BDT for JpsiX^0 modes

Feb. 27th 2019, Annecy/Edinburgh meeting, M. Chefdeville

Status

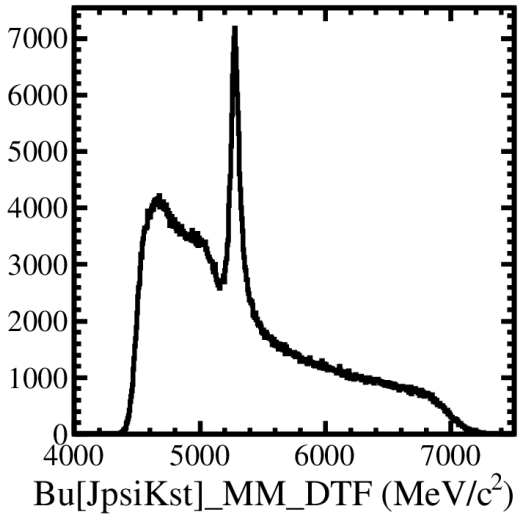
- During last 2 weeks:
  - Finished processing of Run I DiMuon data → Ntuples
  - One week off, resumed work this Monday
  - Started to develop a BDT for neutral modes

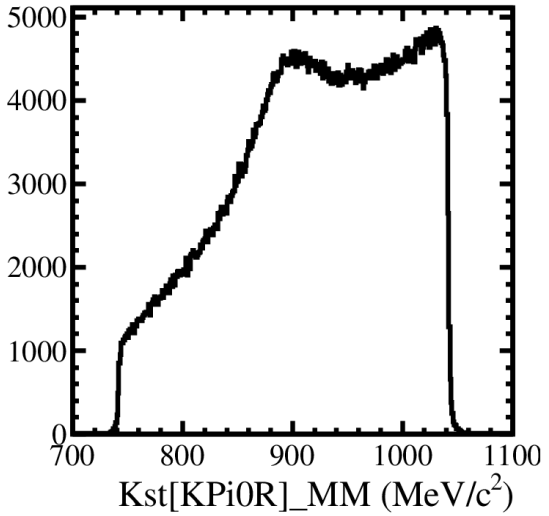
- BDT variables
  - Kinematics & vertex (some previously used in Fisher)
  - Calo variables (CL, cone), and also VtxIsol
    - Signal photons more isolated than bkg (jet-like)
      → potentially strong discriminating power
    - Pile-up dependent
      → will tend to select cleaner photons (i.e. better resolution, less combinatorics)
      → not well reproduced by MC (calibration?)

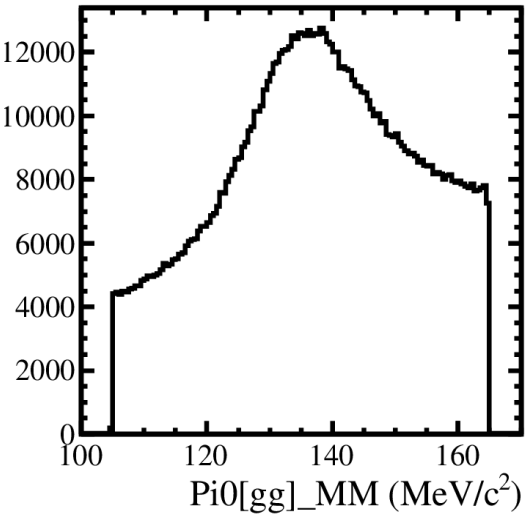
- Calibration strategy: compare efficiency when training data sideband (bkg) VS sPlot-data or MC-truth (signal)
  - Use abundant JpsiK*$^+$[K$^+$pi0]
  - Compare sWeighted data variables to MC variables
  - Assign systematics to using BDT trained on MC-truth

$B^+ \to J/psi\ K^{*+}[K^+pi0]$, Run II

- Selections
  - Ntupling:
    CL(g1,g2)>0.05, PT(pi0)>0.8 GeV/c, PROBNNk(K^+)>0.1,
    $\Delta m(K^{*+})$=150 MeV/c^2, Δm(pi0)=30 MeV/c^2, Δm(Jpsi)=100 MeV/c^2
    Fisher(B^+)>-1.1, DIRA>0.9995, IP<0.2, IPCHI2<20, VTXCHI2/NDOF<10
  - Offline: PT(pi0)>1.5 GeV/c, aligned with Jpsipi0 (will be used for BR meas^t.)







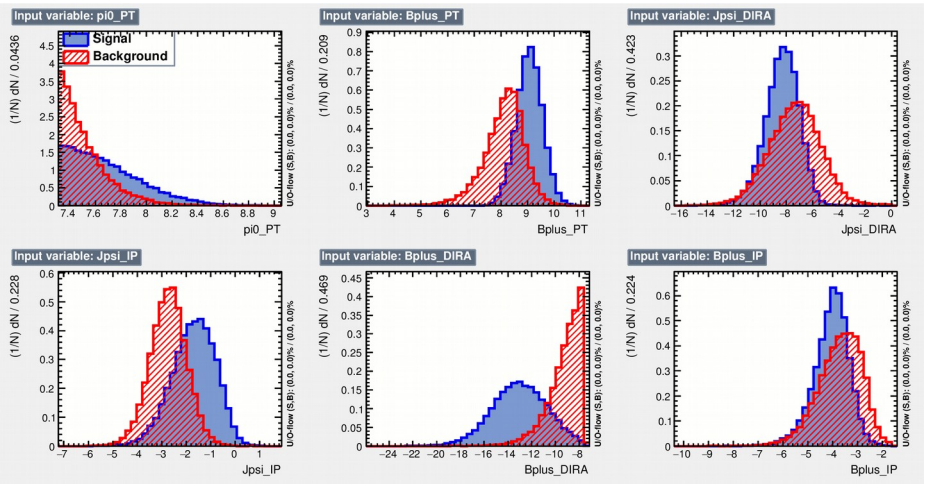
BDT: MC-truth VS data sideband

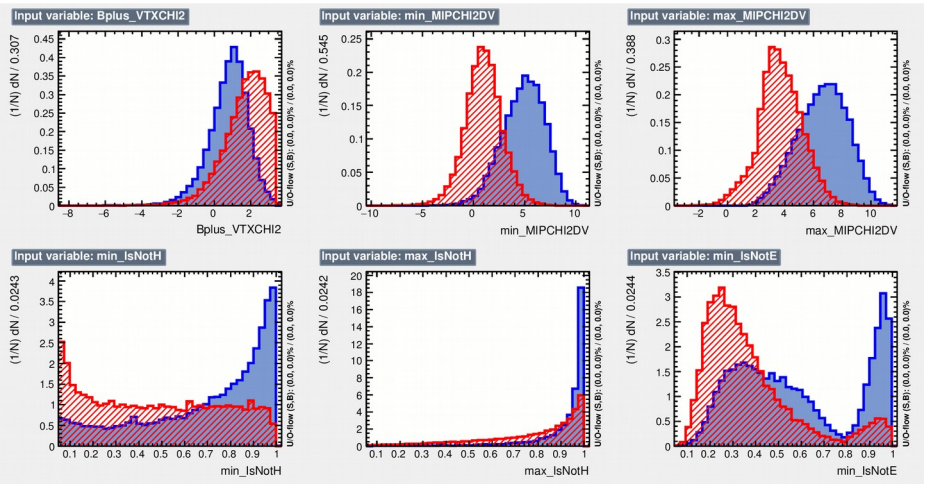
- Samples
  - Signal: MC JpsiK*+ 2015-16 Up-Dw, Sim09e, Pythia8, s28r1 (4M evt)

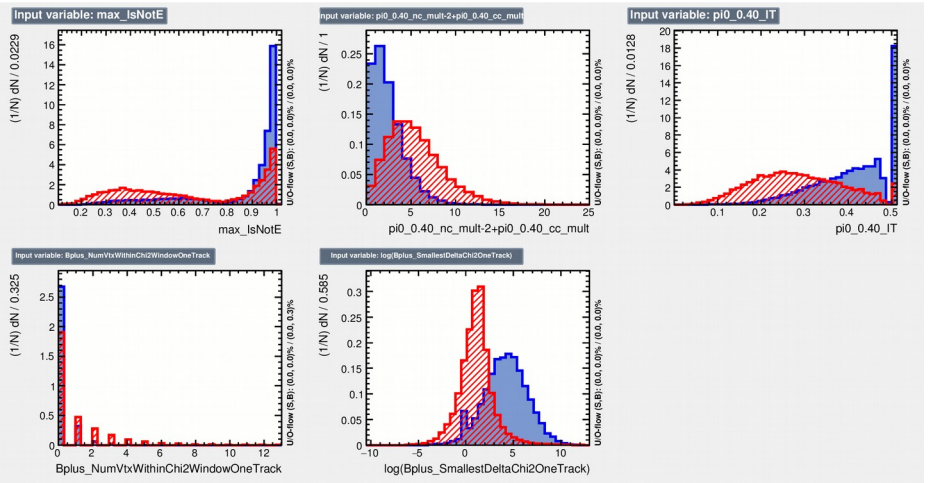
    → 50k evt after selections
  - Background: Run II data sideband (Bplus_MM>6650)

    Cut tuned such as to have similar stat as signal

```
: ------------------------------------------------------------------------------
: Rank : Variable                                : Variable Importance
: ------------------------------------------------------------------------------
:    1 : Bplus_PT                                : 8.746e-02
:    2 : Bplus_DIRA                              : 8.154e-02
:    3 : min_MIPCHI2DV                           : 8.023e-02
:    4 : pi0_PT                                  : 7.177e-02
:    5 : Bplus_IP                                : 6.034e-02
:    6 : max_MIPCHI2DV                           : 5.983e-02
:    7 : Bplus_VTXCHI2                           : 5.951e-02
:    8 : Jpsi_DIRA                               : 5.765e-02
:    9 : log(Bplus_SmallestDeltaChi2OneTrack)    : 5.732e-02
:   10 : min_IsNotH                              : 5.340e-02
:   11 : pi0_0.40_IT                             : 5.339e-02
:   12 : max_IsNotE                              : 5.170e-02
:   13 : pi0_0.40_nc_mult-2+pi0_0.40_cc_mult     : 4.950e-02
:   14 : Jpsi_IP                                 : 4.660e-02
:   15 : min_IsNotE                              : 4.469e-02
:   16 : max_IsNotH                              : 4.325e-02
:   17 : Bplus_NumVtxWithinChi2WindowOneTrack    : 4.183e-02
: ------------------------------------------------------------------------------
```

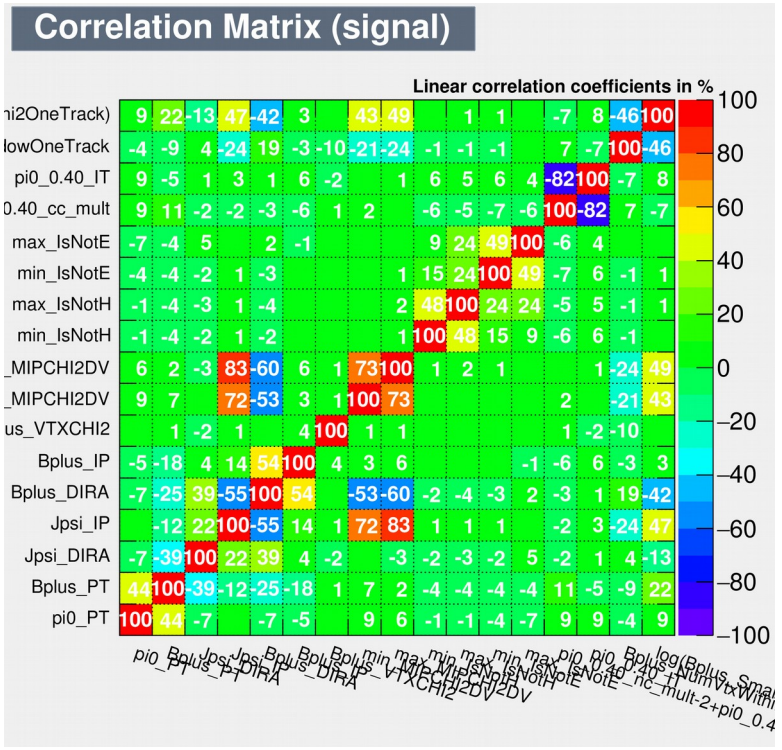


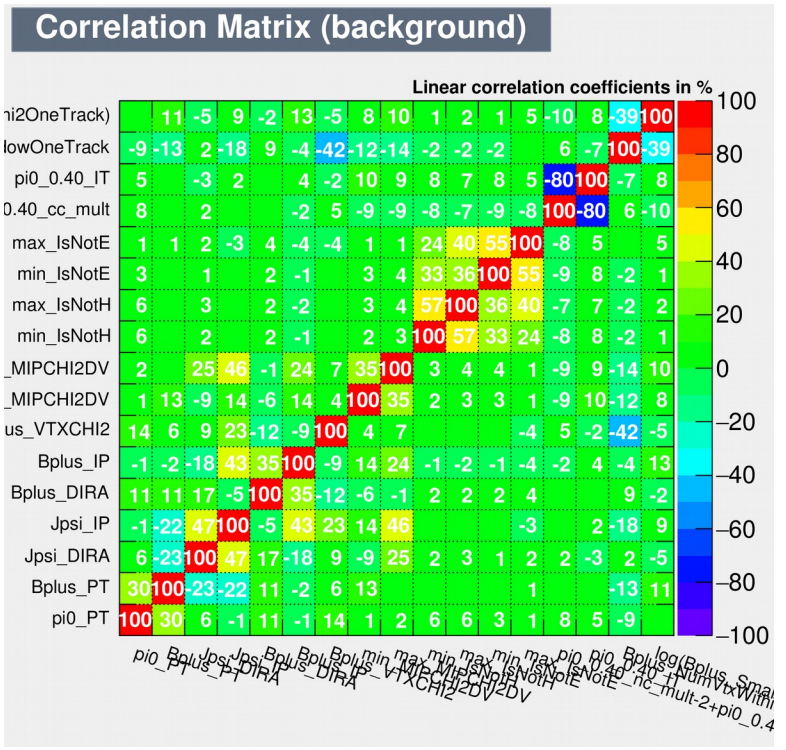




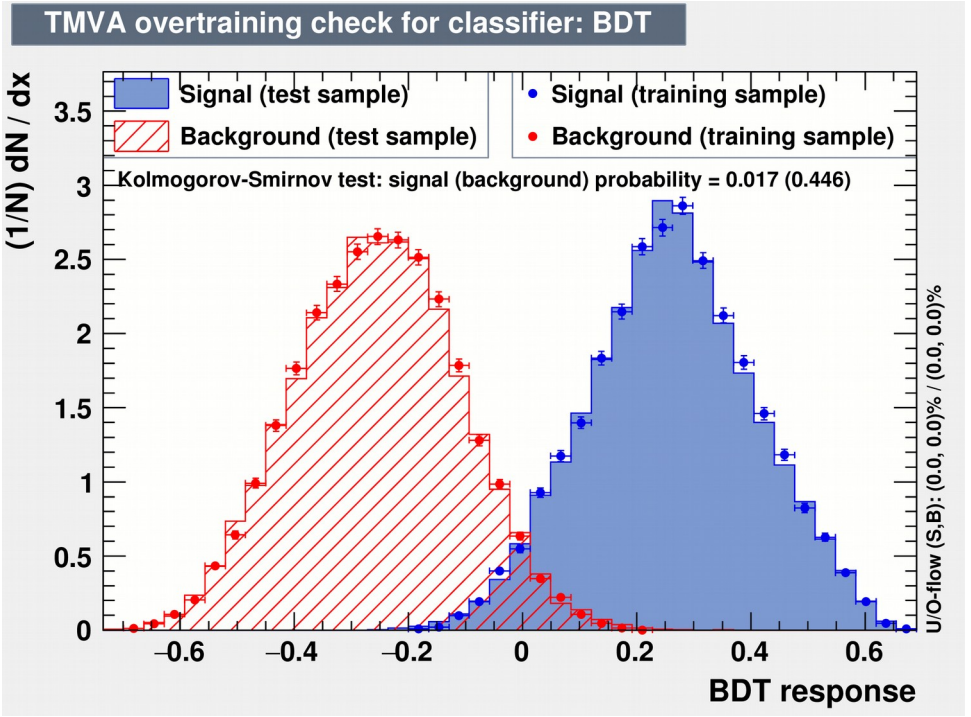
BDT variables

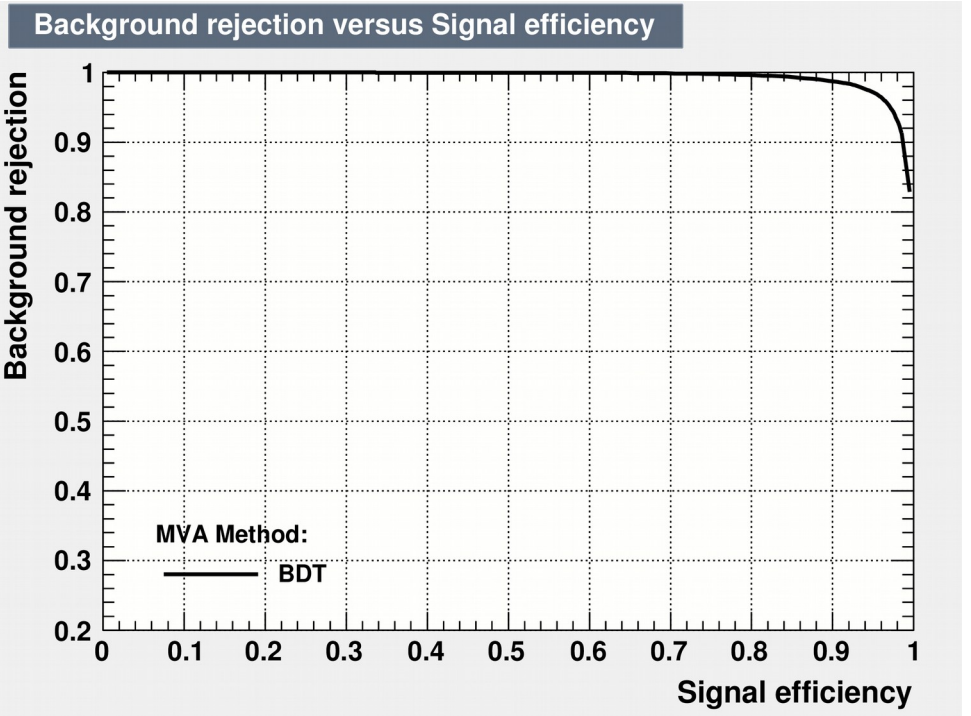
- Correlations
  - pi0 transverse isolation & Nobjects in cone
  - Photon PID (especially for true pi0)
  - IPCHI2 of muons (especially for true Jpsi)
- Will consider removing some variables (performance VS Nvar)





BDT performance

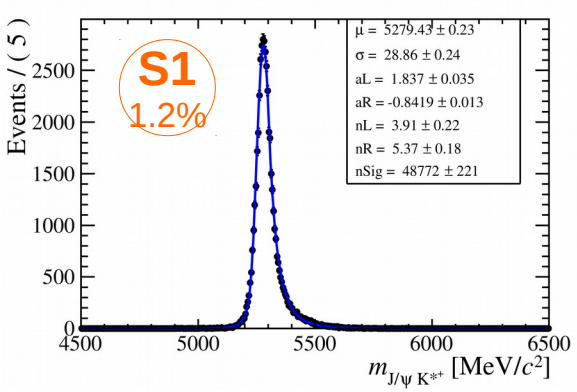


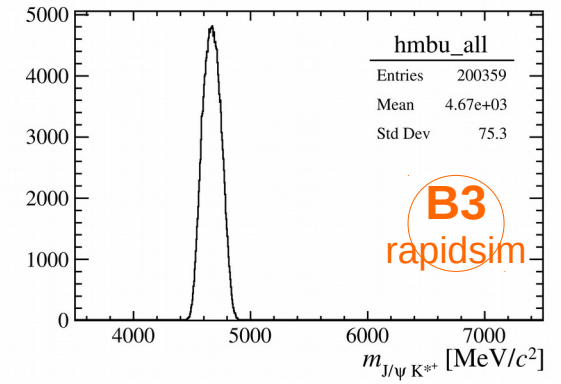


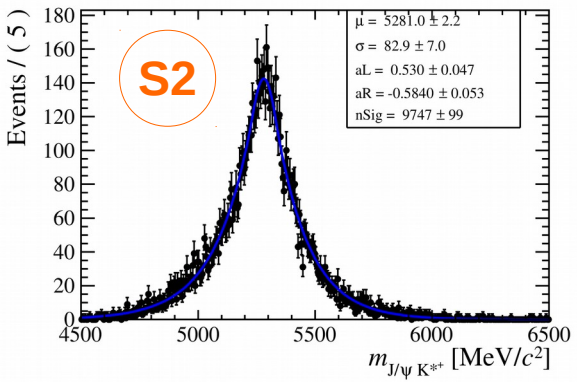
- Now, to compare simulated VS real performance, let's fit the B^+ mass. Then:
  - Determine sWeighted BDT distribution, or
  - Cut & fit & count

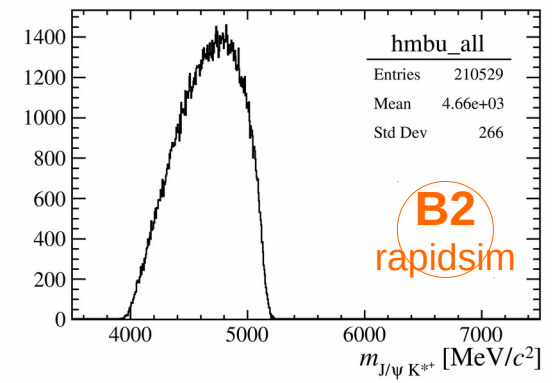
Fit model

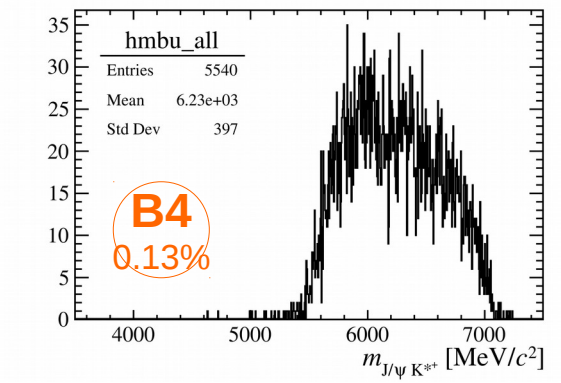
- Signal S1: double-CB with tails fixed on 2016 MC
- S2: 1 signal photon + 1 random photon
  → Gaussian with expo. tails fixed on same MC, S1/(S1+S2) ratio = 0.833, only σ & yield floating
- B1: combinatorial bkg: exponential
- B2: part. reco'ed B^0 → JpsiK*(pi): HistoPdf using RapidSim mass shape (13 TeV, PT cuts)
- B3: part. reco'ed B^+ → Psi(2S)[Jpsi (pipi)]K*: HistoPdf using RapidSim mass shape
- B4: peaking B^+ → JpsiK+ combined with random pi0. Use fullsim sample Sim09e, s28r1, 2016 (4M)
  Ignored for now (yield under investigation)





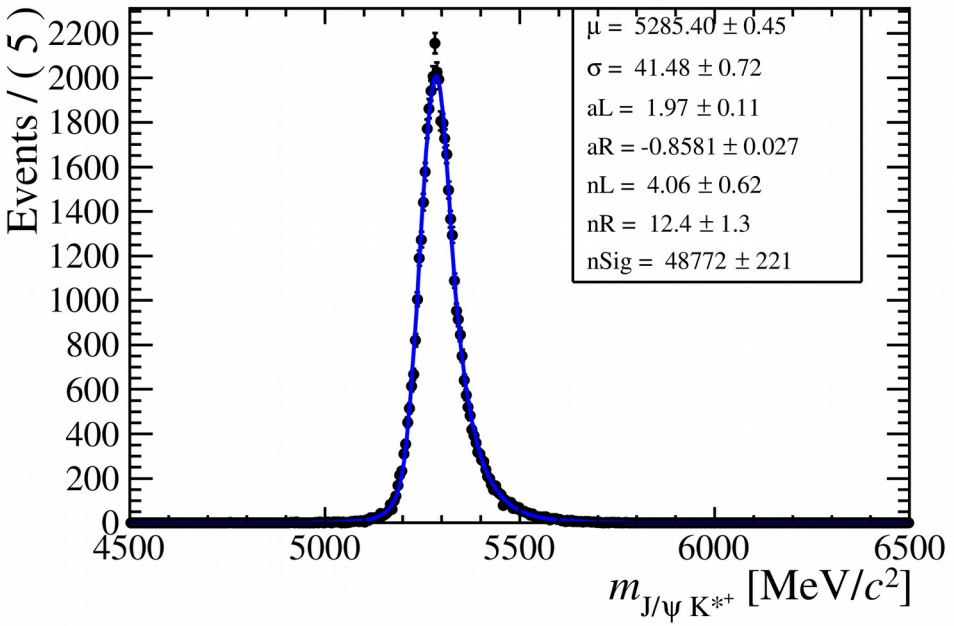


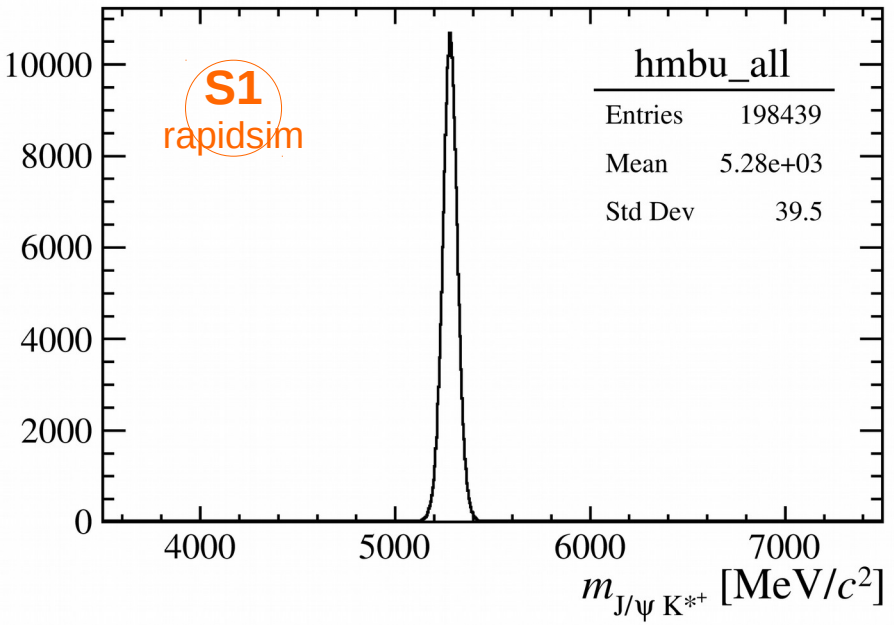




RapidSim, apparté

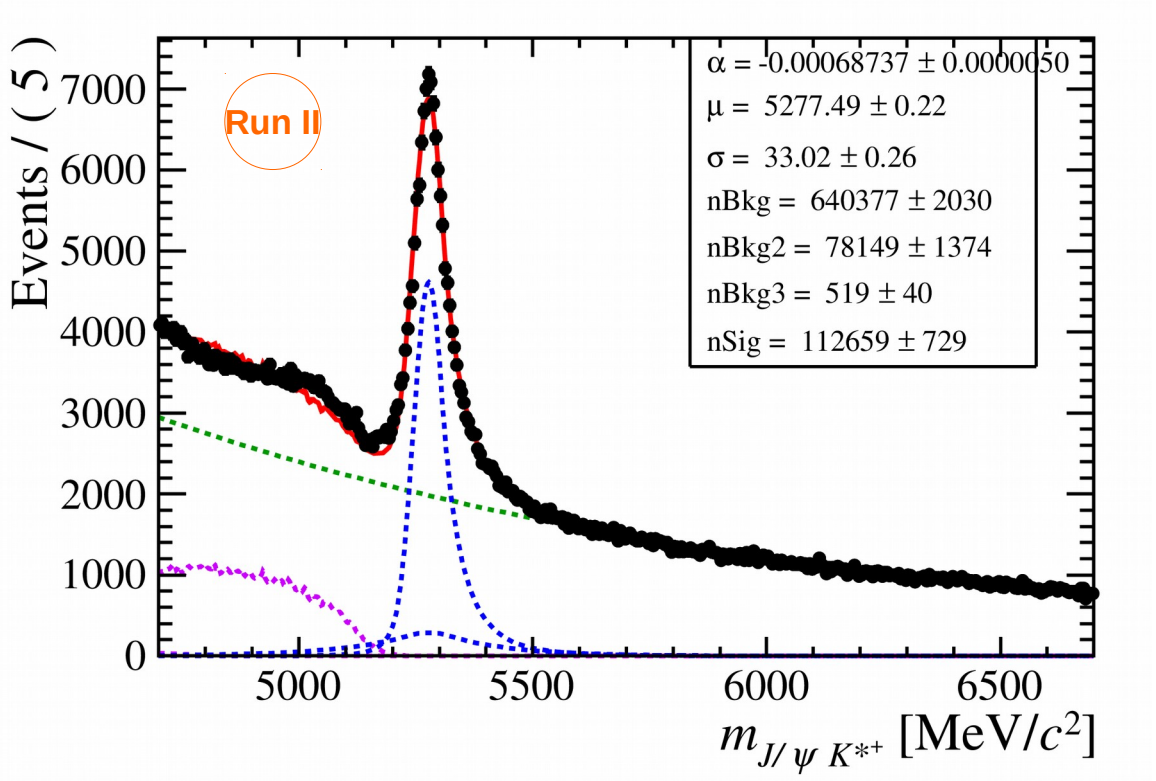
- FullSim, no mass constrain:
  - Signal σ = 41.5 MeV/c^2
- RapidSim, after tuning calo constants on other decays:
  - Signal σ = 39.05 MeV/c^2
- $\rightarrow$ Peak width compatible (but no pile-up tails in RapidSim)





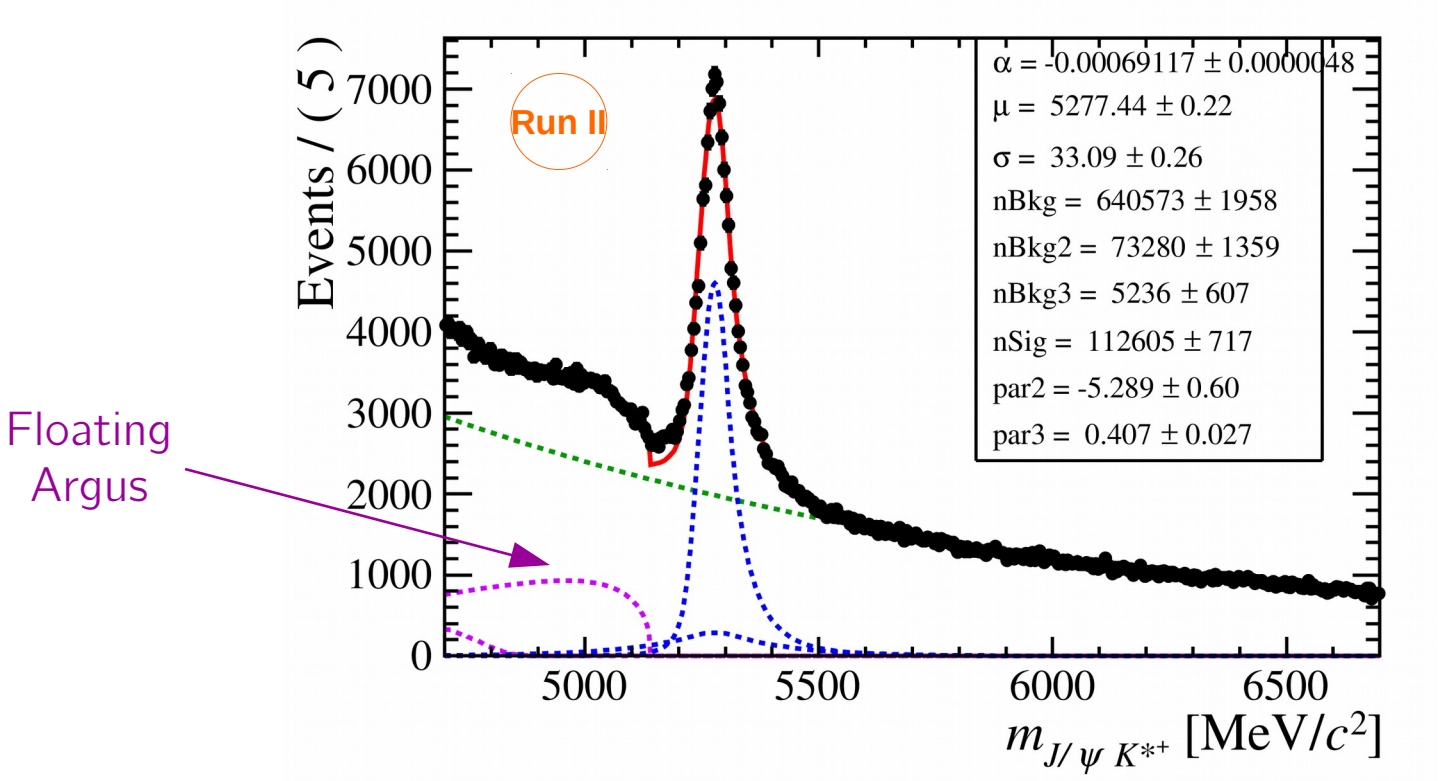
Fit result

- Some parameters:
  - N_S = 93.8x10^3 signal B^+ (good enough for BDT training)
  - Signal σ = 33.0 MeV/c^2 VS 28.9 MeV/c^2 in full-sim MC
- Part. reco'ed RapidSim B^0 shape to be improved (Argus⊗Gauss)



Fit result

- Some parameters:
  - N_S = 93.8x10^3 signal B^+ (good enough for BDT training)
  - Signal σ = 33.0 MeV/c^2 VS 28.9 MeV/c^2 in full-sim MC
- Part. reco'ed RapidSim B^0 shape to be improved (Argus⊗Gauss)



Next

1) Fit B^+ mass @ various BDT cuts
   - Compare efficiency: MC VS real data
   - Quantify discrepancy (can we leave with it?)
2) Alternative: train BDT on sWeighted data signal
   - sPlot B^+ mass
   - Use sWeights in TMVA
3) Apply BDT to other modes
   - If 1) $\rightarrow$ train on signal MC
   - If 2) $\rightarrow$ find a way to export JpsiK* BDT