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


Xiangwei Meng

Institute of High Energy Physics, CAS, Beijing


**July 24, 2009**

# Outline

- Systematic Uncertainties
- Fit validation in RooFit

# Systematic Uncertainties

- Luminosity ~10%
- Acceptance
  - Jpsi polarization: CDF-II & BaBar 2-5%
  - pT spetrum: flat pT distribution needed
  - M.C. simu. stat.

- Signal number: Mass pdf; moumentum scale; misalignment
  pdl pdf: resolution model; Hb lifetime medol;
  background pdf; pdl eff
  fit bias
- Eff. Reconstruction: M.C. Truth & tag&probe corr. 1-3%
- Eff. Trigger: M.C. Truth & tag&probe corr. 1-3%
- Unfolding: Jpsi-> Hb
  - Hb->jpsi decay spectrum: CLEO, BaBar vs. EvtGen
  - eff. For unfolding correlation: limited pT region
  - M.C. simu. stat.

**Luminosity** the design goal for CMS is a systematic accuracy of 5%[2]. For the $J/\psi$ measurement of CDF II[1], that uncertainty was about 6.0% for an integrated luminosity of $39.7pb^{-1}$. We assume uncertainty of 10% for Luminosity determination at LHC start up.

**Acceptance and efficiency**

**Branching fractions** the errors on the branching fractions reported in the PDG[3] is used. The total branching ratio $\mathrm{Br} = (6.88 \pm 0.66) \times 10^{-4}$, is derived from $\mathrm{Br}(H_b \to J/\psi X)$ $(1.16 \pm 0.10) \times 10^{-2}$ and $\mathrm{Br}(J/\psi \to \mu^+\mu^-)$ $(5.93 \pm 0.06) \times 10^{-2}$, so about 9.6% uncertainty.

**$J/\psi$ polarization** kinematic acceptance as a function of $p_T$ depends on the $J/\psi$ spin alignment. For the early data at LHC start up, maybe no available CMS estimate can be refered. The systematic uncertainty is estimated from ref.[5] from the measurements of CDF II for prompt $J/\psi$[1]( inclusive $J/\psi$:$\pm 2 - 5\%$) and of BaBar for $b \to J/\psi$ from $\Upsilon(4S)$[4]: the uncertainty is $1.8 - 7.0\%$. $J/\psi$ mesons from decays of bottom hadrons have a different average spin alignment than an inclusive sample of $J/\psi$ mesons, we estimate an acceptance correction of the same order percentage.

- $p_T$ **spectrum** to estimate the uncertainty from variations of the input transverse momentum spectrum, the acceptance is recalculated using a Monte Carlo sample generated using a flat distribution. The flat distribution is an extreme alternative from the nominal spectrum which is a fast falling function of $p_T$. The fractional change in acceptance is taken as the uncertainty on the input transverse momentum distribution. To study the dependence of the $b$ fractions on the modeling of the $b$-hadron spectrum used in the Monte Carlo, a flat distribution in $p_T$ and $\eta$ of the $b$ production spectrum is used to regenerate the distributions and the fits were repeated. we estimate an acceptance correction of the same order to CDF II[1]: $0 - 5\%$.

- **M.C. statistics** uncertainty on the reconstruction efficiency due to finite data sample size. Monte Carlo samples are used to determine the total efficiencies in the analysis, and a systematic uncertainty with the statis-

- **Trigger efficiency** we assume this will be measured to a precision of 5% with $10pb^{-1}$ of data.

**PDF shapes** The resulting variation in yields is taken as an estimate of the systematic uncertainty coming from imperfect knowledge of the PDF shapes.

- **Resolution model** The double Gaussian resolution function is not an exact description of the resolution function shape but only an approximate parametrization of many different resolution effects. Therefore, to estimate the systematic uncertainty due to the resolution function modeling, the maximum range of values for the ratios of areas and widths of the two Gaussians supported by the data are estimated. We find that the ratios of the second Gaussian to the dominant Gaussian vary from 3.1 to 8.3 in width and 0.83 to 0.94 in area. We have compared a triple gaussian resolution function and taken the differences as a systematic error: $0 - 1.5\%$
- $H_b$ **lifetime model** The mix of hadrons and their respective lifetimes is a contributing factor to the shape of the $J/\psi$ pseudoproper decay time distributions. The $b$ hadron pseudoproper time is described by an exponential decay while in analysis we start from its decay daughter $J/\psi$, the mass, decay length and momentum. The different kinds of $b$ hadrons' pseudo-proper time are different from that calculated from only $J/\psi$ part.
- **Fit function form** consider different assumed functional forms for the PDFs, but the resulting changes in signal yields and lifetimes. we tried a single Gaussian (instead of double) plus an exponential, or three Gaussians and no exponential to the Resolution function from.
- **Fit PDF parameters** we perform ensembles of fits varying the PDF parameters individually according to their statistical uncertainties.

TABLE I: Systematic uncertainties from MC statistics.

| $p_T$ $(J/\psi)$ GeV/c | $\Delta\sigma/\sigma(\%)$ |
|---|---|
| 5-6 | $\pm1.11$ |
| 6-7 | $\pm0.87$ |
| 7-8 | $\pm0.75$ |
| 8-9 | $\pm0.71$ |
| 9-10 | $\pm0.68$ |
| 10-11 | $\pm0.68$ |
| 11-12 | $\pm0.70$ |
| 12-13 | $\pm0.73$ |
| 13-14 | $\pm0.76$ |
| 14-15 | $\pm0.90$ |
| 15-17 | $\pm0.69$ |
| 17-20 | $\pm0.70$ |
| 20-24 | $\pm0.81$ |
| 24-30 | $\pm0.99$ |
| 30-40 | $\pm1.27$ |

TABLE II: Systematic uncertainties from pdl resoultion.

| $p_T$ $(J/\psi)$ GeV/c | $\Delta\sigma/\sigma(\%)$ |
|---|---|
| 5-6 | $\pm-1.$ |
| 6-7 | $\pm1.$ |
| 7-8 | $\pm0.5$ |
| 8-9 | $\pm+0.8$ |
| 9-10 | $\pm+0.7$ |
| 10-11 | $\pm-0.6$ |
| 11-12 | $\pm-0.3$ |
| 12-13 | $\pm-0.3$ |
| 13-14 | $\pm-0.4$ |
| 14-15 | $\pm-1.3$ |
| 15-17 | $\pm-1.4$ |
| 17-20 | $\pm+0.5$ |
| 20-24 | $\pm-1.5$ |
| 24-30 | $\pm+0.6$ |
| 30-40 | $\pm+0.6$ |

**Momentum scale** An uncertainty of average 0.4% on the momentum scale is extracted by comparing

the reconstructed $J/\psi$ mass to the world averaged value. The mass of $J/\psi$ candidate from the dimuon depends linearly on the momentum scale of th two muons, and the number of events will be de termined from a fit to the dimuon invariant mas distribution. A difference in the muon $p_T$ cut be tween data and Monte Carlo will affect the numbe of events that pass the HLT trigger, and also the number of events in a given $p_T$ bin. Thus, an uncertainty in the muon momentum scale affects the total efficiency for each $p_T$ bin. This uncertainty has been evaluated to be at most 1%.

**Misalignment** This alignment scenario corresponding to the assumed precision obtainable with 10 $pb^{-1}$ of data is used: $\sim 0.7-3.5\%$

TABLE III: Sources of systematic uncertainties on the measurement of the $J/\psi$ from $b$-hadrons in inclusive $J/\psi$ decays.

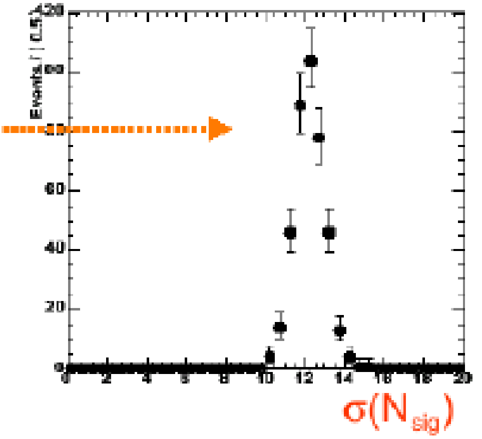
| | Source | Syst. |
|---|---|---|
| Luminosity | | $\sim 10\%$ |
| Data quality | | $\sim 1\%$ |
| Branch fraction | | 9.6% |
| Accept-eff | $J/\psi$ spin alignment | $2-7\%$ |
| Accept-eff | $p_T$ spectrum | $0-5\%$ |
| Accept-eff | M.C. statistics | $0.7-1.3\%$ |
| Yield | mass Pdf | $1-6\%$ |
| Yield | Mis-alignment | $0.7-3.5\%$ |
| Yield | Mis-calibration | $\sim 1\%$ |
| Trig(Tag-Probe Cor) | | $\sim 3\%$ |
| Reco(Tag-Probe Cor) | | $\sim 3\%$ |
| $b$ fraction | Resolution function model | $0-1.5\%$ |
| $b$ fraction | Background function model | $\sim 0-2\%$ |
| $b$ fraction | MC $b$ production spectrum | $\sim 2-7\%$ |
| $b$ fraction | MC $b$ decay spectrum | $\sim 0.5-3\%$ |
| $b$ fraction | MC inclusive $H_b$ lifetime | $\sim 0.5-4\%$ |
| $b$ fraction | lifetime efficiency | 1% |
| $b$ fraction | Fit bias | $0-2\%$ |
| Total | | $14-22\%$ |

# Fit Validation Study – The pull distribution

What about the validity of the error?

Distribution of error from simulated experiments is difficult to interpret…

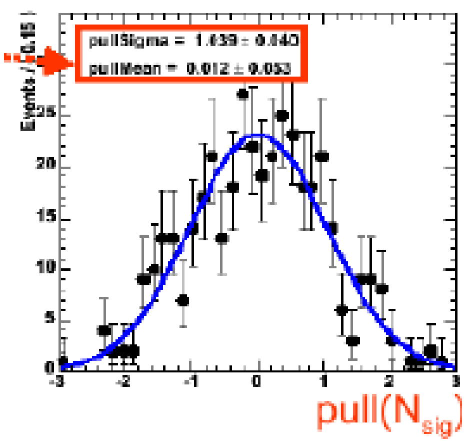
We don't have equivalent of $N_{sig}$(generated) for the error




Solution: look at the ***pull distribution***

Definition:

$$\mathrm{pull}(\mathrm{N_{sig}}) = \frac{N_{sig}^{fit} - N_{sig}^{true}}{\sigma_N^{fit}}$$




Properties of pull:

**Mean is 0 if there is no bias**

**Width is 1 if error is correct**

In this example: no bias, correct error within statistical precision of study

# Pull validation

- RooMCStudy mcstudy(twoDTotModel2,RooArgSet(mass,nnet),Binned(kTRUE),FitModel(twoDTotModel),Silence(),Extended(),FitOptions(Extended()));

- mcstudy->generateAndFit(1000);
  RooPlot* frame1 = mcstudy->plotParam(nSig,Bins(40)) ;
  RooPlot* frame2 = mcstudy->plotError(nSig,Bins(40)) ;
  RooPlot* frame3 = mcstudy->plotPull(nSig,Bins(40)) ;