CMS Analysis Status: Inclusive $b \to J/\psi X$, $J/\psi \to \mu\ \mu$

Xiangwei Meng

Institute of High Energy Physics, CAS, Beijing

June 19, 2009

Outline

- Motivations
- Data samples
- Analysis strategy
- Acceptance & efficiency
- Pdl's PDF correction
- J/Psi production X-section
- To do list

Motivations

- First measurement of b hadrons production x-section at sqrt(s) = 10 and 14 TeV

  --other than inclusive & prompt Jpsi's; the same strategy as CDFII's on inclusive b measurement.

  PHYSICAL REVIEW D 71, 032001 (2005)
- Test of QCD calculations
- Essential measurement of background for many other processes, validation of b-tagging

- The meanwhile study on inclusive b lifetime
  - calibrate resolution function
  - understand calibration & alignment

  CDF note 6023

pp@10TeV

CMS Summer08 M.C. data samples

b2J/psi

1,413,547

- **/BtoJpsiMuMu/Summer08_IDEAL_V9_PAT_v1/USER** **CMSSW_2_2_1**
  Created 22 Jan 2009, 4817138 events, 137 files, 1 block(s), 139.7GB, located at 1 site
- **/BtoJpsiMuMu/Summer08_IDEAL_V11_redigi_v1/GEN-SIM-RECO**
  Created 09 Jan 2009, 2453008 events, 1536 files, 16 block(s), 635.7GB, located at 4 sites

Sum08 Incl b: ~100 pb-1

p-J/psi

1,941,162

- **/JPsi/Summer08_IDEAL_V11_redigi_v1/GEN-SIM-RECO** **CMSSW_2_2_1**
  Created 18 Feb 2009, 1941162 events, 382 files, 2 block(s), 394.1GB, located at 4 site

Sum08 pJpsi: ~16 pb-1

QCD

5,132,579

- **InclusivePPmuX/Summer08_IDEAL_V9_PAT_v1/USER** **CMSSW_2_2_1**
  Created 22 Jan 2009, 10345428 events, 312 files, 1 block(s), 304.5GB, located at 3 sites
- **/InclusivePPmuX/Summer08_IDEAL_V11_redigi_v1/GEN-SIM-RECO**
  Created 18 Dec 2008, 5309035 events, 1702 files, 18 block(s), 1.5TB, 11 sites

Sum08 QCD: ~0.044 pb-1

Early Data with 10 pb-1 analysis

events normalized or scaled to:

- Inclusive b: 253,600(SW227)
- P-J/Psi: 1,150,000(SW227)
- Inclusive PPmuX: 5,132,579(SW227)

Scale: 231.5->1,188,000,000

Monte Carlo Sample: signal

- Inclusive $b \rightarrow J/\Psi X$, $J/\Psi \rightarrow \mu\mu$ analyzed with CMMSW_2_2_7
- EvtGen with inclusive $b \rightarrow J/\Psi X$ $J/\Psi \rightarrow \mu\mu$
  Filter on 2 μ with pT>2.5 GeV/c, $|\eta|<2.5$
- σ gen_tot = 51.56 mb (@10TeV),
  $Br(b \rightarrow J/\Psi X) = 0.0116$
  $Br(J/\Psi \rightarrow \mu\mu)$ 0.0593
  filterEfficiency $\varepsilon = 0.0007139$
- σ eff = σ gen_tot* ε filter*Br
  = 25.36nb

Event Reconstruction & selection

- Global mu: pT> 3 GeV/c, $|\eta|<2.4$
- KalmanVertexFitter on $\mu^+\mu^-$ pairs, i.e. 2mus sharing the same vertex
- Jpsi Mass window: [2.8, 3.4] GeV/c^2

Further on pdl:

- Pdl: [-0.1,0.5]cm, error_Pdl<0.1cm

Measurement of b production Cross-secton

The inclusive b differential cross-section is calculated as

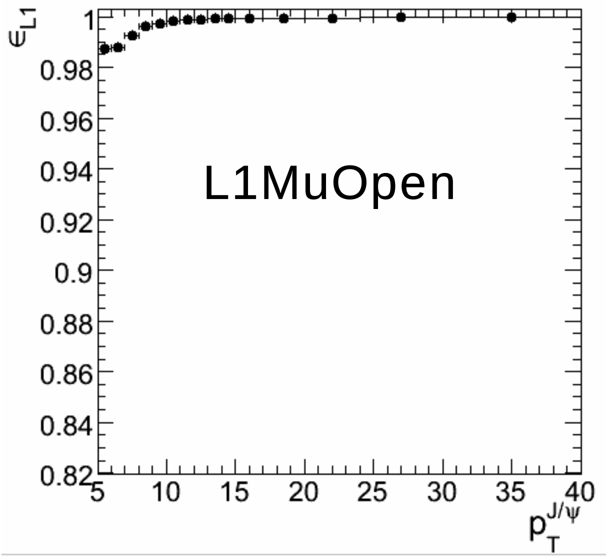
$$\frac{d\sigma}{dp_T^b} \times Br(b \to J/\psi) \times Br(J/\psi \to \mu^+\mu^-)$$

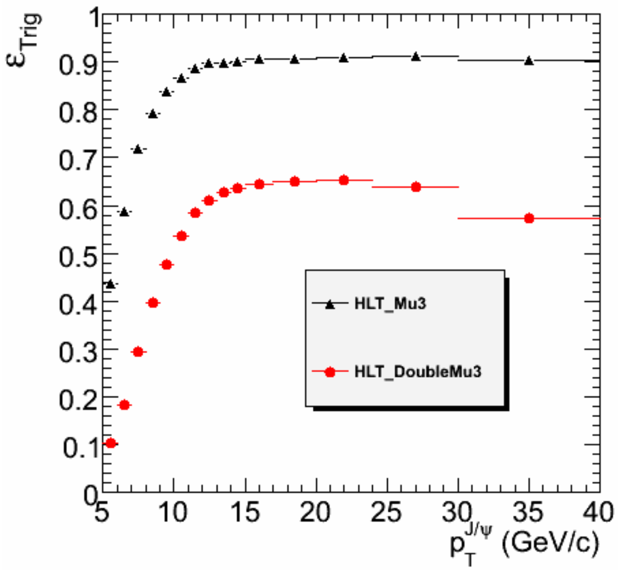
$$= f_b(\Delta p_T^{J/\psi}) \times \frac{N_{\Delta p_T}^{J/\psi}(A, \varepsilon_{Trig}, \varepsilon_{Reco}, \lambda_{Trig}^{corr}, \lambda_{Reco}^{corr})}{\int L dt \cdot \Delta p_T^{J/\psi}}$$

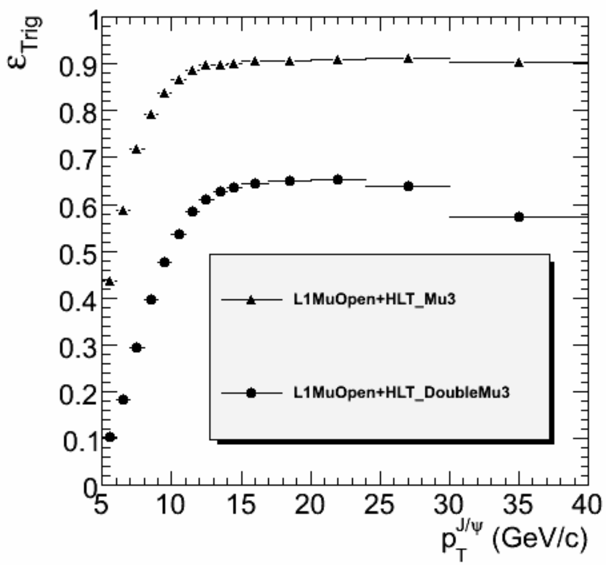
Unfolding: pT Jpsi-> b

- $\int L$ **dt : the integral luminosity**
- **f_b : fraction for J/ Ψ from b**
- **ΔP_T : the size of the reconstructed J/ψ signals**
- $N_{\Delta p_T}^{J/\psi}$ **A: geometric and kinematic P_T bin.**
- **: the number of acceptance of J/ ψ candidates**
- **ε_{Trig}, ,ε_{reco} : trigger and recon efficiency from M.C.**
- $\lambda_{Trig}^{corr}, \lambda_{Reco}^{corr}$ **: correction on trigger and recon efficiency from Tag&probe**

Trig. Efficiency vs. pT : JPsi's





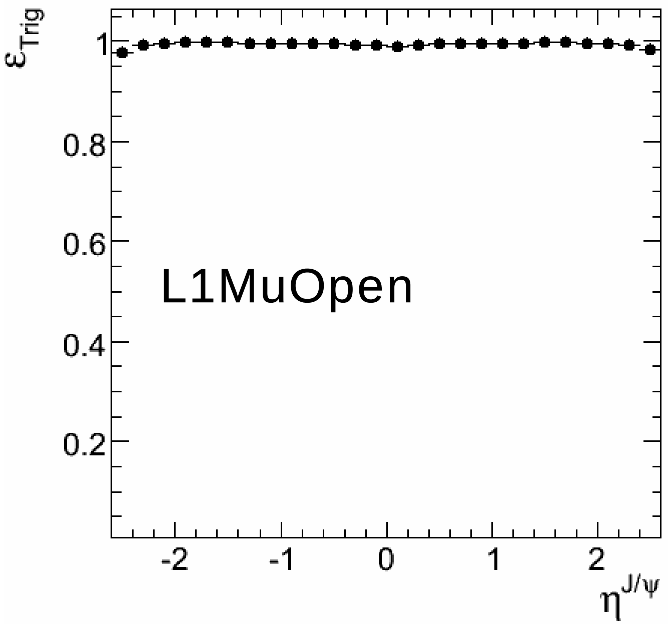


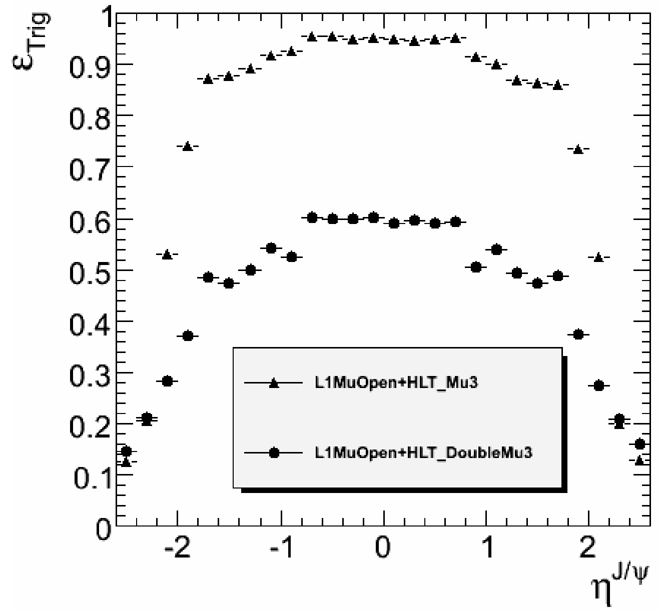
loose L1 mu trigger selection

$$\varepsilon_{L1,L1+HLT}^{w.r.t.Offline} = \frac{N_{L1,L1+HLT}^{Offline+Glb}(p_T^{J/\psi}, |\eta^{J/\psi}| < 2.4)}{N^{Offline+Glb}(p_T^{J/\psi}, |\eta^{J/\psi}| < 2.4)}$$

•The L1T/HLT efficiency is calculated based on the numbers of an offline reconstructed J/psi passes or not corresponding triggers

Trig. Efficiency vs. η : JPsi's



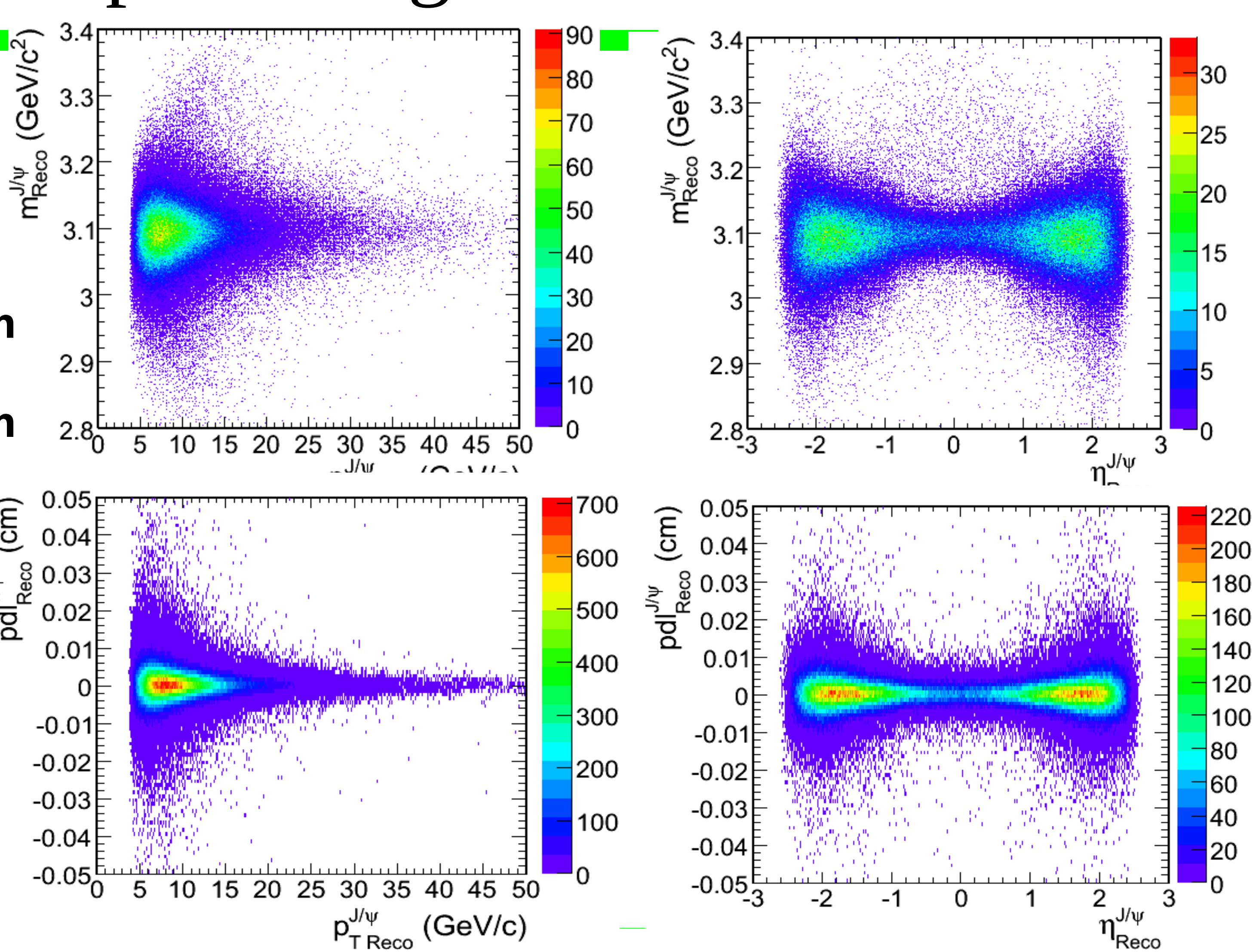


$$\varepsilon^{w.r.t.Offline}_{L1,L1+HLT} = \frac{N^{Offline+Glb}_{L1,L1+HLT}(p_T^{J/\psi} > 5, \eta^{J/\psi})}{N^{Offline+Glb}(p_T^{J/\psi} > 5, \eta^{J/\psi})}$$

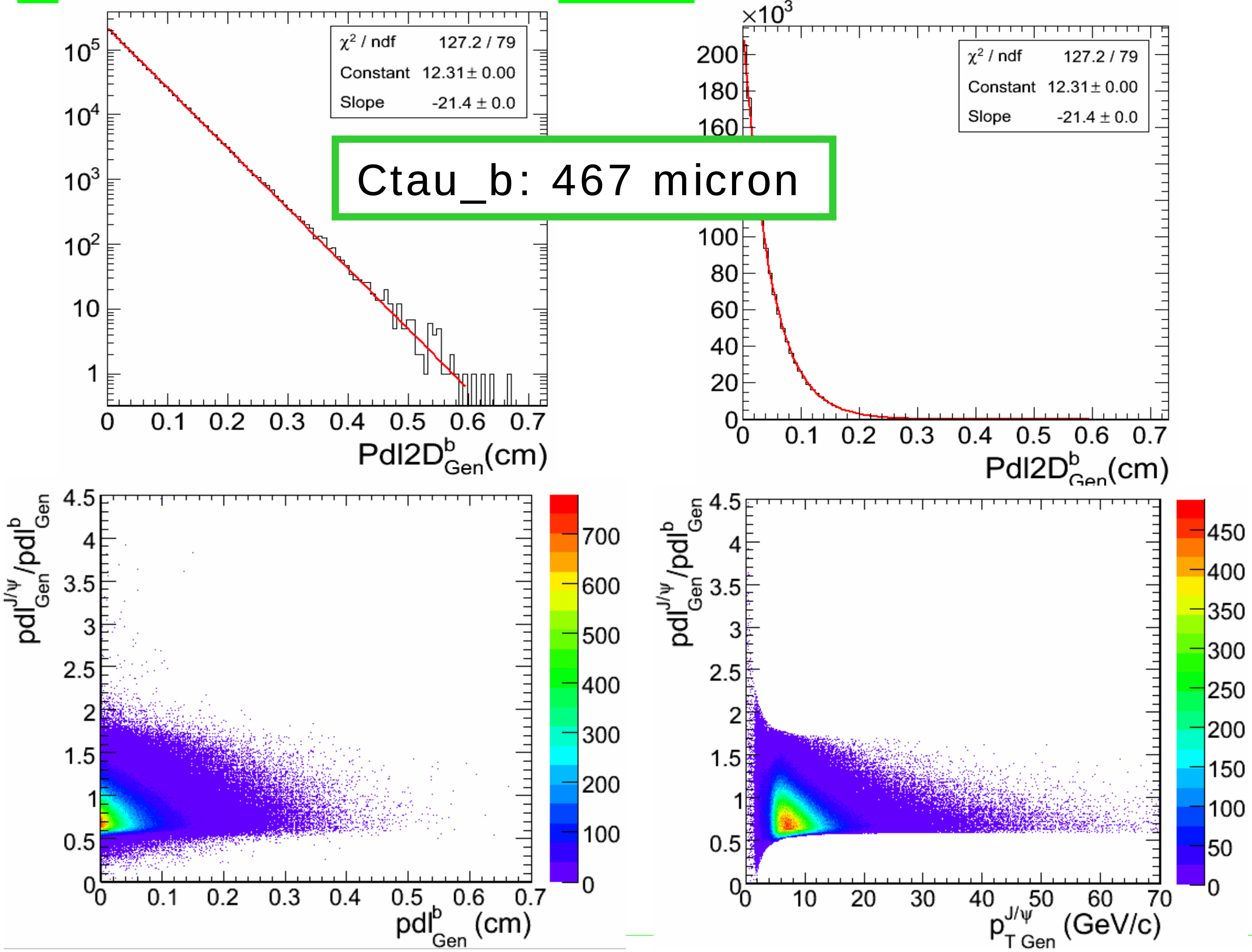
- The L1T/HLT efficiency is calculated based on the numbers of an offline reconstructed J/psi passes or not corresponding triggers

Mass & pdl fitting: resol. Model Sum08

Mass:
3Gaussion
Pdl:
2Gaussion

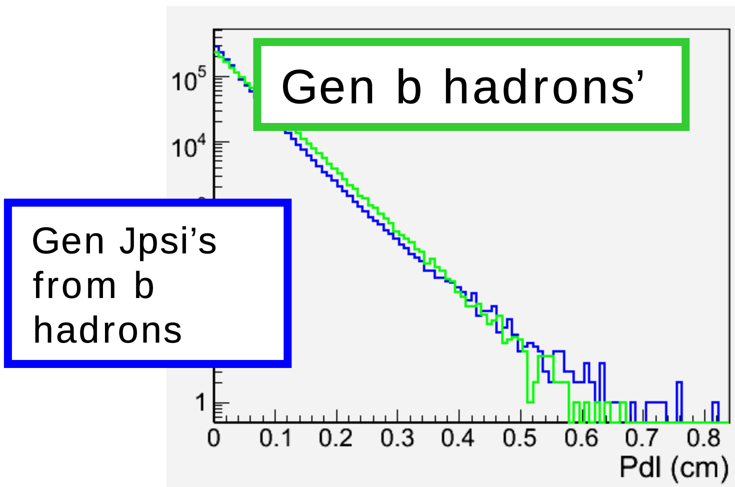


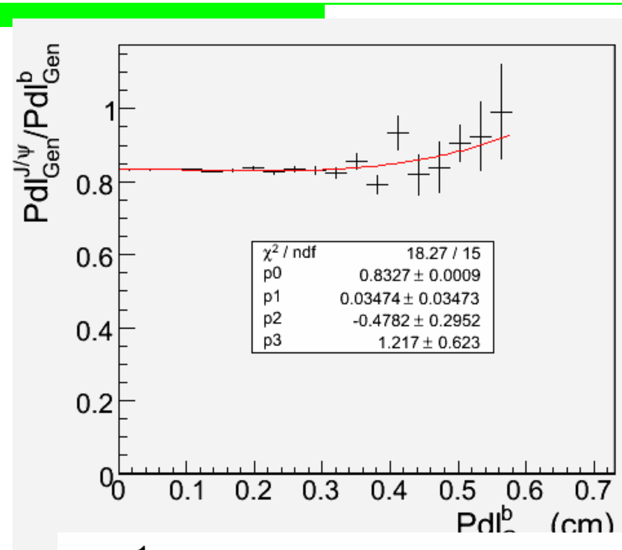
b ctau at Gen: pdl's PDF correction (1)

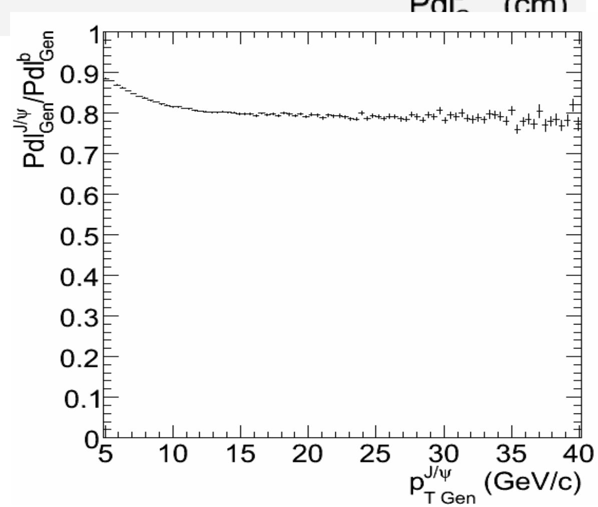


Ctau_b: 467 micron

pdl PDF correction(2): Jpsi's -> b's

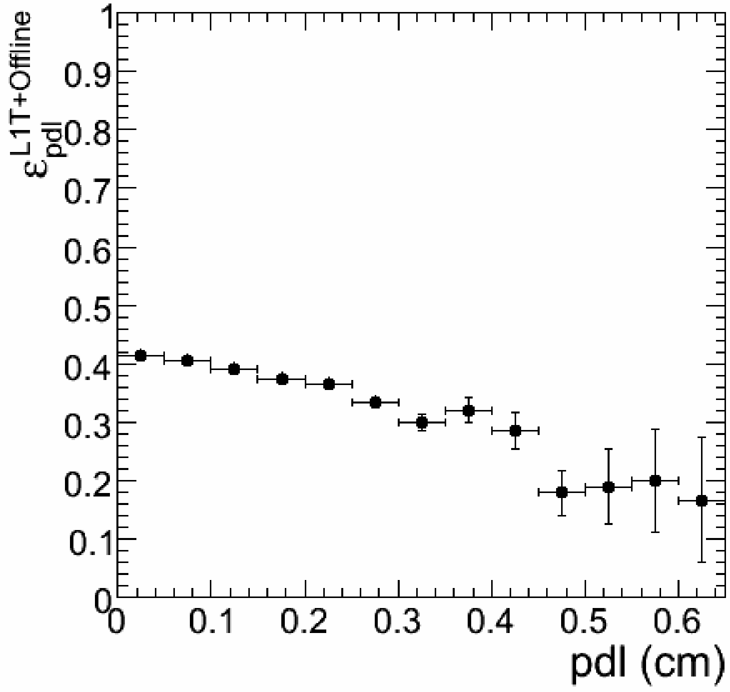


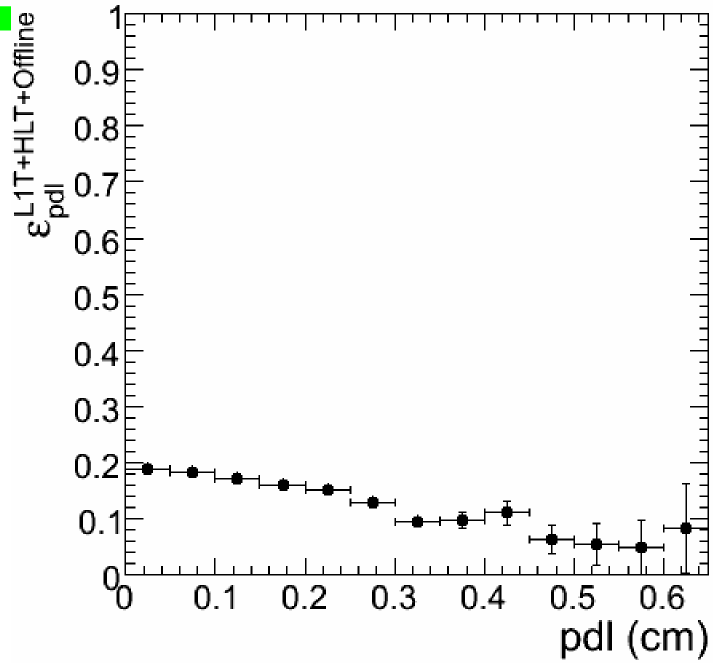




- Used for Pdl's PDF for ctau fitting

Efficiency: Jpsi's Pdl

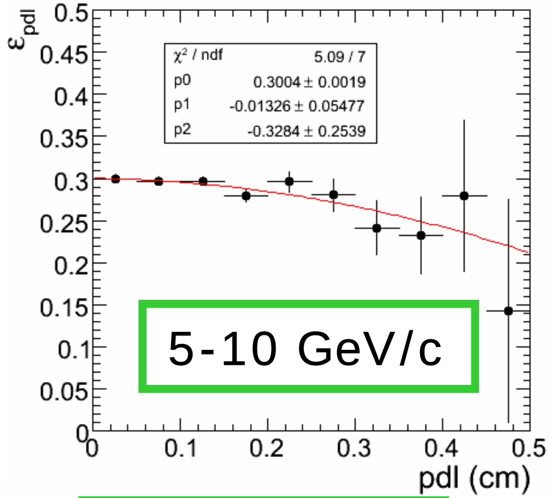


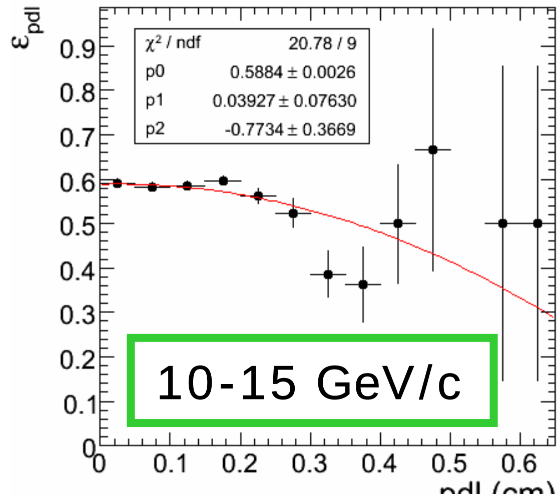


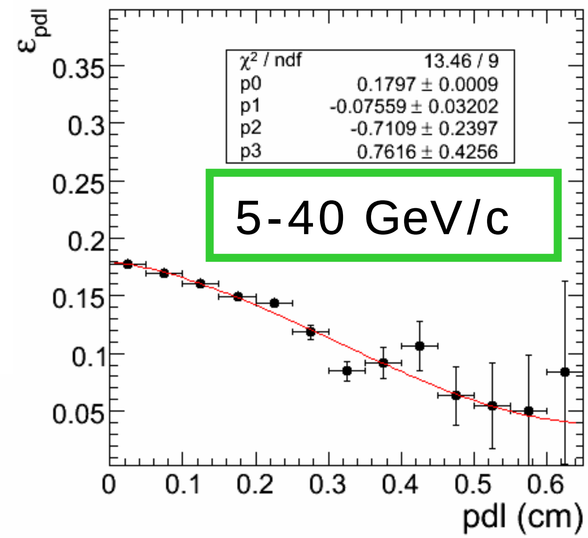
$$\varepsilon_{Pdl} = \frac{N_{Gen}^{Matching\_L1+HLT+Offline}(Pdl^{J/\psi}, p_T^{J/\psi}, |\eta^{J/\psi}| < 2.4)}{N_{Gen}(Pdl^{J/\psi}, p_T^{J/\psi}, |\eta^{J/\psi}| < 2.4)}$$

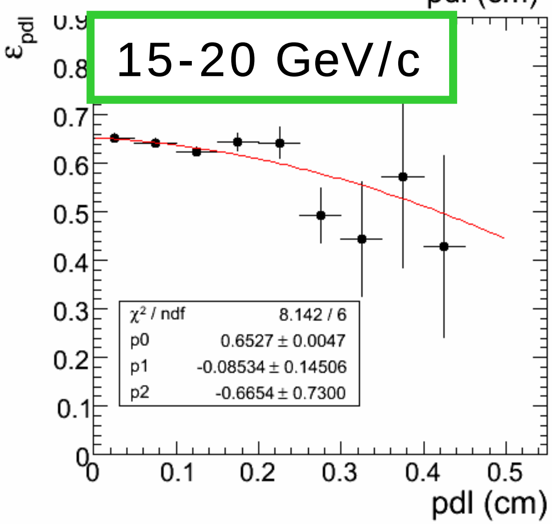
•This efficiency is calculated based on the numbers of J/psis pass L1T+HLT selection and offline reconstruction, and with Truth matching to the generator level and the total number at that level for the distribution.

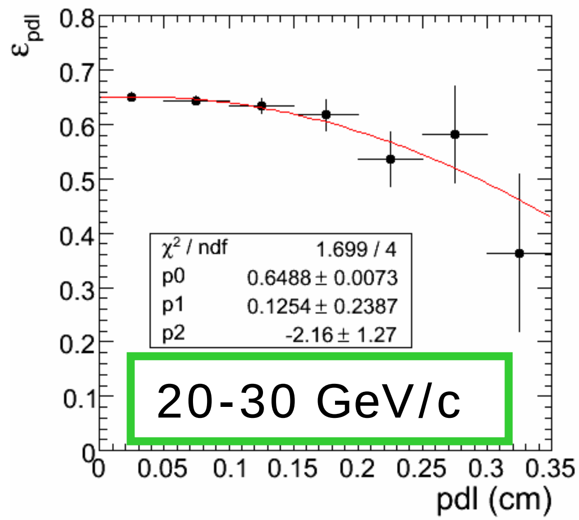
pdl PDF correction(3): efficiency vs. Jpsi pT bins











Correction shape nearly uniform for Pdl's PDF for the ctau fitting

- Used for Pdl's PDF for ctau fitting

UCMLH fit & analysis (mass, Pdl)

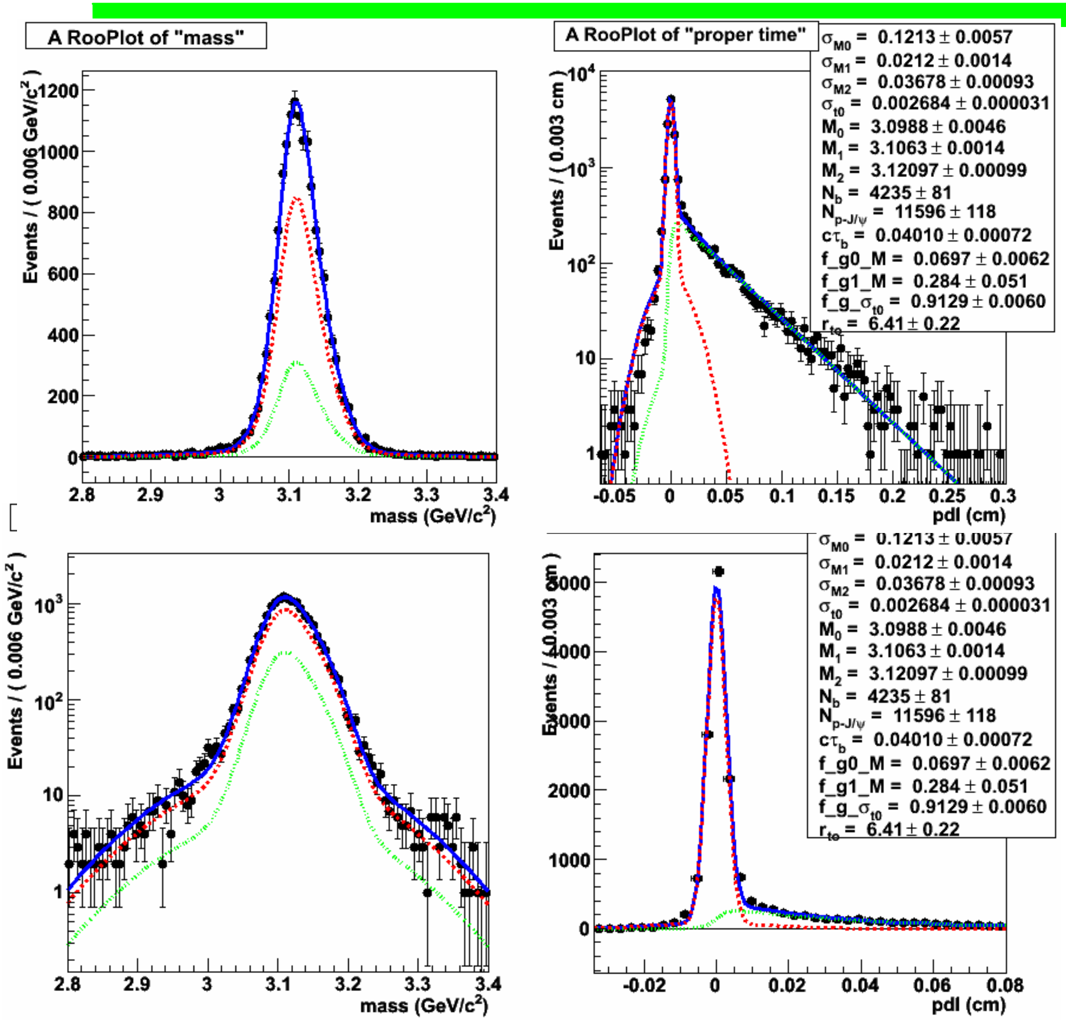
Unbinned combined MLH fit & analysis method: 10pb-1 "data" PT: 5-40 GeV/c

1) Jpsi: prompt, b hadrons, and inclusive production
2) abstraction for b fraction

Fit techniques:

Offline p-jpsi: 11545

b: 4288



a finite experimental resolution on each measurement

$$F_I(\boldsymbol{t}) = \exp(-\boldsymbol{t}/\tau)$$

$$F_R(\boldsymbol{t}) = \exp(-t/\tau) \otimes G(t, \mu, \sigma)$$
$$\equiv \varepsilon(\boldsymbol{t}) \cdot \int dt' \exp(-t'/\tau)\, G(t-t', \mu, \sigma)$$

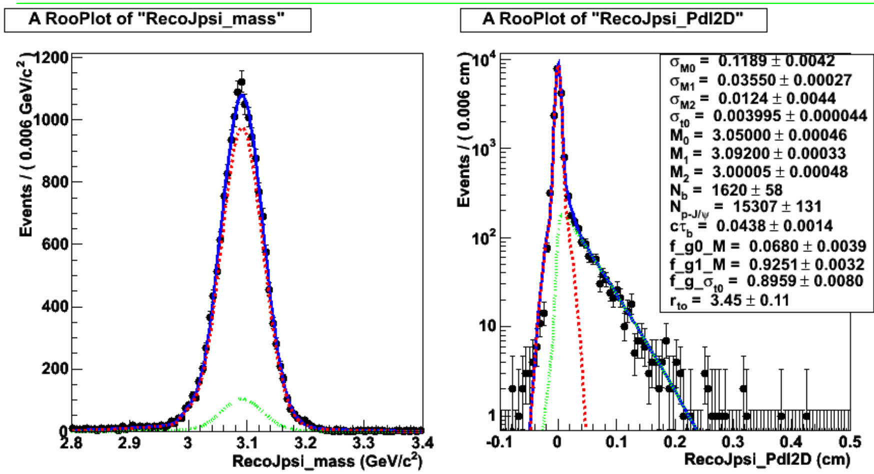
$$F_E(\boldsymbol{t}, \boldsymbol{dt})$$
$$= \varepsilon(\boldsymbol{t}) \cdot \exp(-\boldsymbol{t}/\tau) \otimes G(t, \mu, \boldsymbol{dt})$$

$$F_E(\boldsymbol{t}, \boldsymbol{dt})$$
$$= \varepsilon(\boldsymbol{t}) \cdot \exp(-\boldsymbol{t}/\tau) \otimes G(\boldsymbol{t}, \mu, \boldsymbol{s} \cdot \boldsymbol{dt})$$

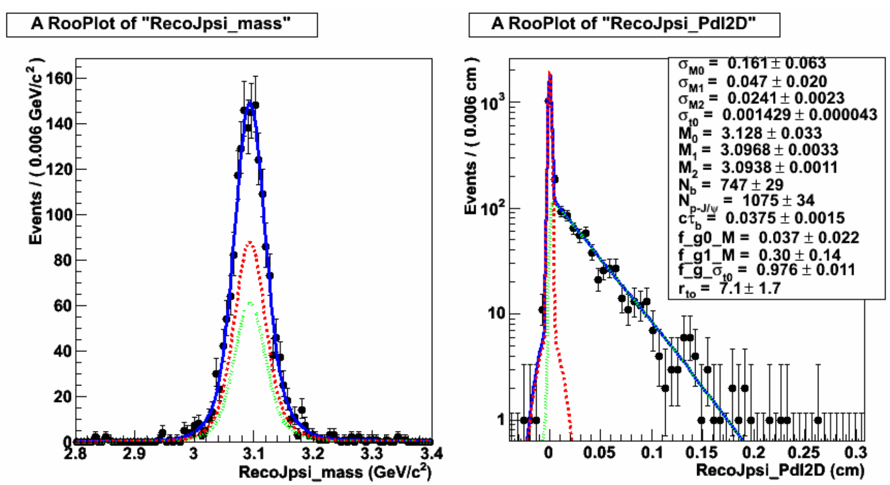
(***t* / *dt***) fitting with new version RooFit package

pT: 9-10GeV

Jpsi Pt bins: 5-6, 30-40GeV/c

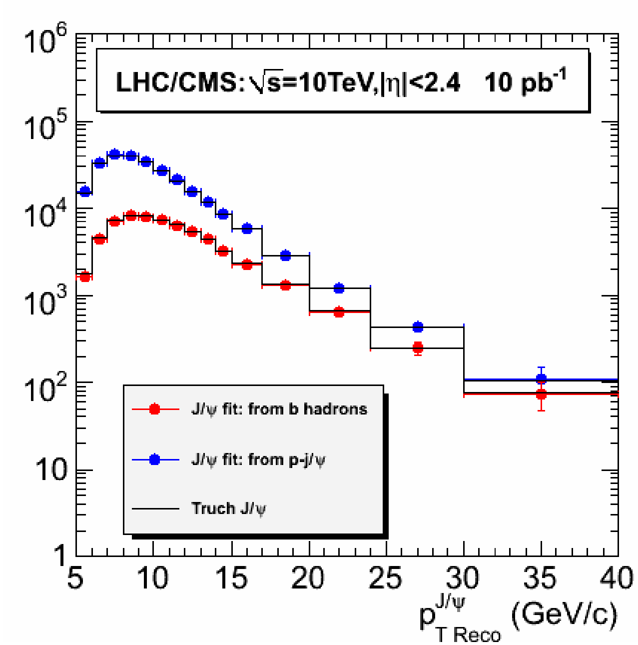


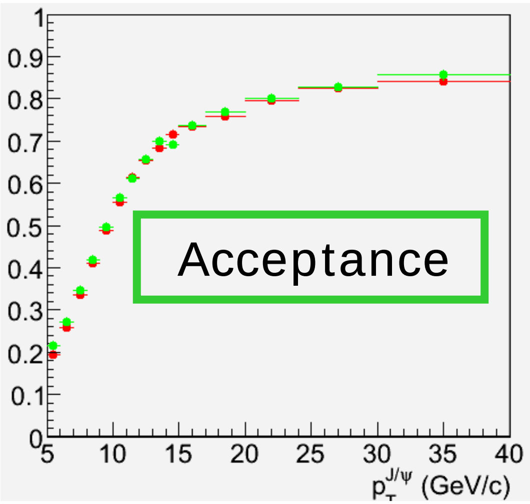
p-jpsi: 15126
b: 1801

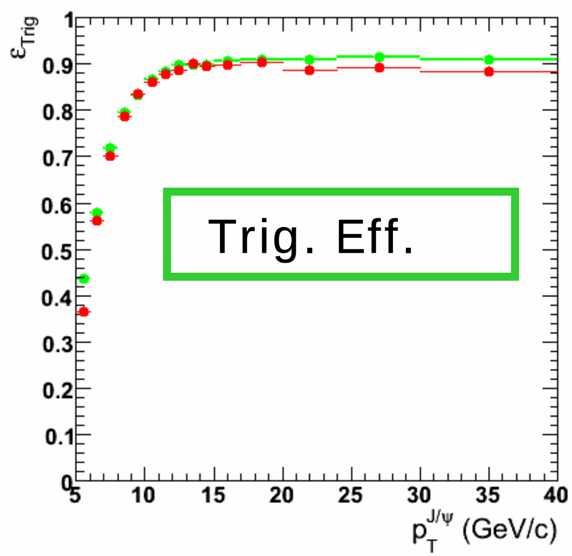


p-jpsi: 1060
b: 762

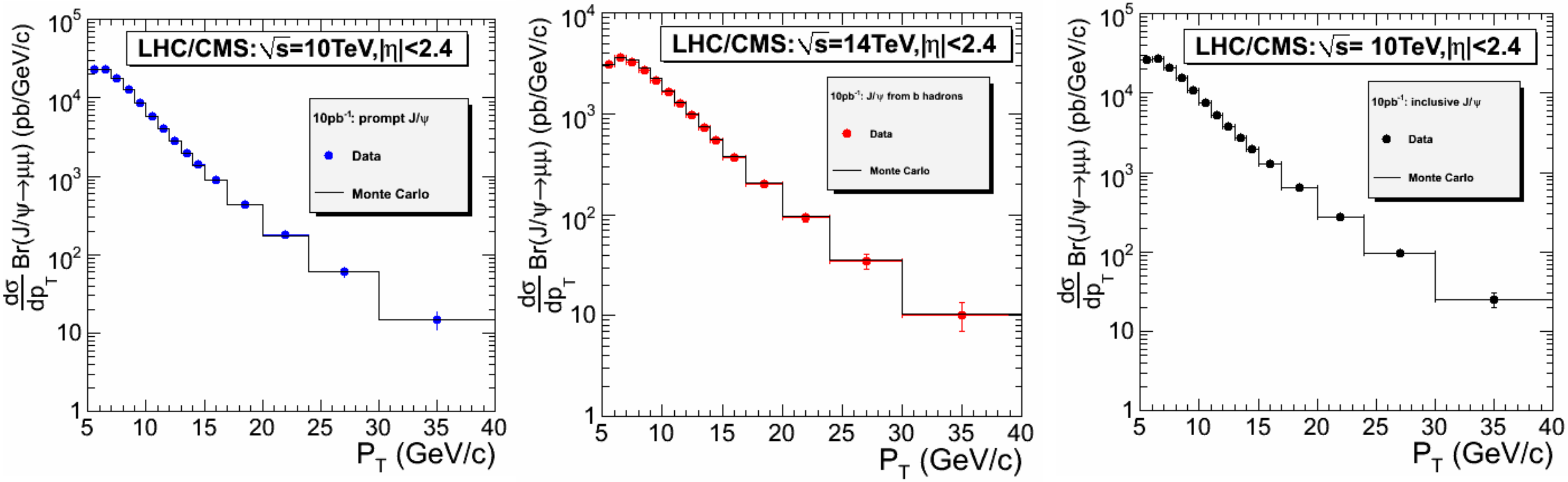
Jpsi: Acceptance, Trig eff. & fraction







J/psi production x-section@pp 10TeV



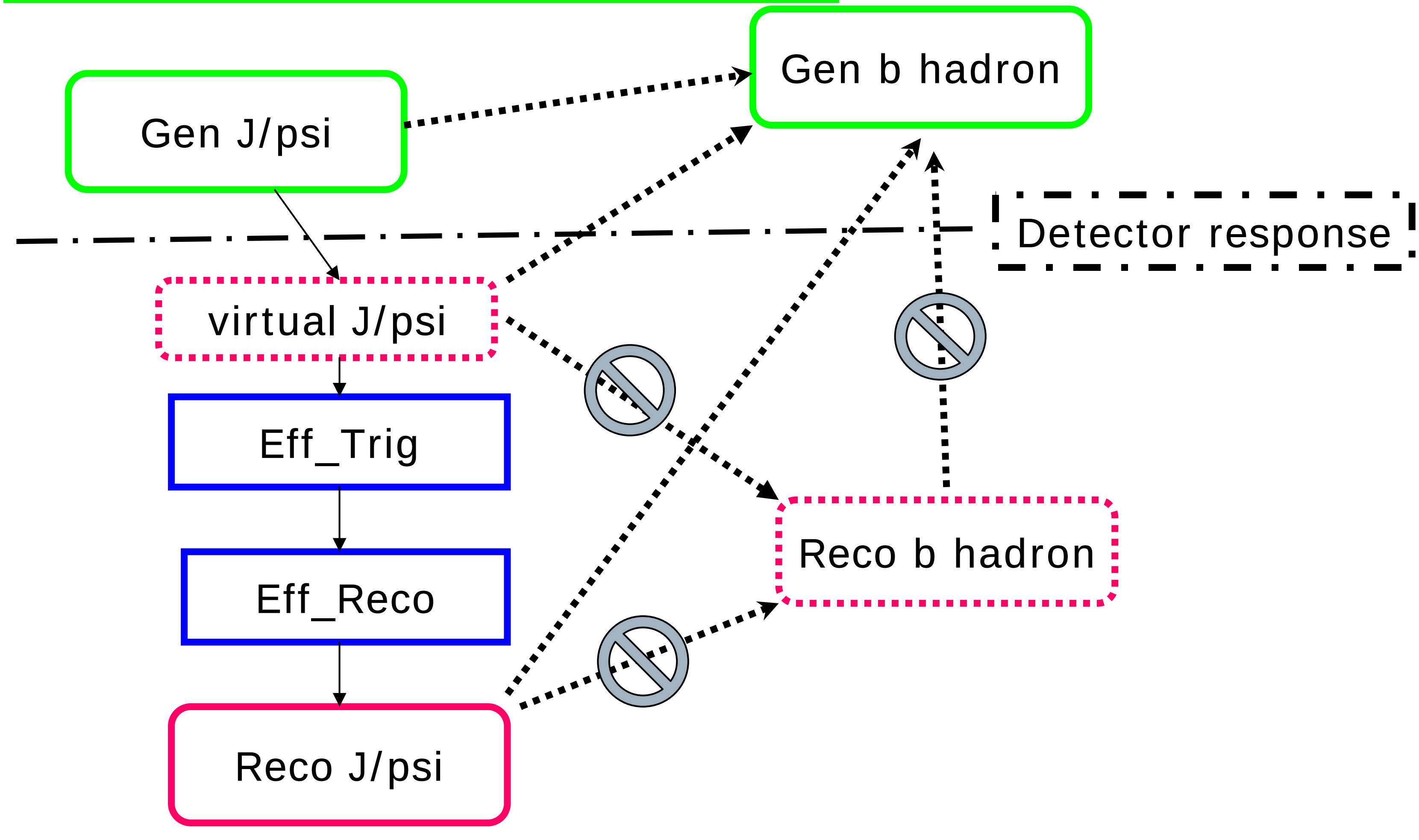
Unfolding

- **The differential *b*-hadron cross section vs. pT(H_b) is extracted from the measured differential ones of H_b ->J/ΨX**
- **Distortions between pT distribution of b hadrons and J/Ψs from them**

Unfolding methods I

- **Bin-to-bin correction: no into account migrations a bin to the others; neglect correlation between adjacent bins.**
- **The matrix method: solve the problem of migrations; singular problem; statistical fluctuations; results unstable.**
- **Regularized unfolding: satisfactory results but technical complications; only with one dimension**

Unfolding


Gen J/psi
Gen b hadron
Detector response
virtual J/psi
Eff_Trig
Eff_Reco
Reco b hadron
Reco J/psi

Unfolding Method II: Bayes'

A Multidimensional unfolding method based on Bayes' theorem by G.D'Agostini, Nucl. Instr. Meth. A362 (1995) 487-498. -- Model independent method

$$P(C_i|E_j) = \frac{P(E_j|C_i)P_0(C_i)}{\sum_{l=1}^{n_C} P(E_j|C_l)P_0(C_l)}.$$

$$\hat{n}(C_i)|_{obs} = \sum_{j=1}^{n_E} n(E_j)P(C_i|E_j)$$

$$\hat{n}(C_i) = \frac{1}{\epsilon_i}\sum_{j=1}^{n_E} n(E_j)P(C_i|E_j) \quad \epsilon_i \neq 0$$

Ci: cause in i-th bin
Ej: effect in j-th bin
P(Ci|Ej): corelation matrix for Ej to Ci

$$\hat{N}_{true} = \sum_{i=1}^{n_C} \hat{n}(C_i),$$

$$\hat{P}(C_i) \equiv P(C_i | n(E)) = \frac{\hat{n}(C_i)}{\hat{N}_{true}},$$

$$\hat{\epsilon} = \frac{N_{obs}}{\hat{N}_{true}}.$$

the unfolding can be performed through the following steps:

1) choose the initial distribution of $P_0(C)$ from the best knowledge of the process under study, and hence the initial expected number of events $n_0(C_i) = P_0(C_i)N_{obs}$; in case of complete ignorance, $P_0(C)$ will be just a uniform distribution: $P_0(C_i) = 1/n_C$;

2) calculate $\hat{n}(C)$ and $\hat{P}(C)$;

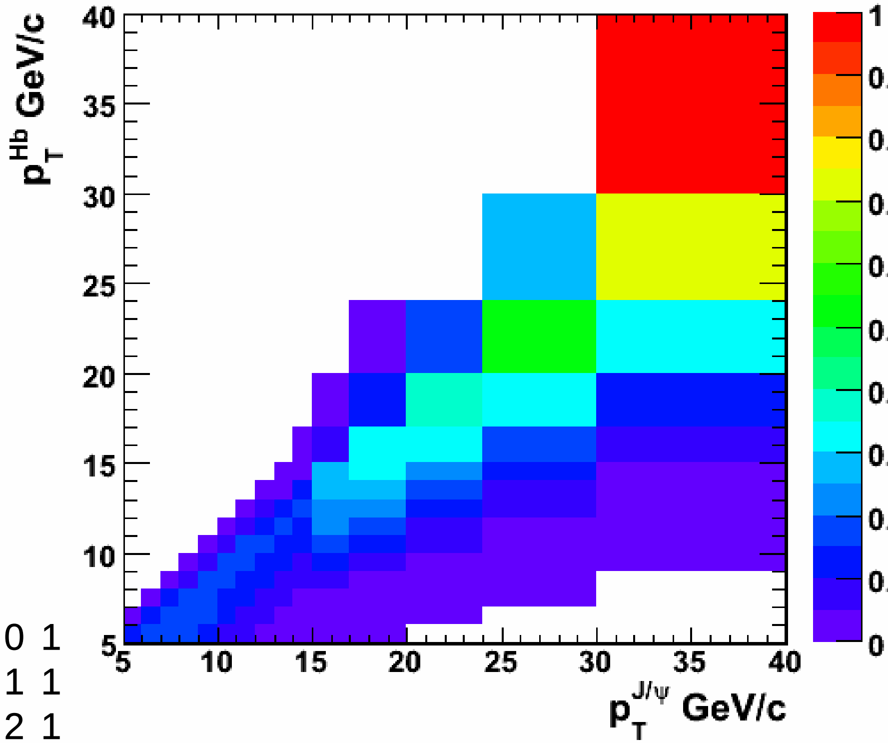
3) make a χ^2 comparison between $\hat{n}(C)$ and $n_0(C)$;

4) replace $P_0(C)$ by $\hat{P}(C)$, and $n_0(C)$ by $\hat{n}(C)$, and start again; if, after the second iteration the value of χ^2 is "small enough", stop the iteration; otherwise go to step 2. Some criteria about the optimum number of iterations will be discussed later.

— $\sum_{i=1}^{n_C} P_0(C_i) = 1$, as usual. Notice that if the probability of a cause is initially set to zero it can never change, i.e. if a cause does not exist it cannot be invented;

— $\sum_{i=1}^{n_C} P(C_i|E_j) = 1$: this normalization condition, mathematically trivial since it comes directly from Eq. (3), tells that each effect must come from one or more of the causes under examination. This means that if the observables contain also a non-negligible amount of background, this needs to be included among the causes;

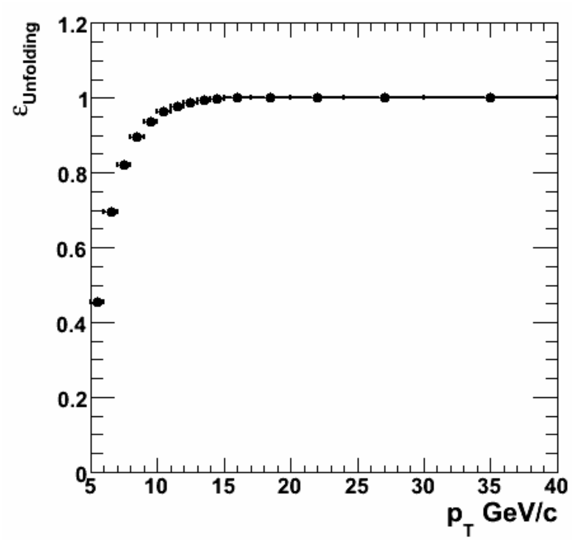
— $0 \le \epsilon_i \equiv \sum_{j=1}^{n_E} P(E_j|C_i) \le 1$: there is no need for each cause to produce at least one of the effects taken under consideration. ϵ_i gives the *efficiency* of detecting the cause C_i in any of the possible effects.

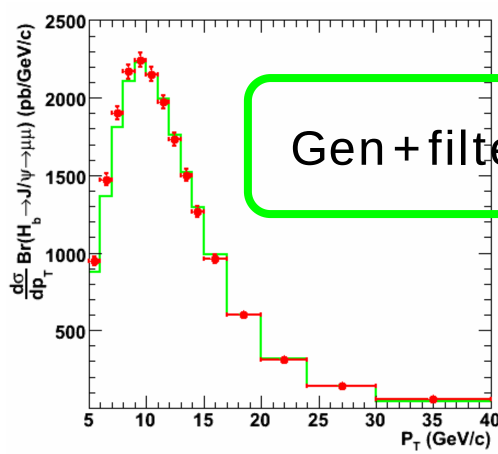


0	P0##:0.0389734	0.0389734
1	P0##:0.0609192	0.0998926
2	P0##:0.0794061	0.179299
3	P0##:0.091277	0.270576
4	P0##:0.0949365	0.365512
5	P0##:0.091491	0.457003
6	P0##:0.0843421	0.541345
7	P0##:0.0743683	0.615713
8	P0##:0.0643528	0.680066
9	P0##:0.0543306	0.734397
10	P0##:0.0825217	0.816919
11	P0##:0.0770165	0.893935
12	P0##:0.053219	0.947154
13	P0##:0.0347848	0.981939
14	P0##:0.0180612	1

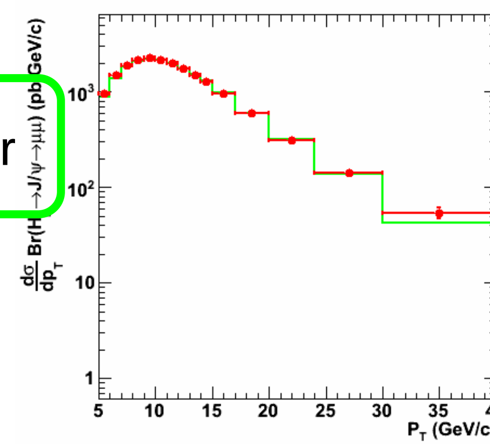
ETotI_P(Ci,E):0 1
ETotI_P(Ci,E):1 1
ETotI_P(Ci,E):2 1
ETotI_P(Ci,E):3 1
ETotI_P(Ci,E):4 1
ETotI_P(Ci,E):5 1
ETotI_P(Ci,E):6 1
ETotI_P(Ci,E):7 1
ETotI_P(Ci,E):8 1
ETotI_P(Ci,E):9 1
ETotI_P(Ci,E):10 1
ETotI_P(Ci,E):11 1
ETotI_P(Ci,E):12 1
ETotI_P(Ci,E):13 1
ETotI_P(Ci,E):14 1

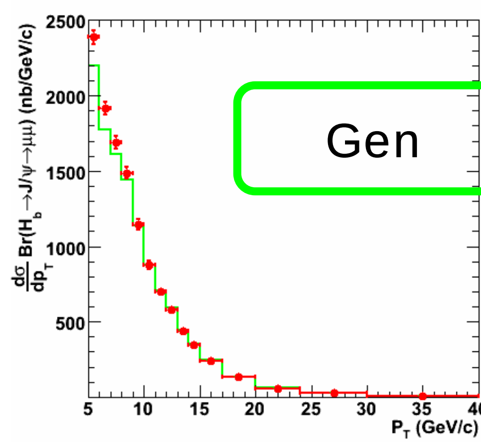
Hb production x-section@pp 10TeV



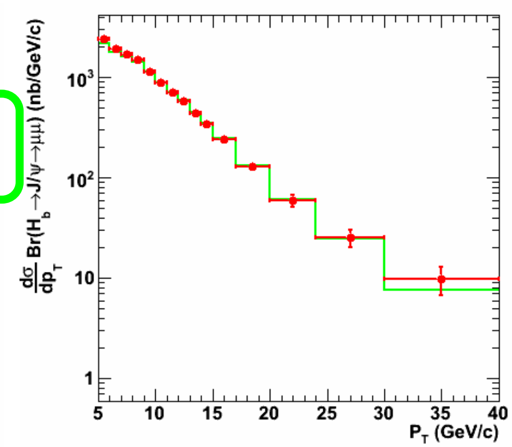


Gen+filter

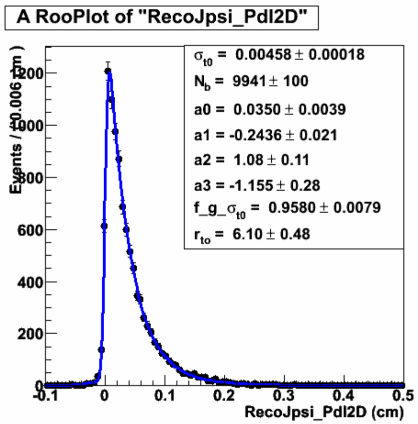


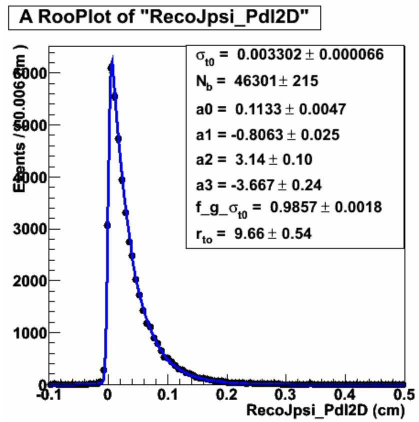


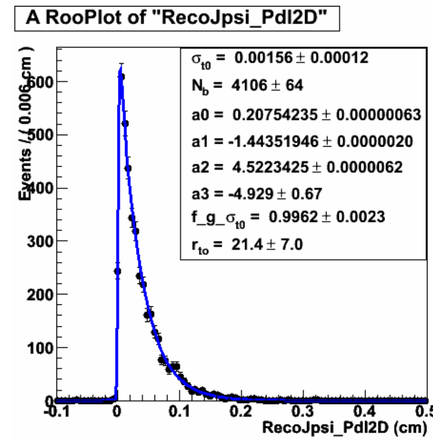
Gen

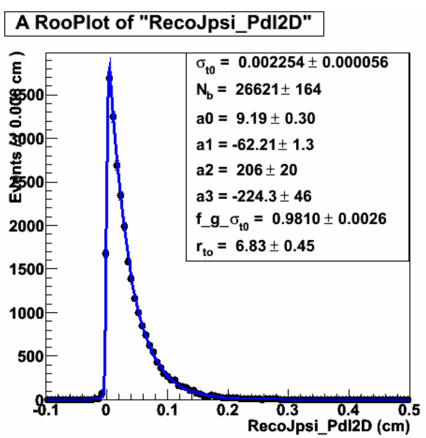


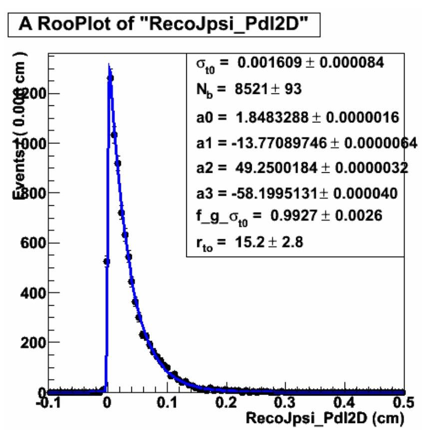
PDL efficiency: fit for jpsi pt bins 5-6,9-10,15-17,24-30, 30-40

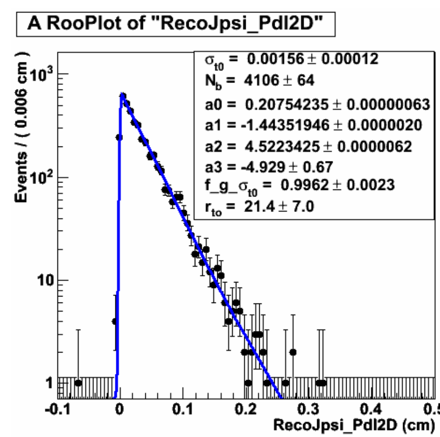












Statistic Error: formula & efficiencv error

Poisson Errors

$$\begin{aligned}\delta\epsilon' &= \epsilon'\sqrt{\left(\frac{\delta k}{k}\right)^2 + \left(\frac{\delta N}{N}\right)^2}\\ &= \frac{k}{N}\sqrt{\frac{1}{k}+\frac{1}{N}}\\ &= \sqrt{\frac{k^2(N+k)}{N^3}}.\end{aligned}$$

$$\sigma_i = N_i' / L_{\text{int}}$$

$$N_i' = N_i / \varepsilon_{Total}$$

$$\varepsilon_{Total} = \varepsilon_{L1} \times \varepsilon_{HLT} \times \varepsilon_{\text{Reco}} \times \varepsilon_{Cuts} = \prod_j \varepsilon_j$$

$$\frac{\Delta\varepsilon_{Total}}{\varepsilon_{Total}} = \sum_j \frac{\Delta\varepsilon_j}{\varepsilon_j}$$

$$\frac{\Delta N_i'}{N_i'} = \frac{\Delta N_i}{N_i} - \frac{\Delta\varepsilon_{Total}}{\varepsilon_{Total}}$$

$$\Delta\sigma_i = \Delta N_i' / L_{\text{int}} = \left(\frac{\Delta N_i}{N_i} - \sum_j \frac{\Delta\varepsilon_j}{\varepsilon_j}\right) \times \frac{N_i}{\prod_j \varepsilon_j} \times \frac{1}{L_{\text{int}}}$$

Binomial Errors

$$\langle k\rangle = \epsilon N,$$

$$\delta\epsilon' = (1/N)\sqrt{k(1-k/N)}.$$

$$\begin{aligned}\sigma_k &= \sqrt{\text{var}(k)}\\ &= \sqrt{\epsilon(1-\epsilon)N}.\end{aligned}$$

Bayes'

$$P(k|\epsilon, N, I) = \frac{N!}{k!(N-k)!}\epsilon^k(1-\epsilon)^{N-k}.$$

$$P(\epsilon|k, N, I) = \frac{\Gamma(N+2)}{\Gamma(k+1)\Gamma(N-k+1)}\epsilon^k(1-\epsilon)^{N-k}.$$

To do List

- anal on CSA07 data @14TeV
- anal on Sum08 data @10TeV
  - Efficiency: Accept., Trig & reco. :M.C. & T.P.
  - Combined MLH fit
  - Jpsi prodcution x-section
  - Unfolding & final b hadrons spectrum

    -- method ok on CSA07 data
  - Average b lifetime
    - --Pdl efficiency
  - Systematic uncertainties
- Prepare PAS and Note draft

backups

Efficiencies from data: Tag&Probe

- Tag-and-Probe
  - Successfully used in experiments: TEVATRON/CDF&D∅
- Current availability of code
  - Egamma: EgammaAnalysis/EgammaEfficiencyAlgos
  - Muon: MuonAnalysis/TagAndProbe

adapt code to use under CMSSW_16X with PAT

Efficiency Measurements: Tag&Probe

- **The overall dimuon efficiencies of the measurement are assumed to be the product of several parts**

$$\varepsilon = \varepsilon_{\text{acceptance}} \times \varepsilon_{\text{trigger}} \times \varepsilon_{\text{offline}}^{2}$$

$$\varepsilon_{\text{trigger}} = \varepsilon_{\text{L1}} \times \varepsilon_{\text{HLT}}$$

$$\varepsilon_{\text{offline}} = \varepsilon_{\text{global}} \times \varepsilon_{\text{isolation}} \times \varepsilon_{\text{id}}$$

$$\varepsilon_{\text{global}} = \varepsilon_{\text{standalone}} \times \varepsilon_{\text{tracker}} \times \varepsilon_{\text{matching}}$$

- Choose a *tag* muon
  - A “high quality” reconstructed muon
- Choose a *probe* track
  - A probable muon in tracker or muon system
- Requiring $M_{\mu\mu}$ consistent with $M_{J/\Psi}$ yileds a high-purity and almost unbiased sample of *probe* muons

Description of Tag and Probe

TAG	**Global muon with p_T > 5GeV** **Associated to a L3 muon**

Probe Type	Description
Golden	Global muon that is also a TAG
Matched	Global muon that is not a TAG
Unmatched	Tracker track AND Standalone muon found, but they are not associated with a Global Muon
Tracker Only	Only a tracker track
Stand Alone Muon	Only a standalone muon

With the five types of probes, we get five combinations of tag-and-probe: GG, GM, GU, GS, GT

Tracking and Matching Efficiencies I

- Standalone, Tracking, and Matching efficiencies calculated with simple event counting

$$\varepsilon_{\text{standalone}} = \frac{2N_{GG} + N_{GM} + N_{GU}}{2N_{GG} + N_{GM} + N_{GU} + N_{GT}}$$

$$\varepsilon_{\text{tracker}} = \frac{2N_{GG} + N_{GM} + N_{GU}}{2N_{GG} + N_{GM} + N_{GU} + N_{GS}}$$

$$\varepsilon_{\text{matching}} = \frac{2N_{GG} + N_{GM}}{2N_{GG} + N_{GM} + N_{GU}}$$

$$\varepsilon_{\text{global}} = \varepsilon_{\text{standalone}} \times \varepsilon_{\text{tracker}} \times \varepsilon_{\text{matching}}$$

Pseudo proper decay length

$$\vec{X} = \vec{x}_B - \vec{x}_{prim}$$

$$L_{xy}^B = \frac{\vec{X} \cdot \vec{p}_T^B}{|\vec{p}_T^B|}.$$

$$\lambda^B = \frac{L_{xy}^B}{(\beta\gamma)_T^B} = L_{xy}^B \cdot \frac{M_B}{p_T^B};$$

$$\lambda_\psi = \frac{L_{xy}^\psi}{(\beta\gamma)_T^\psi} = L_{xy}^\psi \cdot \frac{M_\psi}{p_T^\psi}.$$

$$\lambda = \frac{\lambda_\psi}{\langle F(p_T^\psi)\rangle} = L_{xy}^\psi \cdot \frac{M_\psi}{p_T^\psi \langle F(p_T^\psi)\rangle}$$

$$F(p_T^\psi) = \frac{(\beta\gamma)_T^B}{(\beta\gamma)_T^\psi} = \frac{\lambda_\psi}{\lambda_B}$$

- **Measure the 2-dimensional decay length *Lxy* for the *J/* Ψ meson sample**
- **pseudo proper decay length distribution**
- **Measure the l distribution of the background under the *J/* Ψ by studying the μ + μ - mass sidebands of the *J/* Ψ**
- **Fit the distribution to the sum of background, direct (zero-lifetime) and *B* decay (non-zero lifetime) Contributions and extract the lifetime**

Unfolding methods I

- **Bin-to-bin correction: no into account migrations a bin to the others; neglect correlation between adjacent bins.**
- **The matrix method: solve the problem of migrations; singular problem; statistical fluctuations; results unstable.**
- **Regularized unfolding: satisfactory results but technical complications; only with one dimension**

Unfolding Method II: Bayes'

A Multidimensional unfolding method based on Bayes' theorem by G.D'Agostini, Nucl. Instr. Meth. A362 (1995) 487-498. -- Model independent method

$$P(C_i|E_j) = \frac{P(E_j|C_i)P_0(C_i)}{\sum_{l=1}^{n_C} P(E_j|C_l)P_0(C_l)}.$$

$$\hat{n}(C_i)|_{obs} = \sum_{j=1}^{n_E} n(E_j)P(C_i|E_j)$$

$$\hat{n}(C_i) = \frac{1}{\epsilon_i}\sum_{j=1}^{n_E} n(E_j)P(C_i|E_j) \quad \epsilon_i \neq 0$$

Ci: cause in i-th bin
Ej: effect in j-th bin
P(Ci|Ej): corelation matrix for Ej to Ci

$$\hat{N}_{true} = \sum_{i=1}^{n_C} \hat{n}(C_i),$$

$$\hat{P}(C_i) \equiv P(C_i \mid n(E)) = \frac{\hat{n}(C_i)}{\hat{N}_{true}},$$

$$\hat{\epsilon} = \frac{N_{obs}}{\hat{N}_{true}}.$$

the unfolding can be performed through the following steps:

1) choose the initial distribution of $P_0(C)$ from the best knowledge of the process under study, and hence the initial expected number of events $n_0(C_i) = P_0(C_i)N_{obs}$; in case of complete ignorance, $P_0(C)$ will be just a uniform distribution: $P_0(C_i) = 1/n_C$;

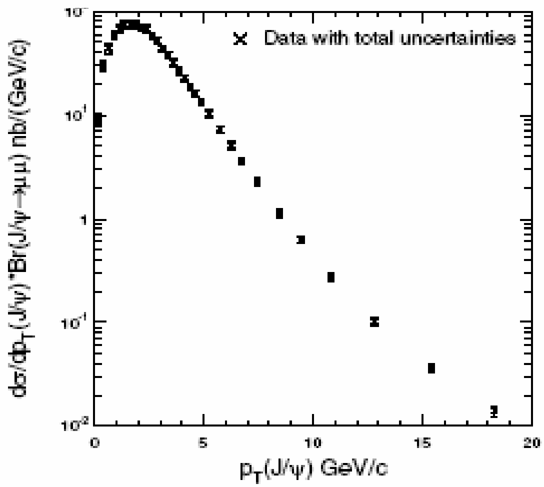
2) calculate $\hat{n}(C)$ and $\hat{P}(C)$;

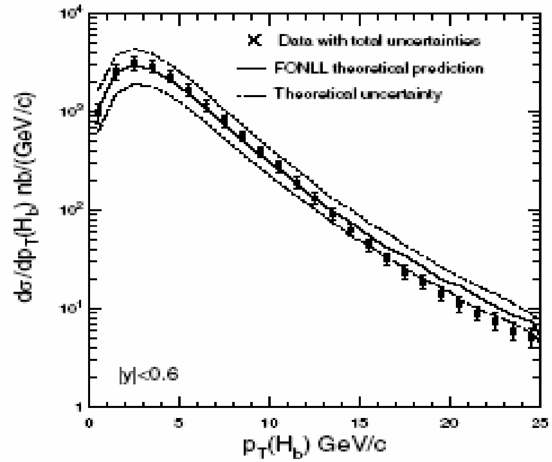
3) make a χ^2 comparison between $\hat{n}(C)$ and $n_0(C)$;

4) replace $P_0(C)$ by $\hat{P}(C)$, and $n_0(C)$ by $\hat{n}(C)$, and start again; if, after the second iteration the value of χ^2 is "small enough", stop the iteration; otherwise go to step 2. Some criteria about the optimum number of iterations will be discussed later.

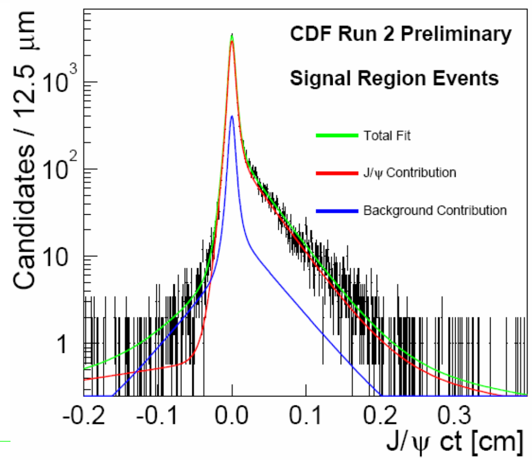
CDFII's result

PHYSICAL REVIEW D **71**, 032001 (2005)





- **PPbar energy: 1.96 TeV.**
- **Integrated luminosity: 39.7 pb^{-1}.**



$$\sigma[p\bar{p}\to J/\psi X, |y(J/\psi)|<0.6] = 4.08 \pm 0.02(\text{stat})^{+0.36}_{-0.33}(\text{syst})\ \mu\text{b}.$$

$$\sigma(p\bar{p}\to H_b X, |y|<0.6) = 17.6 \pm 0.4(\text{stat})^{+2.5}_{-2.3}(\text{syst})\ \mu\text{b}.$$

CDF note 6023

$$\tau_B = 1.526 \pm 0.034\ (\text{stat}) \pm 0.035\ (\text{syst})\ \text{ps}$$

Inv. Mass of J/ψ vs diff. pT bin