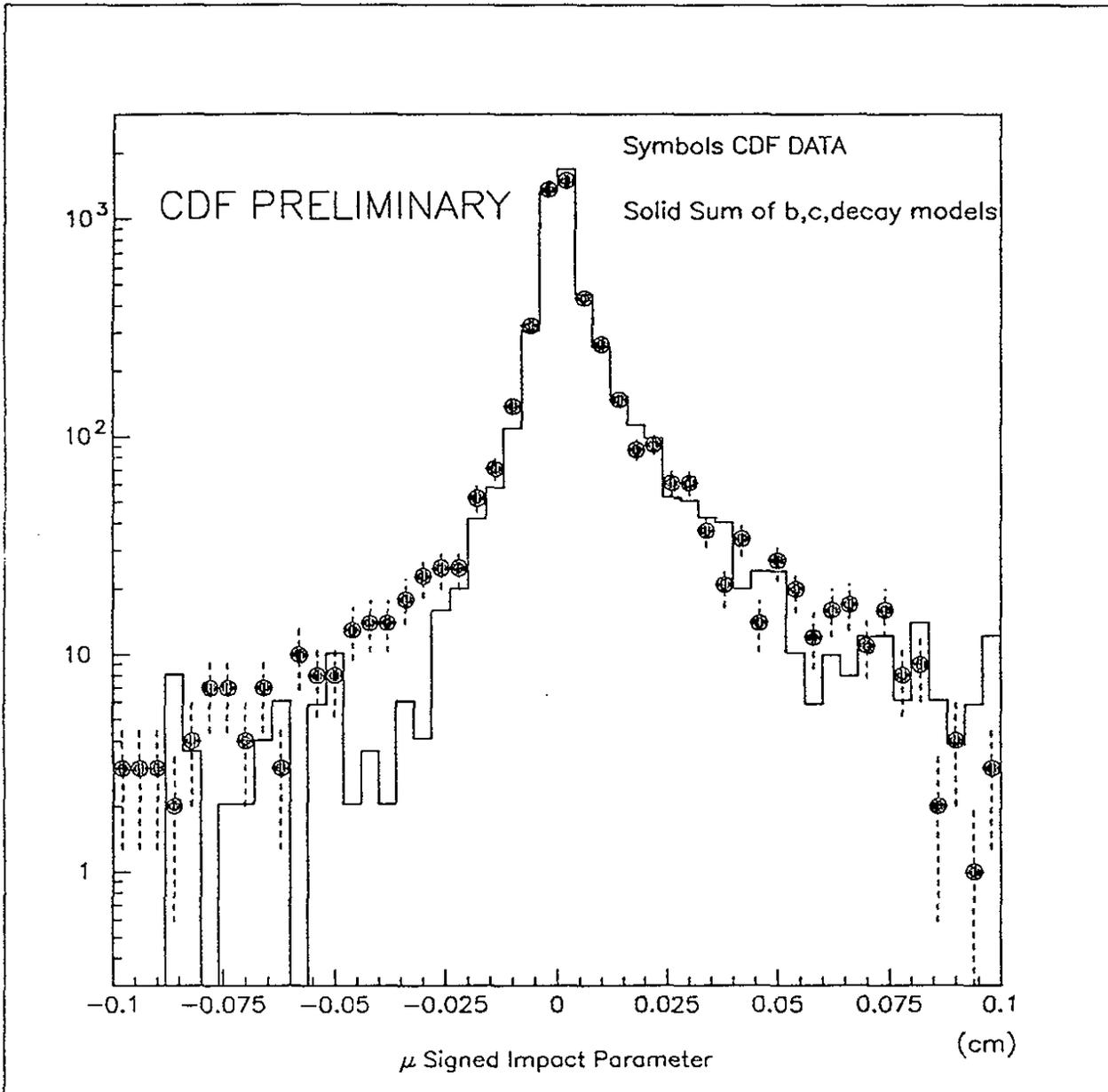


Figure 6.6: Muon signed impact parameter distribution in the data compared to the sum of those expected for  $b$  decays,  $c$  decays, and decays in flight

where the error contains also the systematic effect related to the jet probability analysis. Note that  $F_b(\text{away})/F_c(\text{away}) \sim 2.6$ , while for the  $\mu$  side, from the signed IP method one gets  $F_b(\mu)/F_c(\mu) \sim 2.9$ .

The  $F_b(\text{away})$  values obtained from the two methods are in good agreement. In the following,  $F_b(\text{away})$  extracted with the JETPRB method will be used, since it is measured with a smaller error and it is less dependent on theoretical predictions than method 1.

### 6.1.6 $b$ -Tagging Efficiency ( $\epsilon_b$ )

Having measured  $F_b(\mu)$  and  $F_b(\text{away})$ , the  $b$ -tagging rates of fiducial  $\mu$  and away jets can be converted into separate tagging efficiencies for  $b$ -hadron decays,  $\epsilon_b(\mu)$  and  $\epsilon_b(\text{away})$ .

**Measurement of  $\epsilon_b(\mu)$ .** This measurement has the advantage of a good statistics (2185  $b$  tags), thanks to the large value of  $F_b(\mu)$  and to the good containment of CMU-CMP muons in the SVX geometrical acceptance. Another advantage is that the dynamics of semileptonic decays is well known and well modeled by the Montecarlo program adopted for this analysis. In addition, lepton identification is also well modeled by the detector simulation program of CDF, which has recently measured the  $b$  production cross section using  $b \rightarrow \mu X$  [120] and  $b \rightarrow eX$  [21] decays (more recently, the  $b\bar{b}$  correlated cross section has also been submitted for publication; also in this paper,  $b$ 's are identified by means of their semileptonic decays). However,  $\epsilon_b(\mu)$  has the disadvantage that the  $b$  tagging is applied to a special category of jets, those containing a muon, which fired the trigger (and containing also a neutrino), different from the generic  $b$  jets which are looked for in  $t\bar{t}$  events. In particular, one expects the charged multiplicity of  $\mu$  jets to be lower than for generic  $b$  jets due to the presence of the neutrino, and, for this reason, the  $b$  tagging efficiency for  $\mu$  jets to be lower than for generic  $b$  jets. However, this effect is to some extent compensated by the fact that the  $D - \phi$  algorithm always uses the SVX track associated with the muon in the search for SVs, independently of its IP significance (see section 5.2.2, point D): for  $b$  tagging purposes, one  $b$  track is always lost due the presence of the neutrino, but one is most of the times gained due to the identification of the muon. Another peculiar feature of  $\mu$  jets is that a certain fraction of  $b \rightarrow \mu X$  decays are accompanied by the prongs of the other  $b$  in event of small angle gluon splitting  $g \rightarrow b\bar{b}$ .

**Measurement of  $\epsilon_b(\text{away})$ .** The main advantage of this measurement is the fact that, in contrast to  $\mu$  jets, away jets are generic jets; they are not special jets as a consequence of the trigger and the event selection; away jet tags are, therefore, similar to  $b$  jets in  $t\bar{t}$  events. Also, they represent a sample of isolated  $b$  jets, in which the fraction of small angle gluon splitting is reduced by the requirement that they be found well separated in  $\eta - \phi$  from the  $\mu$  jet, originated by a  $b$  quark in a fraction  $F_b(\mu) \sim 40\%$  of the events in the sample. A disadvantage is the much reduced statistics compared to  $\mu$  jets (204 away jet  $b$  tags), due to the low value of  $F_b(\text{away})$  ( $\sim 0.14$ ) and from the limited  $\eta$  acceptance of the SVX.

In the following sections it will be shown that for fiducial jets  $\epsilon_b$  varies from 22 % to 29 % depending on the kind of jet considered ( $\mu$  or away) and depending on the track quality requirements. In order to prove that the  $D - \phi$  algorithm tags  $b$  hadrons, it will be shown that the  $b$ -tag decay length, corrected for the appropriate kinematical factors, is fully consistent with the most recent measurements of the lifetime averaged over all  $b$ -hadron states ( $c\tau_B = 446 \pm 44 \mu\text{m}$  [50],  $\tau_B \sim 1.49 \text{ psec}$ ).

### 6.1.7 Measurement of $\epsilon_b$ for Muon Jets

The efficiency measurement reported in the following refers to jets of  $E_t(\mu - jet) > 10$  GeV, contained in the SVX fiducial volume. This  $b$ -tagging efficiency is given by the following expression:

$$\epsilon_b(\mu, fid) = \frac{N_f(tag) - N_f(bkg)}{N_f \cdot F_b(\mu)}, \quad (6.10)$$

with  $F_b(\mu)$  given by equation 6.6, where  $N_f(tag)$  is the number of  $b$  tags found among fiducial jets,  $N_f$  and  $N_f(bkg)$  are the number of (*background*)  $b$  tags associated with muons not coming from  $b$ -hadron decays.  $N_f(bkg)$  is expressed as:

$$N_f(bkg) = (1 - F_b(\mu)) \cdot N_f(non - b), \quad (6.11)$$

where  $(1 - F_b(\mu))$  is the fraction of events with muons not coming from  $b$ -hadron decays, and  $N_f(non - b)$  is the estimate of the number of background  $b$  tags due to: (A) tracking errors; (B) residual contaminations of  $K_s \rightarrow \pi\pi$  and  $\Lambda \rightarrow p\pi$  decays, together with tracking errors (a  $D - \phi$  cluster has  $\geq 3$  tracks); (C)  $c$ -hadron decays. The method for calculating  $N_f(non - b)$  is given in section 6.2.1, which reports the study of the inclusive generic jet sample: jets in this sample have a low  $b$  content and are well suited to study the above effects A, B and C. For the moment, it is sufficient to know that such an estimate gives  $N_f(non - b)/N_f = (0.68 \pm 0.05)\%$ . Note that if all jets come from  $b$  quarks ( $F_b(\mu) = 1$ ) no background subtraction would be necessary, since also possible fake  $b$  tags would be accompanied by true semileptonic decays of  $b$  hadrons.

In order to study the variation of the efficiency with track quality selection, we define the sub-sample of jets containing  $\geq 2$  SVX tracks passing the fiducial cuts plus all track quality cuts imposed by the  $D - \phi$  algorithm (see section 5.2.2) except the cuts on  $S_D$ . This corresponds to the requirement that at least two tracks of the jet are well measured by the SVX. One measures, therefore, also the following efficiency:

$$\epsilon_b(\mu, qual) = \frac{N_q(tag) - N_q(bkg)}{N_q \cdot F_b(\mu)}, \quad (6.12)$$

where  $N_q$  is the number of jets with  $\geq 2$  well measured, good quality SVX tracks,  $N_q(tag)$  is the number of  $b$  tags found in the sub-sample  $N_q$ , and  $N_q(bkg) = (1 - F_b(\mu)) \cdot N_q(non - b)$  is the estimated number of  $b$  tags not coming from  $b$  quarks out of the observed  $N_q(tag)$  tags.  $N_q(non - b)/N_q = (0.79 \pm 0.06)\%$ .  $N_q(non - b)$  is estimated in a way similar to the one used for  $N_f(non - b)$  (see section 6.2.1).

Figures 6.7.a and 6.7.b show, respectively, the  $E_t(\mu - jet)$  distribution of fiducial jets ( $N_f = 23244$  jets, before  $b$  tagging) and of corresponding  $b$  tags ( $N_f(tag) = 2185$  jets, after  $b$  tagging); figures 6.8.a and 6.8.b show the same quantity for  $N_q$  ( $N_q = 17559$  jets, before  $b$  tagging) and  $N_q(tag)$  ( $N_q(tag) = 2124$  jets, after  $b$  tagging). After background subtraction ( $N_f(bkg) = 95$  jets,  $N_q(bkg) = 83$  jets) and the normalization by  $F_b(\mu) = 0.40 \pm 0.08$ , one finds the following efficiencies:

$$\epsilon_b(\mu, fid) = 0.22 \pm 0.04, \quad (6.13)$$

$$\epsilon_b(\mu, qual) = 0.29 \pm 0.06. \quad (6.14)$$

The uncertainty on the efficiency is the sum in quadrature of the statistical uncertainty and of the 20% systematic uncertainty on  $F_b(\mu)$  (the latter dominates).

### 6.1.8 Measurement of $\epsilon_b$ for Recoil Jets

In the event sample in which the  $\mu$  jet is fiducial, the away jet is fiducial about 22% of the time. Using the sub-sample in which both the muon and away jets are fiducial and the one in which both, besides being fiducial, also contain  $\geq 2$  good quality tracks, one measures the following away-jet  $b$ -tagging efficiencies:

$$\epsilon_b(\text{away}, \text{fid}) = \frac{N_f^{(A)}(\text{tag}) - N_f^{(A)}(\text{bkg})}{N_f^{(A)} \cdot F_b(\text{away})}, \quad (6.15)$$

$$\epsilon_b(\text{away}, \text{qual}) = \frac{N_q^{(A)}(\text{tag}) - N_q^{(A)}(\text{bkg})}{N_q^{(A)} \cdot F_b(\text{away})}, \quad (6.16)$$

where symbols similar to those adopted for muon jets have been used. In addition,

$$N_f^{(A)}(\text{bkg}) = (1 - F_b(\text{away})) \cdot N_f^{(A)}(\text{non} - b), \quad (6.17)$$

$$N_q^{(A)}(\text{bkg}) = (1 - F_b(\text{away})) \cdot N_q^{(A)}(\text{non} - b), \quad (6.18)$$

where  $N_f^{(A)}(\text{non} - b)/N_f^{(A)} = 1.1 \pm 0.1\%$  and  $N_q^{(A)}(\text{non} - b)/N_q^{(A)} = 1.3 \pm 0.2\%$  (see section 6.2.1). Figures 6.7.c and 6.7.d show, respectively, the  $E_t(\text{away})$  distribution of fiducial jets ( $N_f^{(A)} = 5053$  jets, before  $b$  tagging) and of corresponding  $b$  tags ( $N_f^{(A)}(\text{tag}) = 204$  jets, after  $b$  tagging); figures 6.8.c and 6.8.d show  $E_t(\text{away})$  for jets with  $\geq 2$  well measured SVX tracks ( $N_q^{(A)} = 3153$  jets, before  $b$  tagging) and for their respective  $b$  tags ( $N_q^{(A)}(\text{tag}) = 162$  jets, after  $b$  tagging). After background subtraction ( $N_f^{(A)}(\text{bkg}) = 48$  jets,  $N_q^{(A)}(\text{bkg}) = 35$  jets) and the normalization by  $F_b(\text{away}) = 0.14 \pm 0.02$ , one finds the following efficiencies:

$$\epsilon_b(\text{away}, \text{fid}) = 0.22 \pm 0.04, \quad (6.19)$$

$$\epsilon_b(\text{away}, \text{qual}) = 0.29 \pm 0.05. \quad (6.20)$$

The quoted uncertainty is the sum in quadrature of the statistical uncertainty and of the systematic 14% uncertainty on  $F_b(\text{away})$ .

### 6.1.9 Measurement of $\epsilon_b$ for Muon Jets Using Double $b$ Tags

The efficiencies measured in the previous sections are based on  $b$  tagging rates relative to single muon or away jets (they do not depend on whether or not the other jet is tagged) and on the respective  $b$  fractions ( $F_b$ ). In order to get an estimate of the  $b$ -tagging efficiency as independently as possible from  $F_b(\mu)$ , one can consider the sub-sample of events in which the away jet is  $b$ -tagged, and measure the tagging rate for the muon jet (double  $b$  tags). The explicit formula for the muon jet tagging efficiency obtained with this method (the double-tag method) is derived and discussed in section 6.4 (appendix at the end of this chapter). In events where both muon and away jets are fiducial, one finds:

$$\epsilon^{\text{doubles}}(\mu, \text{fid}) = 0.20 \pm 0.04; \quad (6.21)$$

## Inclusive muons: $b$ -tagging efficiencies

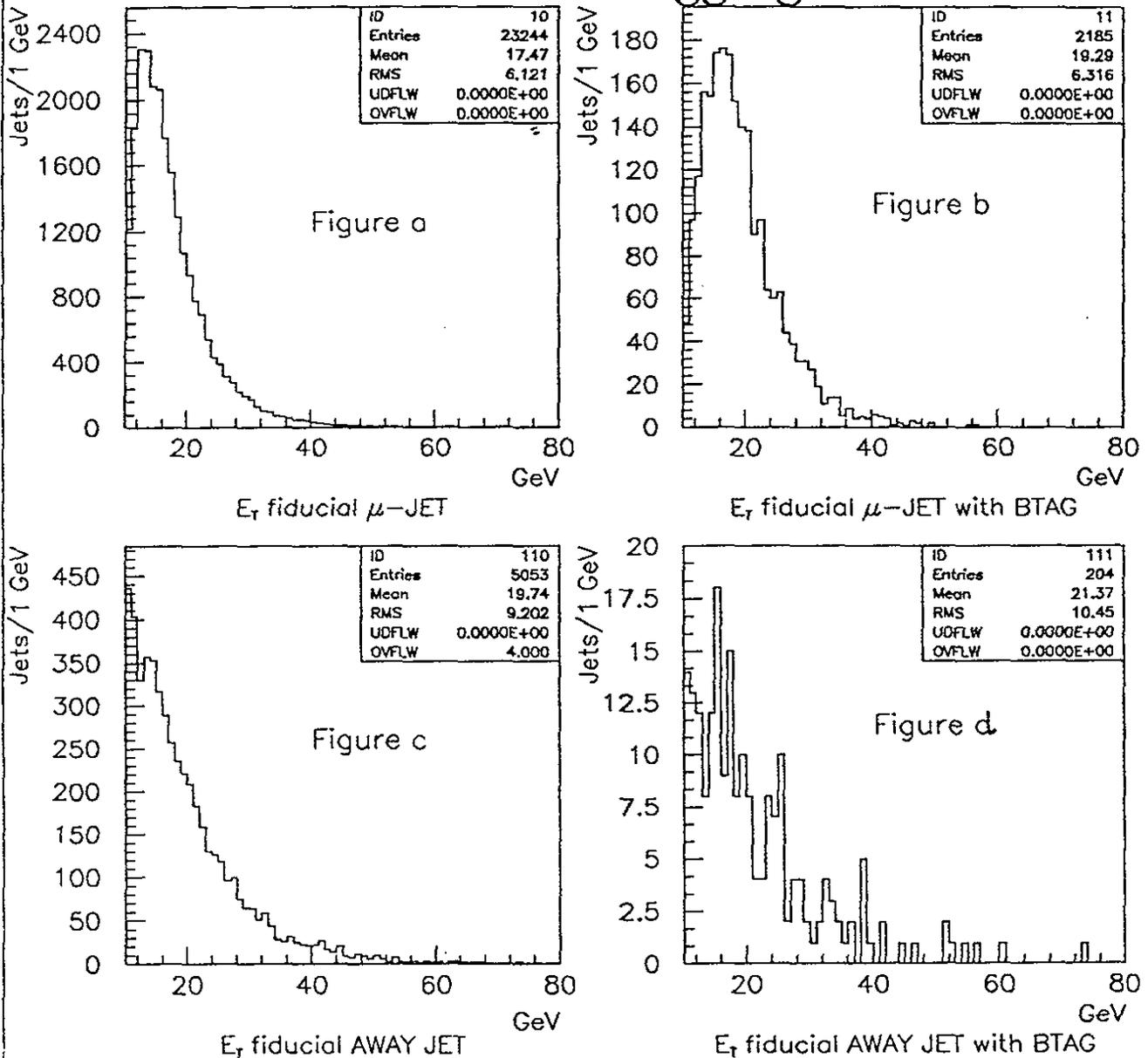


Figure 6.7:  $E_T$  spectrum of muon (*away*) jets contained in the SVX fiducial volume before  $b$  tagging in figure a (figure c) and after  $b$  tagging in figure b (figure d). The background subtraction and normalization by  $F_b(\mu)$  ( $F_b(\textit{away})$ ) has not been performed.

## Inclusive muons: b-tagging efficiencies

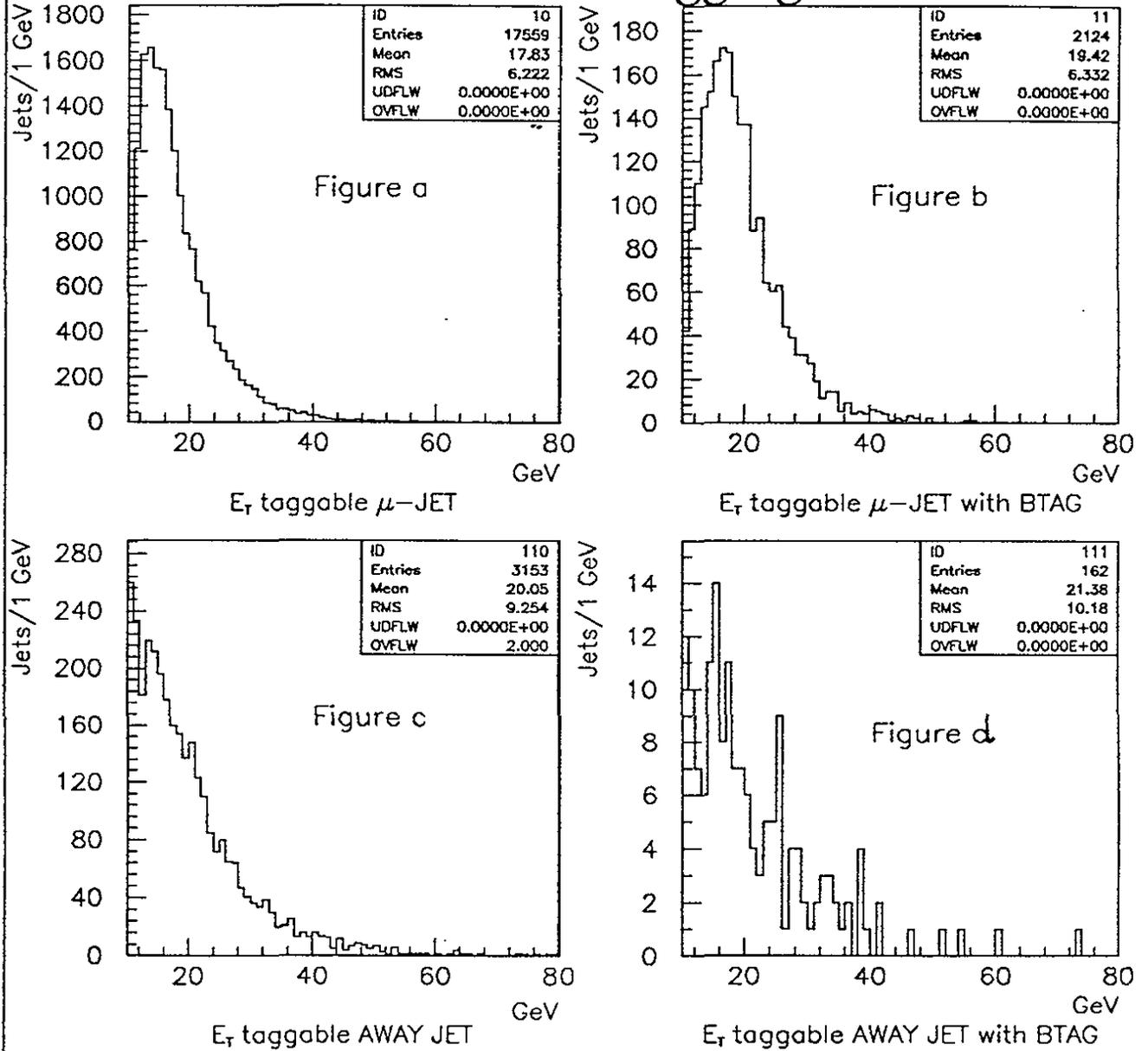


Figure 6.8:  $E_T$  spectrum of muon (*away*) jets with  $\geq 2$  well measured SVX tracks before  $b$  tagging in figure a (figure c) and after  $b$  tagging in figure b (figure d). The background subtraction and normalization by  $F_b(\mu)$  ( $F_b(\textit{away})$ ) has not been performed.

in events with both muon and away jets containing  $\geq 2$  well measured tracks one finds:

$$\epsilon^{doubles}(\mu, qual) = 0.26 \pm 0.07. \quad (6.22)$$

These values are consistent with those obtained using single  $b$ -tagging rates. However, it is not obvious a priori that comparing results obtained with the single-tag and double-tag methods (as well as comparing experimental results of the double-tag method with the corresponding Montecarlo results) is a reasonable thing to do. This is discussed in (the appendix of) section 6.4.

### 6.1.10 Montecarlo $b\bar{b} \rightarrow \mu X$ Sample

The  $b$ -tagging efficiencies measured in the data are compared to those obtained with a Montecarlo event sample, consisting of processes of direct production of  $b\bar{b}$  pairs ( $gg, q\bar{q}, gq \rightarrow b\bar{b}$ ) generated with ISAJET. Further requirements applied are: at least one  $b$  quark of  $P_t(b) > 15$  GeV/ $c$  and at least one  $b$  quark decaying semileptonically,  $b \rightarrow \mu\bar{\nu}_c$ , with the  $\mu$  produced in  $|\eta| < 0.8$ . Gluon splitting, as a third order perturbative diagram in  $\alpha_s$ , is not included;  $b$ -quark fragmentation is modeled using the Peterson function [124], tuned to reproduce the experimental results [125] of  $e^+e^-$  colliders;  $b$  hadrons are decayed using the CLEO Montecarlo program; primary interaction vertices are generated using a gaussian distribution with  $\sigma_x = \sigma_y = \sigma_{beam} = 35\mu\text{m}$  in the transverse plane and  $\sigma_z = 30$  cm, in such a way as to reproduce the widths of the  $p\bar{p}$  interaction region at CDF. Events thus generated are then passed through CDFSIM and the same analysis applied to the data.

The SVX simulation has been tuned [126] to reproduce: (A)  $\langle \epsilon_{hit}(i) \rangle$  and  $\langle \sigma_{hit}(i) \rangle$ , that is, the measured values of the reconstruction efficiency and the single-hit resolution, averaged over each layer ( $i=1,2,3,4$ ); (B) the average number of hits not associated with tracks in the data, on every layer; (C) the average number of malfunctioning strips. An event sample whose time distribution during the 1992-93 data taking period is weighted with the corresponding luminosity was used to tune the SVX simulation. This allows to include, on average, the effect of the Tevatron radiation on the performance of the SVX.

Using SVX tracks, the PV is reconstructed with the VXPRIM algorithm and  $b$  tags are found with the  $D - \phi$  algorithm. Exploiting the fact that in this sample  $F_b(\mu) = F_b(away) = 1$ ,  $b$ -tagging efficiencies are determined for fiducial jets (see table 6.1) and for jets with  $\geq 2$  good SVX tracks (see table 6.2). These tables contain also the corresponding efficiencies measured in the data and the scale factor between them and the efficiencies determined in the Montecarlo:  $F_{scale}(\epsilon_b(fid)) = \epsilon_b(fid)_{data} / \epsilon_b(fid)_{Montecarlo}$  and  $F_{scale}(\epsilon_b(qual)) = \epsilon_b(qual)_{data} / \epsilon_b(qual)_{Montecarlo}$ . The study of double  $b$  tags in the Montecarlo is discussed in the appendix 6.4.

| Efficiency              | Data            | Montecarlo      | $F_{scale}(\epsilon_b(fid))$ |
|-------------------------|-----------------|-----------------|------------------------------|
| $\epsilon_b(\mu, fid)$  | $0.22 \pm 0.04$ | $0.22 \pm 0.01$ | $1.00 \pm 0.19$              |
| $\epsilon_b(away, fid)$ | $0.22 \pm 0.04$ | $0.22 \pm 0.03$ | $1.00 \pm 0.23$              |

Table 6.1: Inclusive muon sample:  $b$ -tagging efficiencies for SVX-fiducial jets in the data and in the Montecarlo.

| Efficiency               | Data            | Montecarlo      | $F_{scale}(\epsilon_b(qual))$ |
|--------------------------|-----------------|-----------------|-------------------------------|
| $\epsilon_b(\mu, qual)$  | $0.29 \pm 0.06$ | $0.27 \pm 0.02$ | $1.07 \pm 0.24$               |
| $\epsilon_b(away, qual)$ | $0.29 \pm 0.05$ | $0.25 \pm 0.03$ | $1.16 \pm 0.24$               |

Table 6.2: Inclusive muon sample:  $b$ -tagging efficiencies for jets with  $\geq 2$  well measured SVX tracks in the data and in the Montecarlo.

For fiducial jets the weighted average of the scale factors is  $F_{scale}(\epsilon_b(fid))=1.00 \pm 0.15$ . Note that  $F_{scale}(\epsilon_b(fid))$  does not vary appreciably with the  $E_t$  of jets in the sample. The efficiencies for jets with  $\geq 2$  well measured SVX tracks have been reported in order to study the performance of the  $D - \phi$  algorithm with the track selection criteria; from now on, however, only quantities related to fiducial jets will be used. The value of  $F_{scale}(\epsilon_b(fid))$  will be used in section 8.5, to estimate the  $b$ -tagging efficiency of the  $D - \phi$  algorithm for top events in the  $e/\mu + jets$  channel. In the following section, the decays length of  $\mu$  and away jet tags is studied, in order to show that they are consistent with the lifetime of  $b$  hadrons.

### 6.1.11 Decay Length of $b$ Tags

The proper lifetime of a  $b$  hadron,  $\tau$ , and its transverse decay length in the PV reference system,  $L_{xy}$ , are related by the equation:

$$c\tau = L_{xy} \frac{M(B)}{P_t(B)}, \quad (6.23)$$

$$L_{xy} = R_B = |\vec{X}_{SV} - \vec{X}_{PV}|, \quad (6.24)$$

where  $M(B)$  and  $P_t(B)$  are the hadron mass and transverse momentum,  $\vec{X}_{PV}$  and  $\vec{X}_{SV}$  are the cartesian coordinates of the event PV and of the SV of the  $b$ -hadron decay. Using the invariant mass and transverse momentum of tracks associated with  $b$  tags,  $M_{SV}$  and  $P_{SV}$ , this equation can be rewritten as:

$$c\tau = L_{xy} \frac{M_{SV}}{P_{SV}} F_{corr}(P_{SV}), \quad (6.25)$$

$$F_{corr}(P_{SV}) = \frac{M(B)/P_t(B)}{M_{SV}/P_{SV}}. \quad (6.26)$$

$F_{corr}(P_{SV})$  is a correction factor which takes into account  $b$  prongs which have not been attached to the SV by the  $D - \phi$  algorithm.  $F_{corr}(P_{SV})$  has been measured using the Montecarlo sample described in the previous section. By correcting for the kinematic factors of equation 6.25, the measurement of  $L_{xy}$  can be converted into an estimate of  $c\tau$ , the proper decay length of  $b$  tags. Figure 6.9 shows the  $c\tau$  distribution for the  $b$  tags used to compute  $\epsilon_b(\mu, fid)$ ; the exponential fit reported in the figure gives a lifetime  $c\tau_{btag} = 444\mu\text{m}$ , to be compared with the  $b$ -hadron lifetime  $c\tau_B = 446 \pm 26\mu\text{m}$  [50]. Figure 6.10 shows the  $c\tau$  distribution of the  $b$  tags used to compute  $\epsilon_b(away, fid)$ : despite the lower statistics, also the  $b$  tags associated with jets recoiling against the inclusive muons are consistent with  $b$ -hadron decays.

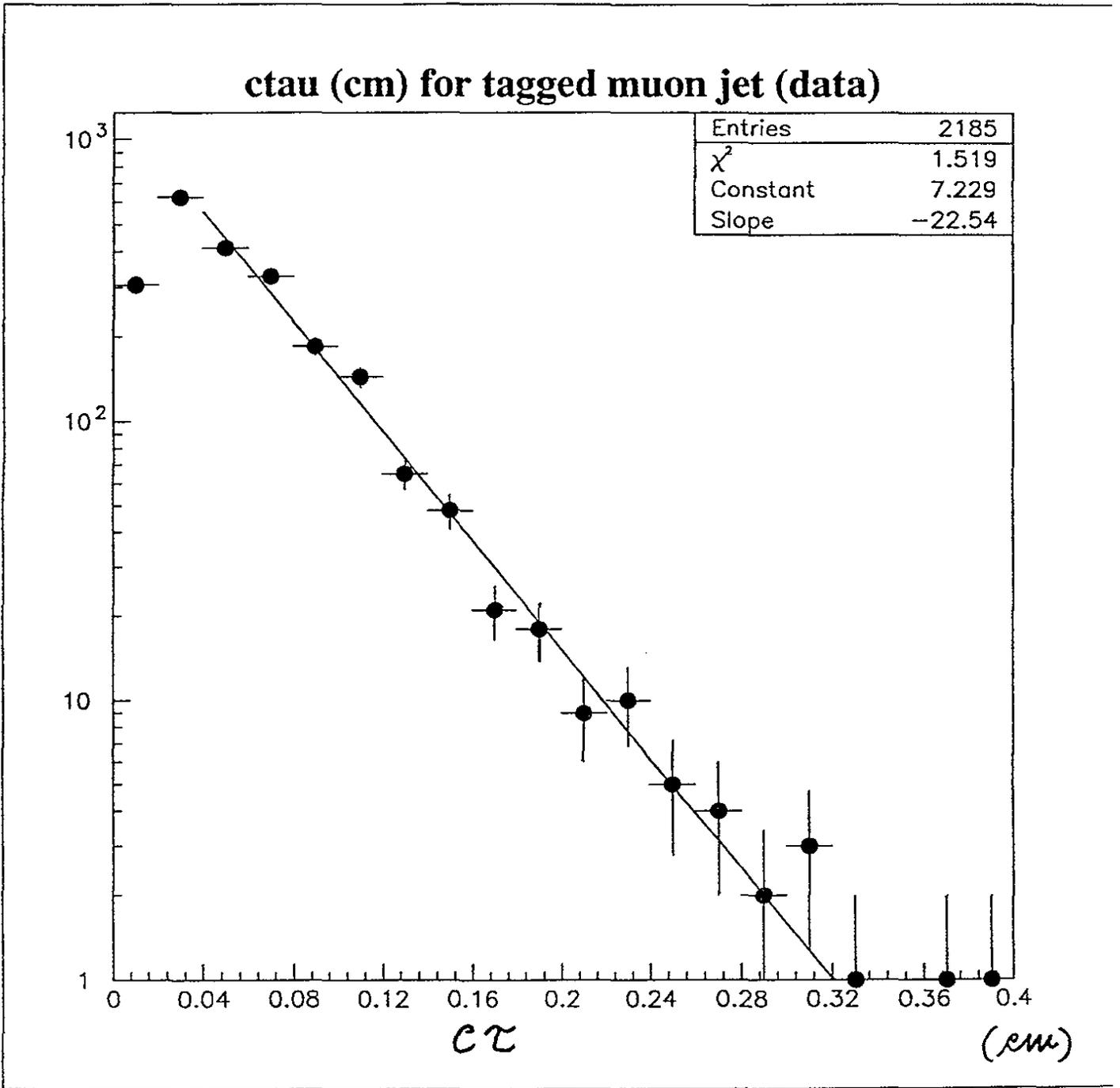


Figure 6.9: Proper lifetime distribution of  $b$  tags associated with SVX-fiducial muon jets. The exponential fit gives  $c\tau_{btag} = 444 \mu\text{m}$ , in agreement with the value expected for  $b$  hadrons.

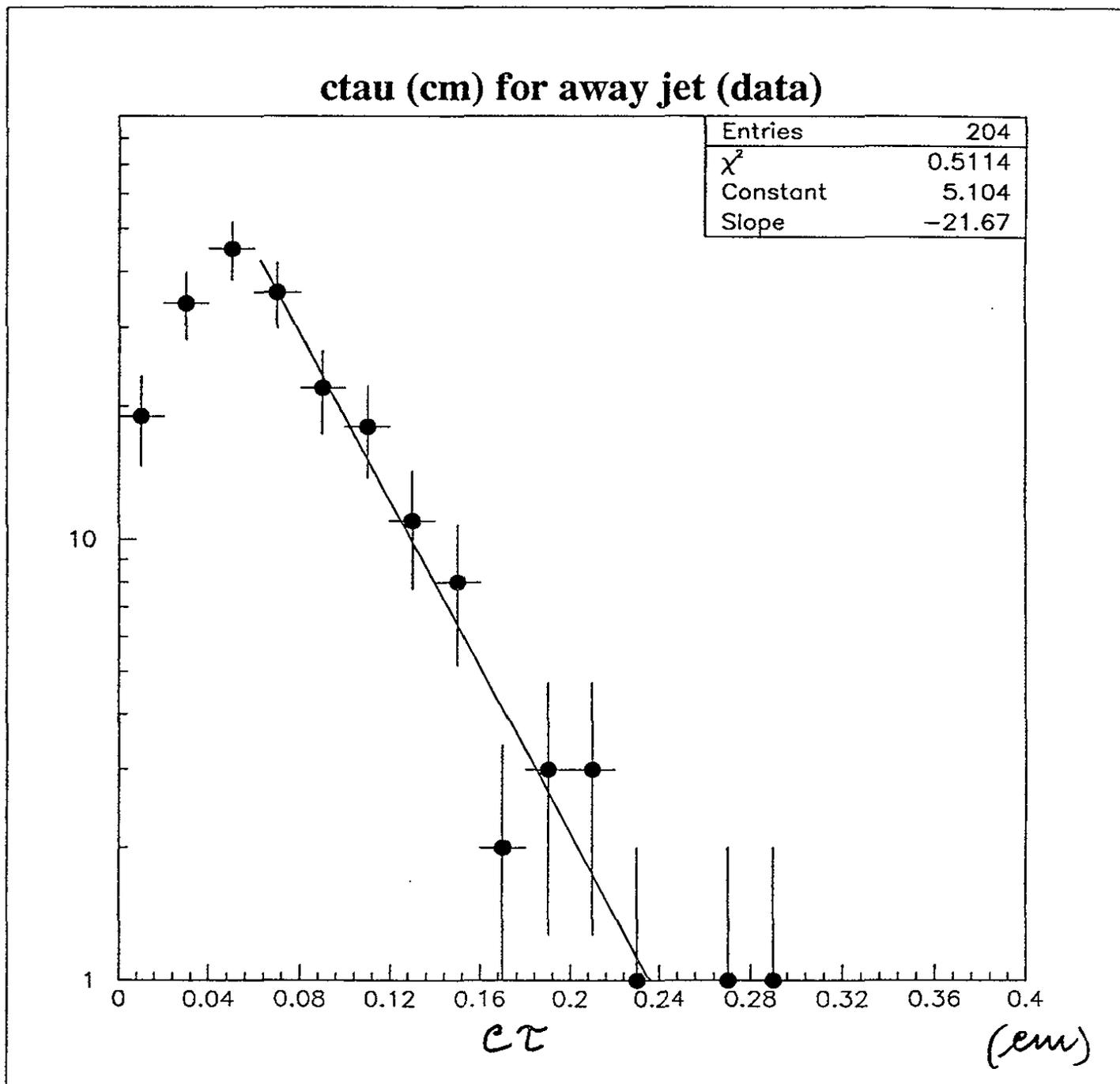


Figure 6.10: Proper lifetime distribution of  $b$  tags associated with SVX-fiducial recoil jets. The exponential fit gives  $c\tau_{btag} = 461 \mu\text{m}$ , in agreement with the value expected for  $b$  hadrons.

## 6.2 The Low $b$ -Content Sample: Inclusive Jets

The inclusive jet samples collected with the trigger thresholds  $E_t > 50$  GeV (“Jet\_50”),  $E_t > 100$  GeV (“Jet\_100”) and with the scalar sum of the calorimeter transverse energies  $> 200$  GeV (“ $\Sigma E_t$ ”) have a low HF content and are, therefore, useful to study the rate of fake  $b$  tags found by the  $D - \phi$  algorithm,  $\epsilon_f$ . This is done by searching for negative  $b$  tags (“negative Decay Length” or -DL tags, see section 5.2.5), which represent unlikely configurations of HF decays. Thus, neglecting the HF contribution, the negative  $b$  tags observed in inclusive jets provide an estimate of the rate of fake (non- $b$ ) SV’s due to tracking errors and residual contaminations from non PV particles, like photon conversions,  $K_s$ ’s and  $\Lambda$ ’s. This fake rate is an important parameter to measure in order to fully characterize the performance of the  $D - \phi$  tagger.

By studying positive  $b$  tags (“positive Decay Length” or +DL tags) in inclusive jets, in addition to completing the characterization of the  $D - \phi$  algorithm, one can get valuable information on the characteristics of “generic” jets, and build some analysis tools (the  $M$  matrices, see below) which are of fundamental importance in the determination of the background to the top signal in  $W +$  multijet events with the  $D - \phi$  algorithm (see section 8.3).

The following sources provide positive  $b$  tags:

- fake SV’s: the same sources (qualitatively and quantitatively) which give negative  $b$  tags, plus other components which may be absent in the negative tags and which are very difficult to model with the Montecarlo; for example, residual contributions of  $K_s$ ’s,  $\Lambda$ ’s and photon conversions might be more likely to populate positive tags (forward decays) than negative tags (backward decays);
- HF decays; this contribution depends on the HF content in generic jets (for example HF’s from gluon splitting,  $g \rightarrow c\bar{c}, b\bar{b}$ ), which is very low compared to the  $b$  fractions ( $F_b$ ) of jets in the inclusive muon sample (see [127]; this ref. estimates  $F_b \sim 2.2\%$  and  $F_c \sim 4.2\%$  for inclusive jets), but still larger than the contribution of fake SV’s; in fact, it will be shown that the ratio between the total numbers of positive to negative tags is  $\sim 2.8$  and that the  $cr$  distribution of positive  $b$  tags contains a large fraction of HF decays.

The knowledge of these two combined sources of  $b$  tags in generic jets is essential to estimate the background in the  $t \rightarrow Wb$  search with  $W +$  multijet events. In particular, since gluon splitting represents the main source of HF’s in background  $W +$  multijet events from QCD, one hopes to measure experimentally the fraction of  $b$  tags from  $g \rightarrow c\bar{c}, b\bar{b}$ , and compare it with the predictions of  $W +$  multijet + HF events from the available Montecarlo programs (see section 8.3).

Next section reports the measurement of the inclusive rate of positive/negative  $b$  tags per jet (that is, the global contribution of the SV sources listed above), and the positive/negative  $b$ -tagging rate is also parametrized as a function of two important variables characterizing jets: the jet transverse energy,  $E_t$ , and the number of SVX tracks with  $P_t > 0.4$  GeV/ $c$ ,  $|Z_{PV} - z_0| < 5$  cm contained in the jet cone of radius = 0.4, called *mult*. Such a probability will be indicated with  $M^+(E_t, mult)$  for positive  $b$  tags and  $M^-(E_t, mult)$  for

negative  $b$  tags. The elements of the  $M$  matrices are computed using the Jet\_50 sample, which has the largest statistics in the energy range populated by the majority of the jets in the inclusive  $W$  events. The independent Jet\_100 and  $\Sigma E_t$  samples are then used to compare the number of positive and negative tags observed in these samples with the predictions obtained from the  $M$  matrix parametrizations: the level of agreement between observations and predictions will determine the systematic error on the elements of the  $M$  matrices.

### 6.2.1 The “Jet\_50” Sample

This sample consists of events with  $|Z_{PV}| < 30$  cm, with at least one jet in the region  $|\eta| < 2.0$  which passed the  $E_t > 50$  GeV trigger. In such a sample we consider all jets (and not just the trigger jet) which satisfy  $|\eta| < 2.0$ ,  $E_t > 15$  GeV and are contained in the SVX fiducial volume (that is, their  $mult \geq 2$ ).  $E_t > 15$  GeV is the cut adopted for the top-quark search in the  $W +$  multijet sample (see chapter 7). The jet four-vectors have not been corrected for mismeasurements due to calorimeter non-linearities, cracks, etc. For this study, a sample of events whose time distribution during the 1992-93 data taking is weighted with the corresponding luminosity has been used. Thus the average effect of the Tevatron radiation on the SVX is taken into account.

Figure 6.11.a shows the jet  $E_t$  spectrum before  $b$  tagging: the effect of the trigger threshold around 50 GeV is evident; the sample contains about 46K jets. Figures 6.11.b and 6.11.c show the  $E_t$  of positive (680 jets) and negative (239 jets)  $b$  tags. To study the dependence of the  $b$ -tagging rate on  $E_t$ , the spectrum is divided in the regions  $15 < E_t < 30$  GeV,  $30 < E_t < 50$  GeV and  $E_t > 50$  GeV. Table 6.3 shows the fraction of positive and negative  $b$ -tagged jets in the full  $E_t$  spectrum, as well as in the three regions mentioned above. For  $E_t > 15$  GeV, the rate of  $D - \phi$  tagging errors (negative  $b$  tags) is  $\epsilon_f = \epsilon_{NEG} = (0.52 \pm 0.03)\%$ , while the fraction of found HF (positive tags - negative tags) is  $\epsilon_{POS} - \epsilon_{NEG} = (0.96 \pm 0.07)\%$ . The ratio of the total number of positive to negative  $b$  tags is 2.8. Positive  $b$  tags are dominated by real HF's.

| $E_t$ Range               | Positive Tagging Rate (%) | Negative Tagging Rate (%) |
|---------------------------|---------------------------|---------------------------|
| $E_t > 15\text{GeV}$      | $1.48 \pm 0.06$           | $0.52 \pm 0.03$           |
| $15 < E_t < 30\text{GeV}$ | $1.72 \pm 0.15$           | $0.7 \pm 0.1$             |
| $30 < E_t < 50\text{GeV}$ | $1.7 \pm 0.1$             | $0.49 \pm 0.07$           |
| $E_t \geq 50\text{GeV}$   | $1.32 \pm 0.07$           | $0.48 \pm 0.04$           |

Table 6.3: Fractions (in %) of fiducial jets with positive and negative  $b$  tags in the inclusive jet sample (“Jet\_50” sample).

The actual composition of positive  $b$  tags in terms of fake SV's (present in the same amount among the negatives) and HF is shown from the behavior of their proper lifetime,  $c\tau$ , reported in figure 6.12, together with the exponential distribution corresponding to the  $c\tau$  measured for  $b$ -tagged recoil jets in the inclusive muon sample.

The variation of  $\epsilon_{POS}$  with time (as a possible systematic effect) has been studied by dividing the whole sample in two parts, one containing events collected in the first half of the data taking and the other one in which the Tevatron radiation caused a partial

## Inclusive Jets: positive and negativetags

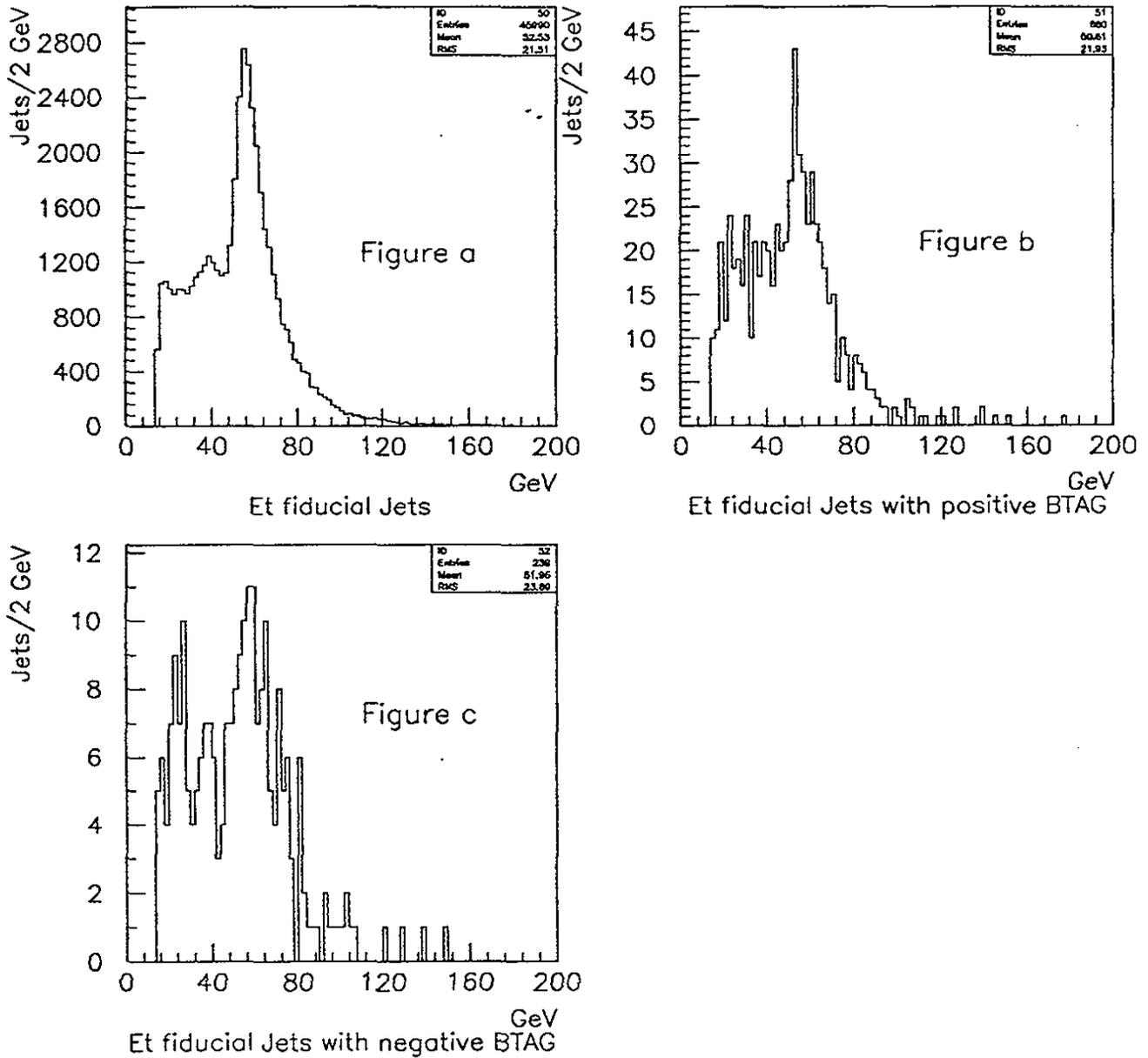


Figure 6.11:  $E_t$  spectrum of inclusive jets (“Jet\_50” sample) before  $b$  tagging (a),  $E_t$  spectrum of positive  $b$ -tagged jets (b),  $E_t$  spectrum of negative  $b$ -tagged jets (c).

### c $\tau$ (cm) for jet tags in JET50 data

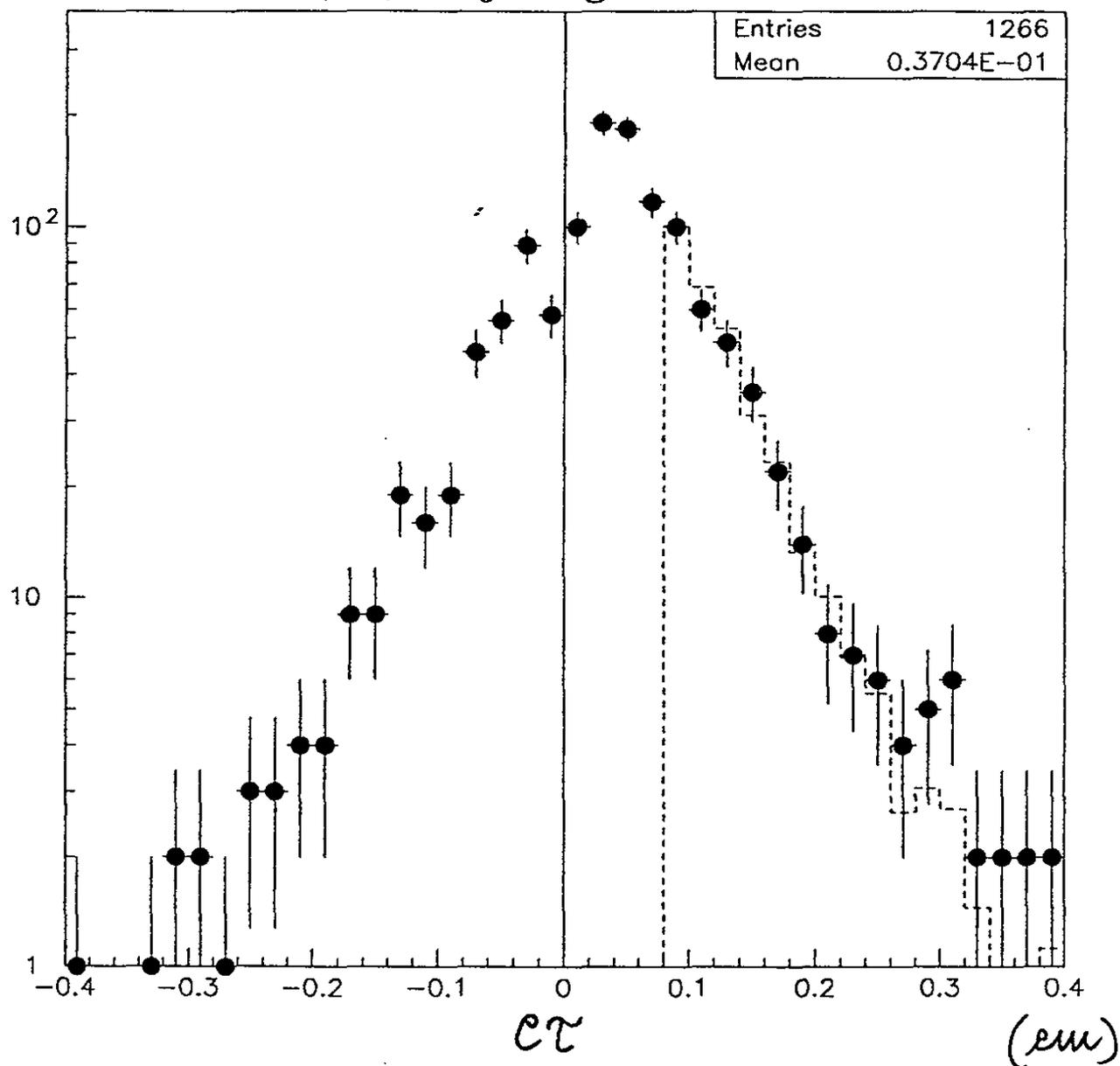


Figure 6.12: Proper lifetime distribution of negative  $b$  tags ( $c\tau < 0$ ) and positive  $b$  tags ( $c\tau > 0$ ). The dashed histogram is obtained by summing: (1) the  $c\tau$  distribution of negative  $b$  tags (which models fake SV's) and (2) an exponential distribution with slope equal to  $c\tau_{btag}(away) = 461\mu\text{m}$ , as measured for recoil jets in the inclusive muon sample, normalized by the difference of the total number of positive  $b$  tags and the total number of negative  $b$  tags (which models the HF component). The good agreement between data points with  $c\tau > 0$  and the histogram demonstrates that the actual composition of  $b$ -tagged inclusive jets in terms of fake SV's and HF's is understood correctly.

worsening of the performance of the first SVX layer ( $\epsilon_{hit}(1^{st}half) \sim 90\%$ ,  $\epsilon_{hit}(2^{nd}half) \sim 63\%$ ). One finds an average value  $\epsilon_{POS} = (1.48 \pm 0.06)\%$  and  $\epsilon_{POS}(1^{st}half) = (1.40 \pm 0.09)\%$ ,  $\epsilon_{POS}(2^{nd}half) = (1.54 \pm 0.07)\%$ . The increase of  $\epsilon_{POS}$  in the second half is due to the increase of  $\epsilon_{NEG}$ , that is to increased fake SV's. The maximum percentage deviation from the average value, 4%, is taken as systematic uncertainty on  $\epsilon_{POS}$ , and it will be propagated on the  $M$  matrices.

The  $b$ -tagging rate increases with the number of SVX tracks contained in jets ( $mult$ ) as shown by figures 6.13 a,b,c (6.14 a,b,c) for positive (negative)  $b$  tags, where a,b,c refer to the three different jet  $E_t$  regions. The contents of the histograms of figures 6.13 a,b,c (6.14 a,b,c) define the values of the  $M^+(E_t, mult)$  ( $M^-(E_t, mult)$ ) matrix elements in the three  $E_t$  bins and for  $mult \geq 2$ . For any given event sample with jets,  $M^+$  and  $M^-$  can be used to assign to jets of a given  $E_t$  and  $mult$ ,  $N(E_t, mult)$ , the corresponding  $b$  tagging probability. The total number of positive  $b$  tags expected (predicted) in the sample,  $N^+$ , is expressed in the following way:

$$N^+ = \sum_{E_t, mult} M^+(E_t, mult) \cdot N(E_t, mult). \quad (6.27)$$

$N^+$  is the predicted number of  $b$  tags due to HF's and fake SV's. Similarly, using the  $M^-$  matrix:

$$N^- = \sum_{E_t, mult} M^-(E_t, mult) \cdot N(E_t, mult), \quad (6.28)$$

one gets the prediction of the expected number of negative  $b$  tags, due to fake SV's. This is the method adopted to calculate the number of background (positive)  $b$  tags associated with  $\mu$  and away jets quoted in section 6.1.6:  $N_f(non - b) = N^+$  (for  $N_q(non - b)$  we used a matrix similar to  $M^+$ , whose matrix elements have been determined by counting the track multiplicity in jets as a function of  $\geq 2$  well measured SVX tracks). Since  $N^+$  also contains an HF contribution, it represents a conservative estimate of the non- $b$  background.

If one assumes that jets in  $W$  events are similar to Jet\_50 inclusive jets, and, in particular, that the HF content of jets in the two samples is comparable, equation 6.27 will allow to estimate the number of positive  $b$  tags expected in the  $W$  sample from sources different from decays of  $t\bar{t}$  pairs. Equation 6.28 will allow to estimate the number of positive and negative  $b$  tags due to fake SV's. The assumption that, for  $b$ -tagging purposes, jets in  $W$  events are similar to inclusive jets will be discussed quantitatively in chapter 8. For the moment, note that figure 6.12 indicates that the composition of positive and negative  $b$  tags in inclusive jets is understood correctly; this justifies the use of the  $M^+$  matrix to predict the background to the  $t\bar{t}$  signal coming from  $W + b\bar{b}$  and  $W + c\bar{c}$  processes.

## 6.2.2 The "Jet\_100" and " $\Sigma E_t$ " Samples

The selection of events of the Jet\_100 sample is similar to the Jet\_50 sample, but the trigger jet has to satisfy  $E_t > 100$  GeV. The  $\Sigma E_t$  sample consists of events collected with a trigger based on the requirement that the sum of transverse energies of all hadronic and electromagnetic towers with  $E_t(tower) > 1$  GeV exceeds 200 GeV. In the analysis described in this section events collected with this trigger of  $\Sigma E_t \geq 300$  GeV are used. Since these two samples contain generic jets like the Jet\_50 sample, they are studied to check the predictive power of the  $M$  matrices. In addition, events in these sample are characterized by a large flow of

## Fraction of positive tags vs SVX multiplicity

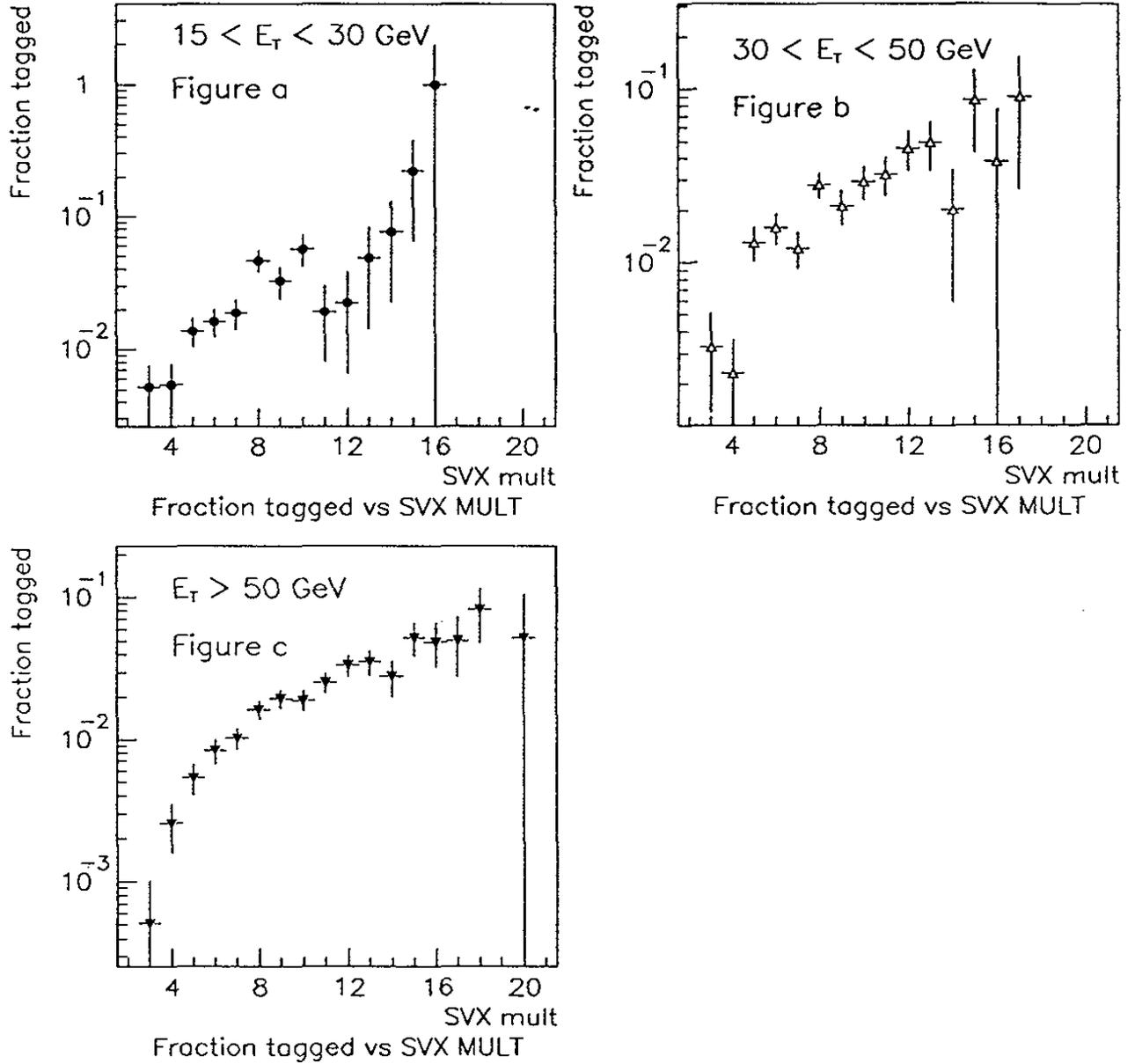


Figure 6.13: Fraction of inclusive fiducial jets associated with positive  $b$  tags as a function of the number of SVX tracks in the jet, in three different jet  $E_t$  bins.

### Fraction of negative tags vs SVX multiplicity

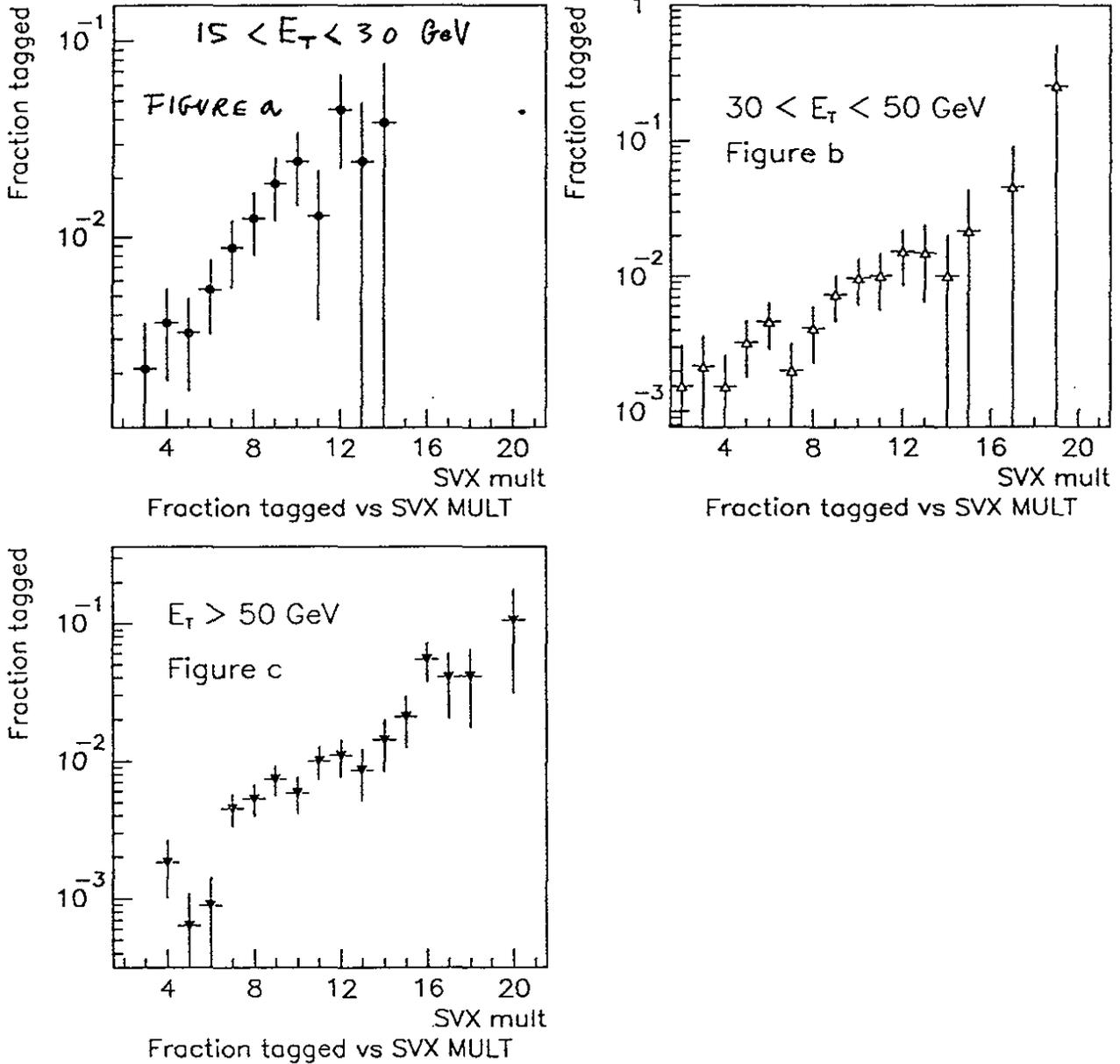


Figure 6.14: Fraction of inclusive fiducial jets associated with negative  $b$  tags as a function of the number of SVX tracks in the jet, in three different jet  $E_t$  bins.

transverse energy, many high-energy, collimated jets and a high track multiplicity, similarly to the  $W + 3, 4$  jets which will be studied in chapter 8 for the top-quark search. In particular,  $\Sigma E_t > 300$  GeV events are the CDF data sample which is by far the most suitable to test the stability and reliability of the SVX tracking, of the  $b$  tagging algorithm and of the  $M$ -matrix predictions. Table table 6.4 reports the total number of positive and negative  $b$  tags observed in the two samples and the number of those predicted by the  $M$  matrices.

| Sample       | # Positive Tags Observed | # Positive Tags Predicted | # Negative Tags Observed | # Negative Tags Predicted |
|--------------|--------------------------|---------------------------|--------------------------|---------------------------|
| Jet_100      | 83                       | 124                       | 42                       | 49                        |
| $\Sigma E_t$ | 359                      | 381                       | 170                      | 155                       |

Table 6.4: Comparison between the number of positive and negative  $b$  tags observed in the “Jet\_100”, “ $\Sigma E_t$ ” samples and the number of tags predicted using the  $M^+$  and  $M^-$  matrices built from the “Jet\_50” sample.

First of all, note that the predictions are conservative, that is, they exceed the number of observed tags, or are consistent with it within the statistical error. For positive  $b$  tags, the maximum deviation occurs for the Jet\_100 sample:  $(124-83)/124 = 33\%$ ; for the  $\Sigma E_t$  sample:  $(381-359)/381 = 6\%$ . For negative  $b$  tags, the maximum deviation occurs for the Jet\_100 sample:  $(49-42)/49 = 14\%$ ; for the  $\Sigma E_t$  sample:  $(170-155)/170 = 10\%$ . Conservatively, these maximum deviations from the central values of the predictions are used to estimate the systematic error on the elements of the  $M$  matrices:

$$\text{systematic error on } M^+ : 33\%, \quad (6.29)$$

$$\text{systematic error on } M^- : 14\%. \quad (6.30)$$

These deviations (systematic errors) are probably related to the fact that the Jet\_50 statistics used to build the  $M$  matrices decreases for  $E_t(\text{jet}) > 100$  GeV (see figure 6.11): therefore, the  $M$  matrices can lose some of their predictive power. For example, for energetic, collimated jets in the Jet\_100,  $\Sigma E_t$  samples, the cut on the number of hits shared with other tracks (and the ad-hoc “refit” applied to tracks passing this cut, but having 1 or 2 shared hits, see section 5.2.2, point B) can limit the tagging rate more than for the lower-energy less-collimated jets which have been used to build the  $M$  matrices. Another possible effect is the following: the Jet\_100 and  $\Sigma E_t$  events contain several jets, some of which, despite being fiducial, are located at the border of the SVX acceptance and are, therefore, tagged with a lower probability; in fact, the same  $M$ -matrix elements are constant in  $|\eta| < 1.0$  and decrease in  $1.0 < |\eta| < 2.0$ ; since the statistics in  $1.0 < |\eta| < 2.0$  are low,  $M$ -matrix elements are basically determined by the tagging rate in  $|\eta| < 1.0$ . This implies that when the  $M$  matrices are applied to other samples, one can get overestimates of the number of expected tags in  $1.0 < |\eta| < 2.0$ . As a confirmation of these hypotheses, note that the other two CDF  $b$  tagging algorithms [53] use  $M$  matrices defined with a more optimized  $E_t, \eta$  binning ( $E_t$  divided into 10 GeV bins and extended to higher energies, two  $\eta$  bins  $|\eta| < 1.0$  and  $1.0 < |\eta| < 2.0$ ) than the one adopted in this analysis. Using these two algorithms and a large variety of samples (events with generic jets, like Jet\_100 and , events with photons + jets) with widely different

characteristics (in terms of jet  $E_t$ , number of jets, total  $\Sigma E_t$ ) one estimates systematic errors of 13 % on  $M^+$  and of 35 % on  $M^-$  [128]. The lower systematic error on  $M^+$  measured for these algorithms w.r.t. the  $D - \phi$  tagger may be due to the adoption of more sensitive  $M$  matrices above 100 GeV  $E_t$ , as explained above; the larger systematic error on  $M^-$ , instead, may be related to the fact that the other two algorithms do not refit tracks with shared hits; this can have an effect in samples with a high flux of transverse energy and high multiplicity of jets and tracks. Further studies are in progress to optimize the  $E_t$  and  $\eta$  grid used by the  $D - \phi$  tagging analysis to build the  $M$  matrices; at the time being, however, the maximum fractional variations of observations w.r.t. predictions are taken as systematic errors: 33 % for  $M^+$  and 14 % for  $M^-$ .

### 6.3 Summary of Results

The  $D - \phi$   $b$  tagger has been characterized in detail using appropriate control data samples and the corresponding Montecarlo samples.

- Events containing inclusive muons of  $P_t > 9$  GeV, enriched in  $b \rightarrow \mu \bar{\nu}_\mu c$  decays.
  - 1)  $F_b$  was measured for the jet containing the muon ( $F_b(\mu) = 0.40 \pm 0.08$  [115]) and for the jet recoiling against the muon ( $F_b(\text{away}) = 0.14 \pm 0.02$  [119]).
  - 2) The measured  $b$ -tagging efficiencies are:
    - 0.22  $\pm$  0.04 for SVX-fiducial muon jets,
    - 0.29  $\pm$  0.06 for SVX-fiducial muon jets well measured in the SVX,
    - 0.22  $\pm$  0.04 for SVX-fiducial recoil jets,
    - 0.29  $\pm$  0.05 for SVX-fiducial recoil jets well measured in the SVX.
  - 3) The measured efficiencies are in good agreement with those measured in the  $b\bar{b} \rightarrow \mu X$  sample obtained from the ISAJET + CLEO + CDFSIM Montecarlo program. For fiducial jets one finds the data/Montecarlo efficiency scale factor  $1.00 \pm 0.15$ . This result is one of the ingredients which are needed in order to normalize to data the  $b$ -tagging efficiency for  $p\bar{p} \rightarrow t\bar{t} \rightarrow e/\mu + \text{multijet}$  events, generated with the same Montecarlo program (see section 8.5).
  - 4) The proper lifetime distribution of  $b$  tags associated with muon and recoil jets is in good agreement with the world-average  $c\tau_B$ . The  $D - \phi$   $b$  tags contain a high fraction of  $b$ -hadron decays.
- Inclusive jet events, with a low (see [127])  $b$ -hadron content.
  - 1) Using negative  $b$  tags, one gets the following estimate of the fake SV rate per jet:  $\epsilon_f = \epsilon_{NEG} = (0.52 \pm 0.03)\%$ .
  - 2) The difference of positive and negative  $b$  tags gives the following estimate of the  $D - \phi$   $b$ -tagged HF component:  $\epsilon_{POS} - \epsilon_{NEG} = (0.96 \pm 0.07)\%$ . In addition, the composition of positive  $b$  tags in terms of real HF's and fake SV's is confirmed by the comparison of their proper lifetime with  $c\tau_B$  and the negative  $b$  tags (see figure 6.12).

- 3) Using positive (negative)  $b$  tags, one calculates the  $b$ -tagging probability matrix of generic jets  $M^+(E_t, mult)$  ( $M^-(E_t, mult)$ ), parametrized as function of jet  $E_t$  and associated SVX track multiplicity.  $M^+$  and  $M^-$  can be used to predict the number of positive and negative  $b$  tags in any sample of generic jets. From the study of the time variation of  $\epsilon_{POS}$  one derives a  $\sim 4\%$  systematic uncertainty on the  $M$  matrices. From the explicit comparison of the number of  $b$  tags observed in independent samples and the number of expected  $b$  tags based on  $M^+$ ,  $M^-$  predictions one determines a further systematic error of 33 % on  $M^+$  and of 14 % on  $M^-$ .

On the basis of these results, the  $D-\bar{\phi}$  algorithm and the  $M$  matrices seem adequate to the top-quark search with tagging of SV's from  $b$ -hadron decays. The next chapter describes the selection of the  $W +$  multijet sample, while in the last chapter the  $b$  tagging will be applied to this sample to try to establish a  $t\bar{t}$  production signal.

## 6.4 Appendix to 6.1.9: Calculation of $\epsilon_b$ for Muon Jets Using Double $b$ Tags

The expression of the  $b$ -tagging efficiency for muon jets obtained with the double-tag method is more complicated than that for the single-tag method (equation 6.10). All variables, formulas and results listed here refer in the same way to fiducial jets and to jets with  $\geq 2$  well measured SVX tracks. We need some definitions: let us consider all possible compositions of muon and away jets pairs, in terms of jets originated by  $b$  quarks (indicated by " $b$ ") and jets not originated by  $b$  quarks (indicated by " $f$ ", which stands for fake  $b$ 's). The possible compositions and the number of events for each type of composition are shown schematically in table 6.5:

| Number of events | $N_1$ | $N_2$ | $N_3$ | $N_4$ |
|------------------|-------|-------|-------|-------|
| $\mu$ jet        | $f$   | $f$   | $b$   | $b$   |
| away jet         | $f$   | $b$   | $f$   | $b$   |

Table 6.5: Events with a muon and an away jet: composition (and corresponding number of events,  $N_i$ , with  $i=1,2,3,4$ ) in terms of jets originated by  $b$  quarks (" $b$ ") and not originated by  $b$  quarks (fake or " $f$ ");  $N_1 + N_2 + N_3 + N_4 = N$ , where  $N$  is the total number of events with fiducial muon and away jets or with muon and away jets both with  $\geq 2$  well measured SVX tracks.

Other useful equations are:

$$F_b(\mu) = \frac{N_3 + N_4}{N}, \quad (6.31)$$

$$F_b(a) = \frac{N_2 + N_4}{N}, \quad (6.32)$$

$$N_1 = N \cdot (1 - F_b(\mu)) - N_2, \quad (6.33)$$

$$N_3 = N \cdot (F_b(\mu) - F_b(a)) + N_2. \quad (6.34)$$

where we redefined  $F_b(\text{away}) = F_b(a)$ .

In the double-tag method, one considers the number of events with tagged away jets,  $N(\text{away tag})$  and tries to measure the efficiency for tagging muon jets,  $\epsilon^{double}(\mu)$ . The number of double tags in this sample,  $N^{double}(\text{tag})$ , is given by the following expression:

$$N^{double}(\text{tag}) = \epsilon_f(\mu) \cdot (N_1 \cdot \epsilon_f(a) + N_2 \cdot \epsilon(a)) + \epsilon^{double}(\mu) \cdot (N_3 \cdot \epsilon_f(a) + N_4 \cdot \epsilon(a)), \quad (6.35)$$

where we redefined  $\epsilon(a) = \epsilon_b(\text{away})$  and where  $\epsilon_f(a)$  and  $\epsilon_f(\mu)$  are the rates of fake  $b$  tags (see section 6.2.1) for away and muon jets, respectively.

Using equation 6.33, the first term of equation 6.35 becomes:

$$\epsilon_f(\mu) \cdot [N \cdot (1 - F_b(\mu)) \cdot \epsilon_f(a) + N_2 \cdot (\epsilon(a) - \epsilon_f(a))]. \quad (6.36)$$

Using equations 6.32 and 6.34, the second term of equation 6.35 becomes:

$$\epsilon^{double}(\mu) \cdot [N \cdot F_b(a) \cdot \epsilon(a) + N \cdot (1 - F_b(a)) \cdot \epsilon_f(a) - N \cdot (1 - F_b(\mu)) \cdot \epsilon_f(a) - N_2 \cdot (\epsilon(a) - \epsilon_f(a))]. \quad (6.37)$$

Since  $N \cdot F_b(a) \cdot \epsilon(a)$  is the number of away tags from real  $b$ 's and  $N \cdot (1 - F_b(a)) \cdot \epsilon_f(a)$  is the number of away tags from fake  $b$ 's, equation 6.37 can be rewritten as:

$$\epsilon^{double}(\mu) \cdot [N(\text{away tag}) + N \cdot (1 - F_b(\mu)) \cdot \epsilon_f(a) - N_2 \cdot (\epsilon(a) - \epsilon_f(a))]. \quad (6.38)$$

With these substitutions, equation 6.35 can be solved to extract  $\epsilon^{double}(\mu)$ :

$$\epsilon^{double}(\mu) = \frac{N^{double}(\text{tag}) - \epsilon_f(\mu) \cdot [N \cdot (1 - F_b(\mu)) \cdot \epsilon_f(a) + N_2 \cdot (\epsilon(a) - \epsilon_f(a))]}{N(\text{away tag}) - [N \cdot (1 - F_b(\mu)) \cdot \epsilon_f(a) + N_2 \cdot (\epsilon(a) - \epsilon_f(a))]} \quad (6.39)$$

Equation 6.39 is the exact expression of the  $b$ -tagging efficiency according to the double-tag method. Note that in the absence of background in away tags ( $\epsilon_f(a) = 0$ )  $\epsilon_b^{double}(\mu)$  would be independent of  $F_b(\mu)$ ; moreover, even for  $\epsilon_f(a)$  not zero ( $\epsilon_f(a) \sim 1\%$ ) the dependence of  $\epsilon_b^{double}(\mu)$  on  $F_b(\mu)$  is weaker than for the single-tag method. In the case of ISAJET Montecarlo samples, by construction  $F_b(\mu) = F_b(a) = 1$ ,  $N_4 = N$ ,  $N_1 = N_2 = N_3 = 0$ ; therefore, for the Montecarlo  $b\bar{b}$  sample, equation 6.39 reduces to:

$$\epsilon^{double}(\mu) = \frac{N^{double}(\text{tag})}{N(\text{away tag})}. \quad (6.40)$$

In the following, the approximations which allow for  $\epsilon^{double}(\mu)$  to be estimated from the data are listed.

For  $F_b(\mu)$ ,  $F_b(a)$ ,  $\epsilon(a)$ ,  $\epsilon(\mu)$ ,  $\epsilon_f(a)$ ,  $\epsilon_f(\mu)$  we use the values determined in the previous sections.  $N(\text{away tag}) = 204$  (162) for fiducial jets (jets with  $\geq 2$  well measured SVX tracks). To estimate  $N_2$ , we use equation 6.32, which can be rewritten as:

$$N_2 = N \cdot F_b(a) - N_4, \quad (6.41)$$

and the following approximation for  $N_4$ :

$$N_4 \sim \frac{N^{double}(\text{tag})}{\epsilon(a) \cdot \epsilon(\mu)}, \quad (6.42)$$

Equation 6.42 is based on the approximation that the background correction to the numerator of equation 6.39) is negligible. In fact, for fiducial jets one finds in this way  $N^{double}(tag) = 31$ ,  $N_4 = 640 \pm 235$ ,  $N_2 = 67$ , and:

$$\epsilon_f(\mu) \cdot [N \cdot (1 - F_b(\mu)) \cdot \epsilon_f(a) + N_2 \cdot (\epsilon(a) - \epsilon_f(a))] \sim 0.68\% \cdot [33.35 + 14.0] \sim 0.3. \quad (6.43)$$

For jets with  $\geq 2$  well measured tracks one finds  $N^{double}(tag) = 30$ ,  $N_4 = 357 \pm 123$ ,  $N_2 = 84$ , and:

$$\epsilon_f(\mu) \cdot [N \cdot (1 - F_b(\mu)) \cdot \epsilon_f(a) + N_2 \cdot (\epsilon(a) - \epsilon_f(a))] \sim 0.79\% \cdot [25 + 23] \sim 0.4. \quad (6.44)$$

Taking conservatively a 100 % uncertainty on  $N_2$ , one measures the following efficiencies:

$$\epsilon_{fid}^{double}(\mu) = 0.20 \pm 0.04, \quad (6.45)$$

$$\epsilon_{qual}^{double}(\mu) = 0.26 \pm 0.07. \quad (6.46)$$

In the Montecarlo sample one finds the efficiencies:

$$\epsilon_{fid}^{double}(\mu) = 0.31 \pm 0.07, \quad (6.47)$$

$$\epsilon_{qual}^{double}(\mu) = 0.39 \pm 0.08. \quad (6.48)$$

The comparison between  $\epsilon^{double}(\mu)$  and  $\epsilon^{single}(\mu)$  (equation 6.10) and the comparison between  $\epsilon^{double}(\mu)$  measured in the data with  $\epsilon^{double}(\mu)$  measured in the Montecarlo deserves a discussion. First of all, note that  $\epsilon^{double}(\mu)$  is the efficiency for tagging muon jets once the away jet has already been tagged, while  $\epsilon^{single}(\mu)$  is an inclusive measurement, much less dependent on the presence of and the characteristics of the away jet. As such,  $\epsilon^{double}(\mu)$  and  $\epsilon^{single}(\mu)$  are, a priori, two different quantities, which may have different numerical values due to kinematical correlations between the muon and the away jet. The size of these correlations has to be studied before attempting a comparison between single and double tags. It is known, for example, that the presence of away jets makes the muon jet  $E_t$  distribution harder w.r.t. the inclusive muon sample. The value of  $F_b(\mu)$  in events with an away jet could be different from that of the inclusive muon sample, in which the value  $F_b(\mu) = 0.40 \pm 0.8$  is measured. In addition, the value of  $\epsilon^{double}(\mu)$  extracted from the formula 6.39 ( $\epsilon^{double}(\mu) = 0.20$ ) is obtained by means of subtraction of a large background ( $\epsilon^{double}(\mu) = 0.15$  without background subtraction). In order to make sure that all contributions of double  $b$  tags are understood correctly, further studies of the kinematic correlations between muon and away jets in the inclusive sample and in the sample with tagged away jets are required.

The comparison of  $\epsilon^{double}(\mu)$  measured in the data and in the Montecarlo is influenced by the lack of statistics of the double tags in the Montecarlo and by the fact that the Montecarlo sample seems not to reproduce the correlations in the variable  $E_t(\mu - jet) - E_t(away)$  in the data: figure 6.15 shows  $E_t(\mu - jet)$  and the difference  $E_t(\mu - jet) - E_t(away)$  for double tags in the data (fiducial jets), while figure 6.16 shows the same quantities in the Montecarlo. The use of jets with  $\geq 2$  well measured tracks provides the same results. For these reasons, to normalize the Montecarlo  $b$ -tagging efficiency to data (that is, to compute  $F_{scale}(\epsilon)$ )  $\epsilon^{double}(\mu)$  will not be used. To study double tags with the Montecarlo it is necessary to tune the kinematic correlations to agree with the data. Double tags are by themselves an interesting subject, but not directly or necessarily related to the top-quark search with  $b$  tagging.

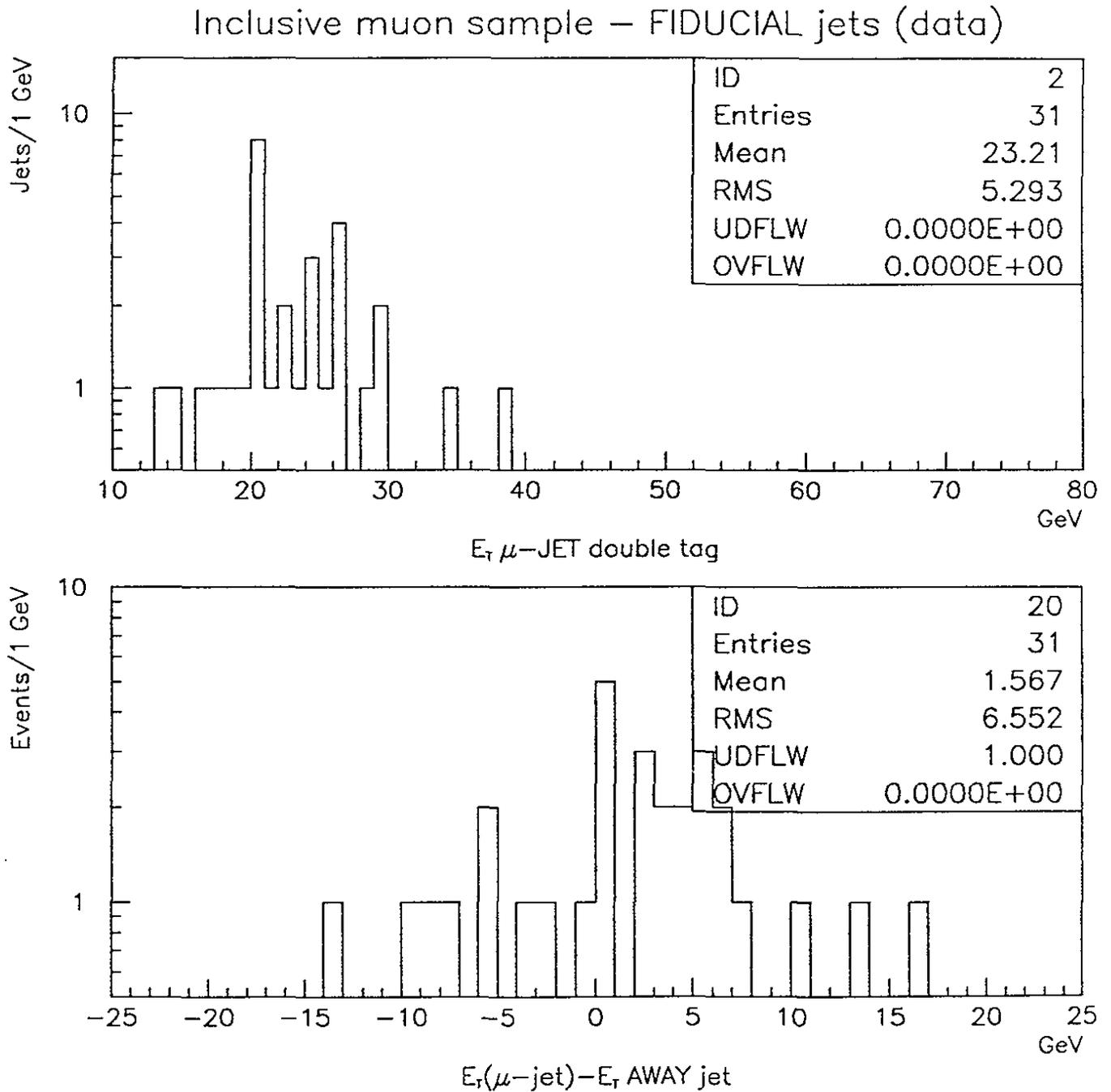


Figure 6.15: Inclusive muon sample: the top (bottom) plot shows the  $E_T(\mu\text{-jet})$  ( $E_T(\mu\text{-jet}) - E_T(\text{away jet})$ ) spectrum for events with double  $b$  tags (data, fiducial jets).

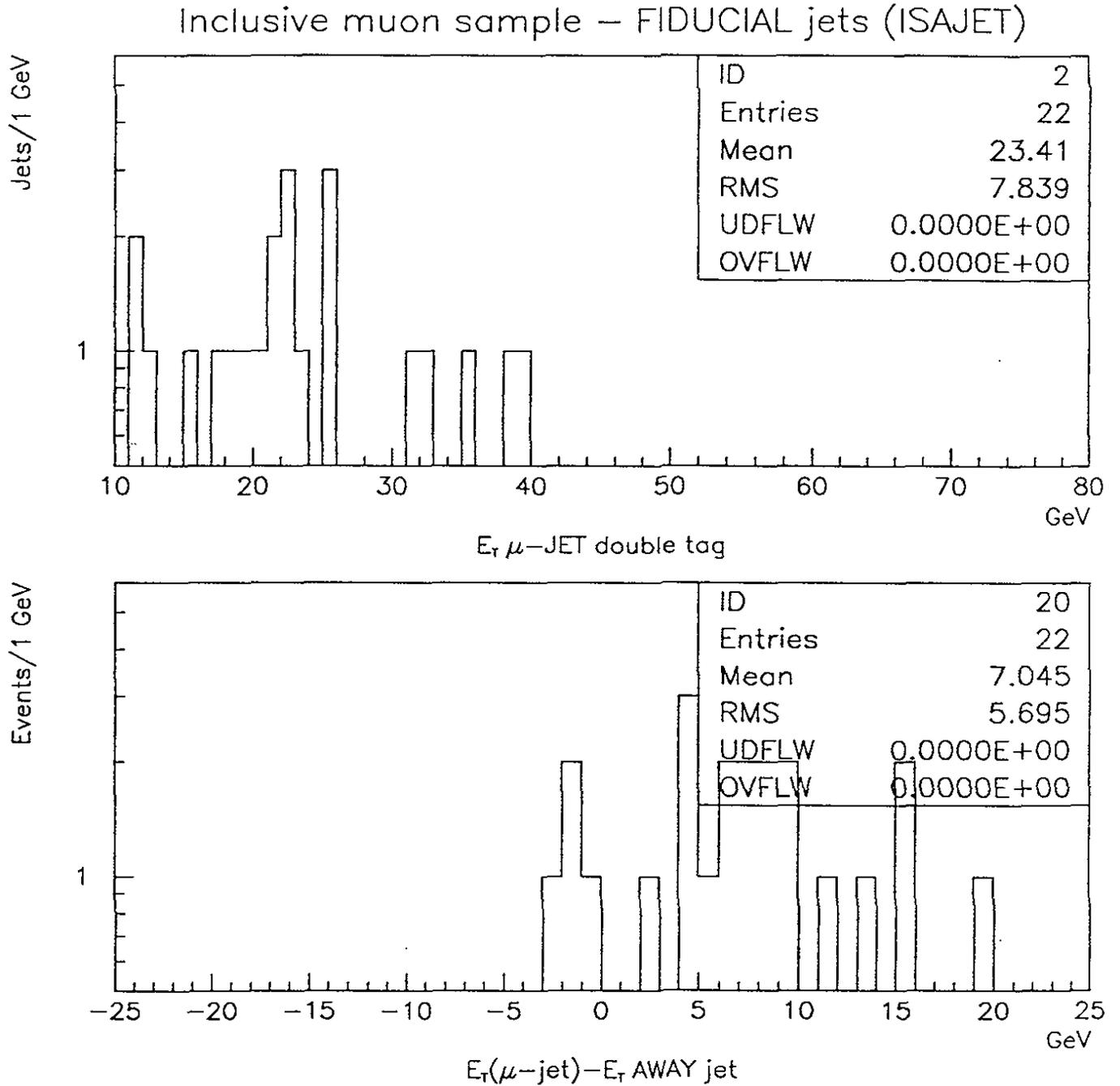


Figure 6.16: Inclusive muon sample: the top (bottom) plot shows the  $E_T(\mu\text{-jet})$  ( $E_T(\mu\text{-jet}) - E_T(\text{away jet})$ ) spectrum for events with double  $b$  tags (ISAJET, fiducial jets).

# Chapter 7

## Inclusive $W + \text{Multijet}$ Selection

### 7.1 $W \rightarrow e + \bar{\nu}_e, \mu + \bar{\nu}_\mu$ Selection

The first step in the search for  $t\bar{t} \rightarrow W^+W^-b\bar{b}$  processes in the lepton + jets channel is the selection of a sample of  $W \rightarrow e + \bar{\nu}_e, \mu + \bar{\nu}_\mu$  events. This selection is based on the identification of electrons and muons of  $P_t > 20 \text{ GeV}/c$  and on the measurement of the missing transverse energy, used to signal the presence of neutrinos.

#### 7.1.1 Identification of High $E_t$ Electrons

For the top-quark search in the  $e + \text{jets}$  channel one considers electrons in the central rapidity region ( $|\eta| < 1.0$ ), which pass the requirements of the trigger described in section 3.2.5.

Electron candidates must have a CTC track whose extrapolation to the calorimeter is associated with an electromagnetic cluster of transverse energy  $E_t(e)$ , formed around a seed tower of  $E_t > 3 \text{ GeV}$  and including contiguous towers of  $E_t > 100 \text{ MeV}$ . The size of the cluster is limited to 3 towers in pseudorapidity ( $\Delta\eta = 0.3$ ) and to 1 tower in azimuth ( $\Delta\phi = 15^\circ$ ). Cuts are then applied to the electromagnetic shower position measured by the CES, in order to select well measured electron candidates, far from the calorimeter edges. The fiducial volume is  $\sim 84\%$  of the solid angle for  $|\eta| < 1.0$ . Electrons from conversions are removed with high efficiency ( $88\% \pm 4\%$ ), by rejecting candidates without an associated VTX track and accompanied by a CTC track with which they form an invariant mass  $< 500 \text{ MeV}/c^2$ . To separate electrons from charged hadrons, the following variables are used: (1) the ratio of the hadronic to the electromagnetic energy of the cluster, HAD/EM; (2) the ratio of the cluster energy to the associated CTC track momentum, E/P; (3) the lateral shower profile in the calorimeter,  $L_{shr}$ , compared to the one measured with test beam electrons; (4) the distance of the track extrapolation from the shower position measured by the CES in the azimuthal view,  $\Delta x$ , and in the longitudinal view,  $\Delta z$ ; (5) the  $\chi^2$  relative to the shower profile in the strip chambers,  $\chi^2_{strip}$ , compared to the one of test beam electrons; (6) the  $z$  distance of the track from the PV, “ $z$ -vertex match”; (7) the transverse energy of the towers contained in a cone of radius  $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.4$  (excluding the contribution of the electron cluster), i.e., the isolation,  $I$ .

For the  $e + \text{jets}$  search, the following electron identification cuts are imposed

1.  $E_t(e) > 20$  GeV,
2. HAD/EM  $< 0.05$ ,
3. E/P  $< 1.5$ ,
4.  $L_{shr} < 0.2$ ,
5.  $|\Delta x| < 1.5$  cm,  $|\Delta z| < 3.0$  cm,
6.  $\chi_{strip}^2 < 10$ ,
7.  $z$ -vertex match  $< 5$  cm,
8.  $I/E_t(e) < 0.1$ .

The identification efficiency is  $(84 \pm 2)\%$ , excluding the loss due to the rejection of photon conversions (5%) and to the isolation cut (13 % for  $t\bar{t}$  events). The efficiency is determined using the  $Z \rightarrow e^+e^-$  sample, obtained with an independent selection based only on calorimeter information; the effect of the photon conversion removal is studied using the inclusive electron sample enriched in  $b \rightarrow e\bar{\nu}X$  decays [21]; the effect of electron non-isolation is modeled with the Montecarlo. These results are used in section 8.4 for the computation of the efficiency for identifying high  $E_t$  electrons in  $t\bar{t} \rightarrow e^\pm + jets$ .

### 7.1.2 Identification of High $P_t$ Muons

For the top-quark search in the  $\mu + jets$  channel we select muons in the region  $|\eta| < 1.0$  satisfying the trigger cuts described in section 3.2.5, which basically require the association of a CTC track (of transverse momentum  $P_t(\mu)$ ) to a track segment in the CMU or CMU-CMP or CMX muon chambers.

To separate muons from hadrons penetrating through the calorimeters, the following variables are used: (1) the energy deposit in the hadronic and electromagnetic calorimeters, which must be consistent with a MIP; (2) the CTC track IP w.r.t. the average beam line,  $D_{beam}$ ; (3) the  $z$  distance of the track from the PV,  $z$ -vertex match; (4) the distance between the track extrapolation and the muon track segment in the  $\phi$  view,  $\Delta x$ ; (5) the difference of the transverse energy of towers in a cone of radius  $R = 0.4$  and the energy of those in a cone of  $R = 0.13$  around the muon track, i.e., the isolation,  $I$ .

For the  $\mu + jets$  search the following muon identification cuts are imposed:

1. a muon track segment in the CMU or CMU-CMP or CMX muon chambers,
2.  $P_t(\mu) > 20$  GeV,
3. EM energy (in the muon tower)  $< 2$  GeV,
4. HAD energy (in the muon tower)  $< 6$  GeV,
5.  $|D_{beam}| < 3$  mm,
6.  $|\Delta x(\text{CMU})| < 2$  cm,  $|\Delta x(\text{CMU-CMP or CMX})| < 5$  cm,

7.  $z$ -vertex match  $< 5$  cm,

8.  $I/P_t(\mu) < 0.1$ .

The identification efficiency is  $(90.6 \pm 1.4)\%$ , excluding the loss due the isolation cut (19 % for  $t\bar{t}$  events). The efficiency is determined using the  $Z \rightarrow \mu^+\mu^-$  sample, obtained with an independent selection; the effect of the muon non-isolation is modeled with the Montecarlo. These results are used in section 8.4 for the computation of the efficiency for identifying high  $P_t$  muons in  $t\bar{t} \rightarrow \mu^\pm + jets$ .

### 7.1.3 Measurement of the Missing Transverse Energy ( $E_t$ )

To signal the presence of high  $P_t$  neutrinos we look for significant unbalance in the transverse energy flow of the event. For this purpose, we use the missing transverse energy ( $E_t$ ), defined as the vector sum of the transverse energies of all calorimeter towers in  $|\eta| < 3.6$ . The  $\eta$  coverage is limited due to the Tevatron focusing magnets, covering part of the forward hadron calorimeter. To be included in the sum, single towers must have an energy exceeding thresholds varying with the calorimeter type: 100 MeV in the central calorimeters (CEM and CHA), 300 in the PEM, 800 MeV in the FHA and 500 MeV in the PHA and FEM. In events with muon candidates the  $E_t$  is corrected by vectorially subtracting the energy deposited in the muon tower and by adding  $P_t(\mu)$ . Note that in this analysis the  $E_t$  value used is not corrected for jet energy measurement errors. This  $E_t$  correction would be essential to measure, for example, the  $W$ -boson mass; on the contrary, it is not strictly necessary for this analysis which is basically a “counting” of the number of top event candidates and of the number of expected background events (see section 8.1). The  $E_t$  correction for jet mismeasurements would allow for the  $E_t$  resolution to be optimized, and consequently, also for the mass resolution of  $(e, E_t)$ ,  $(\mu, E_t)$  pairs to be optimized.

For the  $W +$  multijet search we require  $E_t > 20$  GeV, in order to reject the  $b\bar{b}$  background, where a high-energy  $b$  decays semileptonically. This cut provides, however, a high efficiency for  $t\bar{t}$  processes: figure 7.1 shows the  $E_t$  distribution in  $t\bar{t}$  events of  $M_{top} = 160$  GeV/ $c^2$  selected to contain an electron of  $E_t(e) > 20$  GeV; the efficiency of the  $E_t > 20$  GeV cut is  $\sim 90\%$ , almost insensitive to  $M_{top}$ .

The evidence for  $W \rightarrow e\bar{\nu}_e, \mu\bar{\nu}_\mu$  decays is confirmed by the transverse invariant mass distribution of  $(e, E_t)$  and  $(\mu, E_t)$  pairs shown in figure 7.2. Note that although the  $E_t$  is not corrected for jet energy mismeasurements, the Jacobian peak in the invariant mass around  $M_W$  is clearly visible.

The  $E_t$  resolution depends on the total transverse energy of the event. For minimum bias  $p\bar{p}$  collisions, the resolution of the two components of the vector  $\vec{E}_t$ ,  $E_{tx}$  and  $E_{ty}$ , can be parametrized as [129]  $\sigma(E_{t,x,y}) = 0.47\sqrt{\Sigma E_t}$ , where  $\Sigma E_t$  is the total transverse energy of the event. For  $W +$  multijet events, which have characteristics similar to top events, the  $E_t$  resolution is well reproduced by the full detector simulation CDFSIM. This agreement is indicated in figure 7.3, where, for the final  $W + \geq 3$  jet sample (see next section), the transverse mass distribution of  $(e, E_t)$  and  $(\mu, E_t)$  pairs in the data is compared to the prediction obtained with the VECBOS Montecarlo program (generator of  $W + 3$  and 4 jets) + CDFSIM, and to the prediction obtained with the ISAJET Montecarlo program

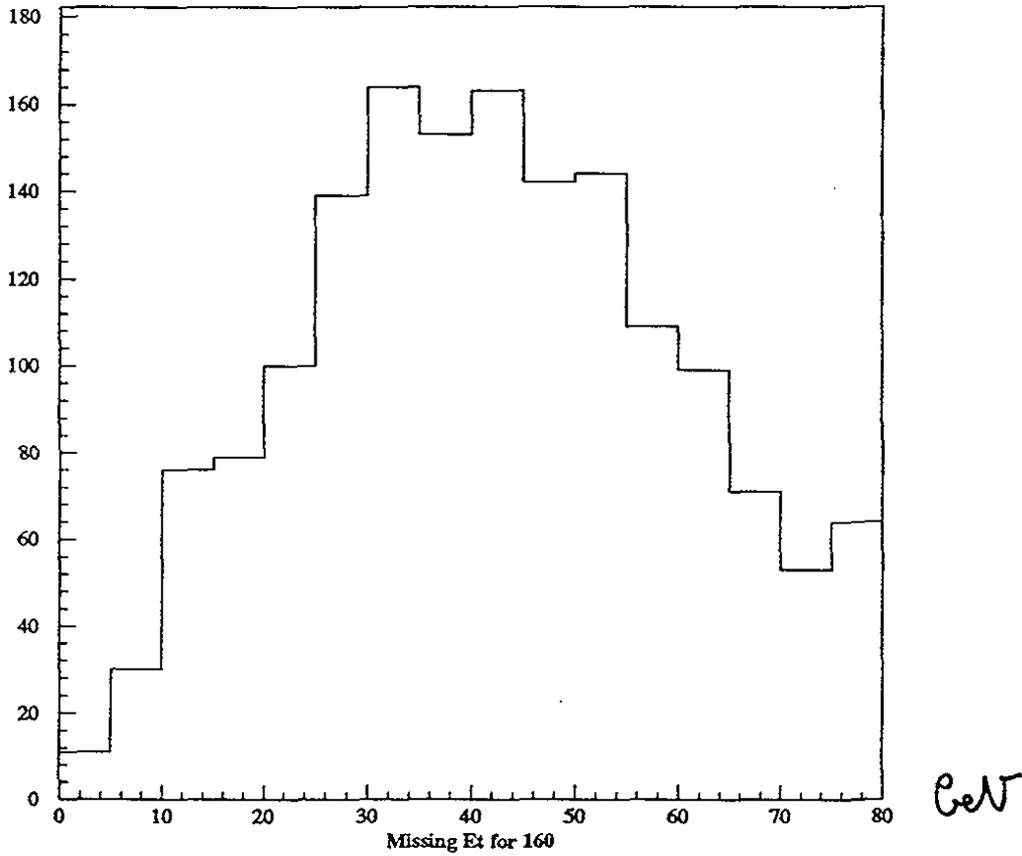


Figure 7.1:  $E_t$  distribution of  $t\bar{t}$  events of  $M_{t\bar{t}} = 160 \text{ GeV}/c^2$  (ISAJET Montecarlo) selected to have an electron of  $E_t(e) > 20 \text{ GeV}$ ; the efficiency of the  $E_t > 20 \text{ GeV}$  cut is  $\sim 90\%$ , almost insensitive to  $M_{t\bar{t}}$  in the range  $120\text{-}180 \text{ GeV}/c^2$ .

(generator of  $t\bar{t}$  events) + CDFSIM. The shape of the invariant mass distribution is sensitive to  $\sigma(\mathcal{E}_{t,x,y})$ ; data and Montecarlo shapes are in good agreement.

## 7.2 Final Event Selection

Besides requiring an electron of  $E_t(e) > 20$  GeV, or a muon of  $P_t(\mu) > 20$  GeV/ $c$ , and  $E_t > 20$  GeV, the final event sample is selected by imposing the following cuts:

- the PV must be within 60 cm from the center of CDF:  $|Z_{PV}| < 60$  cm;
- the electron or muon must be consistent with the primary interaction of maximum  $Q$  found by the VXPRIM algorithm:  $|Z_{PV} - z_0(\text{lepton})| < 5$  cm;
- events with a second lepton such that the invariant mass of the two leptons is in the range 70 and 110 GeV/ $c^2$  are rejected as potential  $Z \rightarrow ee, \mu\mu$  candidates. For  $Z \rightarrow ee$ , the second lepton is any electromagnetic cluster satisfying cuts less tight than those described in the paragraph 7.1.1:
  1.  $E_t > 10$  GeV,
  2.  $E/P < 2.0$ ,
  3.  $\text{HAD}/\text{EM} < 0.12$ ,
  4.  $I/E_t < 0.1$ .

For  $Z \rightarrow \mu\mu$ , also the second muon candidate is defined with cuts looser than those described in paragraph 7.1.2:

1.  $P_t > 15$  GeV,
2.  $|\eta| < 1.1$ ,
3.  $|\Delta x| < 5$  cm (CMU, CMP, CMX); if the track extrapolation hits a region not instrumented with muon chambers (as long as  $|\eta| < 1.1$ ) this cut is completely removed;
4. EM energy  $< 2$  GeV, HAD energy  $< 6$  GeV.

The sample thus selected contains 11949  $W \rightarrow e\bar{\nu}_e$  candidates and 7023  $W \rightarrow \mu\bar{\nu}_\mu$  candidates. Table 7.1 shows the number of events in the sample as a function of  $N_{jets}$ , the number of jets with  $|\eta| < 2.0$  and *observed* transverse energy  $E_t > 15$  GeV. We use the jet energy reconstructed (observed) by the calorimeter, without corrections for calorimeter mismeasurements, energy losses outside the jet definition cone, and underlying-event effects. The same choice is adopted in the corresponding Montecarlo (generator + simulator) samples. For the top-quark search in the lepton + jets channel described in the next chapter, we consider only jets with  $|\eta| < 2.0$  and  $E_t > 15$  GeV. The transverse mass distribution of  $(e, \mathcal{E}_t)$  and  $(\mu, \mathcal{E}_t)$  pairs in the final  $W + \geq 3$  jets sample is compared to VECBOS + CDFSIM and to ISAJET + CDFSIM in figure 7.3.

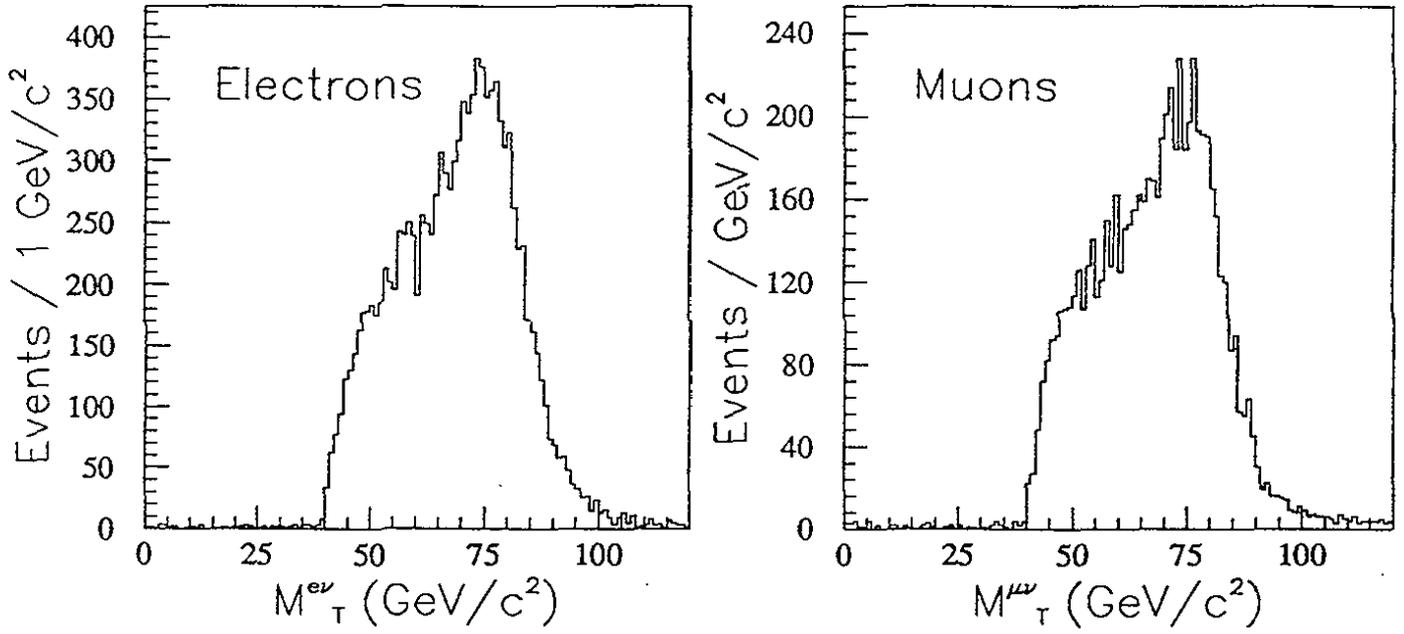


Figure 7.2: Evidence for  $W \rightarrow e\bar{\nu}_e, \mu\bar{\nu}_\mu$  decays confirmed by the transverse mass distribution of  $(e, \cancel{E}_t)$  and  $(\mu, \cancel{E}_t)$  pairs.

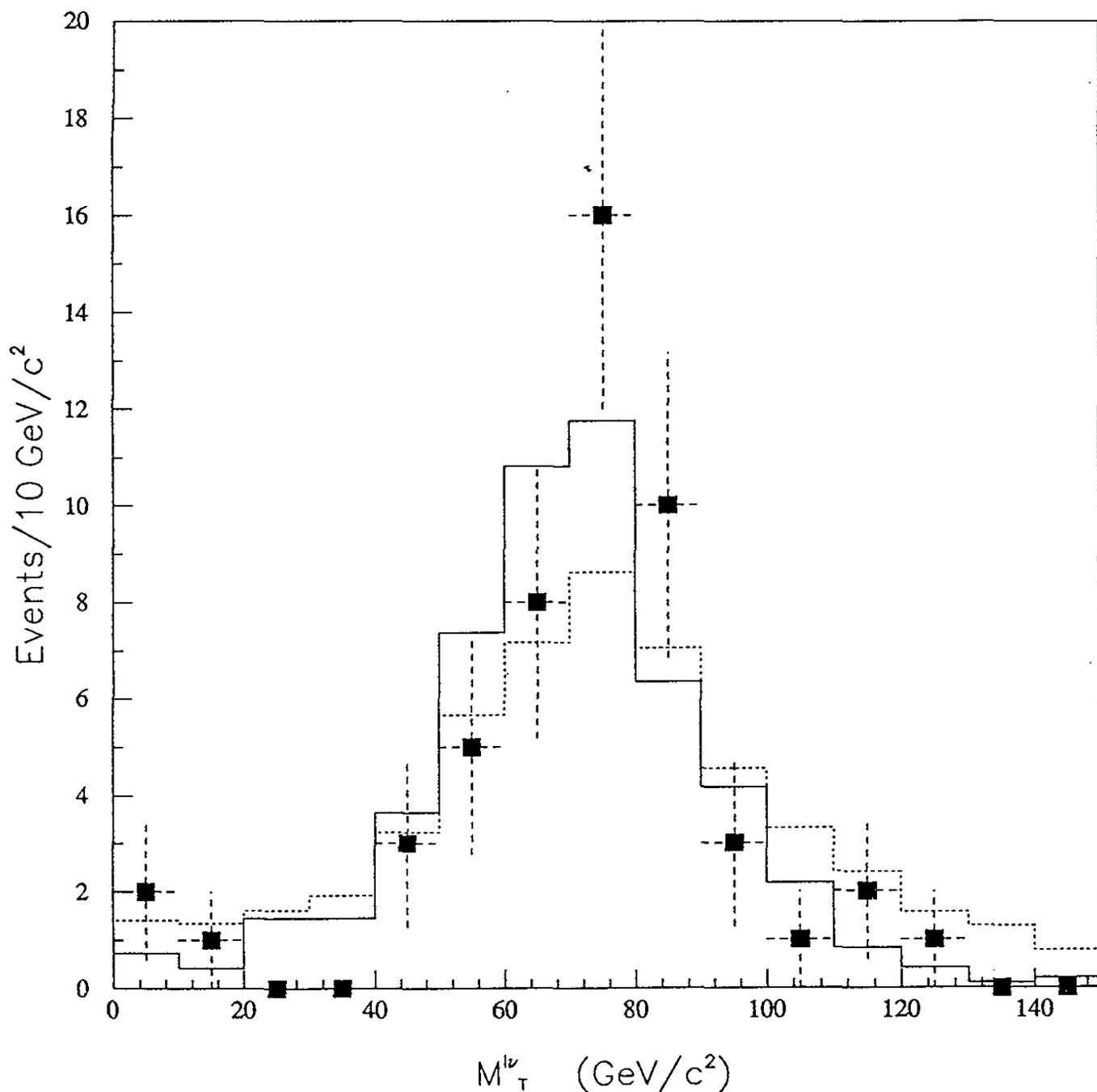


Figure 7.3: Transverse mass distribution of  $(e, \cancel{E}_t)$  and  $(\mu, \cancel{E}_t)$  pairs in the final  $W + \geq 3$  jets sample, compared to the predictions obtained using: (1) the VECBOS Montecarlo program (generator of  $p\bar{p} \rightarrow W + \text{multijet}$  events) + CDF simulation; (2) the ISAJET Montecarlo program (generator of  $p\bar{p} \rightarrow t\bar{t}$  events) + CDF simulation.

| Number of Jets | Electrons | Muons | Electrons + Muons |
|----------------|-----------|-------|-------------------|
| 0 jet          | 10663     | 6264  | 16927             |
| 1 jet          | 1058      | 655   | 1713              |
| 2 jets         | 191       | 89    | 280               |
| 3 jets         | 30        | 13    | 43                |
| 4 jets         | 7         | 2     | 9                 |

Table 7.1: Summary of the number of inclusive  $W \rightarrow e\bar{\nu}_e, \mu\bar{\nu}_\mu$  events as a function of the number of jets in the event,  $N_{jets}$ .

# Chapter 8

## Top-Quark Search with the $D - \phi b$ Tagger

This chapter describes the results of the search for  $t\bar{t}$  events in the  $W +$  multijet decay channel, using the  $D - \phi$  algorithm as the main tool for  $b$  tagging. Section 8.1 shows that by selecting multijet events the signal of  $t\bar{t}$  production is increased w.r.t. background. Section 8.2 describes the sample of  $W + \geq 1$  jet events containing a  $b$ -tag and defines the candidates which represent a potential  $t\bar{t}$  signal. Section 8.3 reports the study of the background sources, that is, processes different from  $t\bar{t}$  production which can give  $W +$  multijet events with  $b$  tags. Section 8.4 contains the study of the significance of the excess of observed candidate events w.r.t. the estimated background; the number of expected  $b$ -tagged  $t\bar{t}$  events in this channel as a function of  $M_{top}$  is given in section 8.6; the consistency of the excess of candidate events with the  $t\bar{t}$  production hypothesis is discussed in section 8.7. Finally, the last section of the chapter summarizes the results of this work and gives the conclusions on the status of the top-quark search as of February 28, 1994, the date on which this thesis was submitted to the Italian Ministry of Scientific Research.

### 8.1 The Counting Experiment

The cross section of background  $W +$  multijet events drops rapidly with the number of jets [45]; this is clearly reflected in the variation of the number of  $W$  events as a function of  $N_{jets}$ , reported in table 7.1. On the contrary,  $t\bar{t}$  decaying in the lepton + jets channel produces on the average three or four jets of  $|\eta| < 2.0$  and  $E_t > 15$  GeV together with the  $W$  boson. Figure 8.1 shows the distribution of  $N_{jets}$  for Montecarlo  $t\bar{t}$  events of 140 and 160 GeV/ $c^2$  mass. A straightforward selection of the  $t\bar{t}$  search sample, which greatly improves the ratio of the top signal to the  $W +$  multijet background, is to require  $N_{jets} \geq 3$ . For  $M_{top} = 160$  GeV/ $c^2$ , about 75% of  $t\bar{t}$  events satisfies this cut. Figure 8.2 shows the cross sections of  $W \rightarrow e/\mu +$  multijet processes observed in the data, compared to the cross section predicted for  $t\bar{t}$  (see sections 8.6, 8.7) and for the background (VECBOS Montecarlo). The cross sections of figure 8.2 are for  $e/\mu +$  jets *before*  $b$  tagging. Note that for  $N_{jets} = 3$ ,  $S/N < 1$  for  $M_{top} > 100$  GeV/ $c^2$ , while for  $N_{jets} \geq 4$ ,  $1 < S/N < 2$  only for  $M_{top} < 150$  GeV/ $c^2$ . This clearly shows the need of  $b$ -tagging to reduce the background level.

The analysis described in this chapter is essentially an experiment of counting of the number of  $W + \geq 3$  jet events with a  $D - \phi$   $b$  tag.

## 8.2 Observed $D - \phi$ Tags in the $W +$ Multijet Sample

The  $D - \phi$  algorithm is applied to the  $W +$  multijet sample described in section 7.2 to tag  $b$  jets. The number of events before and after  $b$  tagging is shown in table 8.1 (positive  $b$  tags only).

| $N_{jets}$ | Events before $b$ tagging | Events after $b$ tagging |
|------------|---------------------------|--------------------------|
| 1          | 1713                      | 3                        |
| 2          | 280                       | 7                        |
| 3          | 43                        | 4                        |
| 4          | 9                         | 1                        |

Table 8.1: Summary of the number of  $W +$  multijet events before and after  $b$  tagging as a function of the number of jets in the event,  $N_{jets}$ .

We observe 5  $b$ -tagged events with  $N_{jets} \geq 3$ : three  $\mu + 3$  jets, one  $e + 3$  jets and one  $e + 4$  jets. No event has two distinct  $b$ -tagged jets. Figure 8.3 shows the  $D - \phi$  representation of the candidate event with  $e + 3$  jets: each SVX track is indicated by a point, whose vertical error bar is the uncertainty on the IP,  $\sigma_D$ . Tracks indicated by open circles are track which do not pass the quality cuts. The  $D - \phi$   $b$  tag in this event is associated to the jet with  $E_t = 101$  GeV,  $\phi = 0.91$ ,  $\eta = 0.07$ , and is formed by the four tracks with high IP values located around  $\phi = 1$ . Figure 8.4 shows the transverse plane view of the event, with tracks crossing the four SVX layers; outside the outer layer, the energy deposited in the central calorimeter is indicated; the event  $D - \phi$  plot is shown in the bottom left corner of the figure. Figure 8.5 shows, for the same event, a magnification of the  $R - \phi$  region of radius = 2 cm in the transverse plane around the VXPRIM PV. The coordinates of the SV associated with the four-track  $b$  tag are shown in the top right corner. The tracks drawn in these event displays are not forced to pass through the respective PV or SV, that is, they are represented using the parameters resulting from the SVX tracking and not from the vertexing fit.

The  $cr$  proper lifetime distribution of the  $b$ -tagged  $W +$  multijet events is shown in figure 8.6: the points represent the  $W + \geq 1$  jet data, while the dashed histogram represents the 5  $t\bar{t}$  candidate events with  $W + \geq 3$  jets. The  $b$  tags in  $W + 1$  and 2 jets should be dominated by sources other than  $t\bar{t}$  decays. The number of events expected for each one of these background sources is calculated in the following sections.

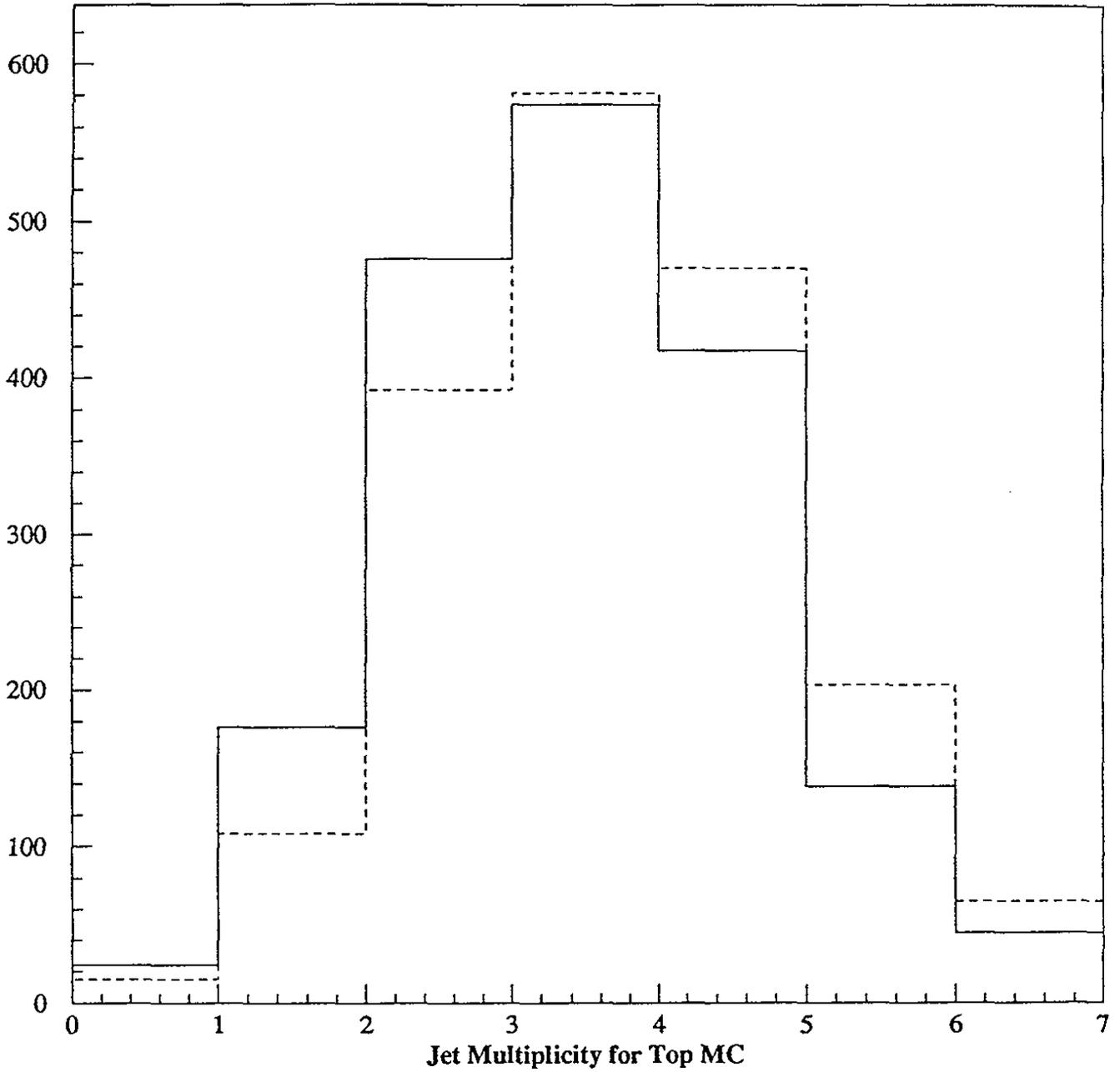


Figure 8.1: Distribution of the number of jets in  $t\bar{t}$  ISAJET Montecarlo events of mass  $M_{top} = 140 \text{ GeV}/c^2$  (solid histogram) and  $M_{top} = 160 \text{ GeV}/c^2$  (dashed histogram).

$$W \rightarrow \mu\nu, e\nu \quad R=0.4, E_T=15 \text{ GeV}/c$$

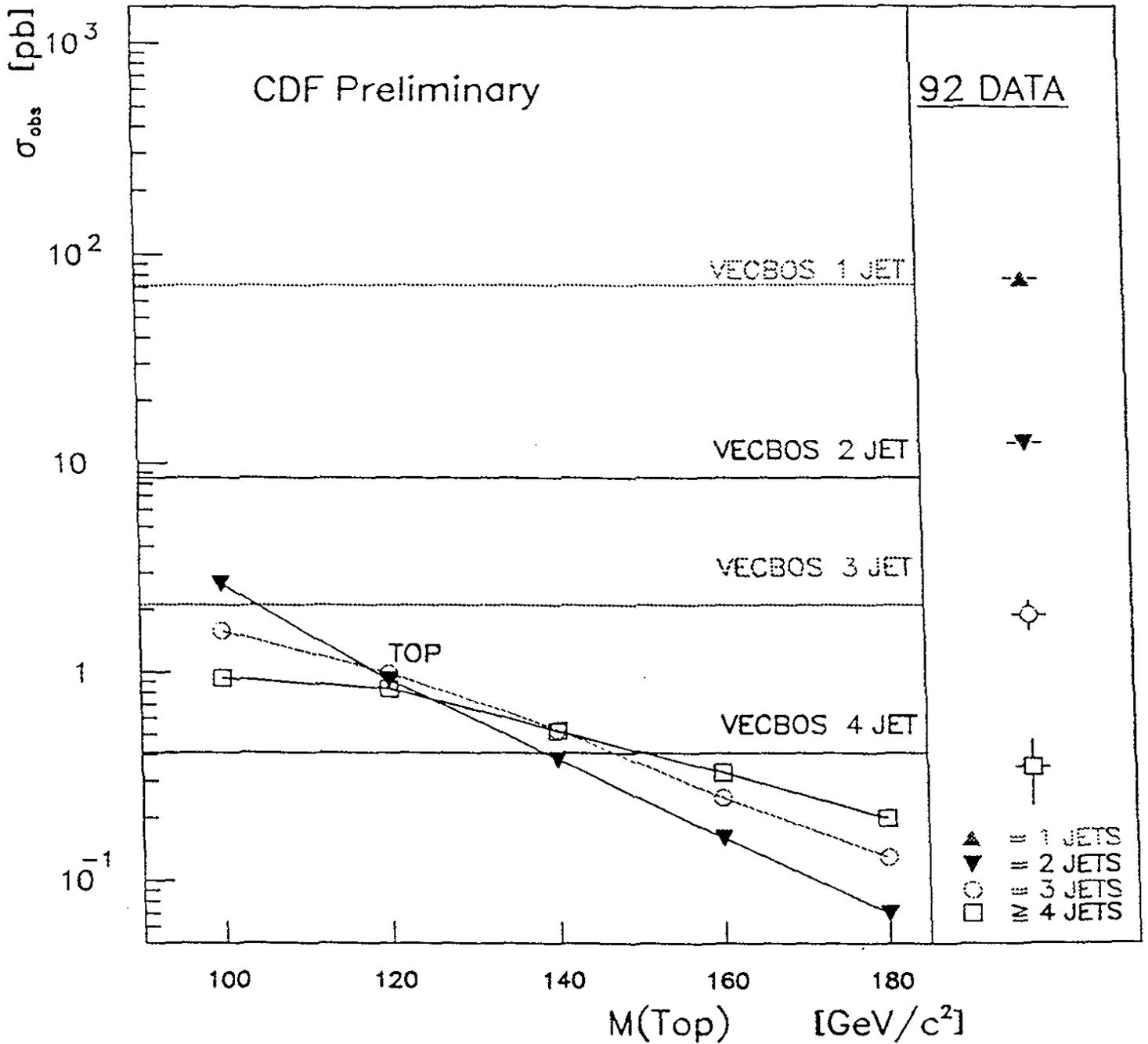


Figure 8.2: Comparison of the  $W \rightarrow e/\mu + \text{multijet}$  cross sections observed in the data, compared to the VECBOS prediction ( $W + \text{multijet background}$ ) and to the ISAJET prediction ( $t\bar{t} \rightarrow W + \text{multijet signal}$ ), as a function of the number of jets and of the top-quark mass.

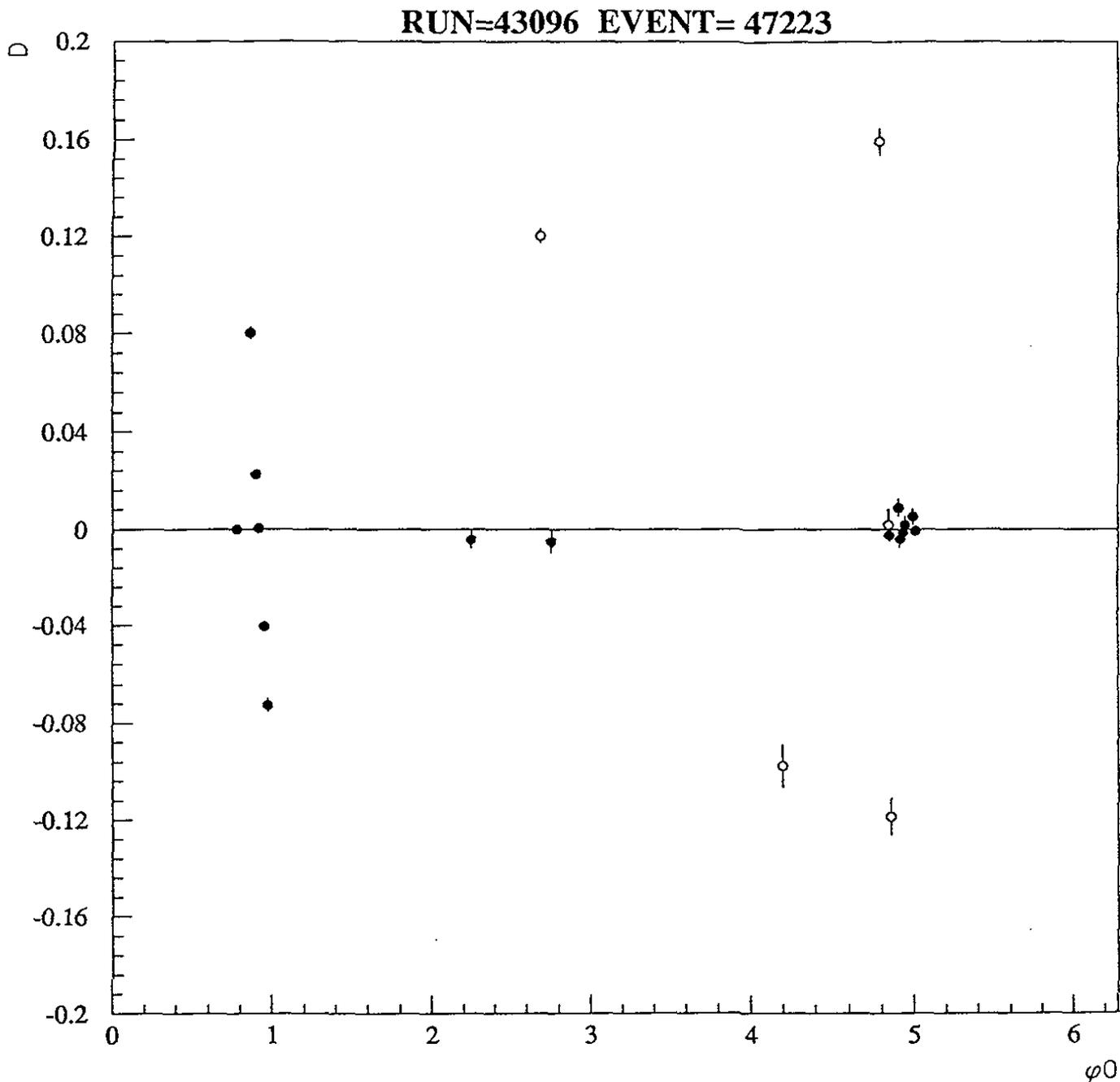


Figure 8.3: Display of the *electron + multijet* candidate event in the impact parameter-azimuth ( $D - \phi$ ) plane, in which each track is a point. The open circles indicate tracks which do not pass the quality cuts.  $D$  is measured in cm,  $\phi$  is measured in radians. The  $D - \phi$  algorithm tags the 4 tracks with large value of the impact parameter around  $\phi = 1$  as a potential  $b$ -hadron decay.

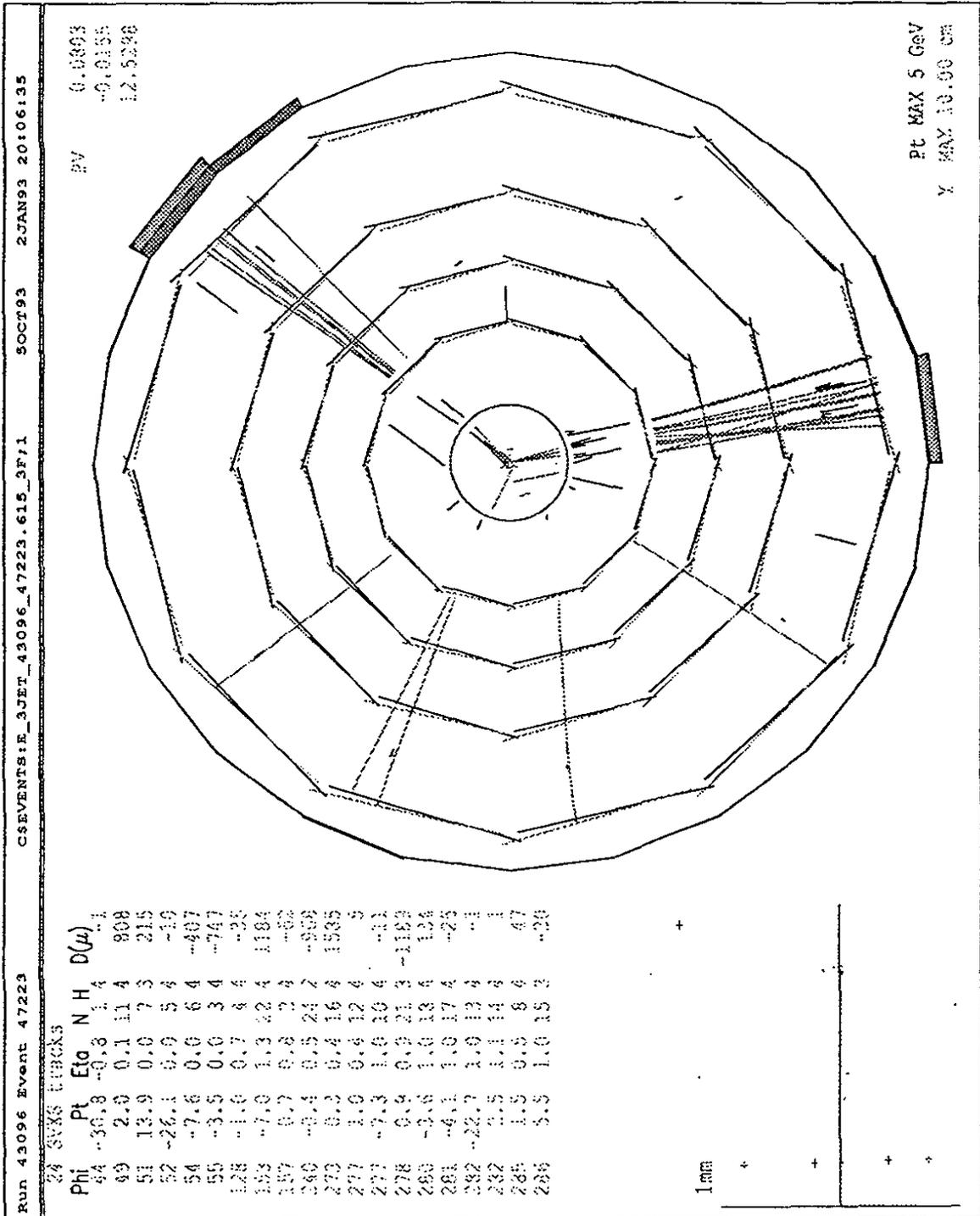


Figure 8.4: *Electron + multijet* candidate event:  $R - \phi$  display, with the tracks crossing the four SVX layers; the energy deposited in the central calorimeter is shown outside the SVX; the primary vertex coordinates w.r.t. the CDF reference system are shown on the top right corner. The azimuth is counted counterclockwise,  $\phi = 0$  being the horizontal semi-axis  $x > 0$  at the right of the event primary vertex.

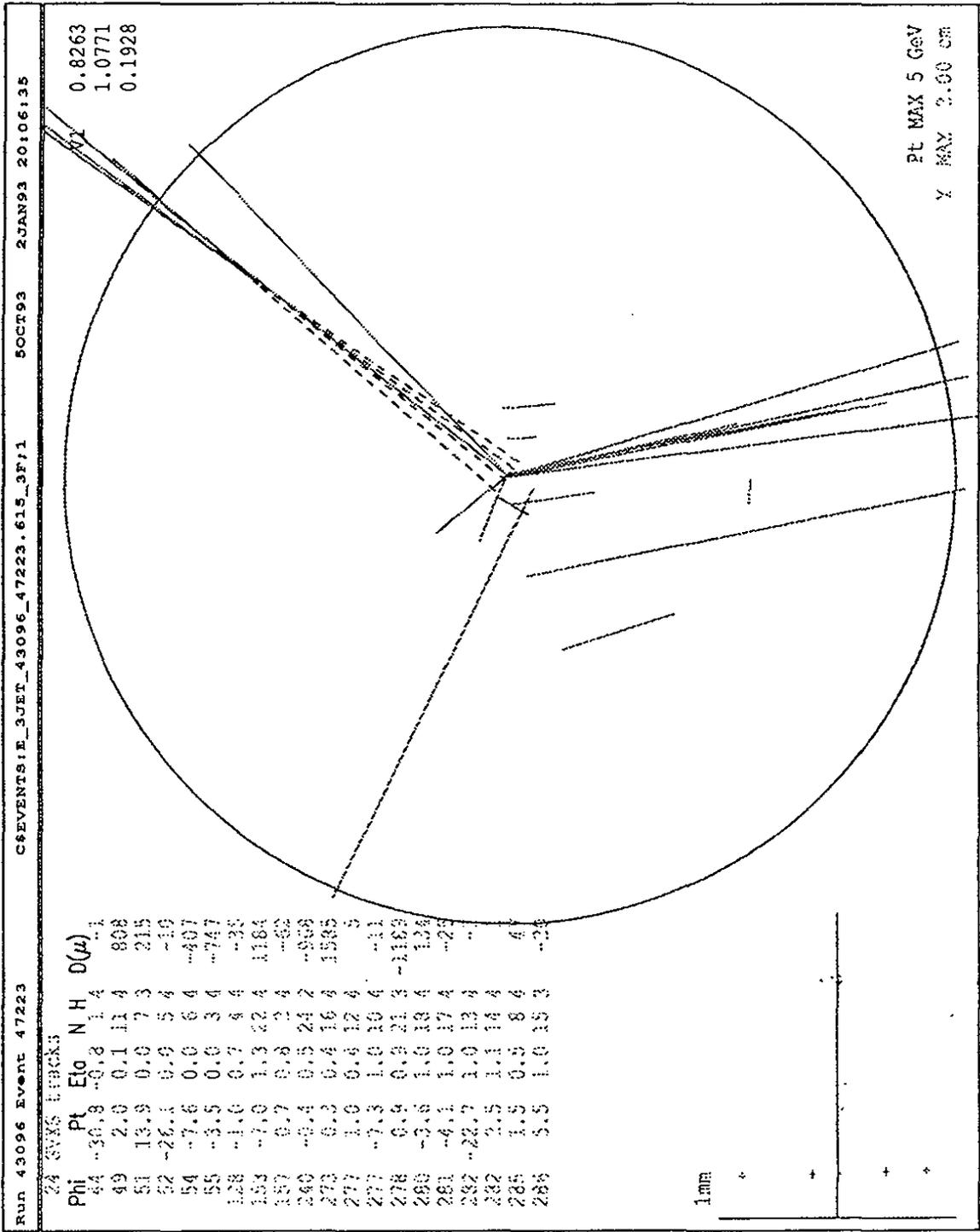


Figure 8.5: *Electron + multijet* candidate event: magnification of the  $R-\phi$  region of radius = 2 cm around the primary vertex. The coordinates of the secondary vertex associated to the  $b$  tag are shown in the top right corner.

## Electron/Muon + Jets + SVX b-Tag

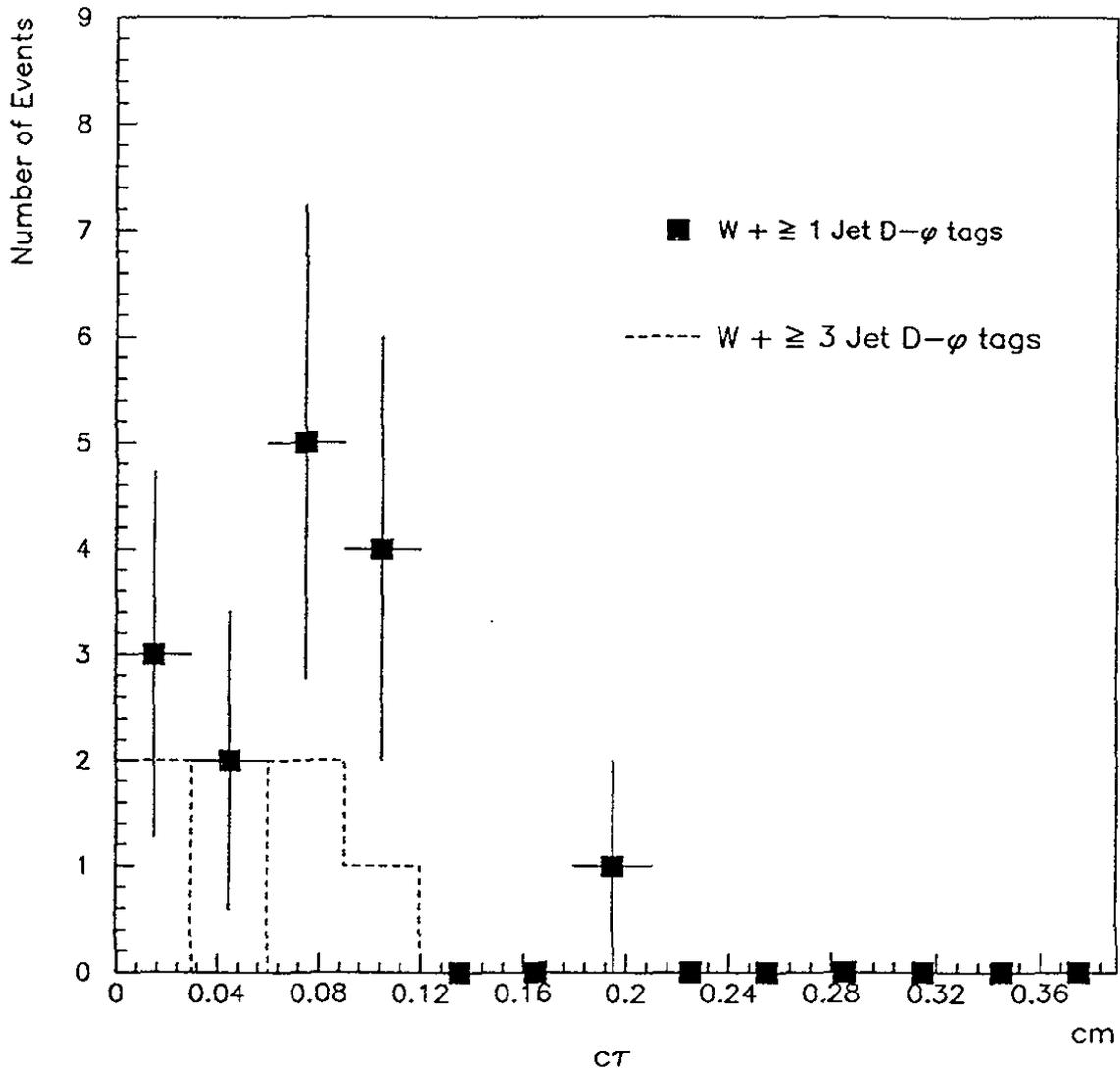


Figure 8.6: Proper lifetime distribution,  $c\tau$ , of positive  $b$  tags in the  $W + \geq 1$  jets sample; the dashed histogram represents the  $b$  tags in  $W + \geq 3$  jets. The data correspond to a statistics of  $21.4 \text{ pb}^{-1}$ .

## 8.3 Background Sources

In the  $W + \text{multijet}$  sample  $b$  tags can originate from sources other than  $t\bar{t}$  decays:

- direct production of  $W$  bosons in association with  $b\bar{b}$  and  $c\bar{c}$  pairs [49];
- fake SV's due to tracking errors in  $W$  events;
- $p\bar{p} \rightarrow W + c$ ; the  $c$ -tagging efficiency is much lower than for  $b$ -tagging, but not negligible;
- production of high-energy  $b\bar{b}$  pairs, in which one  $b$  decays semileptonically, simulating  $W \rightarrow l\nu$  ( $l = e$  or  $\mu$ ), and the other  $b$  gives a  $b$ -tag;
- production of two electroweak bosons,  $WW$  or  $WZ$ , in which  $W \rightarrow c\bar{s}$  or  $Z \rightarrow b\bar{b}$ ;
- $Z \rightarrow \tau\tau$ , in which one  $\tau$  decays semileptonically, simulating a  $W$ , and the other  $\tau$  provides a  $b$  tag; however, the  $\tau$ -tagging efficiency is almost negligible, since  $c\tau_\tau \sim 100\mu\text{m}$ .

The first two are the dominant background sources. In the next two sections these contributions are estimated using two different methods; finally, the other residual backgrounds will be also studied.

### 8.3.1 Estimate of Background from $Wb\bar{b}$ , $Wc\bar{c}$ and Fake Secondary Vertices: Method 1

The first method is based on the assumption that the  $b$ -quark and  $c$ -quark content per jet in the  $W + \text{multijet}$  sample is the same, or less, than in the inclusive jet sample. The matrix  $M^+$  defined in section 6.2.1 is used to assign to any given jet of  $W$  events the  $b$ -tagging probability. This approach has the advantage that the predicted number of  $b$  tags is based on a direct measurement of fake SV's due to tracking errors, of residual contaminations of  $K_s$  and  $\Lambda$  decays and of all contributions to HF production in inclusive, generic jets. In particular,  $M^+$  contains possible sources of fake (non- $b$ ) SV's not present in the negative  $b$  tags in inclusive jets which are used to calculate  $M^-$ .

The presence of HF's in inclusive jets is primarily due to: (1) direct production (like  $gg \rightarrow b\bar{b}$ ), which represents about 70 % of the total HF contribution; (2) flavor excitation; (3) gluon splitting [131]. In  $W + \text{multijet}$  events HF's come only from gluon splitting [49] (the  $W + c$  production is accounted for separately; see below). In addition, the ISAJET and HERWIG [130] Montecarlo programs predict a fraction of gluon jets in  $W + \text{jets}$  lower than in inclusive jets. The HERWIG Montecarlo predicts a  $b\bar{b}$  content per jet in  $W + \text{jets}$  three times lower than in inclusive jets. Therefore, one expects that the number of  $b$  tags from  $W + b\bar{b}$ ,  $W + c\bar{c}$  and fake SV's observed in the  $W$  sample is actually lower than that predicted by using the matrix  $M^+$ . Such a prediction is reported in row (1) of table 8.2.

### 8.3.2 Estimate of Background from $Wb\bar{b}$ , $Wc\bar{c}$ and Fake Secondary Vertices: Method 2

In the second method one calculates the number of expected events with a  $b$  tag from  $Wc\bar{c}$  and  $Wb\bar{b}$  using theoretical and Montecarlo predictions and adds this results to the number of  $b$  tags from fake SV's estimated using the matrix  $M^-$ .

The fraction of  $Wb\bar{b}$  and  $Wc\bar{c}$  events w.r.t. all  $W$  events as a function of the number of jets is calculated with two techniques: (1) using the leading order matrix elements described in [49] and (2) using the HERWIG Montecarlo program. The two estimates provide the same  $b$  and  $c$  fractions per event within the renormalization scale uncertainty (estimated in [49] to be  $\pm 40\%$ ). Multiplying these fractions by the number of observed  $W + n$  jet events ( $n=1,2,3,4$ ), after correcting for the non- $W$  background, and by the expected  $b$ -tagging efficiency (estimated on the corresponding HERWIG-generated samples) one gets the number of expected background events reported in row (2) of table 8.2.

In order to verify that the HERWIG  $b$  and  $c$  fractions in  $W +$  multijet events are reasonable, the HF content predicted by HERWIG for the inclusive jet sample has been compared to what is observed in the Jet.50 sample. The agreement between the number of  $b$  tags predicted by HERWIG and the HF content observed in the data (number of positive - negative  $b$  tags) as a function of the number of jets is good (see figure 8.7). The prediction of the HF fraction in the Jet.50 sample based on the next-to-leading order matrix element calculation [132] provides similar results. This indicates that the prediction of the number of  $b$  tags due to  $Wb\bar{b}$  and  $Wc\bar{c}$  obtained with the second method is accurate. Using the presented information and the error due to the choice of the renormalization scale in the calculation of  $Wb\bar{b}$ ,  $Wc\bar{c}$  processes ( $\pm 40\%$ ), we assign an 80% systematic uncertainty to the HERWIG prediction of the background from  $Wb\bar{b}$  and  $Wc\bar{c}$ .

The number of background events expected from fake SV's (obtained from  $M^-$ ) as a function of the number of jets is shown in row (3) of table 8.2; the sum of these with the number of background events from  $Wb\bar{b}$ ,  $Wc\bar{c}$  is reported in row (4).

### 8.3.3 Other Background Sources

The contributions of the remaining background sources listed at the beginning of section 8.3 are determined separately and added to the previous estimates of the dominant backgrounds reported in rows (1) and (4).

The  $p\bar{p} \rightarrow W + c$  background [133], due to the flavor excitation processes  $sg \rightarrow Wc$  and  $dg \rightarrow Wc$ , is estimated using the HERWIG and VECBOS [134] Montecarlo programs. The results of this calculation are reported in row (5) of table 8.2, where the estimated uncertainty ( $\sim 30\%$ ) reflects mainly the uncertainty on the  $s$ -quark content in the proton, which has been determined by examining a large number of different structure functions [135].

The background from direct production of high-energy  $b\bar{b}$  pairs is strongly suppressed by the value of the  $b\bar{b}$  production cross section, which drops rapidly with  $P_t(b)$ , and by the cut  $E_t > 20$  GeV. The residual contamination of this, as well as other non- $W$  backgrounds in the  $W$  sample is determined directly from the data, by using a standard technique described in [136] to estimate the non- $W$  background before tagging. Form the size of this contamination

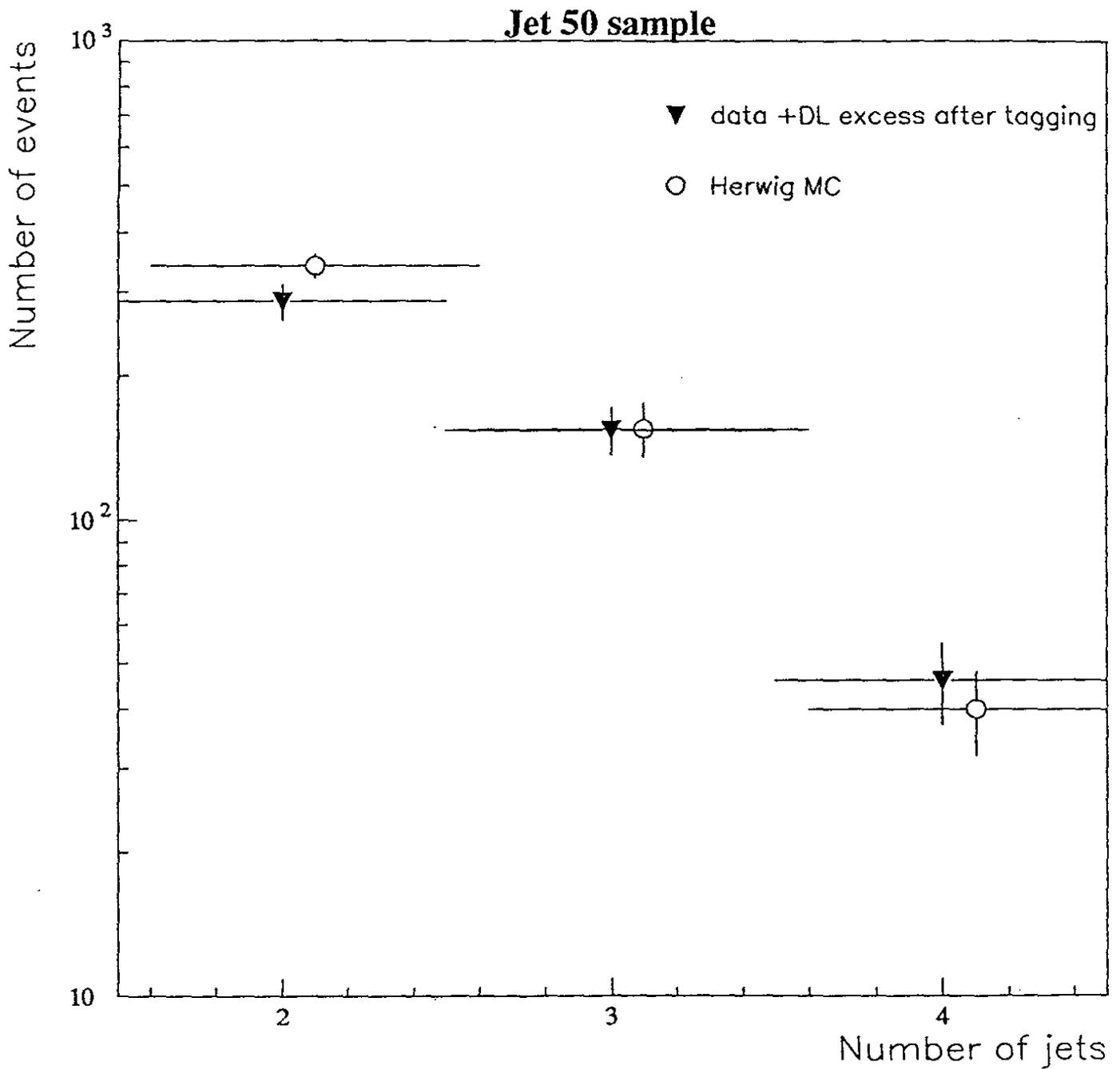


Figure 8.7: Number of events with positive - negative  $b$  tags (+DL excess after tagging) as a function of the number of jets for the inclusive, generic jet sample ("Jet\_50" sample) and for a HERWIG Montecarlo sample of jets containing heavy flavors.

and from the expected  $b$ -tagging efficiency, one finds the background estimate reported in row (7) of table 8.2.

The contribution of  $WW$ ,  $WZ$  and  $Z \rightarrow \tau\tau$  have been estimated with the ISAJET Montecarlo program and turned out to be very very small (see row (6) of table 8.2).

### 8.3.4 Summary of Background Estimate

The total number of background events as a function of  $N_{jets}$ , estimated with methods 1 and 2 (and by adding the residual contributions) is reported in rows (8) and (9) of table 8.2. As expected, method 1 gives a background larger than method 2, since the Montecarlo provides a HF content in  $W + \text{multijet}$  events lower than in inclusive jet events. To be conservative, and to rely on Montecarlo/theoretical predictions as little as possible, we use the value of the background obtained from method 1.

| Source  | $W + 1 \text{ jet}$ | $W + 2 \text{ jets}$ | $W + \geq 3 \text{ jets}$ |
|---|---------------------|----------------------|---------------------------|
| (1) $Wbb, Wc\bar{c} + \text{Fake SV's, Method 1}$ | $12.4 \pm 4.1$      | $3.79 \pm 1.25$      | $1.52 \pm 0.51$           |
| (2) $Wbb, Wc\bar{c}$ Only                         | $2.7 \pm 2.2$       | $0.90 \pm 0.72$      | $0.34 \pm 0.27$           |
| (3) Fake SV's only                                | $4.5 \pm 0.6$       | $1.42 \pm 0.20$      | $0.54 \pm 0.08$           |
| (4) $Wbb, Wc\bar{c} + \text{Fake SV's, Method 2}$ | $7.2 \pm 2.7$       | $2.32 \pm 0.88$      | $0.88 \pm 0.33$           |
| (5) $Wc$  | $1.3 \pm 0.6$       | $0.47 \pm 0.25$      | $0.10 \pm 0.04$           |
| (6) $Z \rightarrow \tau\tau, WW, WZ$              | $0.08 \pm 0.05$     | $0.08 \pm 0.02$      | $0.03 \pm 0.01$           |
| (7) Non- $W$ , included $bb$                      | $0.5 \pm 0.3$       | $0.59 \pm 0.44$      | $0.09 \pm 0.09$           |
| (8) Total Background, Method 1                    | $14.3 \pm 4.1$      | $4.9 \pm 1.3$        | $1.7 \pm 0.5$             |
| (9) Total Background, Method 2                    | $9.1 \pm 2.4$       | $3.5 \pm 0.9$        | $1.1 \pm 0.3$             |
| (10) Events Before Tagging                        | 1713                | 280                  | 52                        |
| (11) Events with $b$ Tag                          | 3                   | 7                    | 5                         |

Table 8.2: Summary of the number of expected background and of observed candidate  $W + \text{multijet}$  events after  $b$  tagging.

## 8.4 Statistical Significance of the Result

The number of  $W + \text{multijet}$  events before and after  $b$ -tagging (rows (10) and (11) of table 8.2), and the background as a function of  $N_{jets}$  (row (8)) is shown in figure 8.8. For  $N_{jets} \geq 3$  we observe  $N_{cand} = 5$  events with a  $b$  tag, with an expected background  $N_{backg} = 1.7 \pm 0.5$  events.

To study the significance of the excess of observed candidate events ( $5 - 1.7$  events) over the expected background ( $1.7 \pm 0.5$  events), we calculate the probability that the background fluctuates, yielding a number of events  $\geq 5$ , i.e., the number of observed candidates. This

is done using Poisson statistics for the background distribution, convoluted with a Gaussian of mean equal to the expected background (1.7 events) and sigma equal to the uncertainty on the expected background (0.5 events). Such a probability is 4 %. If the observed excess of events (3.3 events) were entirely due to associated production of  $W$  bosons and HF's, the corresponding theoretical predictions of table 8.2 for  $Wb\bar{b}$ ,  $Wc\bar{c}$  and  $Wc$  should have been increased by a factor  $\sim 10$ . This excess of candidate events could provide direct evidence of the existence of the top quark.

## 8.5 Estimate of the $b$ -Tagging Efficiency for $t\bar{t} \rightarrow W +$ Multijet Events

The efficiency for  $\geq 1$   $b$  tags in  $p\bar{p} \rightarrow t\bar{t} \rightarrow e/\mu + \geq 3$  jets events,  $\epsilon_{t\bar{t}}(tag, raw)$ , is first estimated using the same ISAJET + CLEO + CDFSIM Montecarlo program used to generate the  $p\bar{p} \rightarrow b\bar{b} \rightarrow \mu X$  sample described in section 6.1.10. To quantify how well this Montecarlo simulates the  $b$ -tagging rate for SVX-fiducial jets in the data, in section 6.1.10 we calculated the scale factor:

$$F_{scale}(\epsilon_b(fid)) = \frac{\epsilon_b(fid)_{data}}{\epsilon_b(fid)_{MC}} = 1.00 \pm 0.15. \quad (8.1)$$

This factor does not vary with the  $E_t$  of jets in the sample. This factor is referred to SVX-fiducial jets, while  $\epsilon_{t\bar{t}}(tag, raw)$  is an efficiency per  $t\bar{t} \rightarrow e/\mu + \geq 3$  jet event and, as such, it includes the inefficiency due to events with PV's outside the SVX and to  $b$  jets which are not SVX-fiducial even when the event PV is inside the SVX. Therefore, the scale factor which is needed to normalize the raw Montecarlo efficiency  $\epsilon_{t\bar{t}}(tag, raw)$  to data can be divided into two parts:

$$F_{scale}(data/MC) = F_{scale}(\epsilon_b(fid)) \cdot R(F_{fid}^{jet}), \quad (8.2)$$

$$R(F_{fid}^{jet}) = \frac{F_{fid}^{jet}(e/\mu + \geq 1jet)}{F_{fid}^{jet}(t\bar{t} \rightarrow e/\mu + \geq 1jet)}. \quad (8.3)$$

where  $F_{fid}^{jet}(e/\mu + \geq 1jet)$  is the fraction of fiducial jets observed in the  $e/\mu +$  multijet sample, and  $F_{fid}^{jet}(t\bar{t} \rightarrow e/\mu + \geq 1jet)$  is the fraction predicted by the Montecarlo for  $t\bar{t} \rightarrow e/\mu +$  multijet events. Having measured the scale factor between the  $b$ -tagging efficiency in the data and in the Montecarlo for fiducial jets, we must then normalize  $\epsilon_{t\bar{t}}(tag, raw)$  by the ratio of the different fractions of fiducial jets observed in the data and in the Montecarlo,  $R(F_{fid}^{jet})$ .  $R(F_{fid}^{jet})$  is not calculated using the inclusive muon sample and the corresponding  $b\bar{b}$  Montecarlo sample, because  $F_{fid}^{jet}$  in the data and in the Montecarlo vary with  $E_t$  ( $R(F_{fid}^{jet})$ , on the other hand, does not vary with  $E_t$ ) and  $t\bar{t}$  and  $b\bar{b}$  events have different jet  $E_t$  spectra. In the data one measures  $F_{fid}^{jet}(e/\mu + \geq 1jet) = 0.545 \pm 0.011$ . Table 8.3 shows that  $F_{fid}^{jet}(t\bar{t} \rightarrow e/\mu + \geq 1jet) \sim 0.64$ , independent of top mass. This table also shows the values of  $R(F_{fid}^{jet})$ , the total scale factor  $F_{scale}(data/MC)$ ,  $\epsilon_{t\bar{t}}(tag, raw)$ , and the efficiency for  $t\bar{t}$  events normalized to data,  $\epsilon_{t\bar{t}}(tag, norm)$ , defined as:

$$\epsilon_{t\bar{t}}(tag, norm) = \epsilon_{t\bar{t}}(tag, raw) \cdot F_{scale}(data/MC). \quad (8.4)$$

## Electron/Muon + Jets + SVX b-Tag

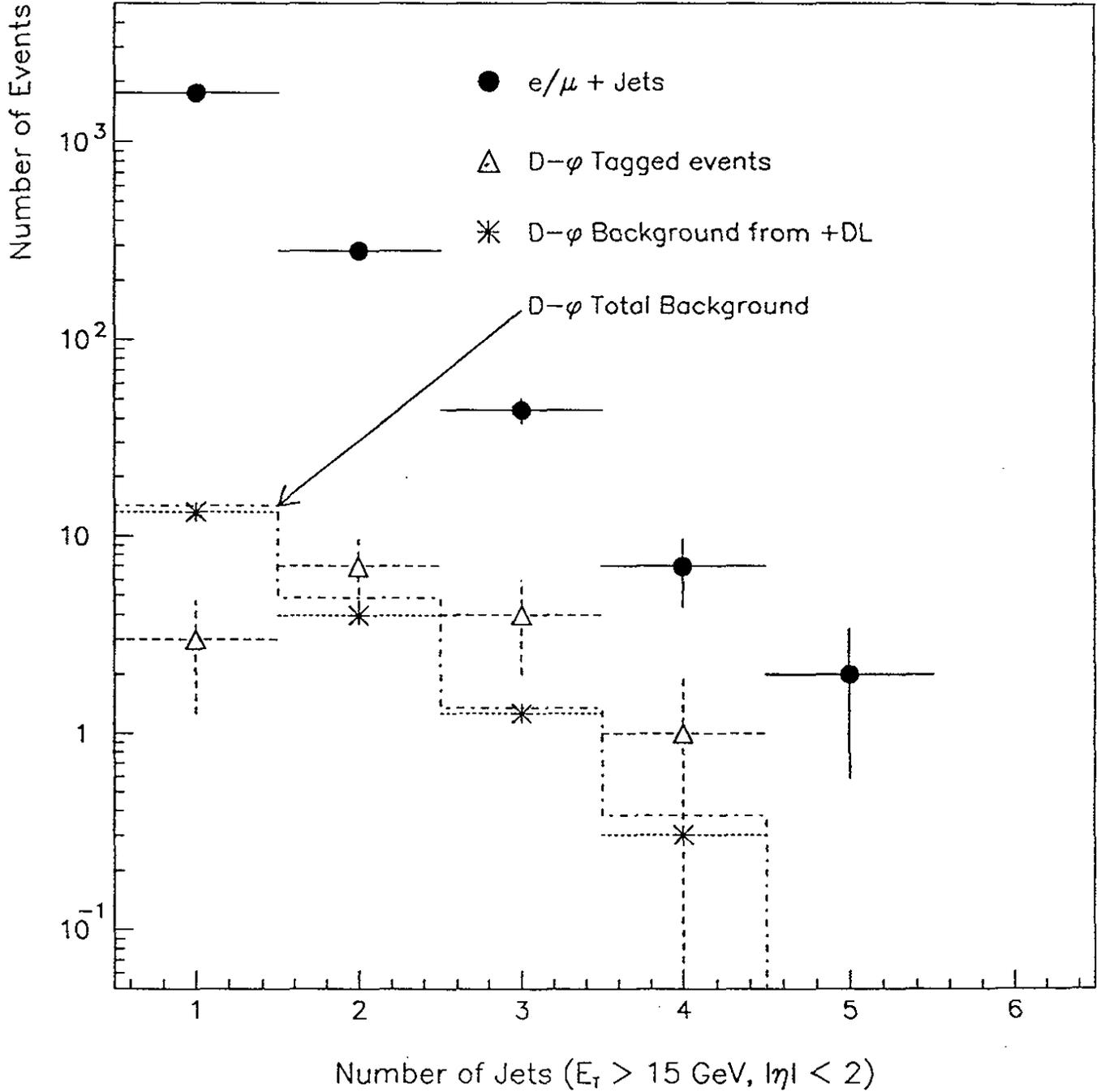


Figure 8.8: Number of jets in the electron/muon + multijet event sample ( $21.4 \text{ pb}^{-1}$  of data) before and after  $b$  tagging. Also shown is the background calculated with the matrix  $M^+$  (background from +DL) and the total expected background.

| Top-quark Mass  | 120 GeV/c <sup>2</sup> | 140 GeV/c <sup>2</sup> | 160 GeV/c <sup>2</sup> | 180 GeV/c <sup>2</sup> |
|---|------------------------|------------------------|------------------------|------------------------|
| $F_{fid}^{jet}(t\bar{t} \rightarrow e/\mu + \geq 1jet)$ | $0.643 \pm 0.008$      | $0.640 \pm 0.007$      | $0.637 \pm 0.007$      | $0.637 \pm 0.006$      |
| $R(F_{fid}^{jet})$                                      | $0.85 \pm 0.02$        | $0.85 \pm 0.02$        | $0.85 \pm 0.02$        | $0.85 \pm 0.02$        |
| $F_{scale}(data/MC)$                                    | $0.71 \pm 0.13$        | $0.71 \pm 0.13$        | $0.71 \pm 0.13$        | $0.71 \pm 0.13$        |
| $\epsilon_{t\bar{t}}(tag, raw)$                         | $0.28 \pm 0.01$        | $0.30 \pm 0.01$        | $0.30 \pm 0.01$        | $0.30 \pm 0.01$        |
| $\epsilon_{t\bar{t}}(tag, norm)$                        | $0.20 \pm 0.04$        | $0.21 \pm 0.04$        | $0.21 \pm 0.04$        | $0.21 \pm 0.044$       |

Table 8.3: Summary of scale factors between data and Montecarlo needed to estimate the  $b$ -tagging efficiency for  $t\bar{t} \rightarrow e/\mu + \geq 3$  jet events.

Note that, if the event PV found by the VXPRIM algorithm is replaced with the PV based on the average  $p\bar{p}$  beam line position (that is,  $\vec{X}_{seed}$ ,  $C_{seed}$  of section 5.1.2),  $\epsilon_{t\bar{t}}(tag, raw)$  decreases by about 20%. This is because a worse PV resolution determines a worse IP resolution and, therefore, a lower  $b$ -tagging efficiency. The excellent PV resolution provided by the VXPRIM algorithm is essential for a high-mass top-quark search in a data sample with a few candidate events.

## 8.6 Estimate of the Expected Number of $b$ -Tagged $t\bar{t} \rightarrow W +$ Multijet Events

To determine whether the excess of observed candidate events over background is consistent with the  $t\bar{t}$  production hypothesis, we estimate the expected number of top events in the  $e/\mu + \geq 3$  jets channel with  $\geq 1$   $b$  tag in the  $21.4 pb^{-1}$  of data. This is done using the top production cross section [43] and a combination of the data, of the ISAJET Montecarlo program (to generate  $p\bar{p} \rightarrow t\bar{t}$  events) + CLEO (for  $b$ -hadron decays) and of the full CDF simulation, in order to calculate the *acceptance* of the  $t\bar{t} \rightarrow e/\mu + \geq 3$  jets selection.

The acceptance is factorized in the following way:

$$A_{top} = BR(t\bar{t} \rightarrow lX) \cdot \epsilon_{lep} \cdot \epsilon_{kin} \cdot \epsilon_{tag}, \quad (8.5)$$

where  $BR(t\bar{t} \rightarrow lX)$  is the branching ratio of the decay of  $t\bar{t}$  pairs to lepton + jets,  $\epsilon_{lep}$  is the high- $P_t$  electron and muon identification efficiency,  $\epsilon_{kin}$  is the efficiency corresponding to the remaining kinematic event selection, and  $\epsilon_{tag}$  is the  $b$ -tagging efficiency.

The lepton acceptance,  $\epsilon_{lep}$ , is the efficiency for an electron or muon from  $W$  decay (the dominant source) or  $b$  decay or  $W \rightarrow \tau\nu \rightarrow e/\mu X$  decay, to have  $E_t(e) > 20$  GeV or  $P_t(\mu) > 20$  GeV/c and to pass the identification and trigger cuts.  $\epsilon_{lep}$  is determined by normalizing the efficiency estimated with the Montecarlo for isolated leptons to the value observed in a  $Z \rightarrow ee, \mu\mu$  sample from the data. The scale factor found is 0.85 for leptons in the fiducial volume. The systematic uncertainty on  $\epsilon_{lep}$  due to this procedure is 5%. Using this acceptance one measures a  $W$  cross section consistent with previous CDF measurements

[136]. The lepton non-isolation in  $t\bar{t}$  events is modeled with the Montecarlo. This gives a systematic uncertainty of 10%.

In events with a lepton passing the previous cuts, the kinematic acceptance,  $\epsilon_{kin}$ , is the efficiency to have  $E_t > 20$  GeV and  $N_{jets} \geq 3$ .  $\epsilon_{kin}$  is determined using the Montecarlo. Varying the jet energy scale by 10 % one gets a 6% systematic uncertainty. The systematic uncertainty related to the initial state radiation model in the  $t\bar{t}$  Montecarlo is estimated to be 7%. This systematic dependence has been studied by turning on/off the gluon radiation in ISAJET.

In top events passing the kinematic selection requirements,  $\epsilon_{tag}$  is the efficiency to have  $\geq 1$   $b$  tag.  $\epsilon_{tag}$  has been estimated in the previous section 8.5 ( $\epsilon_{tag} = \epsilon_{t\bar{t}}(tag, norm)$ ).

The  $t\bar{t}$  production cross section [43] for  $M_{top} = 120, 140, 160, 180$  GeV/ $c^2$ , the product  $BR(t\bar{t} \rightarrow lX) \cdot \epsilon_{lep} \cdot \epsilon_{kin}$ , the value of  $\epsilon_{tag}$ , the value of the total acceptance,  $A_{top}$ , and the expected number of  $t\bar{t}$  events in the  $e/\mu + \geq 3$  jets channel with  $\geq 1$   $b$  tag for the CDF integrated luminosity of  $\mathcal{L} = 21.4$  pb $^{-1}$ ,  $N_{top}$ , are shown in table 8.4.  $N_{top}$  is calculated as:

$$N_{top} = \sigma_{t\bar{t}}(theory) \cdot A_{top} \cdot \mathcal{L}. \quad (8.6)$$

## 8.7 Study of the $t\bar{t}$ Production Hypothesis

If the top-quark mass is in the region 120-180 GeV/ $c^2$ , the excess of candidate events over background ( $N_{cand} - N_{backg} = N_{signal} = 3.3$  events) is consistent with the number of expected  $b$ -tagged  $t\bar{t}$  events (8.3-1.5 events, see table 8.4). The values of  $A_{top}$ ,  $N_{cand}$ ,  $N_{backg}$ , and of the integrated luminosity,  $\mathcal{L}$ , can be used to determine the *observed* cross section for  $t\bar{t}$  production in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.8$  TeV. Such a cross section,  $\sigma_{t\bar{t}}(observed)$ , is calculated using a Likelihood function, which is the convolution of a Poisson distribution for the number of candidate events, and of Gaussian distributions for the number of background events, for the acceptance and for the luminosity. For the luminosity we consider a 10% systematic uncertainty. The result of the likelihood fit is reported in table 8.4. Since the  $t\bar{t}$  acceptance varies with the top-quark mass, the  $t\bar{t}$  cross section also varies with the mass. For completeness, we also show in table 8.4 the cross section obtained by using method 2 as a background estimate.

Note that the difference between  $\sigma_{t\bar{t}}(pb)(observed; method 1)$  and  $\sigma_{t\bar{t}}(pb)(observed; method 2)$  is lower than the uncertainty on both cross sections. The SM-predicted  $t\bar{t}$  cross section and  $\sigma_{t\bar{t}}(pb)(observed; method 1)$  as a function of  $M_{top}$  are compared in figure 8.9: the central values of the two distributions intersect at  $M_{top} \sim 155 - 160$  GeV/ $c^2$ .

# $t\bar{t}$ Production Cross Section

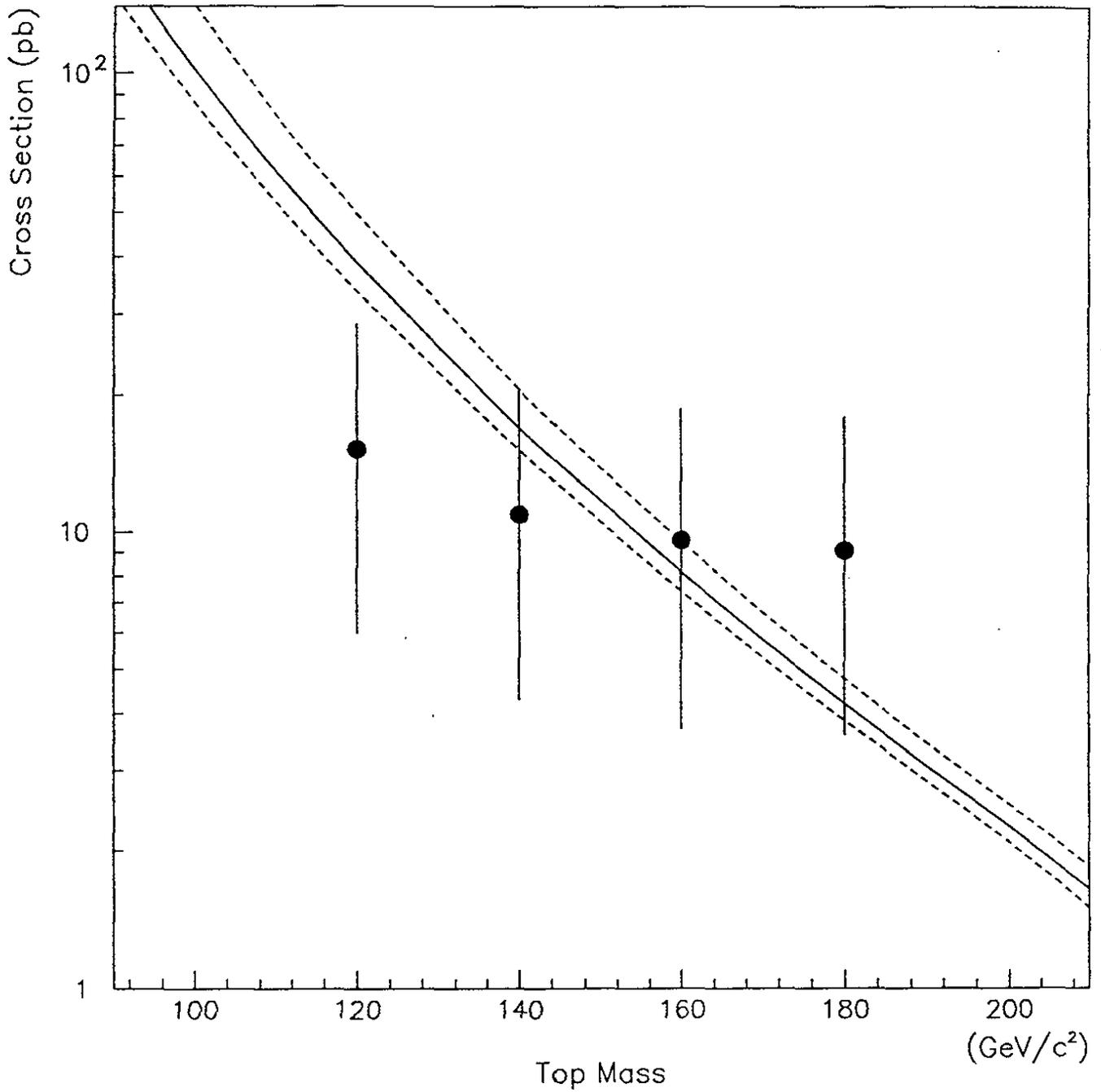


Figure 8.9:  $t\bar{t}$  production cross section measured in the data using method 1 (the most conservative) as background estimate (points) compared to the Standard Model prediction (curves).

| Top-quark Mass  | 120 GeV/c <sup>2</sup> | 140 GeV/c <sup>2</sup> | 160 GeV/c <sup>2</sup> | 180 GeV/c <sup>2</sup> |
|---|------------------------|------------------------|------------------------|------------------------|
| $\sigma_{t\bar{t}}(pb)$ (theory)                                  | $38.9^{+10.8}_{-5.2}$  | $16.9^{+3.6}_{-1.8}$   | $8.2^{+1.4}_{-0.8}$    | $4.2^{+0.6}_{-0.4}$    |
| $BR(tt \rightarrow lX) \cdot \epsilon_{lep} \cdot \epsilon_{kin}$ | $0.051 \pm 0.007$      | $0.067 \pm 0.008$      | $0.077 \pm 0.008$      | $0.081 \pm 0.008$      |
| $\epsilon_{tag}$  | $0.20 \pm 0.04$        | $0.21 \pm 0.04$        | $0.21 \pm 0.04$        | $0.21 \pm 0.04$        |
| $A_{top}$   | $0.010 \pm 0.002$      | $0.014 \pm 0.003$      | $0.016 \pm 0.004$      | $0.017 \pm 0.004$      |
| $N_{top}$   | $8.3 \pm 3.1$          | $4.9 \pm 1.9$          | $2.6 \pm 1.0$          | $1.5 \pm 0.5$          |
| $\sigma_{t\bar{t}}(pb)$ (observed; method 1)                      | $15.2^{+13.4}_{-9.2}$  | $10.9^{+9.7}_{-6.6}$   | $9.6^{+9.1}_{-5.9}$    | $9.1^{+8.4}_{-5.5}$    |
| $\sigma_{t\bar{t}}(pb)$ (observed; method 2)                      | $18.2^{+13.8}_{-9.4}$  | $13.0^{+10.0}_{-6.7}$  | $11.5^{+9.3}_{-5.9}$   | $10.8^{+8.6}_{-5.6}$   |

Table 8.4: Summary of the acceptance, of the  $b$ -tagging efficiency, of the expected number of  $b$ -tagged  $t\bar{t} \rightarrow e/\mu + \geq 3$  jets in  $21.4 pb^{-1}$  of data; also shown is the  $t\bar{t}$  production cross section predicted by the Standard Model and the observed cross section obtained by using both method 1 and method 2 as background estimates.

# Bibliography

- [1] "Evidence for Top Quark Production in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.8$  TeV", CDF Collaboration, Phys. Rev. Lett. **73**, 225 (1994), "Evidence for Top Quark Production in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.8$  TeV", Fermilab-Pub-94/097-E and Phys. Rev. D, **50**, 2966 (1994); the results of this Ph.D. thesis are directly incorporated in section V.B ("Search for  $t\bar{t}$  Using SVX  $b$  Tagging") of Phys. Rev. D, **50**, 2966 (1994).
- [2] S.L. Glashow, Nucl. Phys. **22**, 579 (1961);  
S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967);  
A. Salam, "Elementary Particle Theory" (Nobel Symposium n. 8), Svartholm editor (Almquist and Wiksell, Stockholm, 1968) p. 367;  
S.L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. D **2**, 1285 (1970).
- [3] C.N. Yang and R.L. Mills, Phys. Rev. **96**, 191 (1954);  
M. Gell-Mann, "Acta Physica Austriaca, Suppl.", **IX**, 733 (1972);  
H. Fritzsch and M. Gell-Mann, XVI International Conference High Energy Physics, Batavia, Vol. II p.135 (1972);  
H. Fritzsch, M. Gell-Mann and H. Leutwyler, Phys. Lett. B **47**, 365 (1973).
- [4] LEP Collaborations and Electroweak Working Group, CERN-PPE/93-157, prepared for the "Europhysics Conference on High Energy Physics, Marseille", July 22-28 1993, and for the "1993 International Symposium on Lepton-Photon Interactions at High Energies", Cornell, August 10-15.
- [5] UA1 Collaboration, Z. Phys. C **44**, 15 (1989);  
UA2 Collaboration, CERN-PPE/91-163.
- [6] CDF Collaboration, Phys. Rev. Lett. **65**, 2243 (1990);  
CDF Collaboration, Phys. Rev. D **43**, 2070 (1991); CDF Collaboration, Fermilab preprint, Fermilab-Conf-93/210-E, 1993.
- [7] Particle Data Group, Phys. Rev. D **45** (1992).
- [8] A. Ali, C. Greub and T. Mannel, in the Proceedings of the ECFA Workshop on the Physics of the  $B$  Meson Factory, 1993, edited by R. Aleksan and A. Ali (unpublished).
- [9] CLEO Collaboration, Phys. Rev. D **49**, 5701 (1994).
- [10] G.L. Kane and M. Peskin Nucl. Phys. B **195**, 29 (1982); see also CLEO Collaboration, Phys. Rev. D **35**, 3533 (1987).

- [11] JADE Collaboration, W. Bartel *et al.*, Phys. Lett. B **146**, 437 (1984).
- [12] D. Schaile and P.M. Zerwas, DESY Rep. 91-106.
- [13] M. Kobayashi and K. Maskawa, Progr. Theor. Phys. **49**, 652 (1973).
- [14] N. Cabibbo; Phys. Rev. Lett. **10**, 531 (1963).
- [15] W.A. Bardeen *et al.*, Phys. Rev. D **18**, 3998 (1978).
- [16] J. Huth and M.L. Mangano, Fermilab-Pub-93/019-E, Ann. Rev. Nucl. Part. Sci. **43**, 585 (1993).
- [17] S.D. Ellis, Z. Kunszt, and D.E. Soper, Phys. Rev. Lett. **64**, 2121 (1991).
- [18] CDF Collaboration, Phys. Rev. Lett. **68**, 1104 (1992);  
CDF Collaboration, Phys. Rev. Lett. **70**, 1376 (1993).
- [19] CDF Collaboration, Phys. Rev. D **48**, 998 (1993).
- [20] CDF Collaboration, Phys. Rev. D **45**, 1448 (1992);  
CDF Collaboration, Phys. Rev. Lett. **69**, 2896 (1992).
- [21] CDF Collaboration, Phys. Rev. Lett. **71**, 500 (1993).
- [22] CDF Collaboration, Phys. Rev. Lett. **71**, 2542 (1993).
- [23] CDF Collaboration, Fermilab preprint, Fermilab-Pub-94/072-E.
- [24] UA1 Collaboration, Z. Phys. C **48**, 1 (1990).
- [25] UA2 Collaboration, Z. Phys. C **46**, 179 (1990).
- [26] See, for example, F. Dydak, in the proceedings of the "25th International Conference on High Energy Physics", Singapore, 1990, K. K. Phua and Y. Yamaguchi editors (World Scientific, Singapore, 1991).
- [27] CDF Collaboration, Phys. Rev. D **44**, 29 (1991).
- [28] UA1 Collaboration, Phys. Lett. B **253**, 503 (1991).
- [29] UA2 Collaboration, Phys. Lett. B **276**, 354 (1991).
- [30] UA2 Collaboration, Phys. Lett. B **276**, 365 (1991).
- [31] CDF Collaboration, Fermilab preprint, Fermilab-Conf-93/210-E, 1993.
- [32] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. B **192**, 245 (1987);  
CLEO Collaboration, M. Artuso *et al.*, Phys. Rev. Lett. **62**, 2233 (1989);  
G. Altarelli and P. Franzini. Z. Phys. C **37**, 271 (1988);  
J. Ellis *et al.*, Nucl. Phys. B **304**, 205 (1988).

- [33] "Expression of Interest in a Future Collider Detector in B0", CDF internal note # 2436 (May 3, 1994).
- [34] ALEPH Collaboration, Phys. Lett. B **313**, 299 (1993)
- [35] CDF Collaboration, Phys. Rev. Lett. **68**, 447 (1992).
- [36] CDF Collaboration, Phys. Rev. D **45**, 3921 (1992).
- [37] D0 Collaboration, Phys. Rev. Lett. **72**, 2138 (1994).
- [38] R.K. Ellis, "Status of QCD", presented at the "7th Meeting of the American Physical Society Division of Particles and Fields (DPF 92)", Fermilab, November 1992, C. Albright *et al.* editors, World Scientific, Singapore. Also, M. L. Mangano, private communication.
- [39] P. Nason, S. Dawson and R.K. Ellis, Nucl. Phys. B **303**, 607 (1988); B **327**, 49 (1989); B **335**, 260 (1990).
- [40] G. Altarelli, M. Diemoz, G. Martinelli and P. Nason, Nucl. Phys. B **308**, 724 (1988).
- [41] W. Beenakker, H. Kuijf, W.L. van Neerven and J. Smith, Phys. Rev. D **40**, 54 (1989); W. Beenakker, W.L. van Neerven, R. Meng, G.A. Schuler and J. Smith, Nucl. Phys. B **351**, 507 (1991).
- [42] R.K. Ellis, Phys. Lett. B **259**, 492 (1991).
- [43] E. Laenen, J. Smith and W. van Neerven, Fermilab-PUB-93/270-T (1993), ITP-SB-93-55, THU-93/23.
- [44] B. Denby *et al.*, CDF Collaboration, "Top Search in the Multijet Channel at CDF", presented at the "9th Topical Workshop on Proton-Antiproton Collider Physics", Tsukuba, Japan, October 1993.
- [45] CDF Collaboration, Phys. Rev. Lett. **70**, 4042 (1993).
- [46] CDF Collaboration, Phys. Rev. Lett. **64**, 147 (1990); CDF Collaboration, Phys. Rev. D **43**, 664 (1991).
- [47] CDF Collaboration, Phys. Rev. Lett. **64**, 142 (1990).
- [48] F.A. Behrends *et al.*, Fermilab-Pub-92/196-T, (1992), to be published.
- [49] M.L. Mangano, "Production of  $W$  Plus Heavy Quark Pairs", Nucl. Phys. B **405**, 536 (1993).
- [50] W. Venus, proceedings of the "XVI International Symposium on Lepton-Photon Interactions", 10-15, August 1993; see also CDF Collaboration, Phys. Rev. Lett. **71**, 3421 (1993).
- [51] CDF Collaboration, Phys. Rev. Lett. **72**, 3456 (1994)

- [52] C. Campagnari, B. Farhat, P. Sphicas, A. Yagil, internal CDF notes # 1961, 2098, 2150, 2161, 2174, 2245.
- [53] S. Dell’Agnello *et al.*, “The  $D - \phi$  Clustering  $b$ -Tagger”, internal CDF note # 1882; D. Gerdes *et al.*, “ $B$  Identification Using Jet Probability”, internal CDF note # 2023; R. Hughes *et al.*, “Documentation on the JETVTX  $B$ -Tagging Module”, internal CDF note # 2068.
- [54] M. Contreras, internal CDF note # 1806.
- [55] J. Gunion *et al.*, The Higgs Hunter’s Guide (Addison-Wesley, New York, 1990); S.L. Glashow and E.E. Jenkins, Phys. Lett. B **196**, 233 (1987); V. Barger, J.L. Hewett and R.J.N. Phillips, Phys. Rev. D **41**, 3421 (1990).
- [56] CDF Collaboration, “Search for the Top Quark Decaying to a Charged Higgs in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.8$ ”, Phys. Rev. Lett. **72**, 1977 (1994).
- [57] “Search for the Top Quark Decaying to a Charged Higgs in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.8$  TeV”, submitted to Phys. Rev. Lett., FERMILAB-PUB-94/195-E.
- [58] ALEPH Collaboration, Phys. Rep. **216**, 253 (1992); DELPHI Collaboration, Phys. Lett. B **241**, 449 (1990); L3 Collaboration, Phys. Lett. B **294**, 457 (1992); OPAL Collaboration, Phys. Lett. B **242**, 299 (1990).
- [59] Nucl. Instr. Meth. A **271**, 387 (1988).
- [60] “Proposal for an Upgraded CDF Detector”, internal CDF note # 1172.
- [61] Nucl. Instr. Meth. A **268**, 75 (1988).
- [62] Nucl. Instr. Meth. A **268**, 268 (1988).
- [63] Nucl. Instr. Meth. A **267**, 272 (1988).
- [64] Nucl. Instr. Meth. A **267**, 280 (1988).
- [65] Nucl. Instr. Meth. A **267**, 257 (1988).
- [66] Nucl. Instr. Meth. A **267**, 301 (1988).
- [67] M.E.B Franklin *et al.*, proceedings of ‘7<sup>th</sup> Topical Workshop on Proton-Antiproton Collider Physics’, p. 420, Fermilab, 1988, published by World Scientific.
- [68] Nucl. Instr. Meth. A **267**, 249 (1988).
- [69] Internal CDF note # 1248.
- [70] Nucl. Instr. Meth. A **268**, 46 (1988).

- [71] "A proposal to Upgrade the CDF Muon System", internal CDF note # 773; see also "Proposal for an Upgraded CDF Detector", internal CDF note # 1172.
- [72] Nucl. Instr. Meth. A **268**, 33 (1988).
- [73] Internal CDF notes # 1500 and 1964
- [74] Internal CDF note # 1614.
- [75] C.E. Pagliarone, Laurea thesis (1993), University of Pisa.
- [76] D. Amidei *et al.*, Nucl. Instr. Meth. A **269**, 51 (1988).
- [77] G. Ascoli *et al.*, Nucl. Instr. Meth. A **269**, 63 (1988).
- [78] CDF Collaboration, Phys. Rev. Lett. **71**, 1685 (1993).
- [79] F. Bedeschi *et al.*, IEEE Trans. on Nucl. Sci. NS-**33**, p. 140 (1985). See also F. Bedeschi *et al.*, internal CDF note # 601.
- [80] S.A. Kleinfelder *et al.*, IEEE Trans. on Nucl. Sci. NS-**35**, p. 171 (1988).
- [81] C. Haber *et al.*, proceedings of "7<sup>th</sup> Topical Workshop on Proton-Antiproton Collider Physics", p. 443, Fermilab, 1988, published by World Scientific.
- [82] W.C. Carithers, *et al.*, Nucl. Instr. Meth. A **289**, 388 (1990).
- [83] IEEE Trans. on Nucl. Sci. NS-**37**, p. 1120 (1990).
- [84] F. Bedeschi *et al.*, IEEE Trans. on Nucl. Sci. NS-**36**, p. 35 (1989).
- [85] S. Dell'Agnello *et al.*, Nucl. Instr. Meth. A **315**, 188 (1992), presented by S. Dell'Agnello at "5th Pisa Meeting on Advanced detectors", La Biodola, Italy, May 1991; S. Dell'Agnello *et al.*, internal CDF note # 1423; G.F. Tartarelli, Laurea thesis (1990), University of Pisa.
- [86] "The SVX Laser Test", S. Dell'Agnello *et al.*, internal CDF note # 1480.
- [87] "The CDF Silicon Vertex Detector: an Overview and Test Results", W. C. Carithers *et al.*, proceedings IEEE Nucl. Sci. Symp. (1992), p. 475.
- [88] B. Barnett *et al.*, Nucl. Instr. Meth. A **315**, 125 (1992).
- [89] "The CDF Silicon Vertex Detector", presented by S. Dell'Agnello for the CDF Collaboration at the SIF (Italian Physical Society) conference, L'Aquila, Italy (1991).
- [90] "Construction of the CDF Silicon Vertex Detector", J. Skarha, S. Dell'Agnello *et al.*, Published Proceedings IEEE Nucl. Sci. Symp. (1992).
- [91] "Commissioning of the Control and Data Acquisition Electronics for the CDF Silicon Vertex Detector", S.M. Tkaczyk *et al.*, IEEE Nucl. Sci. Symp., p. 753 (1991).

- [92] "Control and Data Acquisition Electronics for the CDF Silicon Vertex Detector", K.J. Turner, *et al.*, IEEE Nucl. Sci. Symp., p. 888 (1991).
- [93] "The SVX of the Collider Detector at Fermilab", FERMILAB-PUB-94/024-E, to be published in Nucl. Instr. Meth..
- [94] "Prospects of Physics at CDF with the SVX", presented by S. Dell'Agnello for the CDF Collaboration at the "3rd Topical Seminar on Heavy Flavors", San Miniato, Italy, June 1991, Nucl. Phys. B. (Proc. Suppl.) 27 246 (1992).
- [95] "Top-Quark Search at CDF with the SVX", presented by S. Dell'Agnello for the CDF Collaboration at the SIF conference, Pavia, Italy (1992).
- [96] Internal CDF note # 1606.
- [97] Internal CDF note # 1673.
- [98] Nucl. Instr. Meth. A 241, 115 (1985);  
DELPHI internal note 84-18 Prog 5 (1984).
- [99] F. Bedeschi, internal CDF note # 371 (1985);  
H. Wenzel, internal CDF note # 1790 (1992).
- [100] P. Derwent, internal CDF note # 2188.
- [101] W. Yao *et al.*, internal CDF note # 1948.
- [102] F. Bedeschi, internal CDF note # 817.
- [103] F. Bedeschi, internal CDF note # 818.
- [104] M. Schub, F. Bedeschi, internal CDF note # 823.
- [105] A. Byon, internal CDF note # 822.
- [106] H. Wenzel, internal CDF note # 1924.
- [107] S. Dell'Agnello *et al.*, "A Primary Vertex Finding Package", internal CDF note # 1789.
- [108] S. Dell'Agnello, F. Bedeschi, "Combining the SVX Beam Line Measurement with the Primary Vertex Finder", internal CDF note # 2055.
- [109] S. Dell'Agnello, F. Bedeschi, "Handling Multiple Primary Vertices with the Primary Vertex Finder", internal CDF note # 1084.
- [110] S. Dell'Agnello *et al.*, "The  $D - \phi$  Clustering  $b$ -Tagger", internal CDF note # 1882.
- [111] S. Dell'Agnello *et al.*, internal CDF note # 2022.
- [112] F.E. Paige and S.D. Protopopescu, Report # BNL-38034, 1986 (not published). We used version 6.43.

- [113] P. Avery, K. Read, and G. Trahern, Report # CSN-212, 1985 (not published).
- [114] J. Freeman, in the "Proceedings of the Workshop on Detector Simulation for the SSC", Argonne, Illinois, 1987, L.E. Price editor (ANL Report No. ANL-HEP-CP -88-51, Argonne, 1978), p. 190.
- [115] P. Derwent *et al.*, internal CDF note # 1962.
- [116] Collaboration ALEPH, CERN-PPE/93-108 (to be published in Phys. Lett. B); see also internal ALEPH note 92-133.
- [117] D. Gerdes *et al.*, internal CDF note #2023.
- [118] M.L. Mangano *et al.*, internal CDF note # 1783; M.L. Mangano *et al.*, Nucl. Phys. B **373**, 295 (1992).
- [119] P. Derwent *et al.*, private communication.
- [120] Collaboration CDF, Phys. Rev. Lett. **71**, 2396 (1993).
- [121] S. Dell'Agnello, F. Bedeschi, "Performance of the  $D - \phi$   $b$ -Tagger on Control Samples and Lepton + Jet Events", internal CDF note # 2093.
- [122] S. Dell'Agnello, F. Bedeschi, *et al.*, "Top-Quark Search in the Lepton+Jet Channel with the  $D - \phi$   $b$ -tagger", internal CDF note # 2288.
- [123] D. Amidei *et al.*, "Top Search with SVX  $b$ -Tag", internal CDF note # 2095.
- [124] C. Peterson, D. Schlatter, I. Schmitt, and P.M. Zerwas, Phys. Rev. D **27**, 105 (1983).
- [125] J. Chrin, Z. Phys. C **36**, 165 (1987); D. Decamp *et al.*, Phys. Lett. B **244**, 551 (1990). We used  $\langle z \rangle = 0.83 \pm 0.03$  or  $\epsilon = 0.006 \pm 0.002$ .
- [126] G. Houk and B. Harral, internal CDF note # 2211.
- [127] P. Kohen *et al.*, internal CDF note # 2212.
- [128] N.M. Shaw *et al.*, internal CDF note # 2353.
- [129] CDF Collaboration, Phys. Rev. D **43**, 2070 (1991).
- [130] G. Marchesini and B.R. Webber, Nucl. Phys. B **310**, 461 (1988).
- [131] M.L. Mangano and P. Nason, Phys. Lett. B **285**, 160 (1992); A.H. Mueller and P. Nason Nucl. Phys. B **266**, 265 (1986).
- [132] M.L. Mangano, P. Nason and G. Ridolfi, Nucl. Phys. B **373**, 295 (1992).
- [133] U. Baur *et al.*, FSU-HEP-930816 (1993).
- [134] F.A. Berends, W.T. Giele, H. Kuijf and B. Tausk, Nucl. Phys. B **357**, 32 (1991).

- [135] H. Plochow-Besch, CERN-PPE/92-123 (1992); J. Botts *et al.*, Phys. Lett. B **304**, 159 (1993); A.D. Martin, W.J. Stirling e R.G. Roberts, Phys. Rev. D **47**, 867 (1993).
- [136] CDF Collaboration Phys. Rev. D **44**, 29 (1991);  
CDF Collaboration Phys. Rev. Lett. **69**, 28 (1992).