PRESTACIONES DEL ALGORITMO JET PROBABILITY PARA LA SELECCIÓN DE SUCESOS $t\bar{t}$ EN EL DETECTOR CDF Y MEDIDA DE LA SECCIÓN EFICAZ DE PRODUCCIÓN $p\bar{p} \rightarrow t\bar{t}$

Enrique Palencia Cortezón
Instituto de Física de Cantabria

XXX Reunion Bienal de la RSEF, Orense, Septiembre 2005

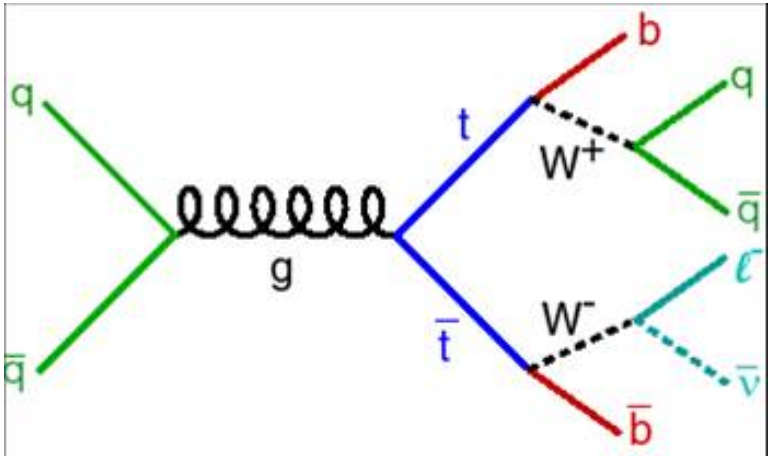
Índice

- Motivación
  - ¿Por qué Top?
  - ¿Por qué etiquetado de sabor fuerte?
  - ¿Por qué Jet Probability?
- El Tevatron y CDF
- Descripción del Algoritmo Jet Probability
  - Algoritmo
  - Eficiencia
  - MisTag Rate
- Medida de la Sección Eficaz de Producción $t\bar{t}$

Motivación

- ¿Por qué Top?
  - Partícula fundamental conocida más pesada $\Rightarrow$ papel especial en EWSB
  - Podría ser sensible a física más allá del Modelo Estandard
  - Debido a su pequeña vida media, decae antes de hadronizar
  - Último quark en ser descubierto (1995, FNAL): estudio detallado con el aumento de estadística

- ¿Por qué etiquetado de sabor fuerte?
  - En colisiones $p\bar{p}$, el top se produce pincipalmente en pares $t\bar{t}$ via procesos de aniquilación $q\bar{q}$ y, en menor medida, mediante fusión de gluones; y decae a un boson W y un quark b (BR $\sim$ 100%)
  - La señal del top presenta 2 quarks b y solo $\sim$ 1% de los principales fondos tienen sabor fuerte $\Rightarrow$ S/F aumenta



- ¿Por qué Jet Probability?
  - JP proporciona una variable continua $\Rightarrow$ forma más flexible de entender la composición de la muestra optimizando el corte de JP
  - JP se puede optimizar de forma diferente dependiendo del tipo de análisis

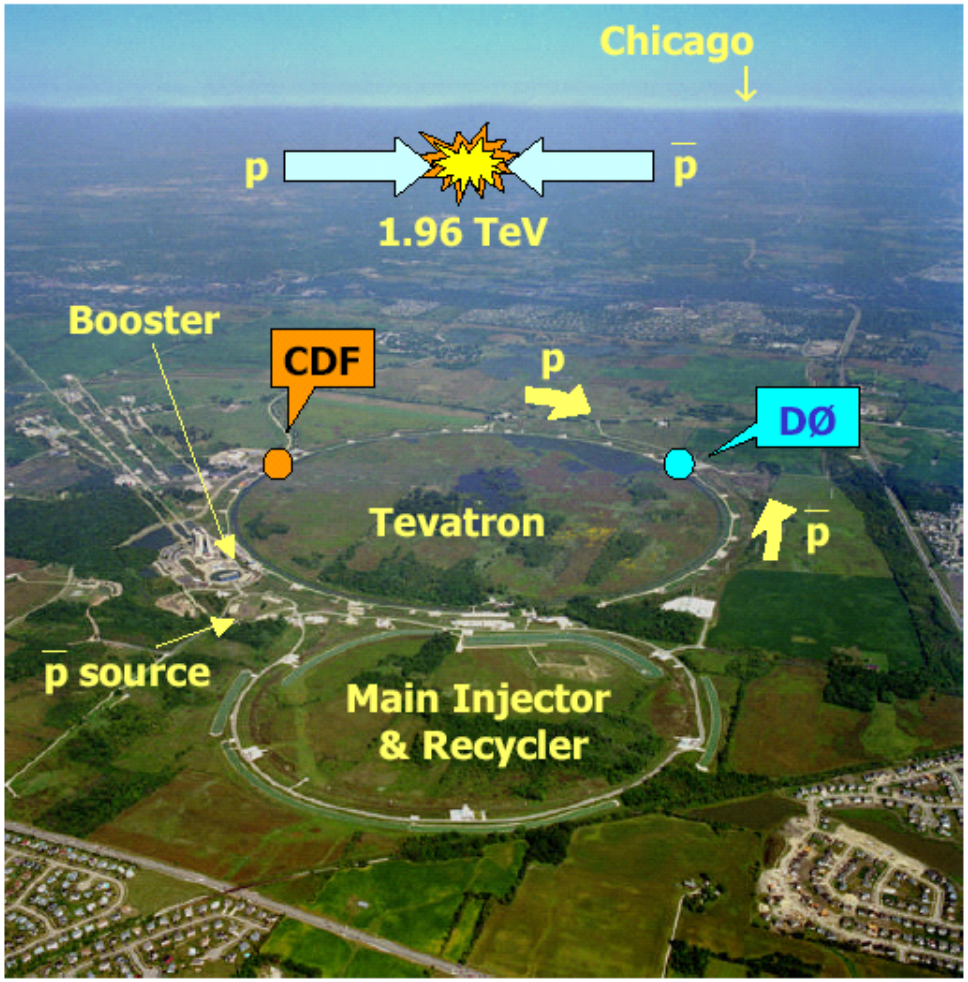
El Tevatron

- Colisionador $p\bar{p}$ a más alta energía
  - Energía de los haces = $980\ GeV$
  - $\sqrt{s} = 1.96\ TeV$ (Run I $\rightarrow 1.8\ TeV$)

- Colisiones cada $396\ ns$ (Run I $3.5\ \mu s$)

- Run I: 1992 - 1996 (¡quark *top*!)

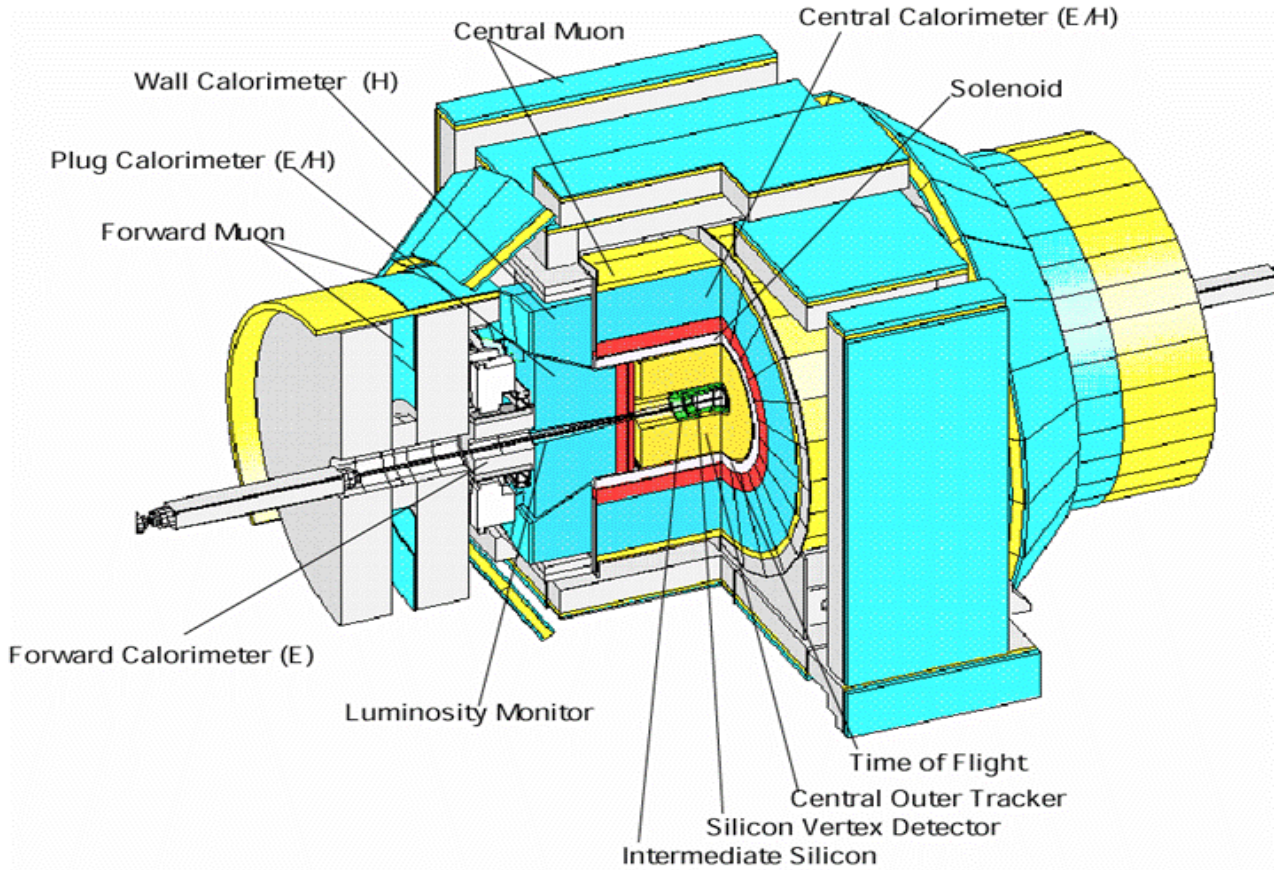
- Run II: 2001 - actualidad
  - Numerosas mejoras: *Main Injector*
  - $\mathcal{L}_{int}$: $100\ pb^{-1}$ (Run I) $\longrightarrow > 1\ fb^{-1}$ (Run II)



- Otros descubrimientos: quark *bottom* (1977) y ν_τ (2000)

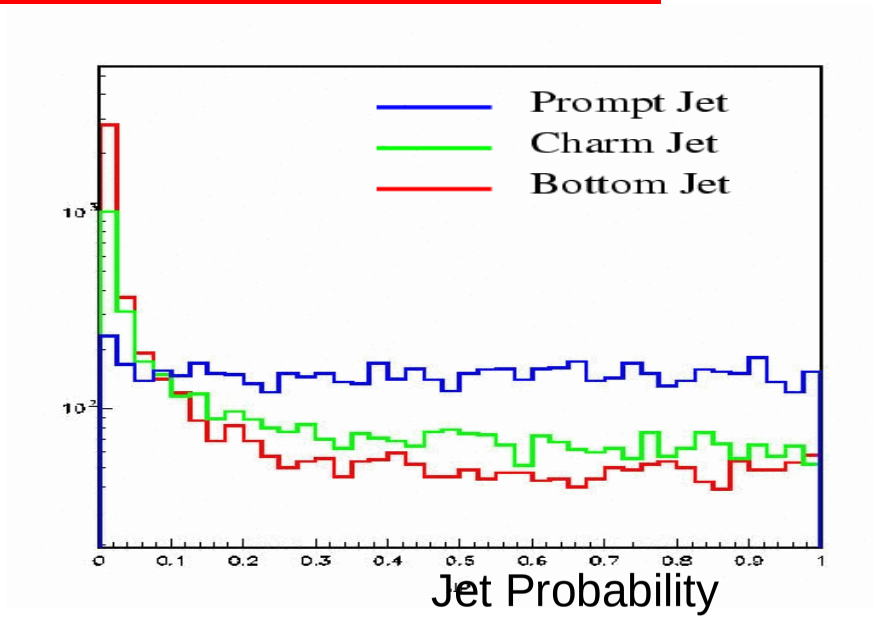
CDF

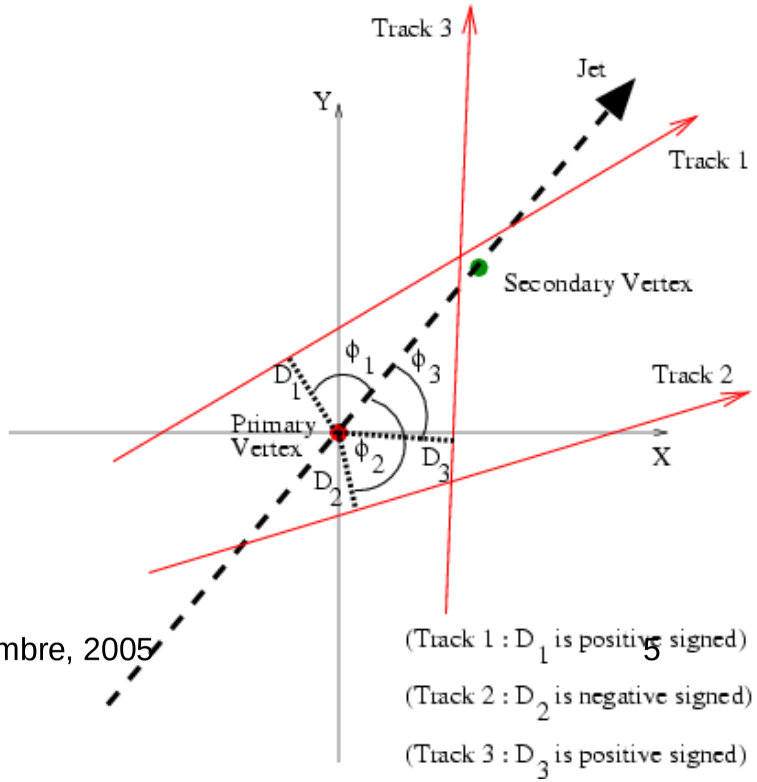
- Detector de partículas de propósito general. Simetría cilíndrica
- 3 subsistemas: reconstrucción de trazas (inmerso en un solenoide de $1.4\ T$), calorimetría y muones
- Mejoras del Run II:
  - Nuevo detector de Silicio
  - Detector de TOF
  - Calorímetros *plug*
  - Detectores *forward* de μ
  - Electrónica DAQ y triggers
  - SVT trigger en L2



El Algoritmo Jet Probability

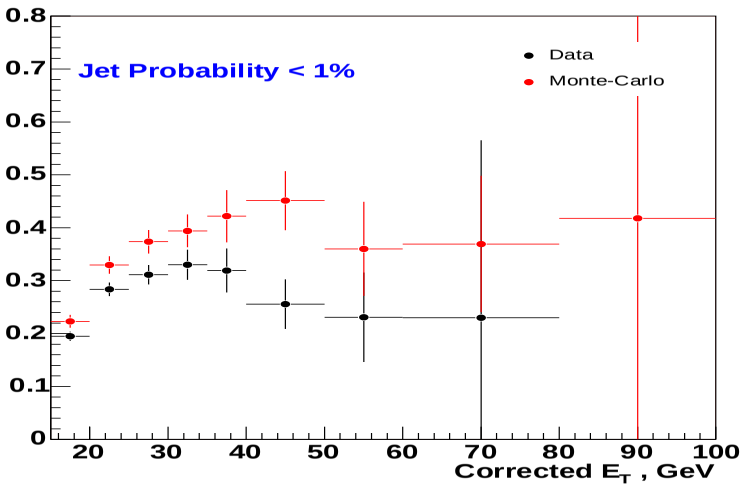
- Los hadrones con sabor fuerte tienen una vida media larga y viajan una distancia detectable antes de desintegrarse, dando lugar a vértices (y trazas) desplazados del vértice primario
- JP es un algoritmo que sirve para determinar si un jet se ha producido en el vértice primario de la interacción o si proviene de la hadronización de un quark con sabor fuerte
- Físicamente, representa la probabilidad de que un jet se haya producido en el vértice primario de la interacción
- Probabilidad uniforme para light quark or gluon jets y con un pico en 0 para jets que contienen trazas desplazadas debido a desintegraciones de HF
- Se utiliza un parámetro de impacto con signo: D>0 si el punto de máxima aproximación al vértice primario está en la misma dirección que la dirección del jet (cos $\phi > 0$)
- + (-) Jet Probability: solo trazas con parámetro de impacto positivo (negativo)

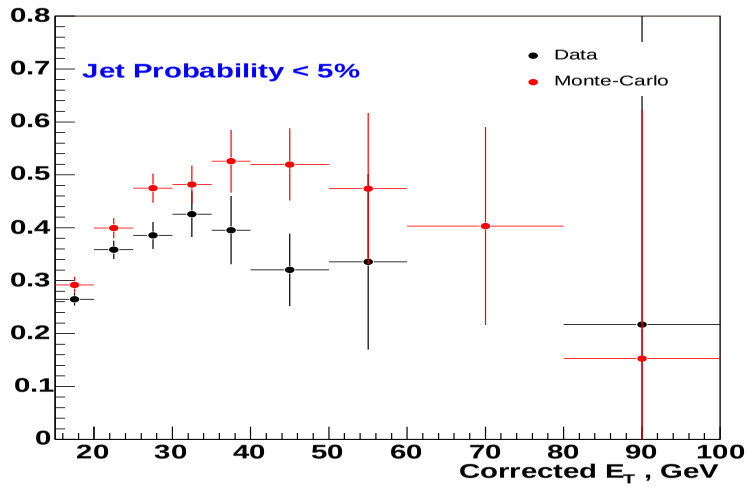




Eficiencia de Jet Probability

- Para la medida se utiliza una muestra de datos inclusiva de electrones de $\geq$ 8 GeV pues está enriquecida con HF debido a los desintegraciones semileptónicas de B
- Puesto que también utilizamos MC en el análisis, se mide la eficiencia en MC y, después, se calcula el Factor de Escala ($\epsilon^{datos}/\epsilon^{MC}$)
- Eficiencia de JP vs E_T^{jet} en una muestra inclusiva de electrones





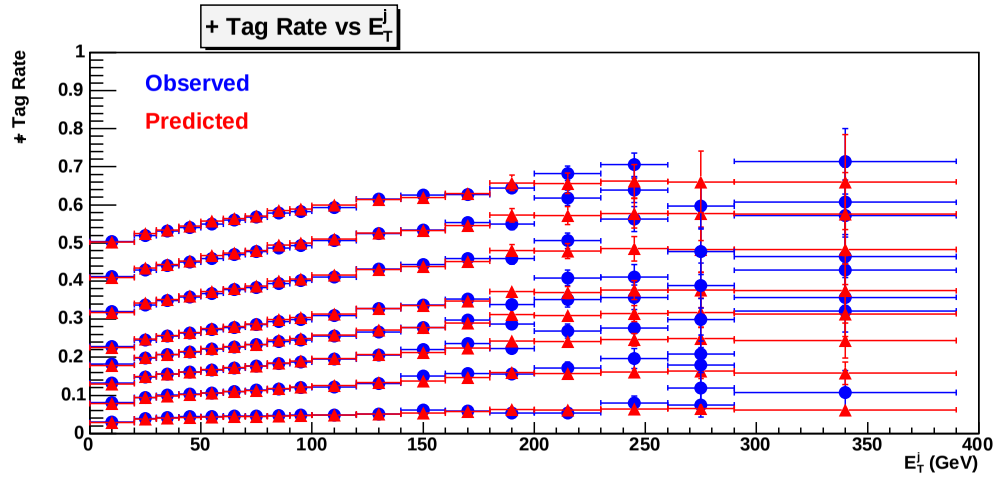
- Eficiencias de etiquetar un jet de sabor fuerte con E_T >15 GeV (318 pb^{-1})

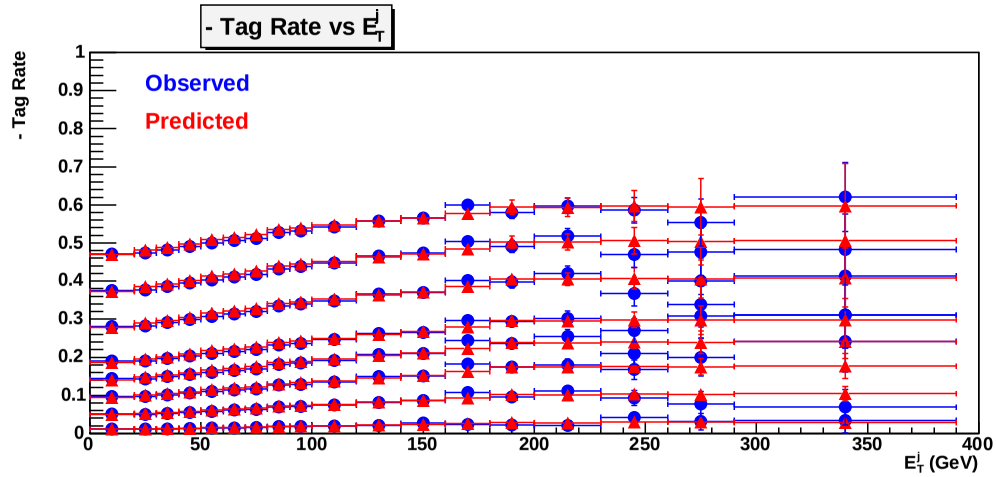
	JP $<$ 1%	JP $<$ 5%
ϵ^{datos}	0.258 $\pm$ 0.018	0.334 $\pm$ 0.026
ϵ^{MC}	0.316 $\pm$ 0.021	0.392 $\pm$ 0.026
Factor de Escala	0.817 $\pm$ 0.070	0.852 $\pm$ 0.072

Jet Probability: MisTag Rate

- Fondos debido a jets producidos por quarks ligeros o por gluones y que son etiquetados, erróneamente, como jets de sabor fuerte
- Mistag rate: probabilidad de identificar un jet ligero como jet de sabor fuerte
- Se determina, a partir de muestras de jets inclusivas en datos con diferentes umbrales de trigger (20, 50, 70 y 100 GeV), utilizando trazas con parámetro de impacto negativo
- Se parametriza como una matriz de 6 dimensiones de las variables: $(E_T,\, N_{trk},\, \sum E_T^j,\, \eta,\, Z_{vtx},\, \phi)$
- Cross check: observado (de la muestra de trigger de multijet) vs predicción (de jets inclusivos en datos)
- Resultados con 318 pb^{-1}

	JP $<$ 1%	JP $<$ 5%
- tag rate promedio (%)	1.22 ± 0.08	5.30 ± 0.25
+ tag rate promedio (%)	3.54 ± 0.18	9.20 ± 0.26





Medida de la Sección Eficaz de $t\bar{t}$: Selección de Datos

- Medida de contaje: $\sigma_{t\bar{t}} = \frac{N_{obs} - B_{bkg}}{\epsilon_{t\bar{t}} \times \int L dt}$

- Muestra de datos: todos los datos del Run II con Si (hasta Sept. 04)

	CEM (Central electrons, $\|\eta\| < 1$)	CMUP (Central muons, $\|\eta\| < 0.6$)	CMX (Extension muons, $0.6 < \|\eta\| < 1$)
Lum (pb^{-1})	318.5 $\pm$ 18.8	318.5 $\pm$ 18.8	305.2 $\pm$ 18.0

- Selección de sucesos:

  - 1 leptón (electrón o muón) aislado de alto p_T

  - alta $\not\!\!E_T$

  - $\geq$ 3 jets energéticos

  - $\geq$ 1 (o 2) jet(s) identificados (jets con +JP $<$0.01 (o 0.05))

Jet Multiplicity	1 jet	2 jet	3 jet	$\geq$ 4 jets
Pretag Events				
CEM	16897	2657	182	105
CMUP	8169	1175	83	44
CMX	4273	610	35	17
Total	29339	4442	300	166
Single Tagged Events, JP $<$ 1%				
CEM	207	106	33	36
CMUP	92	58	13	24
CMX	51	27	6	8
Total	350	191	52	68

Medida de la Sección Eficaz de $t\bar{t}$: Acept. y Fondos

- JP tagging efficiencies para sucesos $t\bar{t}$ (muestra de Monte Carlo PYTHIA con $M_{Top} = 178\ GeV/c^2$)

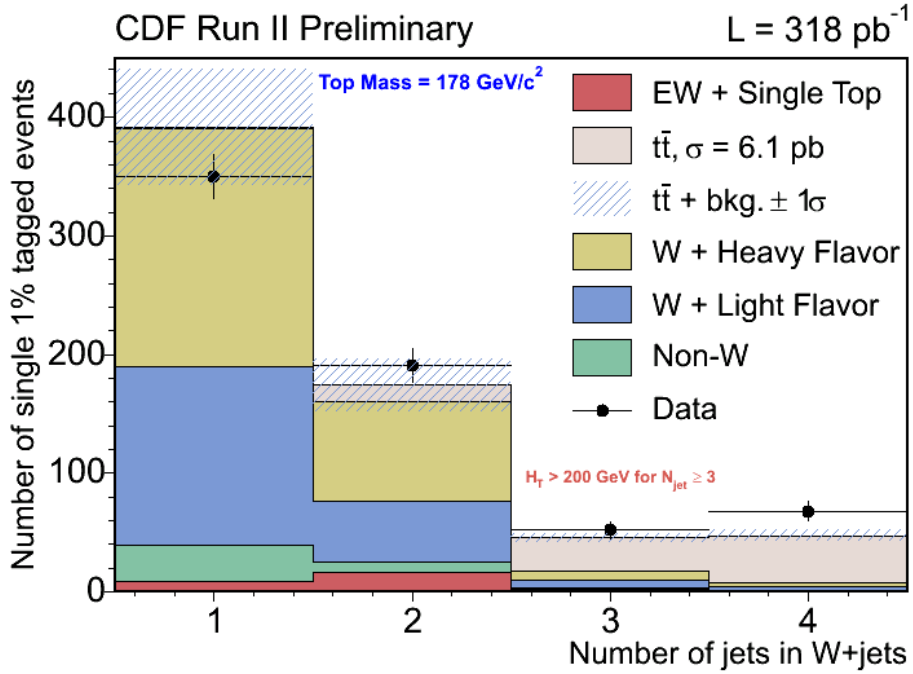
Quantity	CEM	CMUP	CMX
Single tag, JP<1% (SF = 0.817 $\pm$ 0.070)			
Acc. No Tag	3.665 $\pm$ 0.017 $\pm$ 0.279	1.919 $\pm$ 0.011 $\pm$ 0.147	0.751 $\pm$ 0.008 $\pm$ 0.057
Tag Eff.	54.70 $\pm$ 0.24 $\pm$ 3.59	54.05 $\pm$ 0.30 $\pm$ 3.54	55.21 $\pm$ 0.52 $\pm$ 3.57
Average Tag Eff.	54.5 $\pm$ 0.17 $\pm$ 3.57		
Acc. with Tag	2.00 $\pm$ 0.01 $\pm$ 0.20	1.04 $\pm$ 0.01 $\pm$ 0.10	0.41 $\pm$ 0.01 $\pm$ 0.04
$\epsilon_{t\bar{t}} \int$L dt	6.38 $\pm$ 0.04 $\pm$ 0.74	3.30 $\pm$ 0.03 $\pm$ 0.39	1.32 $\pm$ 0.02 $\pm$ 0.15

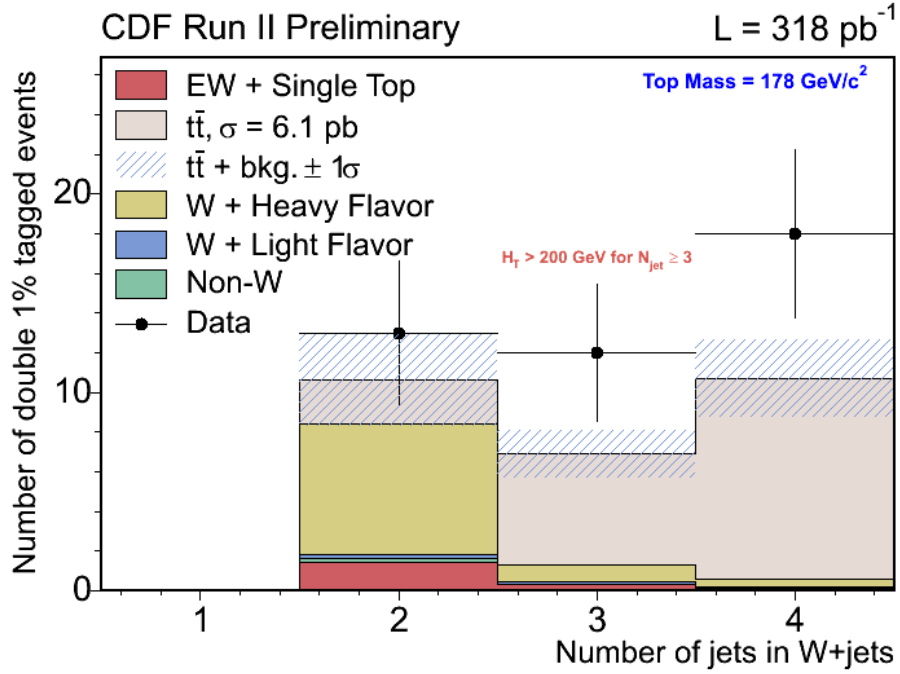
- Estimación de los fondos :
  - Mistags: se predicen utilizando la matriz de tag rate negativa
  - non-W: se derivan a partir de las regiones complementarias de $\not\!\!E_T$ vs lepton isolation
  - W+HF: se estiman utilizando W+HF MC para

  $\rightarrow$ extraer las fracciones de HF a partir de $\frac{W+HF}{W+Jets}$ MC y las eficiencias de b-tag

  $\rightarrow$ y se normalizan a W+jets pretag data
  - Diboson, Z $\rightarrow \tau\tau$ and single top se obtienen a partir de MC

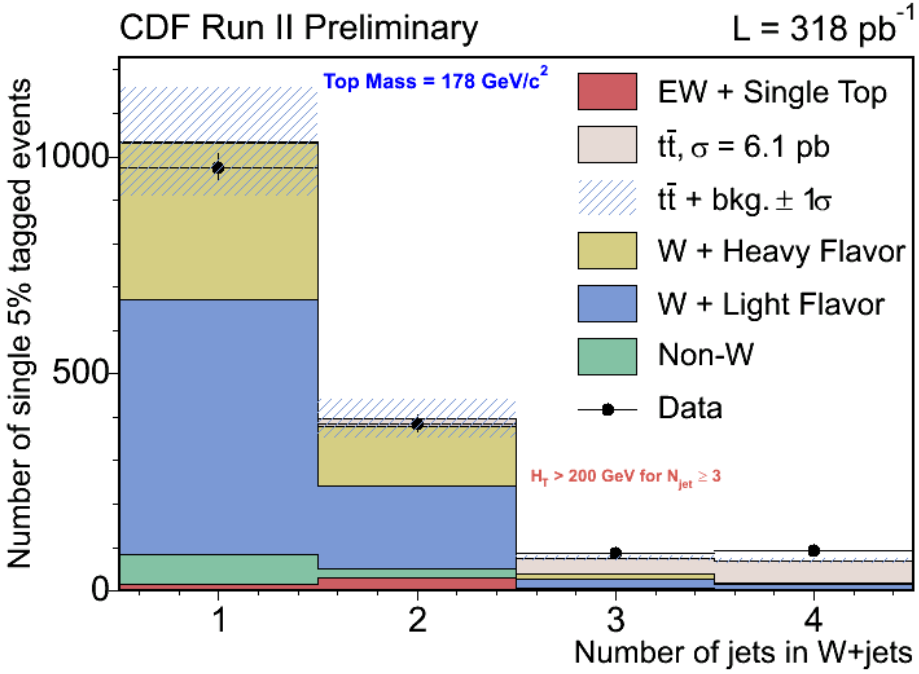
Medida de la Sección Eficaz de $t\bar{t}$, JP $<$ 1%

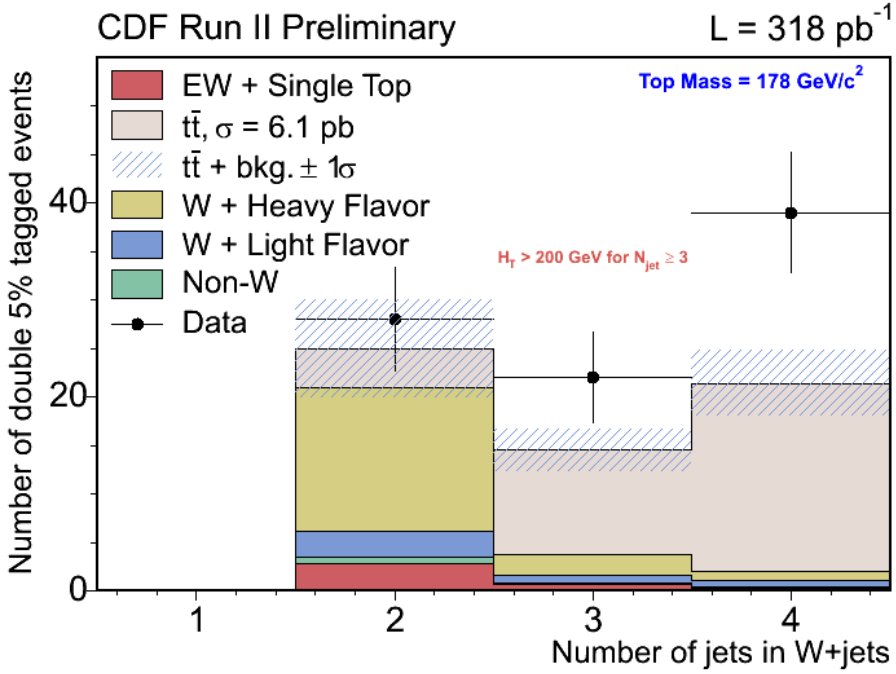




Tight Cut (JP $<$ 1%)	Single Tag	Double Tag
$\sigma_{t\bar{t}}$ (pb)	8.9 $\pm$ 1.0 (stat) $\pm$ 1.1 (syst)	$11.1^{+2.3}_{-1.9}$ (stat) $\pm$ 1.9 (syst)

Medida de la Sección Eficaz de $t\bar{t}$, JP $<$ 5%





Loose Cut (JP $<$ 5%)	Single Tag	Double Tag
$\sigma_{t\bar{t}}$ (pb)	$9.6^{+1.0}_{-0.9}$ (stat) $\pm$ 1.2 (syst)	$11.6^{+1.7}_{-1.5}$ (stat) $\pm$ 1.9 (syst)

Resumen

- Se ha descrito un algoritmo de etiquetado (JP) basado en la información del parámetro de impacto de las trazas y que proporciona una variable continua capaz de discriminar jets con contenido de sabor fuerte.

- Se caracteriza el algoritmo midiendo, a partir de datos, su eficiencia y mistag rates. La eficiencia obtenida para sucesos $t\bar{t}$ (JP $<$ 1%) es del 54.5 $\pm$ 3.57 %.

- Se ha utilizado este algoritmo en la muestra de Lepton+Jets para calcular la sección eficaz de producción $t\bar{t}$

$$\sigma = 8.9 \pm 1.0(stat) \pm 1.1(syst) \text{ pb } (M_{Top} = 178\ GeV/c^2)$$

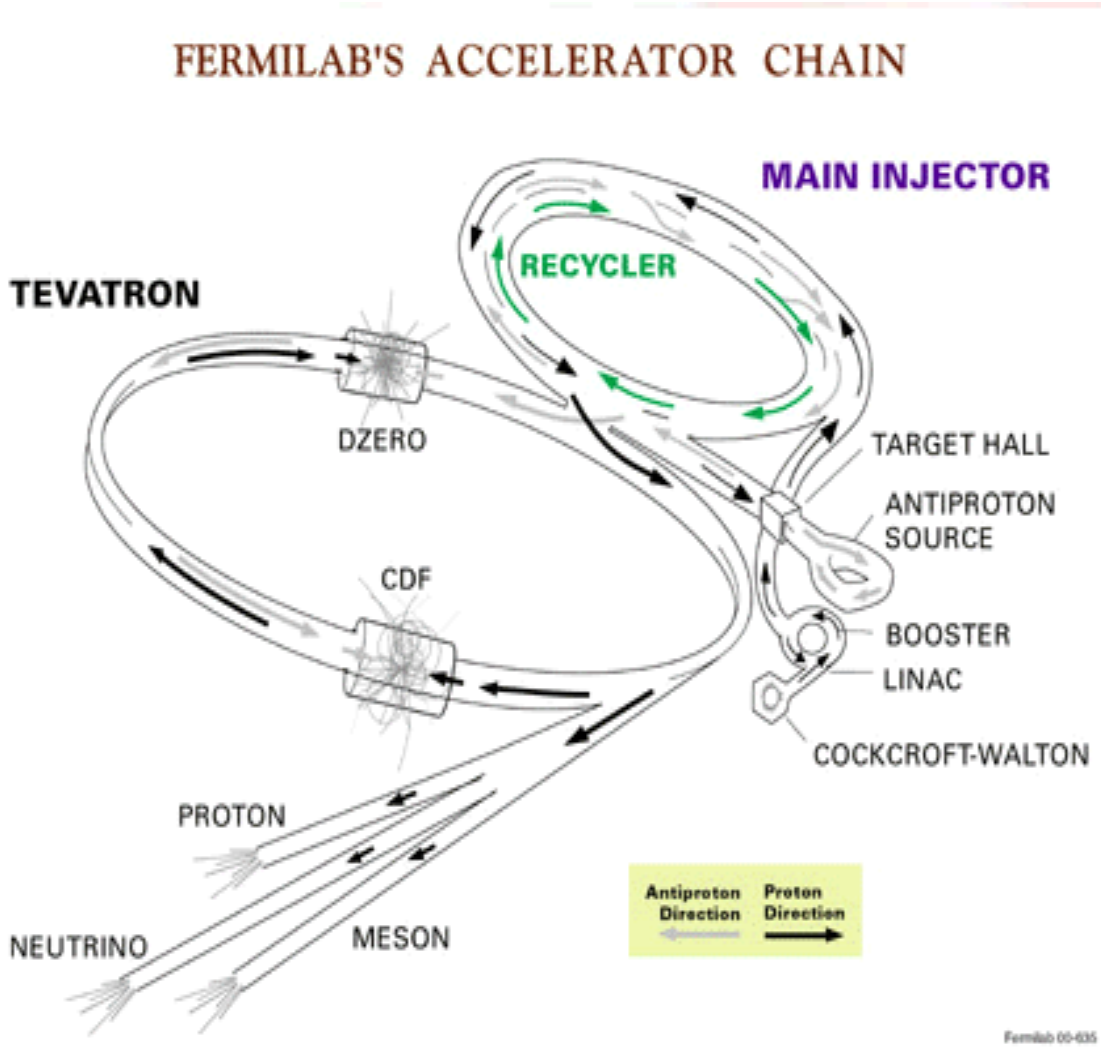
- El valor obtenido es compatible con otras medidas (y con el teórico) con un error total del 17%.

BACK-UP SLIDES

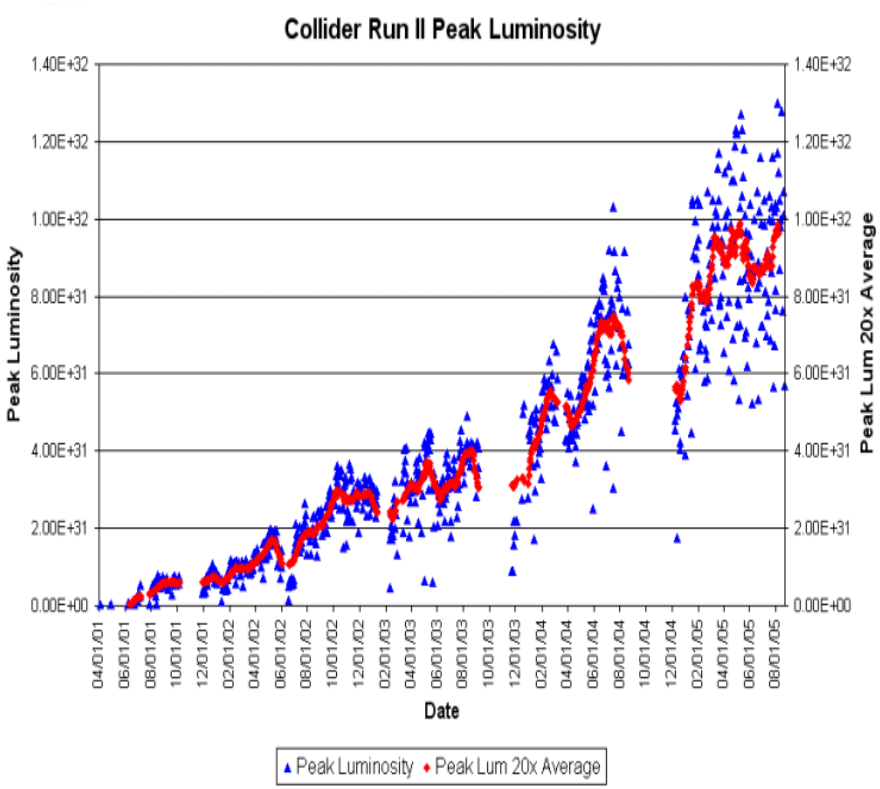
- El Tevatron: Cadena de Aceleradores
- El Tevatron: Luminosidad (Run II)
- CDF
  - Sistema de Reconstrucción de Trazas
  - Detectores de Silicio
  - Cámara de Trazas (COT)
  - Sistema de Calorimetría
  - Sistema de Muones
- El Algoritmo Jet Probability
- Eficiencia de Jet Probability
- Deducción de la Expresión de Jet Probability
- Definición de la Matriz de Tag Rate
- Medida de la Sección Eficaz de $t\bar{t}$: Resumen de Fondos

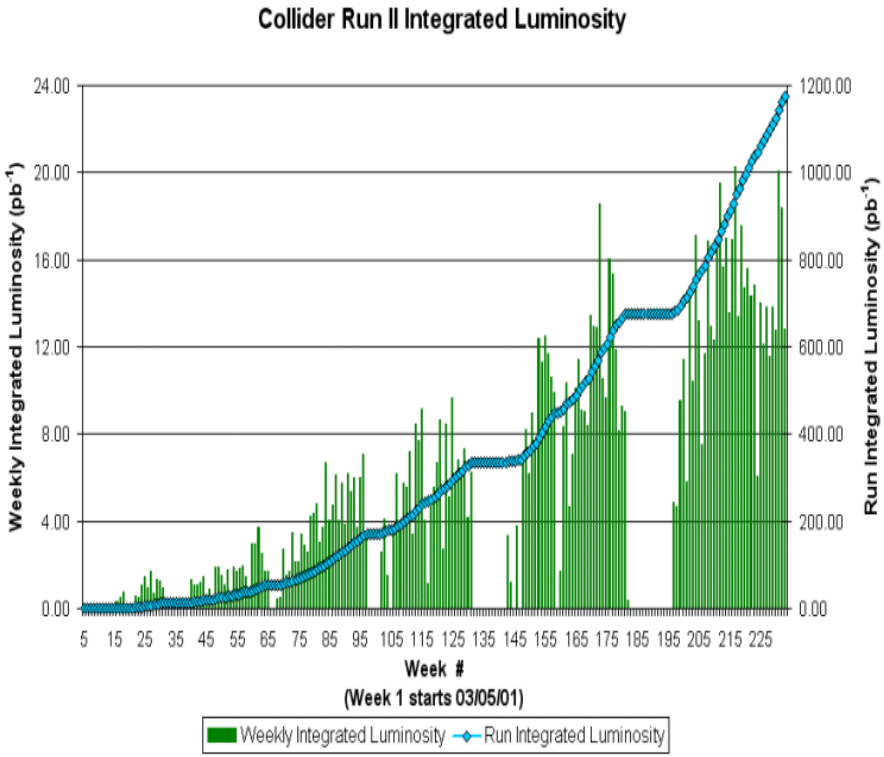
El Tevatron: Cadena de Aceleradores

- Cockcroft-Walton: H gas $\rightarrow$ iones H^-, $150\ KeV$.
- Linac: Acelerador lineal, $150\ m$,
  - $H^- \rightarrow 400\ MeV$.
- Booster: Sincroton circular, R $\sim 37\ m$.
  - Separa los e^- de los protones
  - $8\ GeV$.
- MI: Sincroton circular, R $\sim 475\ m$.
  - 36 paquetes de p ($\bar{p}$) con $3 \cdot 10^{11}\ p$ y $3 \cdot 10^{10}\ \bar{p}$ por paquete
  - $p \rightarrow 120\ GeV$
  - p + Ni (fijo) $\rightarrow \bar{p}$ ($8\ GeV$)
  - $20\ \bar{p}$ por cada $10^6\ p$
  - $p, \bar{p} \rightarrow 150\ GeV$
- Tevatron:
  - $p, \bar{p} \rightarrow 980\ GeV$



El Tevatron: Luminosidad (Run II)

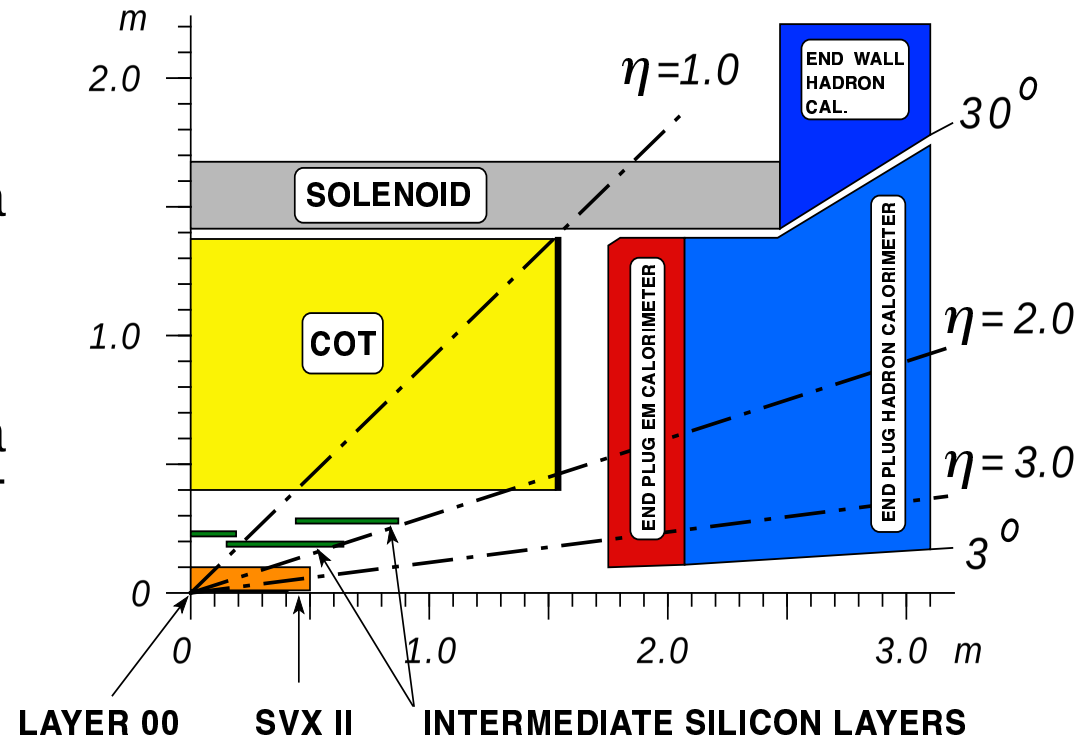




- $\mathcal{L}_{inst}^{max} \sim 1.3 \cdot 10^{32}\ cm^{-2}s^{-1}$, $\mathcal{L}_{int} \sim 1.2\ fb^{-1}$ ($\sim 900\ pb^{-1}$ en cinta)
- $\mathcal{L}_{int} \sim 4.4$ (objetivo básico)- 8.5 (diseño) fb^{-1} en 2009?

CDF: Sistema de Reconstrucción de Trazas

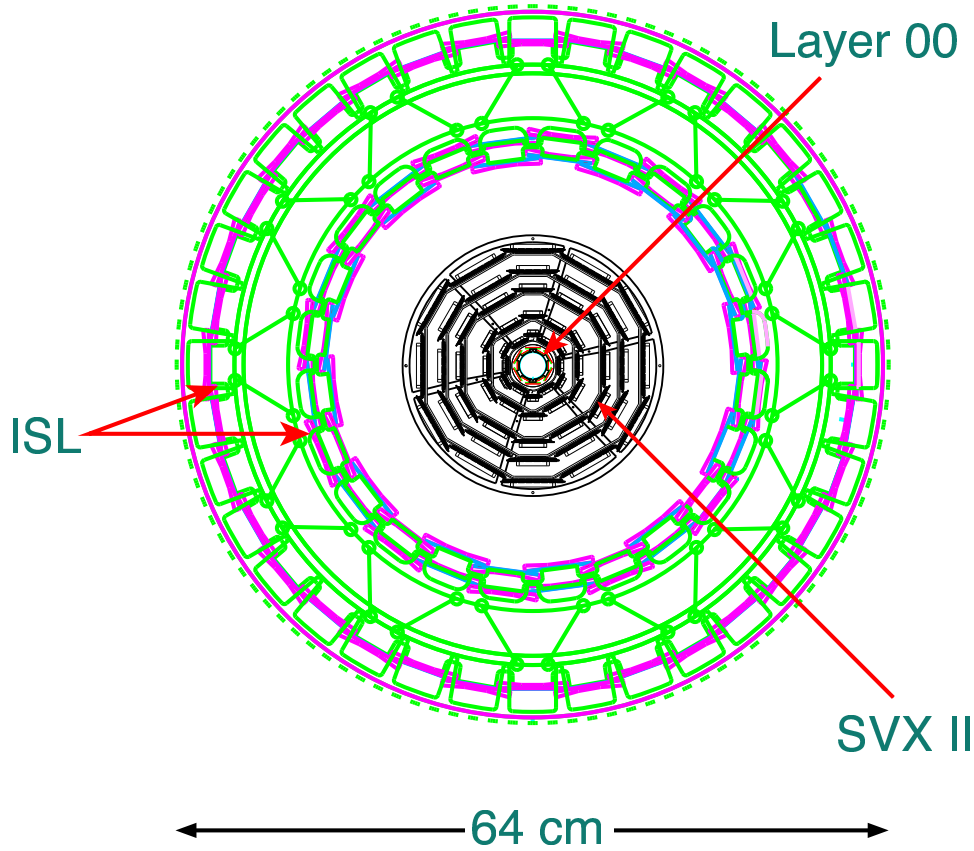
- Información conjunta de la COT + Detectores de silicio
- Algoritmos OI: reconstrucción *de fuera hacia dentro*
- Información del silicio es añadida a la dada por la COT
- 95% de eficiencia en la reconstrucción de trazas COT ($p_T > 400\ MeV$)
- Si OI falla, otros algoritmos (IO,...)



CDF: Detectores de Silicio

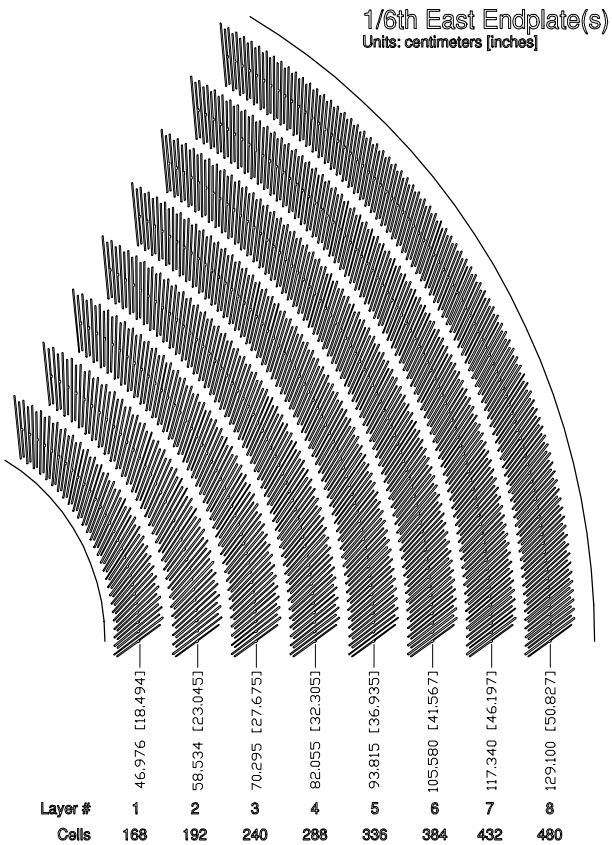
- Uniones p-n
- Medidas muy precisas de la posición y cantidad de carga depositada por las partículas $\longrightarrow$ trayectoria $\longrightarrow$ reconstrucción de vértices

- Capa 00:
  - 1 capa a un R $\sim 1.7\ cm$
  - Resolución $\sim 15\ \mu m$
- SVX II:
  - 5 capas concéntricas con $2.5\ cm < R < 10.6\ cm$
  - Resolución $\sim 20\ \mu m$
- ISL:
  - 1 capa central en R $\sim 22\ cm$
  - 2 *forward* en R $\sim$ 20 y $28\ cm$



CDF: Cámara de Trazas (COT)

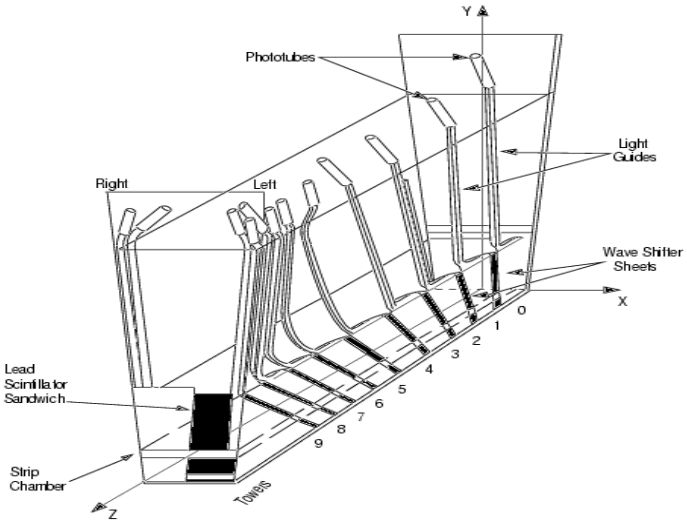
- Cámara de deriva cilíndrica ($L = 3.1\ m$, $R_{min} = 40\ cm$, $R_{max} = 137\ cm$)
- Mezcla de Argon, Etano y CF_4 (50:35:15) $\longrightarrow v_{deriva} \sim cte$
- 96 capas con hilos en 8 supercapas
- Capas axiales y estereo $\longrightarrow$ info 3D
  - muy precisa en $r - \phi$ (p_T)
  - ... pero no tanto en $r - z$ (p_z)
- Resolución $\sim 140\ \mu m$
- $\frac{\sigma_{p_T}}{p_T^2} = 0.15\%$



CDF: Sistema de Calorimetría

- Mide la energía de las partículas que lo atraviesan
- Capas sucesivas de material centelleador y absorbente (Pb, Fe)
- Calorímetros electromagnéticos seguidos por los hadrónicos
- Geometría proyectiva: torres en η y ϕ hacia la región de interacción

  - Cobertura: 2π en ϕ y $|\eta| < 3.6$



- Resolucion CEM $= 1.7\% + \frac{13.5\%}{\sqrt{E_T}}$
- Dos regiones: central ($|\eta| < 1.1$) y *plug* ($|\eta| < 3.6$)
- 48 módulos en forma de cuña de 15° en ϕ con 10 torres de 0.1 en η

CDF: Sistema de Muones

- Muones de alto momento dejan trazas en el sistema de trazas pero depositan muy poca energía en el calorímetro
- Cámaras de deriva y centelleadores en la parte más externa de CDF
- Cuatro sistemas: CMU ($|\eta| < 1.0$), CMP, CMX e IMU
- CMU: inmediatamente después del calorímetro hadrónico. Argon + Etano al 50% con un poco de alcohol
- CMP también cámaras de deriva
- CMX: cámaras de deriva + capas centelleadoras
- IMU: instalado en el Run II, $1.0 > |\eta| > 1.5$

El Algoritmo Jet Probability

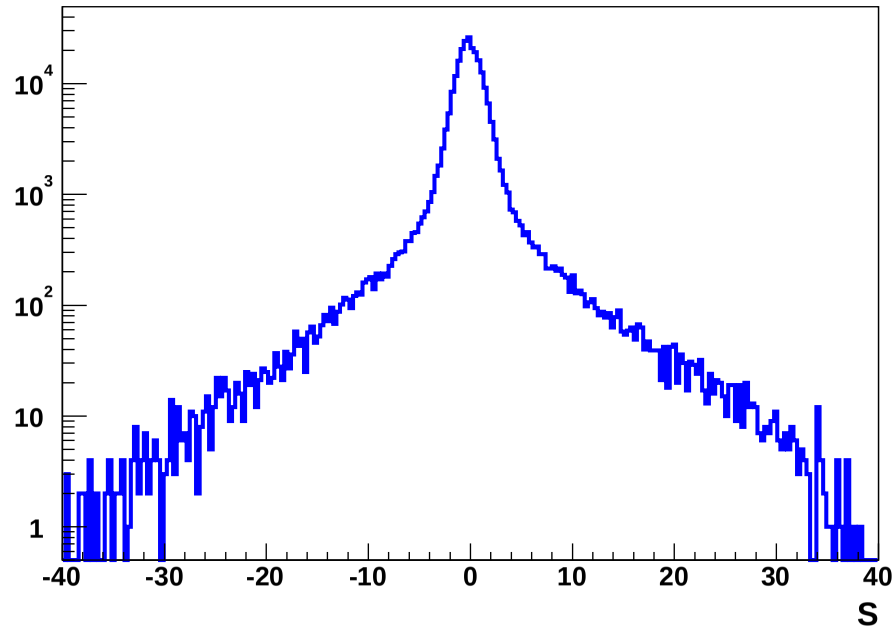
- Significancia del parámetro de impacto de una traza: $S = D/\sigma_D$

- Se ajusta la distribución de la significancia del parámetro de impacto de las trazas para obtener una función de resolución, $R(S)$ (diferente para datos y MC)

- La parte negativa de $R(S)$ se utiliza para determinar la probabilidad ($P_{tr}(S_0)$) de que la significancia del parámetro de impacto (S_0) de una traza dada sea debida únicamente a efectos de resolución del detector

$$P_{tr}(S_0) = \frac{\int_{-\infty}^{-|S_0|} R(S)dS}{\int_{-\infty}^{0} R(S)dS}$$

- La probabilidad de que un jet sea consistente con una hipótesis de vida media nula se define como



$$\prod_{l=1}^{N_{tr}} P_{tr} \times \sum_{k=0}^{N_{tr}-1} \frac{(-ln \prod_{l=1}^{N_{tr}} P_{tr})^k}{k!}$$

Deducción de la Expresión de Jet Probability

- Sea un jet con 2 trazas con parámetro de impacto positivo cuyas probabilidades son P_1 y P_2 y sea $K \equiv P_1 \cdot P_2$

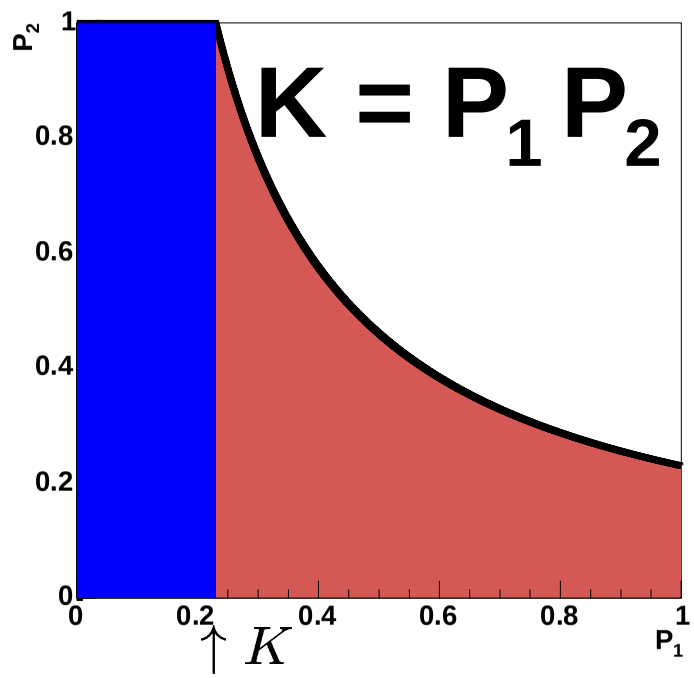
$$0 \leq P_i \leq 1 \text{ i=1,2} \Longrightarrow 0 \leq K \leq 1$$

- El área debajo y a la izquierda de la curva de probabilidad constante K es el conjunto de combinaciones, para las 2 trazas, de tener una probabilidad menor o igual que K. Y este área se define como Jet Probability, P_{jet}

◇ $P_{jet} = A + B, \; A = K \cdot 1$

◇ $B = \int_{x=K}^{x=1} f(x)dx = \int_{x=K}^{x=1} \frac{K}{x} dx = -K LnK$

$\Longrightarrow P_{jet} = K(1 - LnK)$



- En general, se demustra inductivamente que

$$P_{jet} = \prod_{l=1}^{N_{tr}} P_{tr} \times \sum_{r=0}^{N_{tr}-1} \frac{(-ln \prod_{l=1}^{N_{tr}} P_{tr})^r}{r!}$$

Eficiencia de Jet Probability

- Para la medida se utiliza una muestra de datos inclusiva de electrones de $\geq 8\ GeV$ pues está enriquecida con HF debido a los desintegraciones semileptónicas de B

- Método de etiquetado único: $\epsilon = \frac{N_{ej}^{+} - N_{ej}^{-}}{N_{ej}} \cdot \frac{1}{F_B}$

  - Desventaja: confia en la correcta determinación de la fracción de sabor fuerte en la muestra

- El cálculo del contenido de sabor fuerte en el jet (F_B) ha de ser corregido por la contribución de charm (estimada en MC): $F_B = F_b(1 + \lambda_{c/b})$

  - F_b a partir de decaimientos $D^0 \rightarrow K\pi$: $F_b = \frac{N_{D^0}}{N_{ej}} \cdot \frac{1}{\epsilon_{D^0}}$

  - F_b a partir de cascadas de muones: se seleccionan hadrones b con 2 decaimientos semileptónicos (b $\rightarrow$ c $\rightarrow$ X) emitiendo un par e-μ con carga opuesta:

$$F_b = \frac{1}{\epsilon_\mu} \frac{N_{ej}^{\mu}(OS) - N_{ej}^{\mu}(SS)}{N_{ej}}$$

Eficiencia de Jet Probability

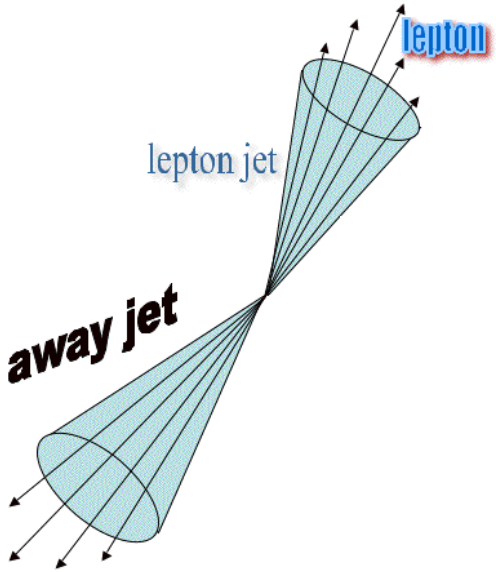
- Método de etiquetado doble: Muestra de sucesos con 2 jets. Como los quarks con sabor fuerte se producen principalmente en pares, el contenido de sabor fuerte en un jet se ve enriquecido pidiendo que el "otro" jet (away jet) sea identificado

$$\epsilon = \frac{(N_{a+}^{e+} - N_{a+}^{e-}) - (N_{a-}^{e+} - N_{a-}^{e-})}{N_{a+} - N_{a-}} \cdot \frac{1}{F_B^a}$$

- estadística limitada en comparación con el método de etiquetado único

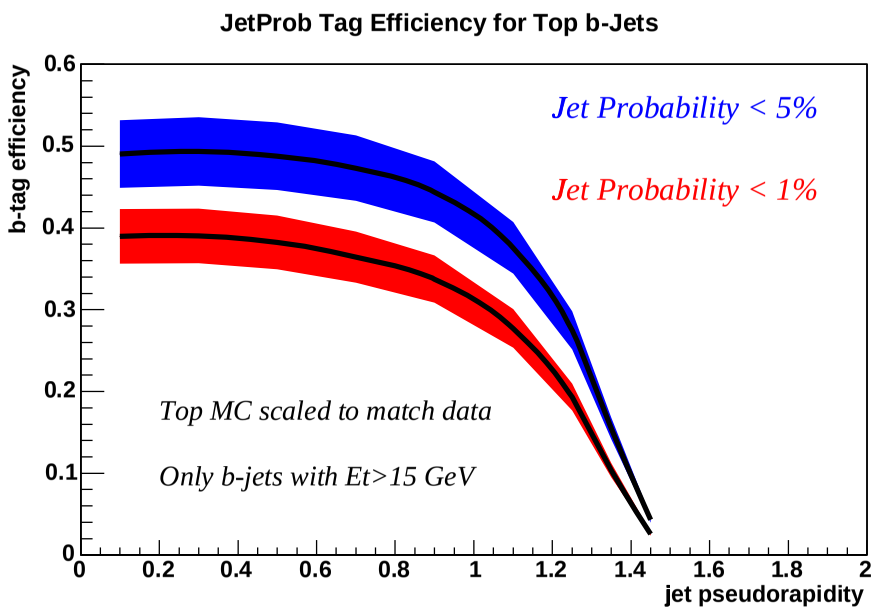
- es necesario corregir por la fracción de sucesos donde el away jet es identificado pero el jet "señal" es LF

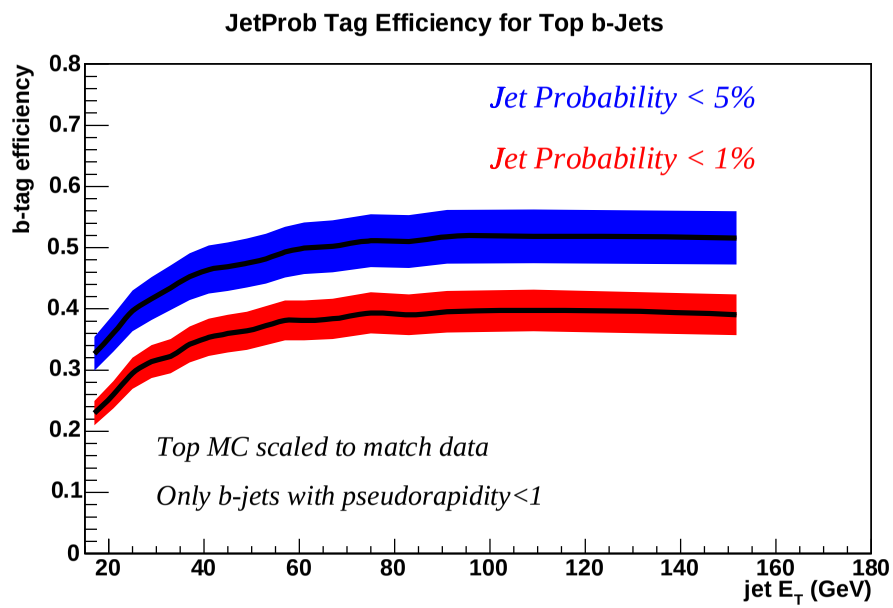
- depende en menor medida de F_B, dando lugar a un error más pequeño ya que F_B no se conoce muy bien



Eficiencia de Jet Probability

- Eficiencia de etiquetar b-jets en una muestra de top Monte Carlo. Las bandas representan el error sistemático en los factores de escala datos/MC





Definición de la Matriz Tag Rate

Bin	E_T (GeV)	Trk. Mult.	$\sum E_T^{\text{jets}}$ (GeV)	$\lvert\eta\rvert$	$\lvert Z_{\text{vtx}}\rvert$ (cm)	ϕ
1	[0,20)	2	[0,80)	[0,1.0)	[0,10)	$[\frac{-\pi}{12}, \frac{\pi}{12})$
2	[20,35)	3	[80,140)	≥ 1.0	[10,20)	$[\frac{\pi}{12}, \frac{3\pi}{12})$
3	[35,50)	4,5	[140,220)		[20,40)	$[\frac{3\pi}{12}, \frac{5\pi}{12})$
4	[50,65)	6,7	≥ 220		[40,50)	$[\frac{5\pi}{12}, \frac{7\pi}{12})$
5	[65,80)	8,9			[50,60)	$[\frac{7\pi}{12}, \frac{9\pi}{12})$
6	[80,100)	10-13			≥ 60	$[\frac{9\pi}{12}, \frac{11\pi}{12})$
7	[100,120)	≥ 14				$[\frac{11\pi}{12}, \frac{13\pi}{12})$
8	[120,150)					$[\frac{13\pi}{12}, \frac{15\pi}{12})$
9	[150,180)					$[\frac{15\pi}{12}, \frac{17\pi}{12})$
10	≥ 180					$[\frac{17\pi}{12}, \frac{19\pi}{12})$
11						$[\frac{19\pi}{12}, \frac{21\pi}{12})$
12						$[\frac{21\pi}{12}, \frac{23\pi}{12})$

Medida de la Sección Eficaz de $t\bar{t}$: Resumen de Fondos

Jet Multiplicity	1 jet	2 jet	3 jet	$\geq$ 4 jets
MC Derived Backgrounds				
WW	2.2 ± 0.3	5.0 ± 0.6	0.7 ± 0.1	0.28 ± 0.06
WZ	1.0 ± 0.1	2.0 ± 0.2	0.23 ± 0.04	0.09 ± 0.02
ZZ	0.027 ± 0.006	0.09 ± 0.01	0.012 ± 0.004	0.007 ± 0.002
Single top $W - g$	4.1 ± 0.5	4.9 ± 0.6	0.7 ± 0.09	0.20 ± 0.03
Single top W^*	1.3 ± 0.2	4.2 ± 0.5	0.60 ± 0.07	0.14 ± 0.02
$Z \to \tau^- \tau^+$	0.74 ± 0.3	0.4 ± 0.2	0.04 ± 0.04	0.0 ± 0.0
Total	9.3 ± 1.1	16.6 ± 1.9	2.3 ± 0.3	0.71 ± 0.09
W + Heavy Flavour				
W$b\bar{b}$	83.0 ± 23.4	47.2 ± 13.2	6.0 ± 1.6	3.3 ± 0.9
W$c\bar{c}$	31.2 ± 9.3	17.5 ± 5.2	2.3 ± 0.7	1.2 ± 0.4
Wc	86.4 ± 21.4	19.2 ± 5.0	1.4 ± 0.4	0.6 ± 0.2
Total	200.6 ± 42.2	83.8 ± 19.8	9.6 ± 2.4	5.2 ± 1.4
Others				
Mistag	149.7 ± 17.9	52.1 ± 6.2	8.6 ± 1.0	6.8 ± 0.8
non W	30.5 ± 15.6	8.6 ± 4.6	0.9 ± 0.6	0.5 ± 0.5
Total Background	390.2 ± 49.1	161.1 ± 22.0	21.5 ± 2.8	13.2 ± 1.7
$t\bar{t}$ (6.1 pb)	1.7 ± 0.4	14.0 ± 1.8	27.5 ± 3.3	39.6 ± 4.6
Data	350	191	52	68