Observation of $D_{s0}(2317)^+ \to D_s^{*+}\gamma$ at Belle (II)

52nd B2GM

Yubo Li[1], Chengping Shen[2]
1. Xi'an Jiaotong University
2. Fudan University

2025/10/15

Motivation:

$$D^*_{s0}(2317)^{\pm}$$

Observation of a narrow meson decaying to $D_s^+\pi^0$ at a mass of 2.32-GeV/c^2 #8

BaBar Collaboration • B. Aubert (Annecy, LAPP) et al. (Apr, 2003)

Published in: *Phys.Rev.Lett.* 90 (2003) 242001 • e-Print: hep-ex/0304021 [hep-ex]

- Mass of $\mathrm{D}_{s0}(2317)^+$ is much lower than the quark model predictions of the lowest $c\bar{s}$ mesons with the corresponding J^P quantum number

  - ➢ modifying the $c\bar{s}$ quark model
  - ➢ D^*K hadronic molecule
  - ➢ compact tetraquarks
  - ➢ chiral partners of the ground state D_s^* meson

Partial decay widths:
unique in discriminating between various models

Can we find any other decay channel?

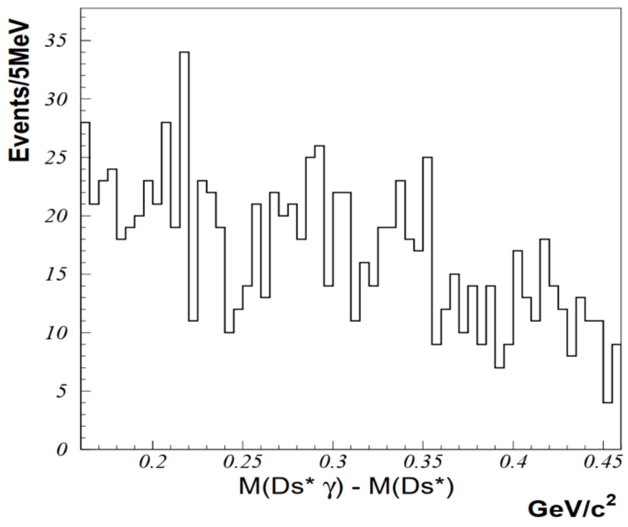
	Mode	Fraction (Γ_i/Γ)	Scale Factor/ Conf. Level	P(MeV/c)
Γ_1	$D_s^+\pi^0$	$(100^{+0}_{-20})\%$		298
Γ_2	$D_s^+\gamma$	$< 5\%$	CL=90%	323
Γ_3	$D_s^*(2112)^+\gamma$	$< 6\%$	CL=90%	
Γ_4	$D_s^+\gamma\gamma$	$< 18\%$	CL=95%	323
Γ_5	$D_s^*(2112)^+\pi^0$	$< 11\%$	CL=90%	
Γ_6	$D_s^+\pi^+\pi^-$	$< 4\times10^{-3}$	CL=90%	194
Γ_7	$D_s^+\pi^0\pi^0$	not seen		205

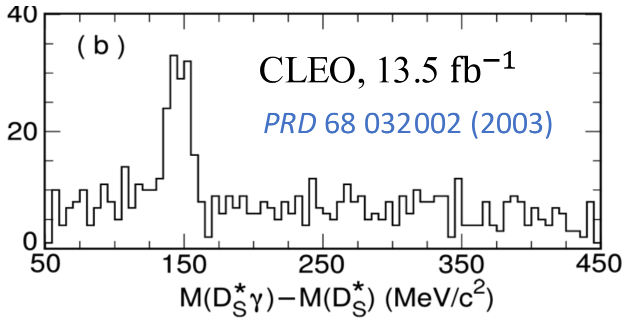
Motivation:

- The width of $D_s(2317)^+$ is unknown: $\Gamma(D_s(2317)^+) < 3.8$ MeV @95% CL.
- Precise absolute branching fraction is unknown: $\mathcal{B}(D_s(2317)^+ \to D_s^+\pi^0) = (100^{+0}_{-20})\%$
  - Now, Belle can only measure the branching fraction ratio $\mathcal{R} = \frac{\mathcal{B}(D_s(2317)^+ \to D_s^{*+}\gamma)}{\mathcal{B}(D_s(2317)^+ \to D_s^+\pi^0)}$
- The radiative decay of $D_{s0}(2317)^+$ has not been founded.
  - $\mathcal{R} < 0.059$ @ 90% CL from CLEO, $\mathrm{N}(\mathrm{D_s}(2317)^+ \to \mathrm{D_s^{*+}}\gamma) = -6.5 \pm 5.2$

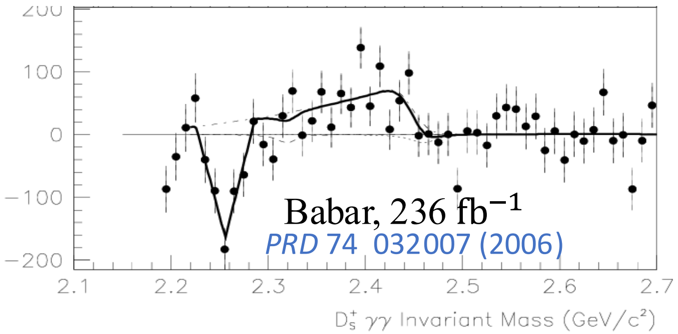
$D_{s0}(2317)^+$ Decay mode	J^P transition	Possible wave
$D_s^+\gamma$	$0^+ \to 0^-1^-$	Forbidden
$D_s^{*+}\gamma$	$0^+ \to 1^-1^-$	$L = 0, 2$

Belle, 87 fb^{-1} *PRL* 92 012002 (2004)





CLEO, 13.5 fb^{-1}
PRD 68 032002 (2003)



Babar, 236 fb^{-1}
PRD 74 032007 (2006)

Motivation:

- Limited predictions on both the decay widths of $D_s(2317)^+ \to D_s^{*+}\gamma$ and $D_s(2317)^+ \to D_s^+\pi^0$

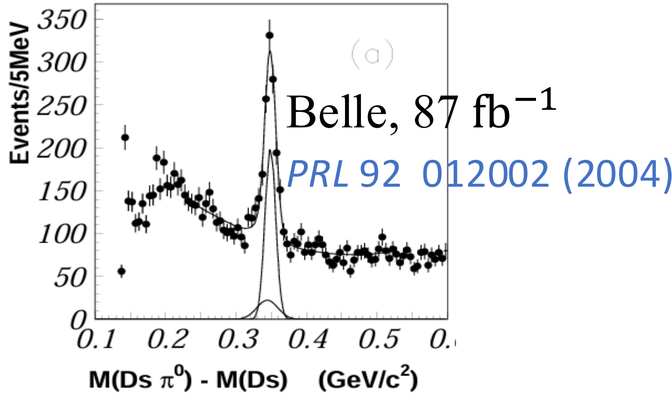
Table I: The theoretical predictions of the $\mathcal{B}(D_{s0}^*(2317)^+ \to D_s^{*+}\gamma)/\mathcal{B}(D_{s0}^*(2317)^+ \to D_s^+\pi^0)$ ($\mathcal{R}$).

Reference	$D_{s0}^*(2317)^+$ Structure	Approach	Radiative Decay Width (keV)	Hadronic Decay Width (keV)	$\mathcal{R}(\%)$
Godfrey [40]	$c\bar{s}$ state	quark model	1.9	$\sim$10	19
Ke [41]		light front quark mode	1.71 ± 2.9	-	$> (0.45 \pm 0.11)$
Bardeen [18]		chiral theory	1.74	21.5	8.1
Faessler [42]	molecule state	effective Lagrangian	$0.55 \sim 1.41$	79.6 ± 33	$0.5 \sim 3.0$
Fu [44]		heavy quark flavor symmetry	3.7 ± 0.3	132 ± 7	2.8 ± 0.9
Lutz [45]		chiral theory	1.94 or 6.47	140	1.35 or 4.57

Larger value

Larger value

$$\Bigg\}$$ **Around 3%**



$$N^{\text{fit}}(D_{s0}(2317) \to D_s\pi^0) = 761 \pm 44 \ (87\ \text{fb}^{-1})$$

$$\boldsymbol{N^{\text{expected}}(D_{s0}(2317) \to D_s^*\gamma) \approx 150\sim600} \ (1.4\ \text{ab}^{-1})$$

Based on Belle measurement ,
assuming the same detection efficiency.

Decay chain: $e^+e^- \to c\bar{c} \to D_{s0}(2317)^+ + \text{ anythings}$

$D_{s0}(2317)^+ \to D_s^+[\phi\pi, K^*K]\,\pi^0$ **Reference channel**

$\to D_s^{*+}[D_s^+\gamma]\gamma$ **Same final states**

Datasets:

- All Belle data: 980 fb^{-1}
- All Belle II data: 428 fb^{-1}

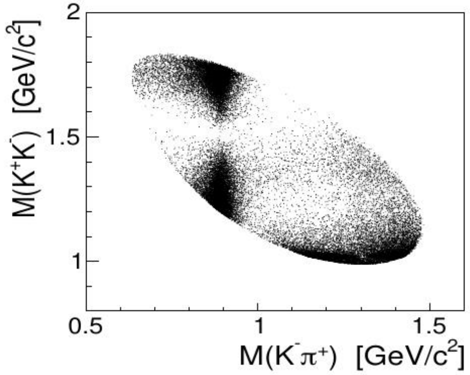
  LS1 dataset

Generic MC:

- Belle generic MC:
  - 4 streams of $B\bar{B}$, (non-)$B_s\bar{B}_s$ and $q\bar{q}$ samples
- Belle II generic MC:
  - All $q\bar{q}$ and $B\bar{B}$ MC15rd samples

    699 fb^{-1} for proc13

    745 fb^{-1} for prompt



Pre-selection:

Tracks:

- |dr| < 0.5 cm, |dz| = 3 cm
- π : binaryPID(π/K) > 0.6
- K: binaryPID(K/π) > 0.6

Photon:

- $E > 100\ MeV$, E9/E25 > 0.8
- π^0 veto: the invariant mass of any combination of one signal γ_{sig} with one γ_{roe} in rest of event should not in the region of $\left|M(\gamma_{sig}\gamma_{roe}) - m(\pi^0)\right| < 15\ \text{MeV}/c^2$

Signal MC:

- $e^+e^- \to c\bar{c}$: Pythia
- $D_s(2317)^+ \to D_s^+\pi^0, D_s^{*+}\gamma$: PHSP (S-Wave)
- $D_s^{*+} \to D_s^+\gamma$: VSP
- $D_s^+ \to KK\pi$: Dalitz

$$D_s(2317)^+ \to D_s^{*+}\gamma$$

Most selection criteria are optimized based on signal mode

Optimized using Punzi FoM $\frac{\varepsilon}{5/2+\sqrt{B}}$

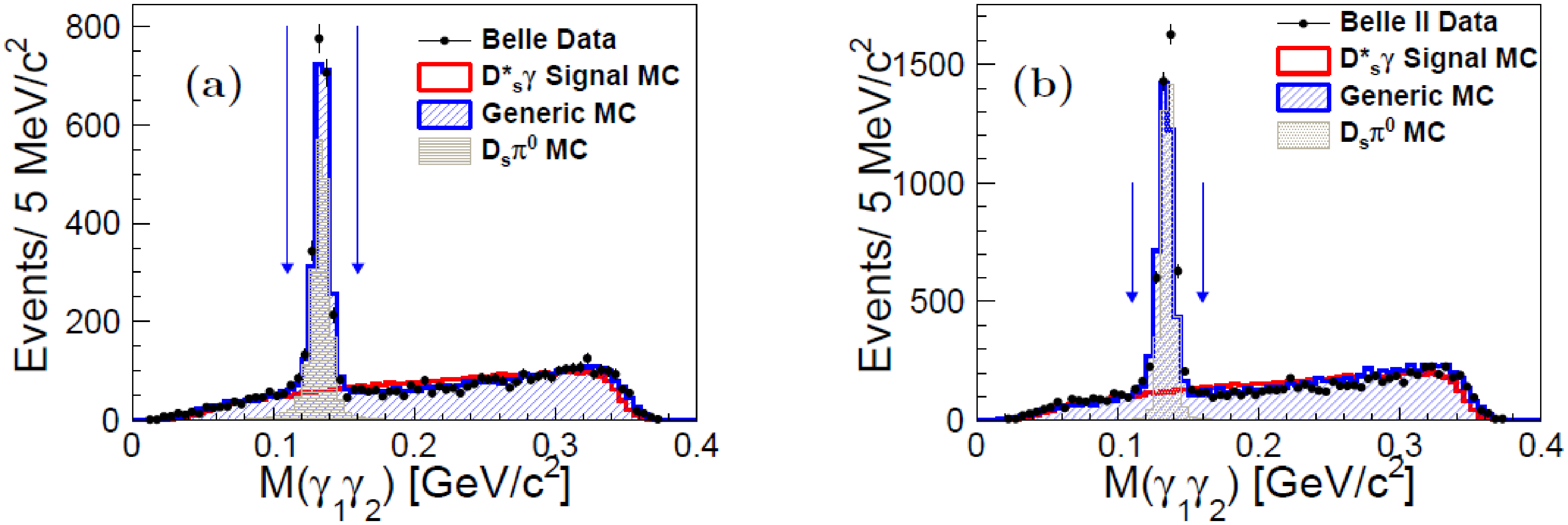
ε: Detection efficiency

B: Background number of $D_s(2317)^+$ sideband of data $2.35 < M(D_s^*\gamma) < 2.40\ \mathrm{GeV}/c^2$

γ selection: $D_s(2317)^+ \to [D_s^{*+} \to D_s^+\gamma_2]\gamma_1$

Here and after,

- Generic MC distributions are stacked and scaled to data yields
- Signal MC distributions are randomly scaled
- All selection criteria are applied, excluding the plotted variable
- Distributions are mainly from the signal region $2.29 < M(D_s^*\gamma/D_s\pi^0) < 2.34\ \mathrm{GeV}/c^2$



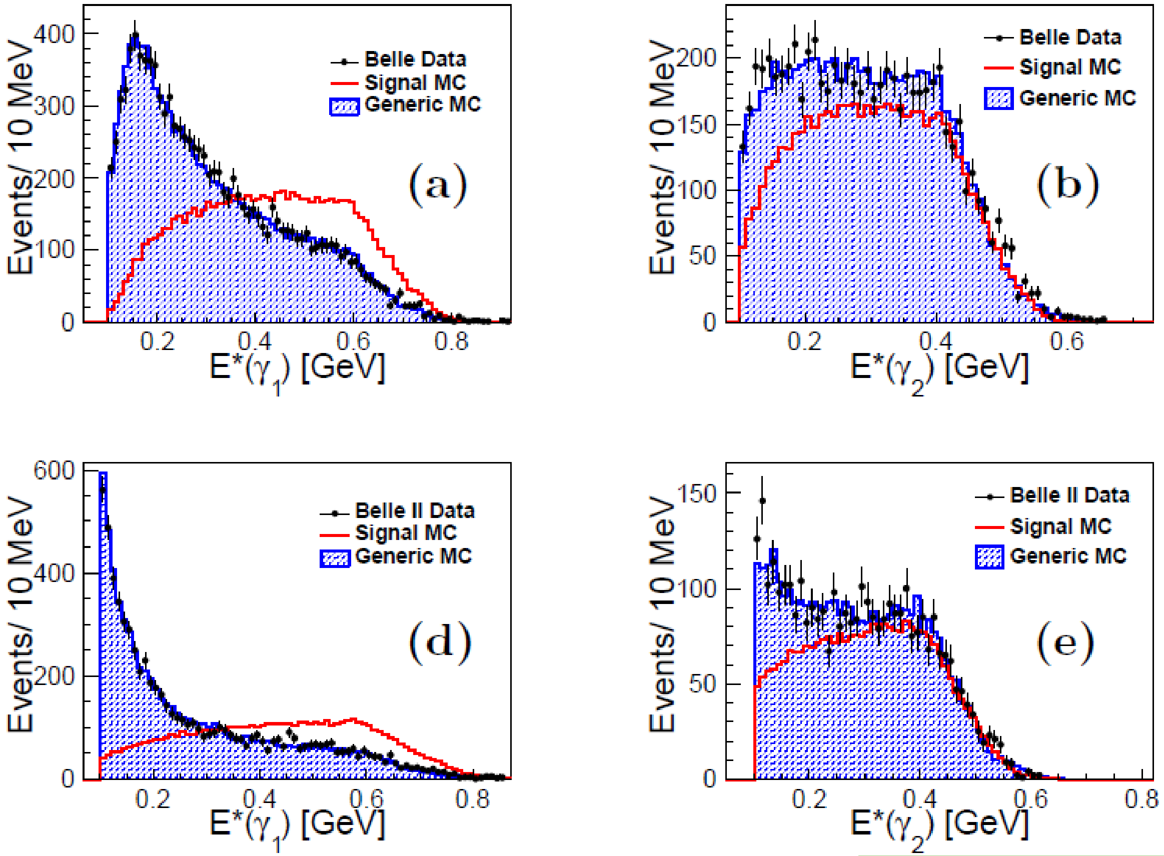
To suppress the contributions from the $\mathbf{D_s(2317)^+ \to D_s^+\pi^0}$ decay

Veto events in the region:

$$0.100 < |M(\gamma_1\gamma_2)| < 0.160\ \mathrm{GeV}/c^2$$

γ selection:

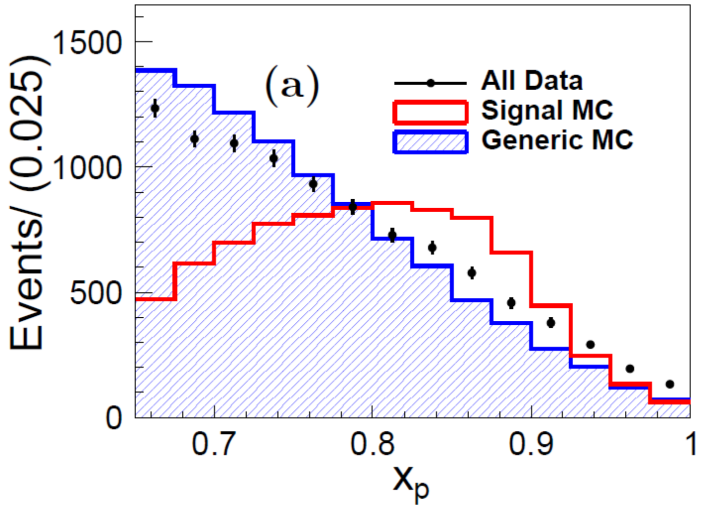
$D_s(2317)^+ \to [D_s^{*+} \to D_s^+ \gamma_2]\gamma_1$

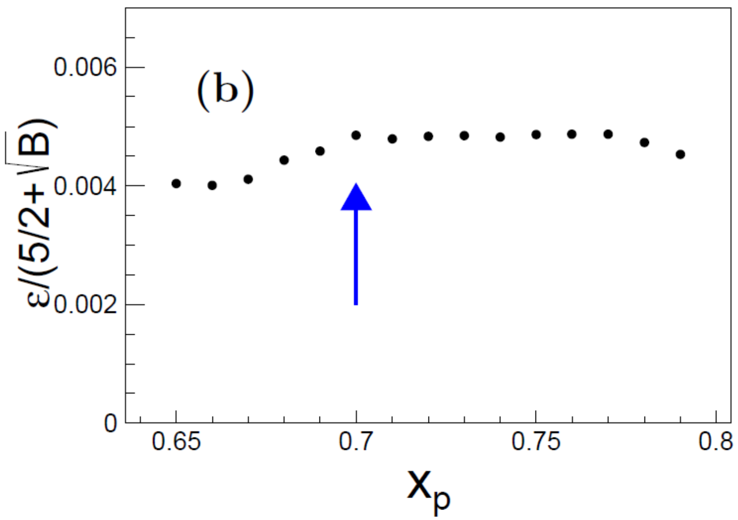


Distributions of the energy of the photons in e^+e^- center of mass system (c.m.s.)

Apply:
$E^*(\gamma_1) > 0.22$ GeV
$E^*(\gamma_2) > 0.10$ GeV

x_p Selection



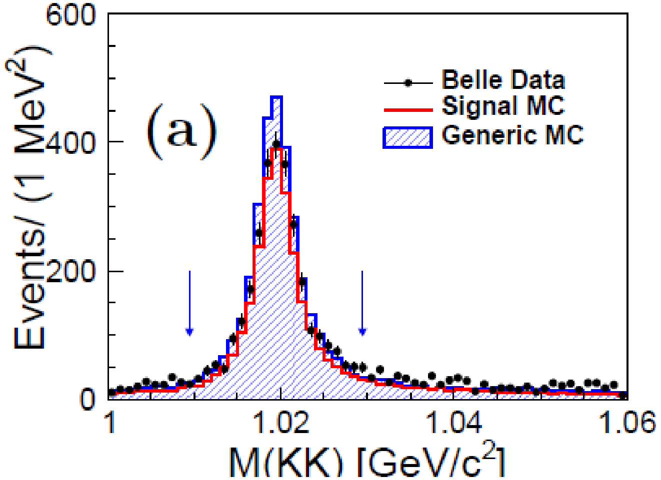


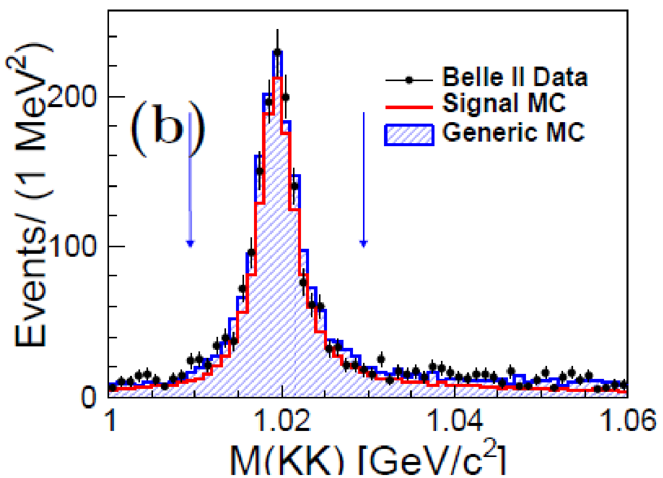
$$xp = \frac{p^*_{D_s^*\gamma}}{\sqrt{E_{beam}^2 - M_{D_s^*\gamma}^2}}$$

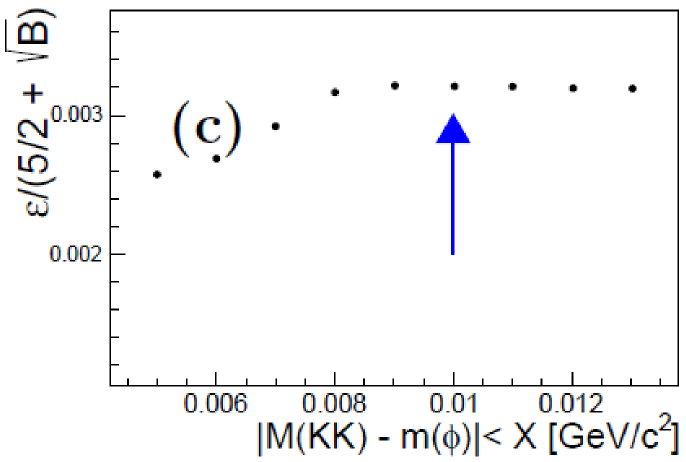
- Only background from $e^+e^- \to q\bar{q}$ survive.
- Apply cut: $x_p > 0.7$ according to Punzi FoM.
- The x_p distributions of signal MC samples are reweighted according to real data of $D_s(2317)^+ \to D_s^+\pi^0$ mode. (shown in page 19)

ϕ selection:

Mass window:
$\left|M_{K^+K^-} - m_\phi\right| < 10\ \mathrm{MeV}/c^2$

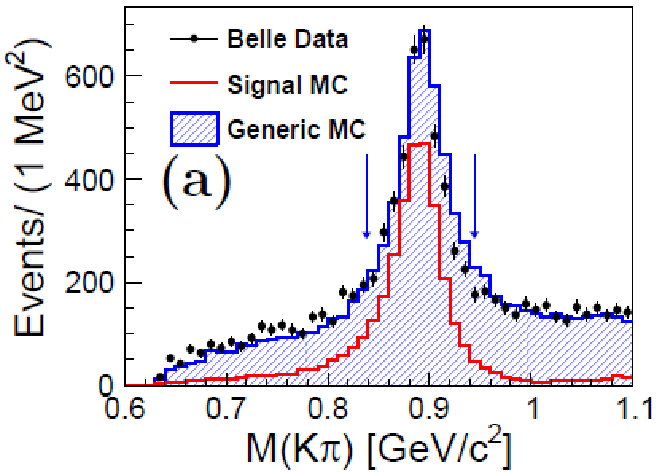


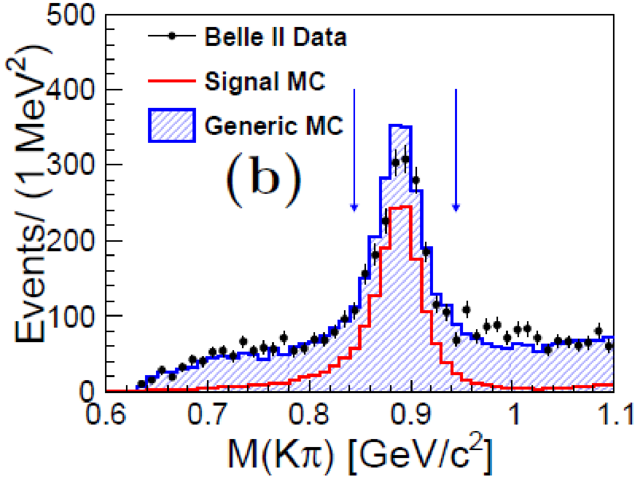


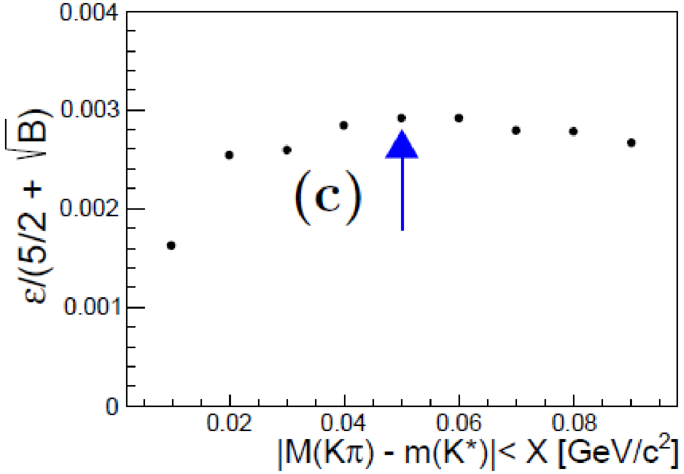


$\bar{K}^{*0}$ selection:

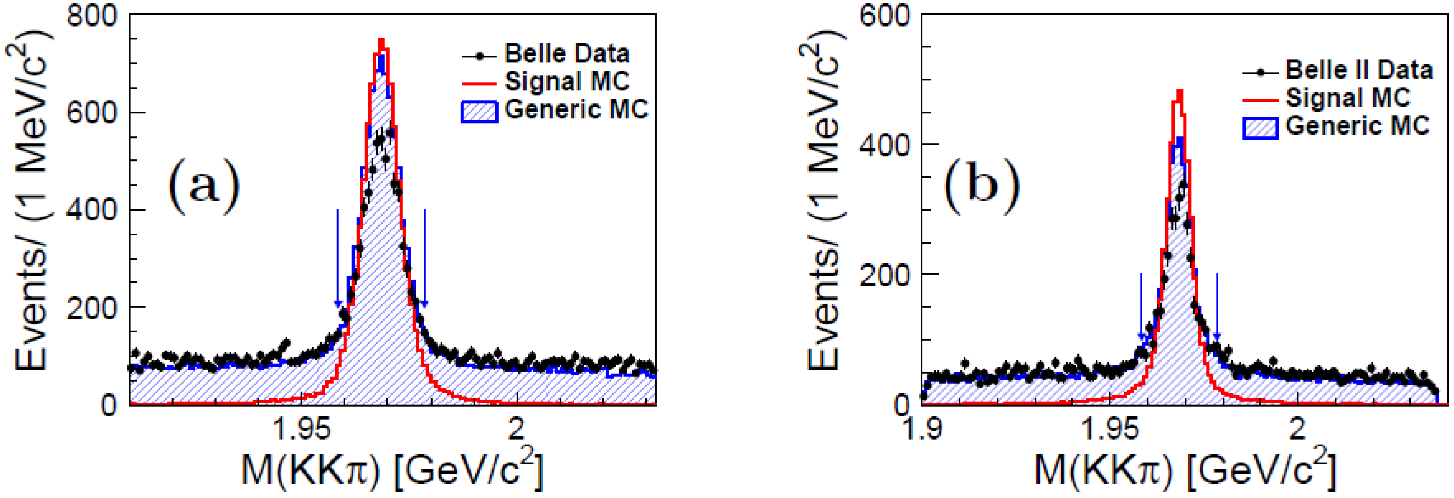
Mass window:
$|M(K\pi) - m(\bar{K}^*)| < 50\ \mathrm{MeV}/c^2$







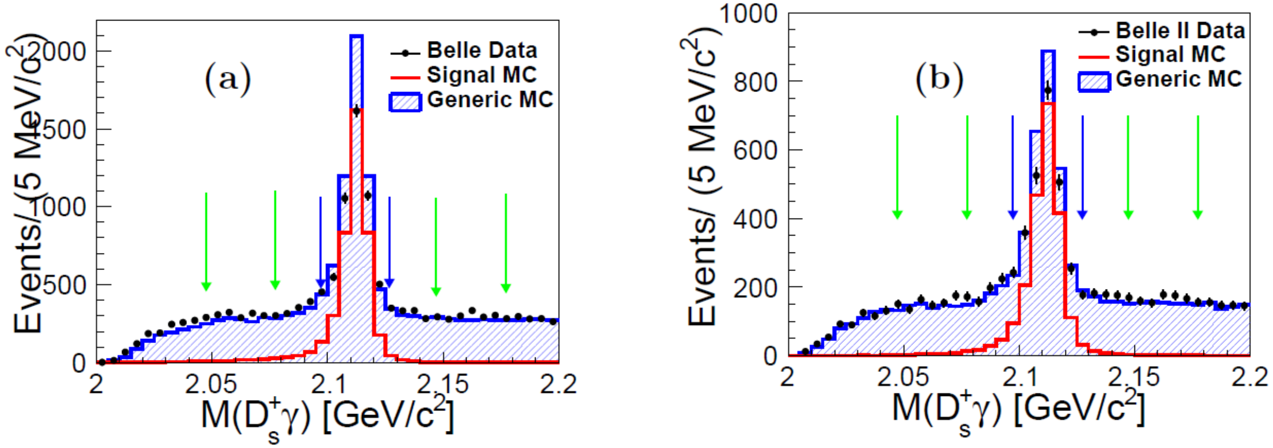
D_s^+ and D_s^{*+} selection:

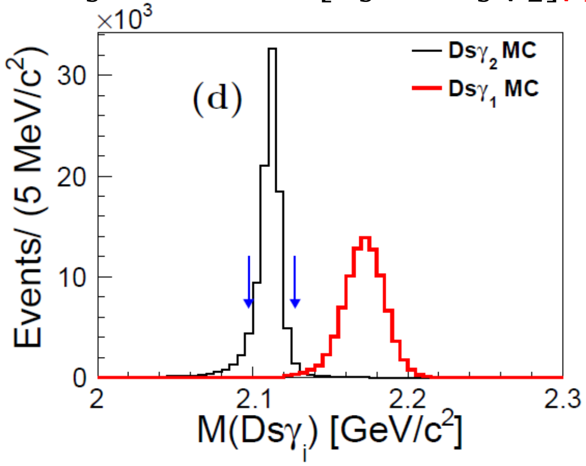


Mass window :
$|M(K^+K^-\pi^+) - m(D_s^+)| < 10\ \mathrm{MeV}/c^2$

To reduce the resolution of $m(D_s^+)$: $M(D_s^+\gamma) = M^{rec}(D_s^+\gamma) - M(K^+K^-\pi^+) + m(D_s^+)$

$D_s(2317)^+ \to [D_s^{*+} \to D_s^+\gamma_2]\gamma_1$



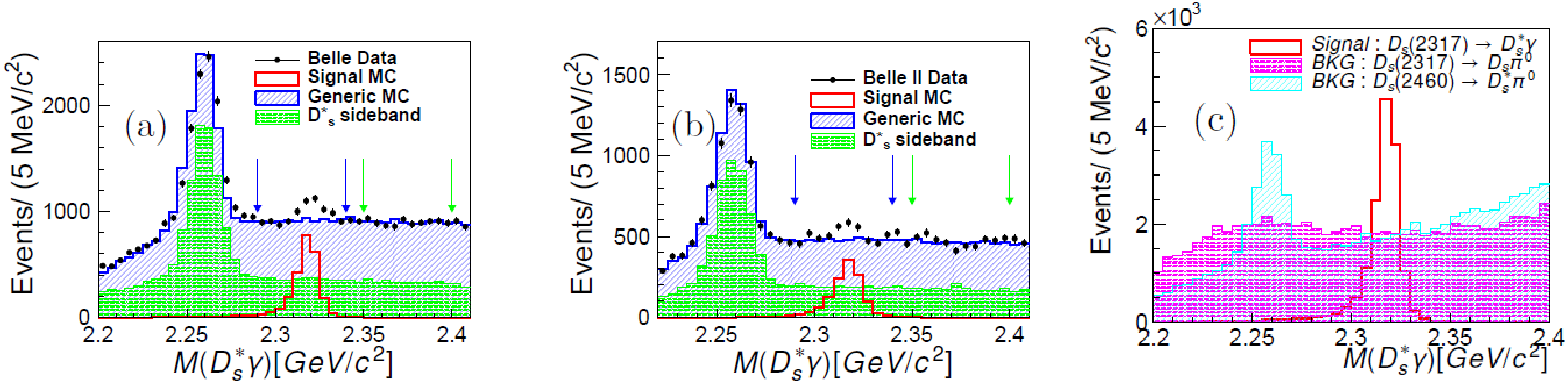


Sideband:
$|M(D_s^+\gamma) - m(D_s^{*+}) \pm 0.05| < 0.015\ \mathrm{GeV}/c^2$

Signal window:
$|M(D_s^+\gamma) - m(D_s^{*+})| < 0.015\ \mathrm{GeV}/c^2$

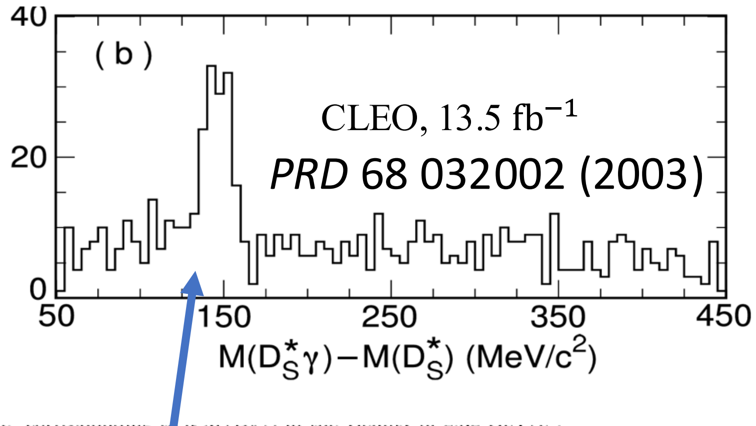
$D_s(2317)$ distributions:

$D_s(2317)^+ \rightarrow [D_s^{*+} \rightarrow D_s^+ \gamma_2]\gamma_1$



To reduce the resolution of $\mathbf{M}(\boldsymbol{D_s^{*+}})$:

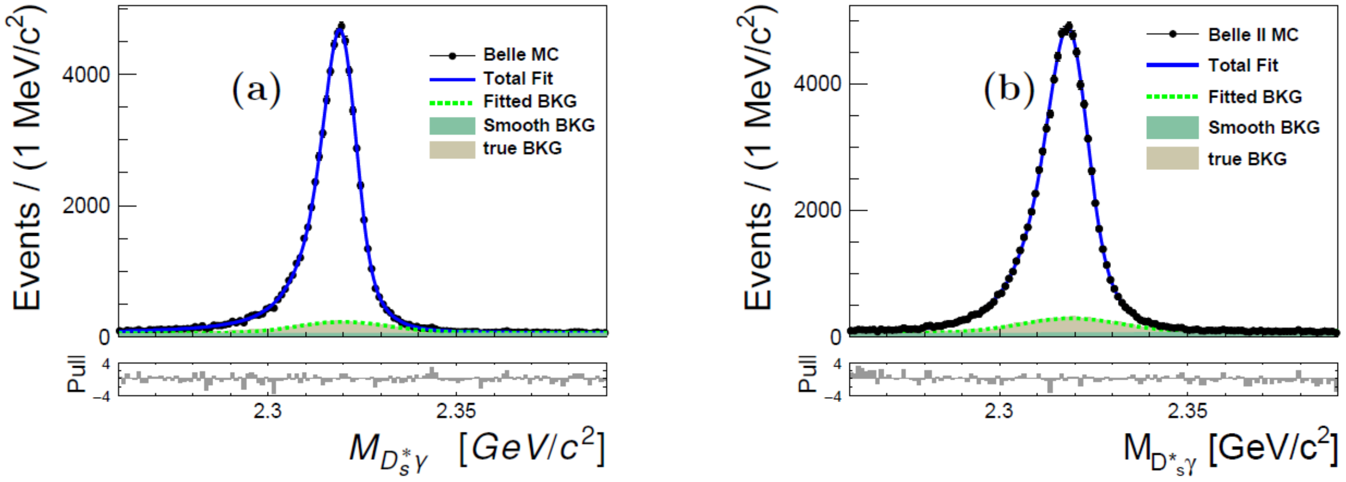
$$M(D_s^{*+}\gamma) = M^{rec}(D_s^{*+}\gamma_1) - M^{rec}(D_s^{+}\gamma_2) + m(D_s^{*+})$$



The same conclusion holds for the $D_s^*\gamma$ final state, shown in Fig. 3(b), where we combine selected D_s^* candidates with photons of energy above 150 MeV. The peak in the $\Delta M(D_s^*\gamma)$ spectrum in Fig. 3(b) near 150 MeV/c^2 is due to real $D_s^{*+} \rightarrow D_s^+\gamma$ decays in which a random photon has been combined with the D_s^+ candidate to form the D_s^* candidate, and the actual photon from this transition is combined with this system to form the D_{sJ}^* candidate. There

- **Clear $D_s(2317)$ signal peaks**
- **Flat background:** from both D_s^* sideband and generic MC
- Peaks around 2.26 GeV/c^2 are from $D_s^* + \gamma_{\text{free}}$ (same as CLEO)

$D_s(2317)$ MC distributions



- Signal: Crystal Ball convoluted with triple-Gaussian
- Broken Signal: asymmetry Gaussian
- Smooth Background: 2nd order polynomial
- All parameters free to float in the fits

Experiments	Belle	Belle II
N^{fit}	$(6.44 \pm 0.02) \times 10^4$	$(8.00 \pm 0.03) \times 10^4$
Truth yields	$64,508$	$80,154$

The IO checks are OK for the signal MC samples

$$D_s(2317)^+ \to D_s^+ \pi^0$$

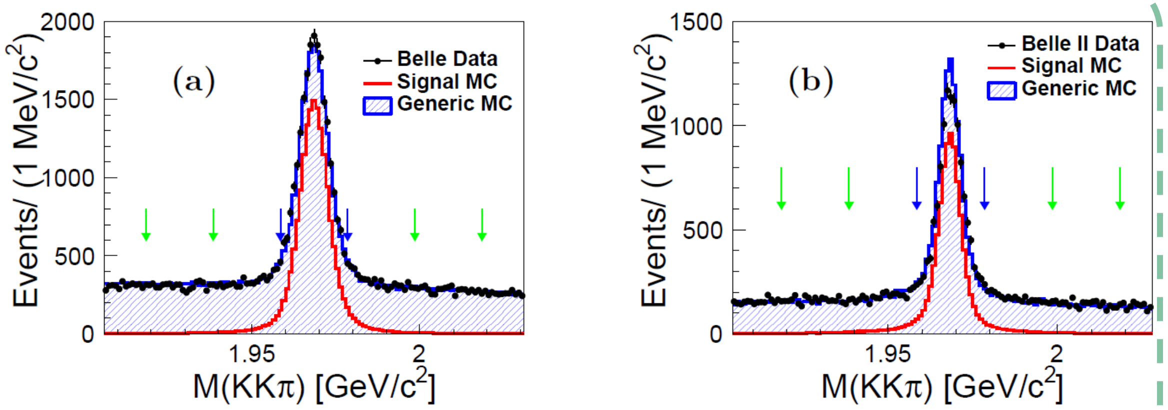
Reference channel:

- Used to validation of the signal selections
- Correct the x_p distribution

Main event selections:

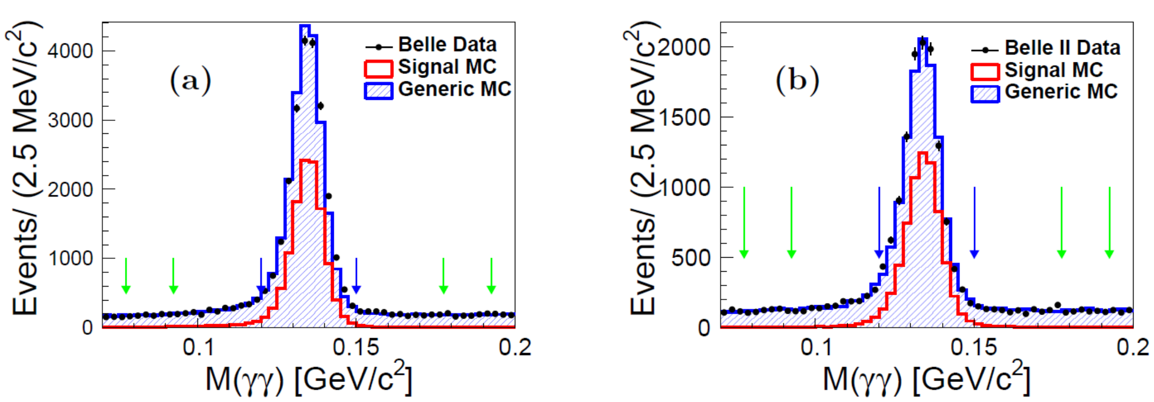
- $x_p = p^*_{D_s^+\pi^0} / \sqrt{E^2_{beam} - M^2_{D_s^+\pi^0}}$ >0.7
- The D_s^+ Selection is the same as the $D_s^{*+}\gamma$ mode

- **The photon selection:**
  - $E^* > 0.22$ GeV for one photon while $E^* > 0.1$ GeV for another
  - Veto $|M(\gamma_{roe}\gamma_i) - m(\pi^0)| < 0.015\ \mathrm{GeV}/c^2$ for signal γ_i
  - π^0 mass window: $|M(\gamma\gamma) - m(\pi^0)| < 0.015\ \mathrm{GeV}/c^2$



Signal window: $|M(K^+K^-\pi^+) - m(D_s^+)| < 0.01\ \mathrm{GeV}/c^2$

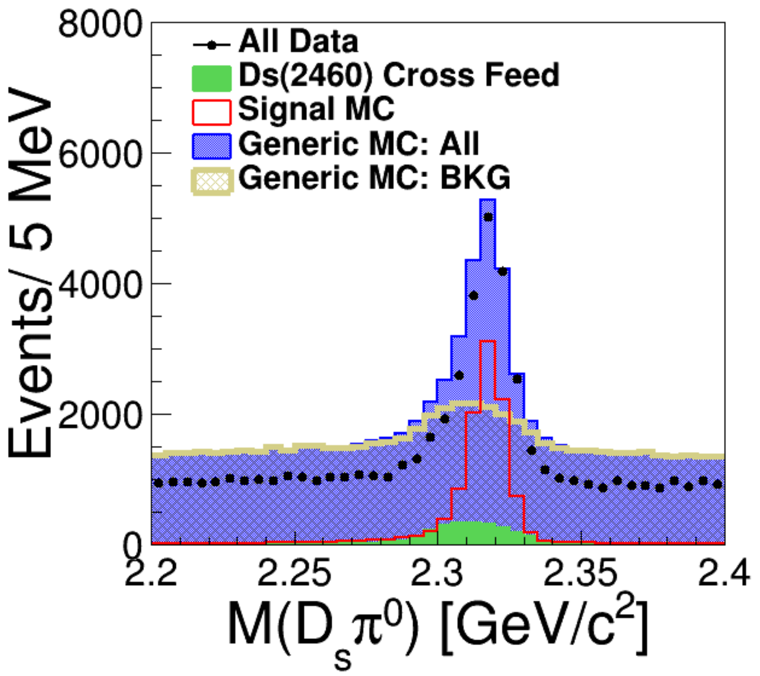
Sideband: $|M(K^+K^-\pi^+) - m(D_s^+) \pm 0.04| < 0.01\ \mathrm{GeV}/c^2$

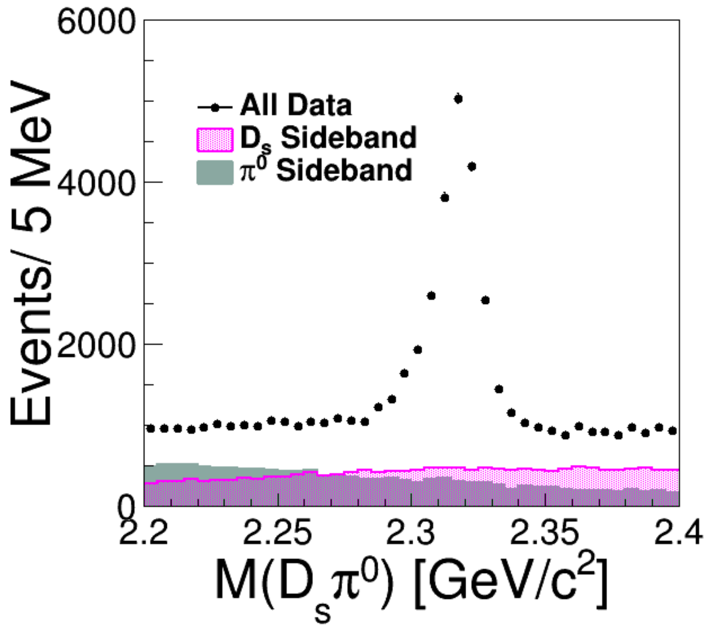


Signal window: $|M(\gamma\gamma) - m(\pi^0)| < 0.015\ \mathrm{GeV}/c^2$

Sideband: $|M(\gamma\gamma) - m(\pi^0) \pm 0.05| < 0.0075\ \mathrm{GeV}/c^2$

$M\left(D_s^+\pi^0\right)$ distributions: combined Belle and Belle II samples



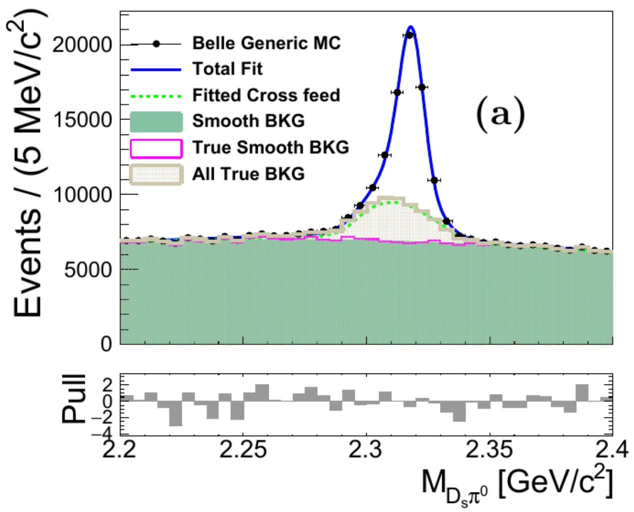


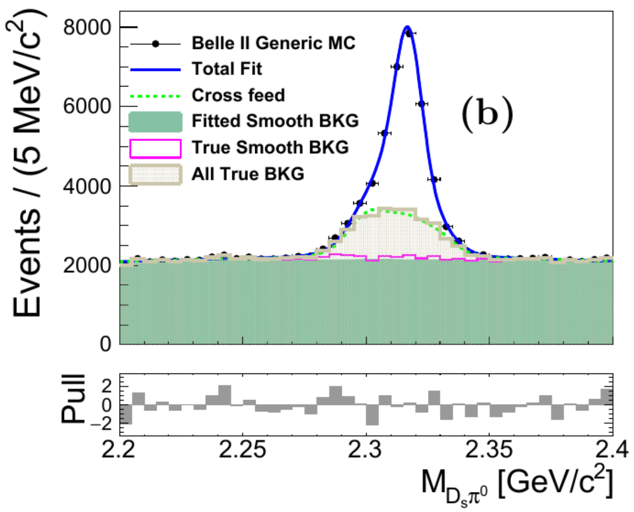
To reduce the resolution of $M(D_s^+)$ and $M\left(\pi^0\right)$:

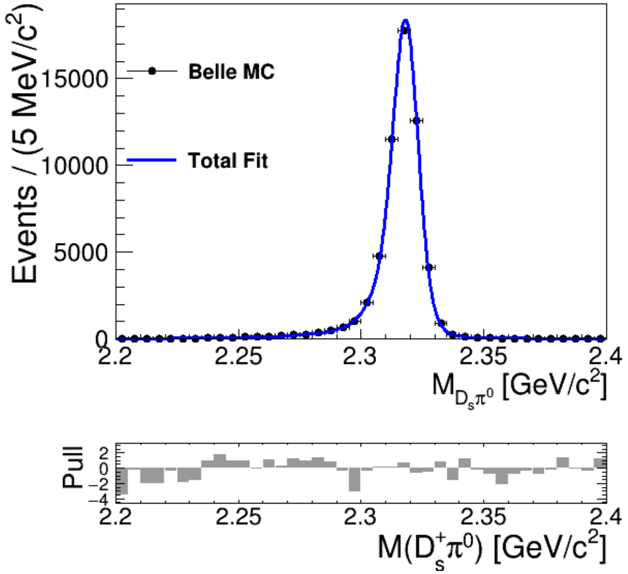
$$M(D_s^+\pi^0) = M^{rec}(D_s^+\pi^0) - M^{rec}(KK\pi) + m(D_s^+) - M^{rec}(\gamma\gamma) + \mathrm{m}(\pi^0)$$

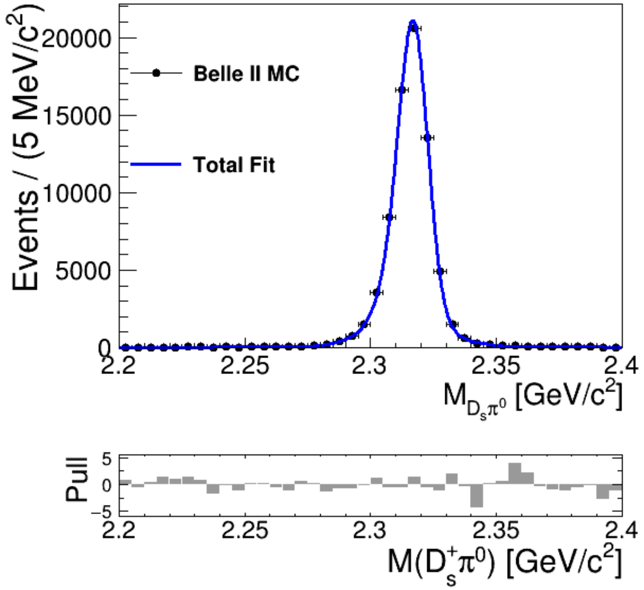
- **Flat background** from D_s^+ sideband
- **Flat background** from π^0 sideband: negligible $D_s(2317)^+ \to D_s^{*+}\gamma$ contribution
- **Small wide Peak BKG**: cross feed from $D_s(2460)^+ \to D_s^{*+}\pi^0$, (i.e. missing one photon)

$D_s(2317)$ signal extraction: generic MC









Signal MC: Crystal Ball convoluted by triple-Gaussian (**parameters fixed** to those gotten from fits to signal MC samples, except for the mean values)

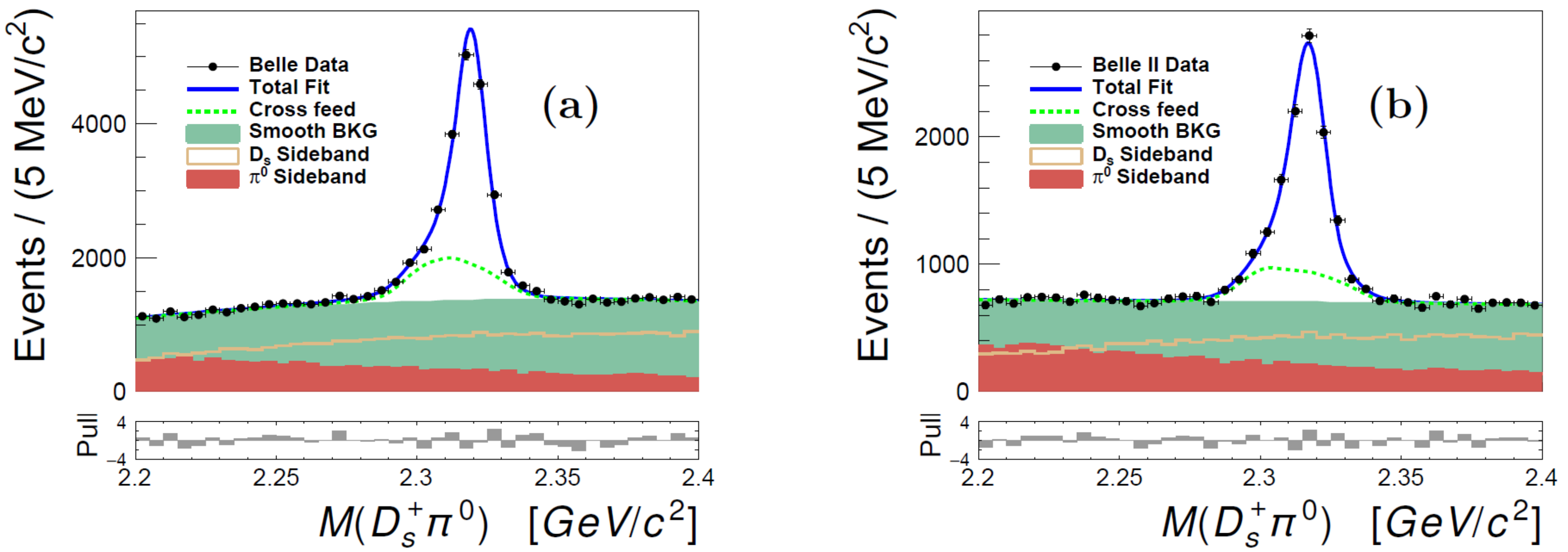
Cross feed: MC histogram shape.

Smooth BKG: 2nd order polynomial (**parameters floated**)

	Fitted yield of $D_s(2317)$	True number of event
Belle	$(4.546 \pm 0.055) \times 10^4$	45,724
Belle II	$(1.821 \pm 0.031) \times 10^4$	17,917

Signal extraction: Data

Unbinned maximized likelihood fits to $M(D_s^+\pi^0)$ spectra



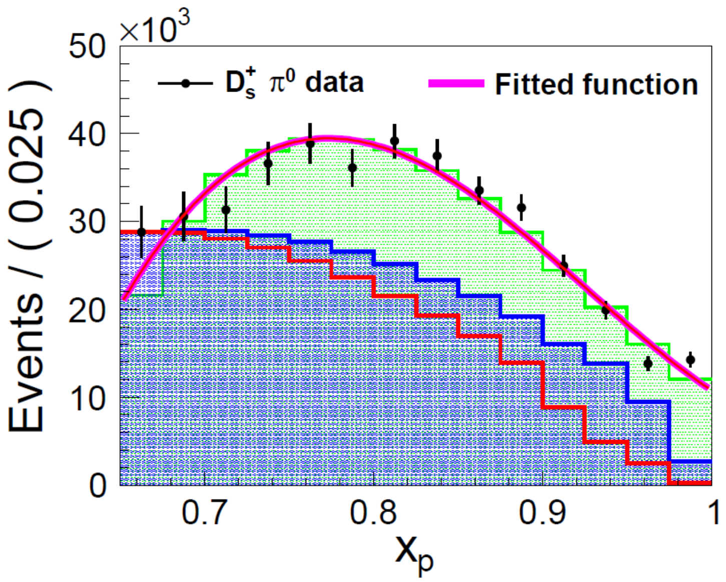
Same method as the previous fit to generic MC sample

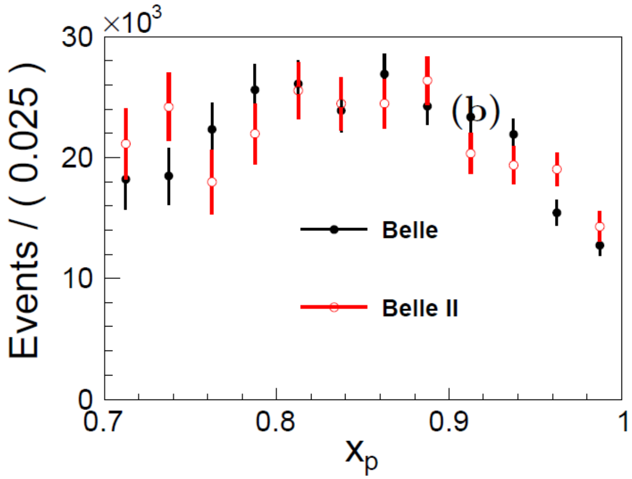
Experiments	Belle	Belle II
N^{fit}	$(1.082 \pm 0.023) \times 10^4$	$(6.108 \pm 0.163) \times 10^3$
Significant	$> 10\ \sigma$	

The correction of x_p distribution

Dots:
Efficiency-corrected x_p from combined Belle and Belle II data

Histograms:
Simulated x_p distributions from
- ◆ Original Belle MC samples
- ◆ Original Belle II MC samples
- ◆ Reweighted MC samples

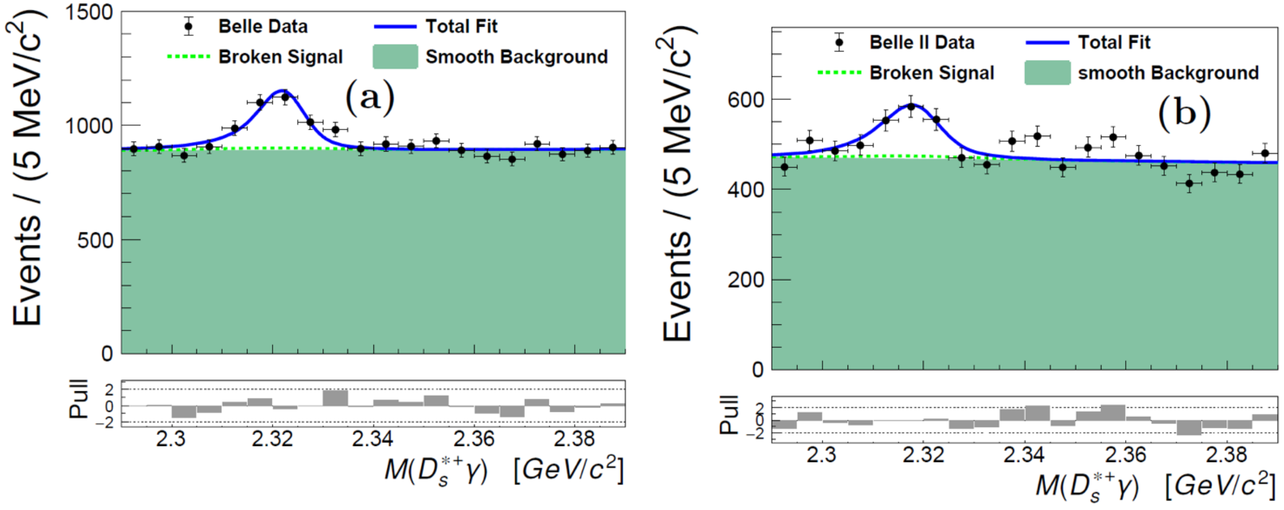




- Yields of $D_s(2317)$ in each bin are extracted with the same fit method.
- Efficiency-corrected x_p distributions from Belle II are scaled to Belle's yields
- The simulated x_p distributions of signal MC are wrong
  - ➢ corrected by reweighting method: consistent with data
  - ➢ **To get correct detection efficiencies and signal shapes.**

Branching fraction ratio extraction:

Unbinned maximized likelihood Simultaneous Fit to $M(D_s^{*+}\gamma)$ spectra from Belle and Belle II



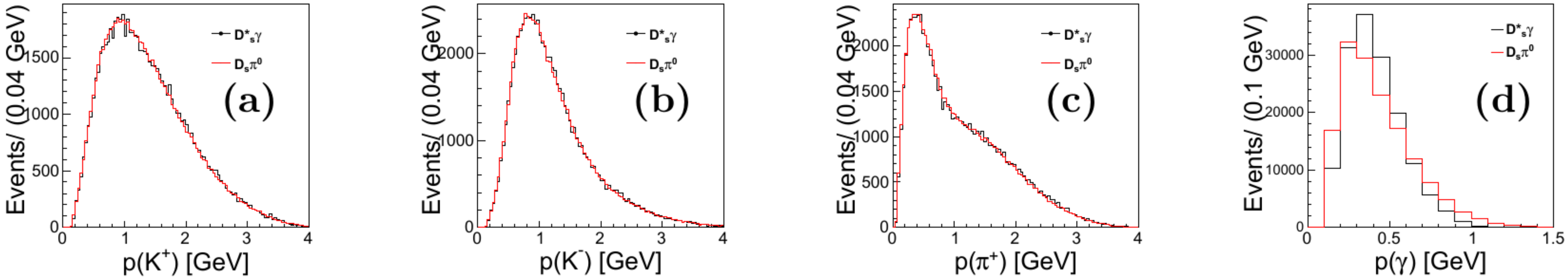
Experiment	Decay mode	N^{fit}	ε	$\mathcal{R}$ separate fits	$\mathcal{R}$ simultaneous fit
Belle	$D_s^{*+}\gamma$	741.5 ± 82.4	4.16%	$(7.43 \pm 0.83)\%$	$(7.14 \pm 0.70 \pm 0.23)\%$
	$D_s^{+}\pi^0$	$(1.082 \pm 0.023) \times 10^4$	4.59%		
Belle II	$D_s^{*+}\gamma$	348.2 ± 69.0	4.60%	$(6.43 \pm 1.27)\%$	
	$D_s^{+}\pi^0$	$(6.108 \pm 0.163) \times 10^3$	5.17%		

$$\mathcal{R}(\text{exp}) = \frac{N^{\text{fit}}_{\text{radiative}}\varepsilon_{\text{hadronic}}}{N^{\text{fit}}_{\text{hadronic}}\varepsilon_{\text{radiative}}}$$

- Signal shapes are the same the previous fits to MC (page 13)
  - Only mean values of CB functions are free to be float
  - Ratios between signal and broken signal yields are fixed
- $N_{\text{exp}}(D_s^{*+}\gamma) = \mathcal{R}\dfrac{N^{\text{fit}}_{\text{exp}}(D_s^{+}\pi^0)\varepsilon_{\text{exp}}(D_s^{*+}\gamma)}{\varepsilon_{\text{exp}}(D_s^{+}\pi^0)}$, exp represents Belle or Belle II.
- Signal shapes, ratios between signal and broken shape are fixed.
- Background: 1$^{\text{st}}$ order of polynomial.
- $\mathcal{R} = [7.14 \pm 0.70(stat.) \pm 0.23(syst.)]\%$
- Significance: 10.1σ, with syst. Uncertainty included.

Systematic Uncertainty Study

- The systematic uncertainties due to $\boldsymbol{D_s}$ and **photon** selection, and cut on $\boldsymbol{x_p}$ are cancelled



- **x_p reweighting:**
  - ➢Change the order of the polynomial used to fit x_p distribution
  - ➢Redo the reweighting to get the new detection efficiency ε_{exp}
  - ➢Fit $\mathcal{R}$ again

$$N_{\text{exp}}(D_s^{*+}\gamma) = \mathcal{R}\frac{N_{\text{exp}}^{\text{fit}}(D_s^{+}\pi^{0})\varepsilon_{\text{exp}}(D_s^{*+}\gamma)}{\varepsilon_{\text{exp}}(D_s^{+}\pi^{0})}$$

Uncertainties on $\mathcal{R}$:
Difference in Fit $\mathcal{R}$

- **MC statistic:**
  - ➢Vary ε_{exp} by $\pm\sqrt{(1-\varepsilon_{\text{exp}})/(\varepsilon_{\text{exp}}N)}$, N is the number of the used signal MC events.
  - ➢Fit $\mathcal{R}$ again

Systematic Uncertainty Study

$$N_{\rm exp}(D_s^{*+}\gamma) = \mathcal{R}\frac{N_{\rm exp}^{\rm fit}(D_s^+\pi^0)\varepsilon_{\rm exp}(D_s^{*+}\gamma)}{\varepsilon_{\rm exp}(D_s^+\pi^0)}$$

- **$D_{s0}(2317) \to D_s\pi^0$ mode:**

In the fits to $M(D_s^+\pi^0)$ spectra, By

1. Changing the fit region, order of background polynomial
2. Enlarging the resolution by 1σ (gotten from fits to signal MC)
3. Using the shapes of $M(D_s\pi)$ from $D_s(2460)^+ \to D_s^{*+}\pi^0$ MC directly

➢ Differences in the fitted signal yields: Uncertainties ($\Delta_{\rm exp}^{N(\mathrm{D}_s^+\pi^0)}$) of $\mathrm{N}_{\rm exp}^{\rm fit}(\mathrm{D}_s^+\pi^0)$

Then, in the simultaneous fit to $M(D_s^{*+}\gamma)$ spectra,

By varying the $N_{\rm exp}^{\rm fit}(D_s^+\pi^0)$ based on Gaussian($N_{\rm exp}^{\rm fit}(D_s^+\pi^0)$, $\Delta_{\rm exp}^{N(D_s^+\pi^0)}$) in the fits

Widths of the distributions of fitted values of $\mathcal{R}$:
Uncertainties on $\mathcal{R}$:

Systematic Uncertainty Study

- $D_{s0}(2317) \to D_s^* \gamma$ **mode:**

  In the simultaneous fit to $M(D_s^{*+}\gamma)$ spectra, by:

  1. Changing the fit region, order of background polynomial
  2. Enlarging the resolution by 1σ (gotten from fits to signal MC)
  3. Changing resolutions of broken signal by 2σ
  4. Changing ratios of broken signal to signal yields by 2σ

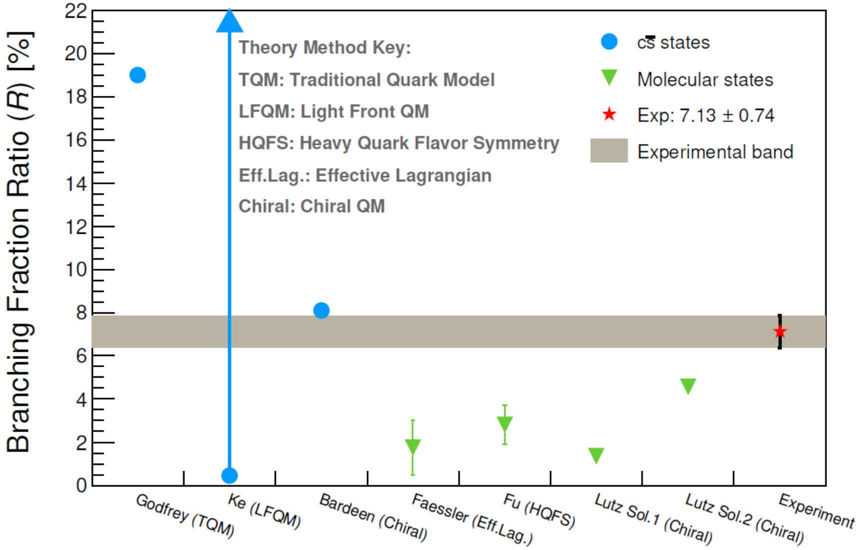
  ➢ Fit $\mathcal{R}$ again

Uncertainties on $\mathcal{R}$:
Difference in Fit $\mathcal{R}$

Source	$D_s^+\pi^0$	$D_s^{*+}\gamma$
Fit region and background pdf	0.8	1.3
Fixed PDF parameters	0.7	2.5
Cross-feed or broken signal	0.6	0.7
x_p reweighting	0.5	
MC sample size	0.5	
Sum	3.2	

Summary

- Using all Belle data (983 fb^{-1}) and Belle II data (427 fb^{-1})
- **Observed $D_{s0}(2317)^+ \rightarrow D_s^{*+}\gamma$ decay for the first time with 10.1σ**
- $\mathcal{R} = \frac{\mathcal{B}(D_s(2317)^+ \rightarrow D_s^{*+}\gamma)}{\mathcal{B}(D_s(2317)^+ \rightarrow D_s^{+}\pi^0)} = [7.14 \pm 0.70(stat.) \pm 0.23(syst.)]\%$
  - **uncertainty dominated by Statistical error of $D_s(2317)^+ \rightarrow D_s^{*+}\gamma$**
  - higher than the molecular assumption for $D_{s0}(2317)^+$
  - possible interpretation: $D_{s0}(2317)^+$ could be the mixture state of pure $c\bar{s}$ state and molecular state



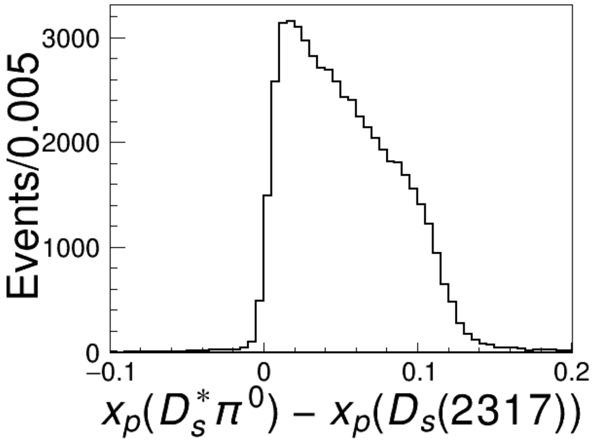
Draft: https://docs.belle2.org/files/4716/BELLE2-PUB-PH-2025-058/6/BELLE2-PUB-PH-2025-058.pdf
Note: https://docs.belle2.org/files/4511/BELLE2-NOTE-PH-2025-027/11/BELLE2-NOTE-PH-2025-027.pdf

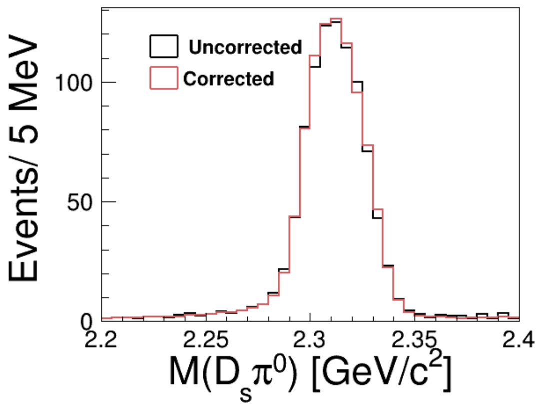
Back up:

Source	$D_s^{*+}\gamma$	$D_s^+\pi^0$
Track	1. $\lvert dr\rvert < 0.5$ cm and $\lvert dz\rvert < 3$ cm 2. $\mathcal{L}_\pi/(\mathcal{L}_K+\mathcal{L}_\pi) > 0.6$ for pions 3. $\mathcal{L}_K/(\mathcal{L}_K+\mathcal{L}_\pi) > 0.6$ for kaons	
Photon	1. $E^*(\gamma_1) > 0.22$ GeV, $E^*(\gamma_2) > 0.1$ GeV 2. $0.10 < \lvert M(\gamma_1\gamma_2)\rvert < 0.16$ GeV$/c^2$ 3. Veto: $\lvert M(\gamma_i\gamma_{\rm roe}) - m(\pi^0)\rvert < 0.015$	1. One of $E^*(\gamma) > 0.22$ GeV, another $E^*(\gamma) > 0.1$ GeV 2. $\lvert M(\gamma\gamma) - m(\pi^0)\rvert < 0.015$ GeV$/c^2$ 3. Veto: $\lvert M(\gamma\gamma_{\rm roe}) - m(\pi^0)\rvert < 0.015$
D_s^+	1. $\lvert M(K^+K^-\pi^+) - m(D_s^+)\rvert < 0.01$ GeV$/c^2$ 2. $\bar{K}^{*0}K^+$: $\lvert\cos\theta(\bar{K}^{*0})\rvert > 0.20$ and $\lvert M(K^-\pi^+) - m(K^*)\rvert < 0.05$ GeV$/c^2$ 3. $\phi\pi^+$: $\lvert\cos\theta(\phi)\rvert > 0.1$ and $\lvert M(K^+K^-) - m(\phi)\rvert < 0.01$ GeV$/c^2$	
D_s^{*+}	$\lvert M(D_s^+\gamma) - m(D_s^{*+})\rvert < 0.015$ GeV$/c^2$	

Signal extraction: Data

Correct the $D_s(2460)^+ \to D_s^{*+}\pi^0$ using the p^* distribution from Babar *PRD* 74 (2006) 032007





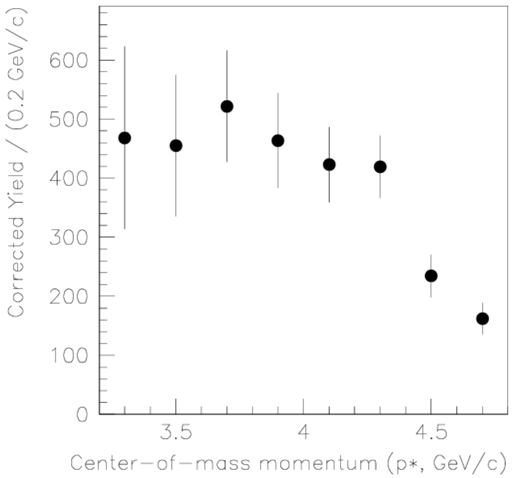
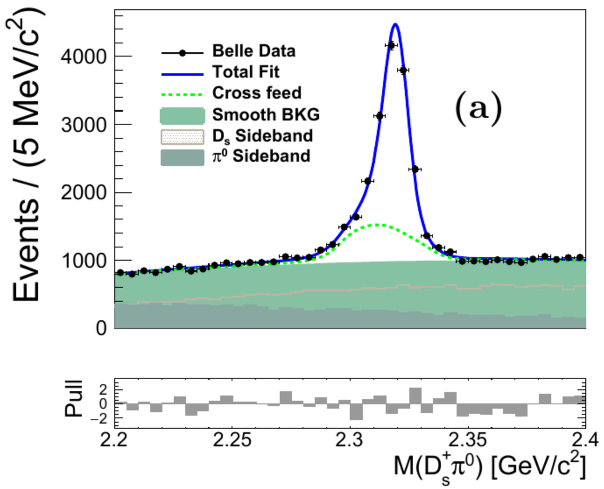
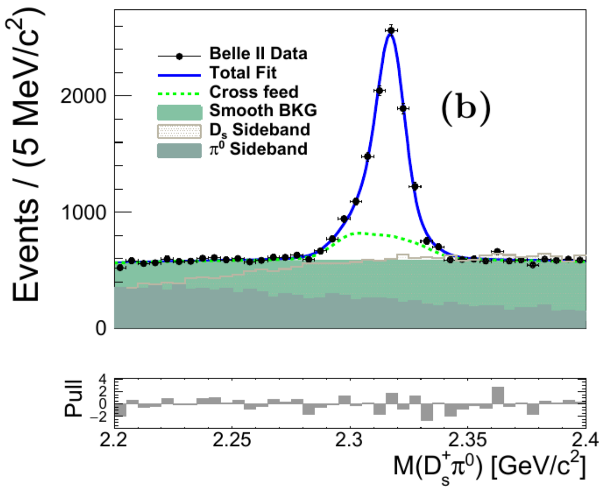


FIG. 15: Corrected $D_{sJ}(2460)^+ \to D_s^+\gamma$ yield as a function of p^*.

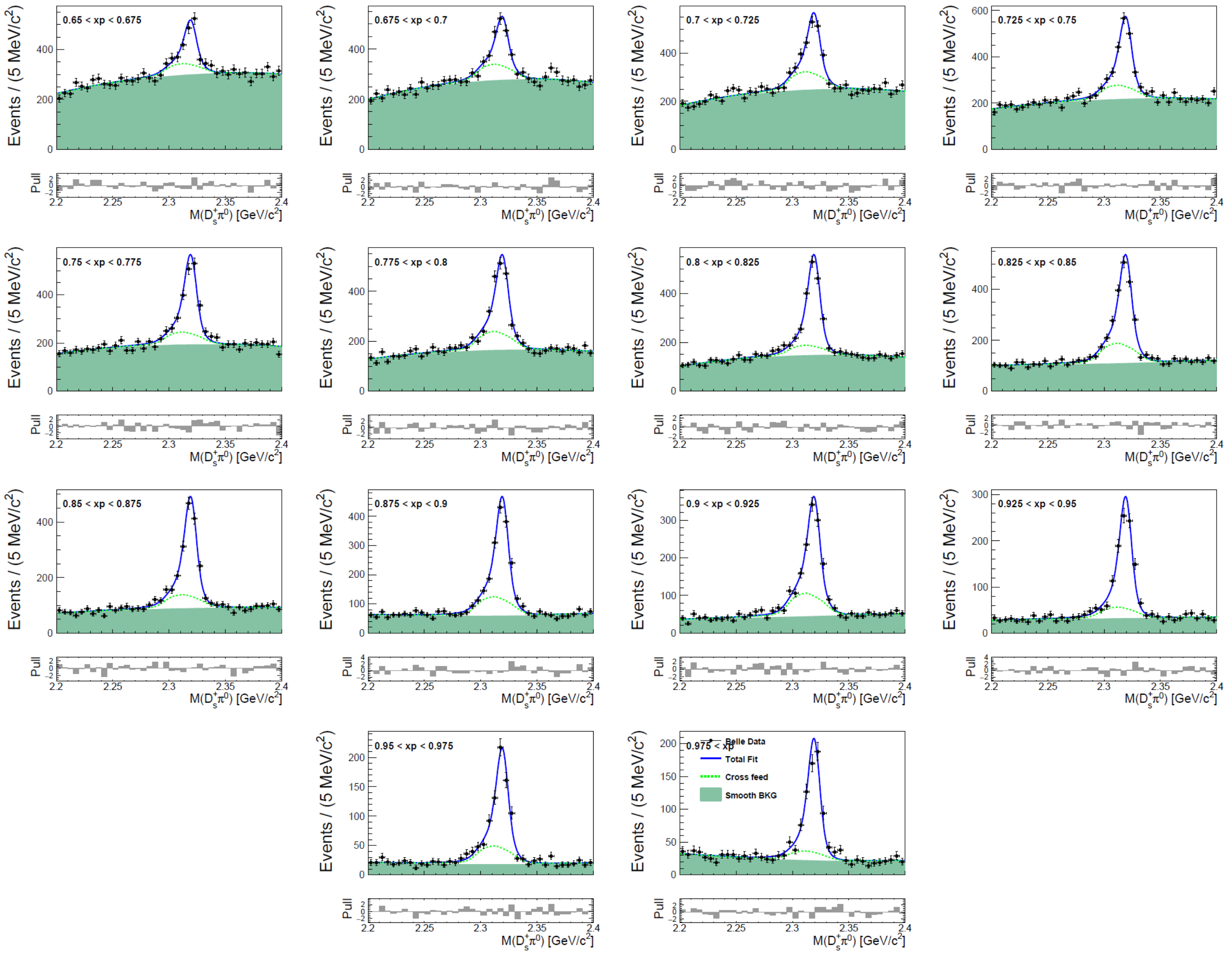


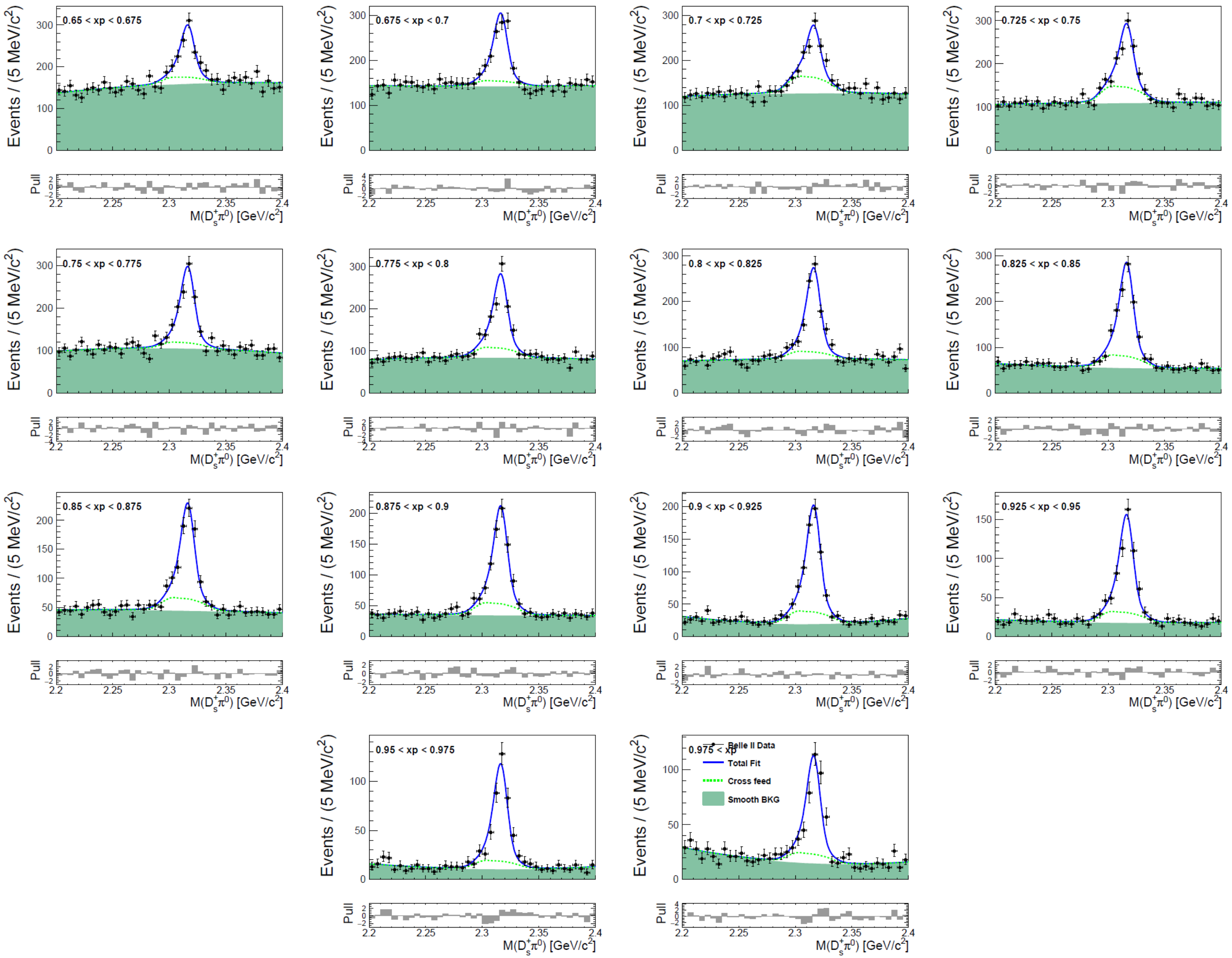


Same method as the previous fit to generic MC sample

$$N_{\text{fit}}^{\text{B1}} = (9.65 \pm 0.20) \times 10^3$$

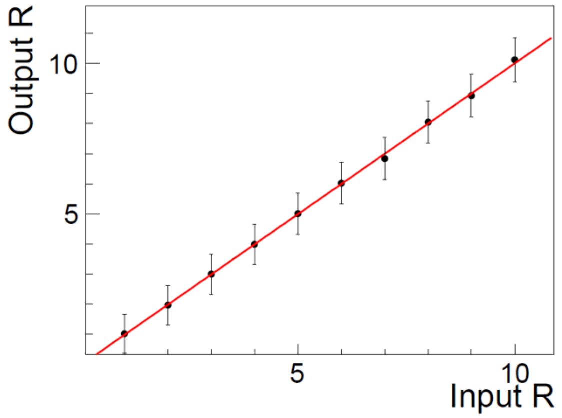
$$N_{\text{fit}}^{\text{B2}} = (5.99 \pm 0.16) \times 10^3$$

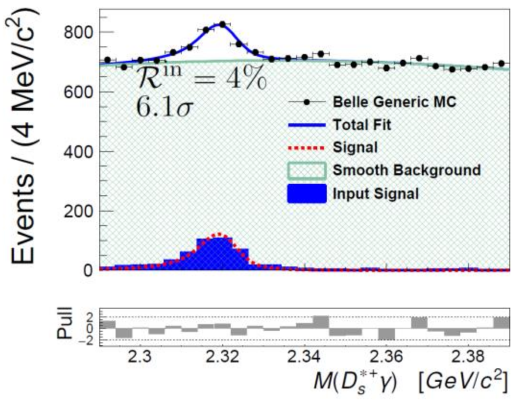

0.65 < xp < 0.675
0.675 < xp < 0.7
0.7 < xp < 0.725
0.725 < xp < 0.75
0.75 < xp < 0.775
0.775 < xp < 0.8
0.8 < xp < 0.825
0.825 < xp < 0.85
0.85 < xp < 0.875
0.875 < xp < 0.9
0.9 < xp < 0.925
0.925 < xp < 0.95
0.95 < xp < 0.975
0.975 < xp
Events / (5 MeV/c^2)
Pull
M($D_s^+\pi^0$) [GeV/c^2]
Belle Data
Total Fit
Cross feed
Smooth BKG

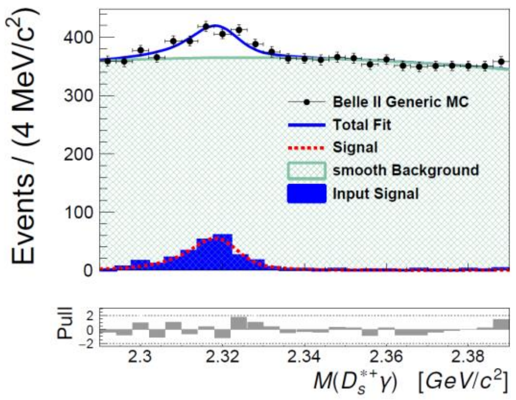

0.65 < xp < 0.675
0.675 < xp < 0.7
0.7 < xp < 0.725
0.725 < xp < 0.75
0.75 < xp < 0.775
0.775 < xp < 0.8
0.8 < xp < 0.825
0.825 < xp < 0.85
0.85 < xp < 0.875
0.875 < xp < 0.9
0.9 < xp < 0.925
0.925 < xp < 0.95
0.95 < xp < 0.975
0.975 < xp
Events / (5 MeV/c^2)
Pull
M($D_s^+\pi^0$) [GeV/c^2]
Belle II Data
Total Fit
Cross feed
Smooth BKG

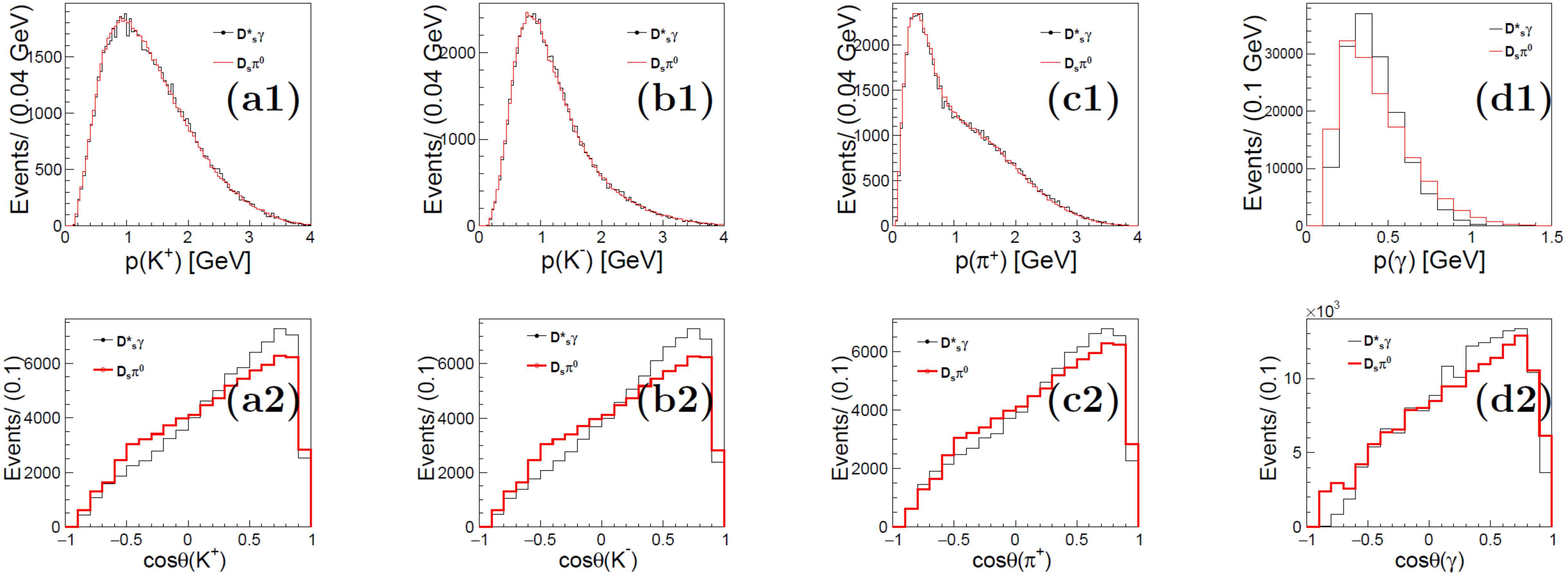
Validation of using mock data samples

$\mathcal{R}^{\text{in}}$ (%)	$\mathcal{R}^{\text{out}}$ (%)	Significance (σ)	$\mathcal{R}^{\text{in}}$ (%)	$\mathcal{R}^{\text{out}}$ (%)	Significance (σ)
1	1.02 ± 0.65	0.16	6	6.02 ± 0.69	8.5
2	1.97 ± 0.66	3.0	7	6.84 ± 0.70	9.9
3	3.00 ± 0.67	4.4	8	8.05 ± 0.71	11.2
4	3.99 ± 0.68	5.8	9	8.93 ± 0.72	12.5
5	5.01 ± 0.68	7.2	10	10.12 ± 0.74	13.8



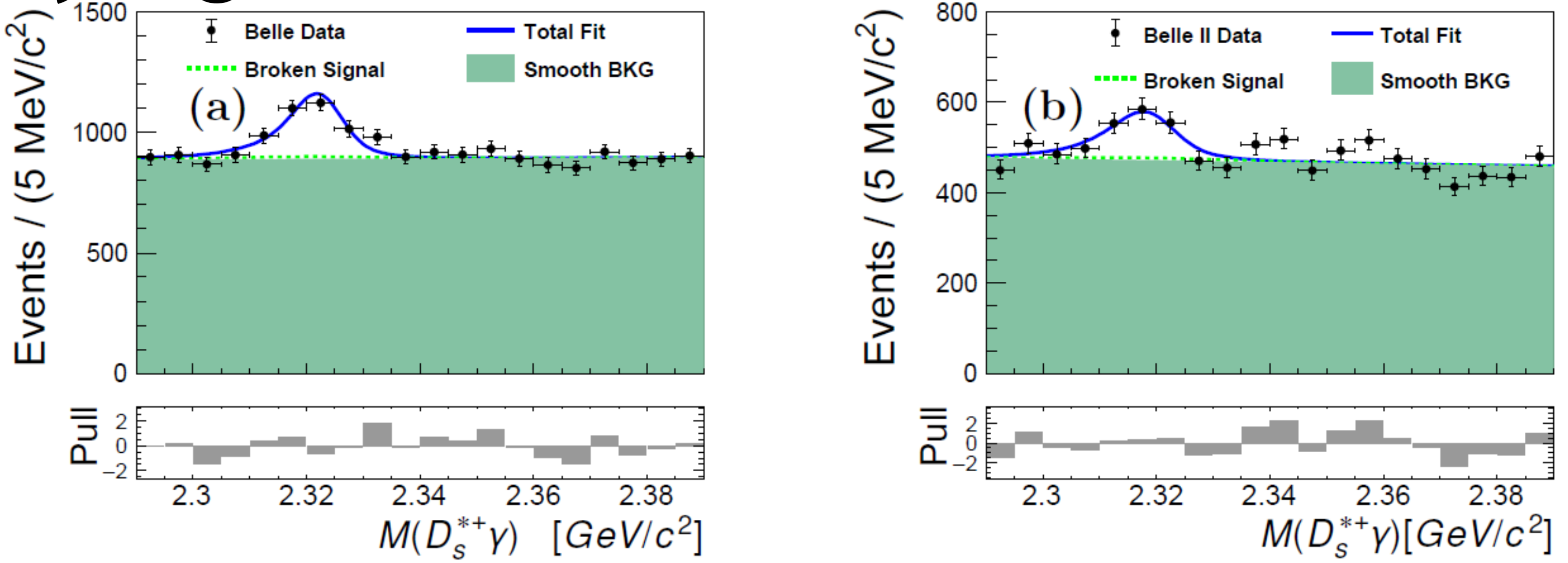





(a1)
(b1)
(c1)
(d1)
(a2)
(b2)
(c2)
(d2)
D*sγ
Dsπ0
Events/ (0.04 GeV)
Events/ (0.1 GeV)
Events/ (0.1)
p(K+) [GeV]
p(K-) [GeV]
p(π+) [GeV]
p(γ) [GeV]
cosθ(K+)
cosθ(K-)
cosθ(π+)
cosθ(γ)
×10³

$D_s(2317)$ signal extraction:

Unbinned maximized likelihood fits to $M(D_s^*\gamma)$ spectra



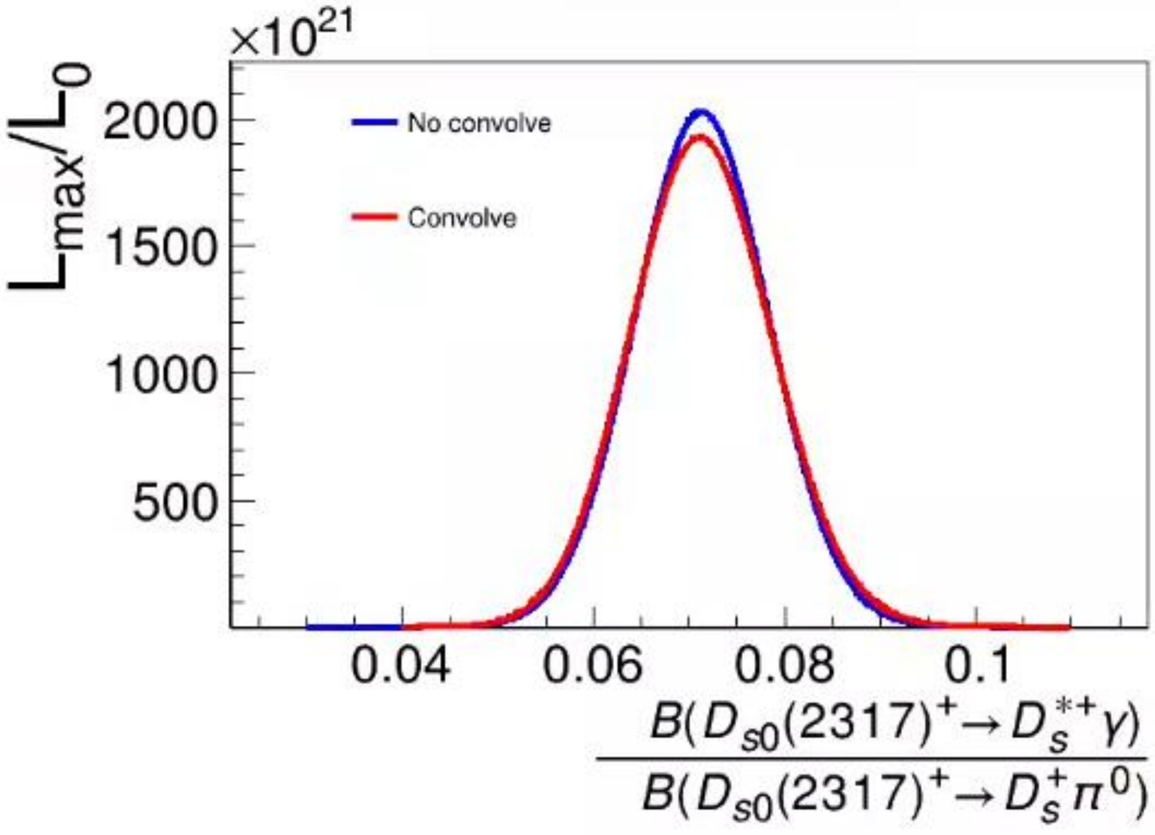
- Fit components are the same as the fits to signal MC samples
- Signal Pdfs parameters fixed to those from fits to MC samples
  - Only mean values of CB functions are free to be float
  - Ratios between signal and broken signal yields are fixed

Experiments	Belle	Belle II
N^{fit}	741.5 ± 82.4	348.2 ± 69.0
$-2\ln(\frac{\mathcal{L}(0)}{\mathcal{L}(\max)})$	85.7	26.7
Significant	9.0σ	$4.8\ \sigma$

$\Delta d.o.f. = 2$

$\mathcal{L}(0)$: maximized likelihood without signal component

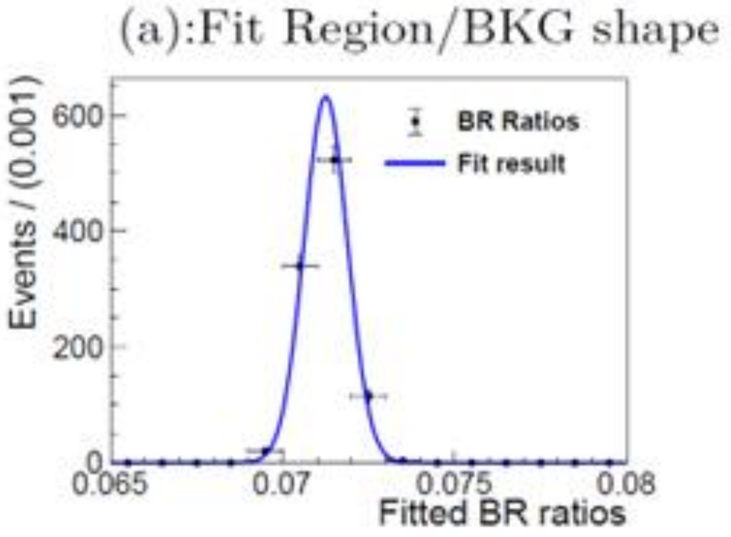
$\mathcal{L}(\max)$: maximized likelihood with signal component

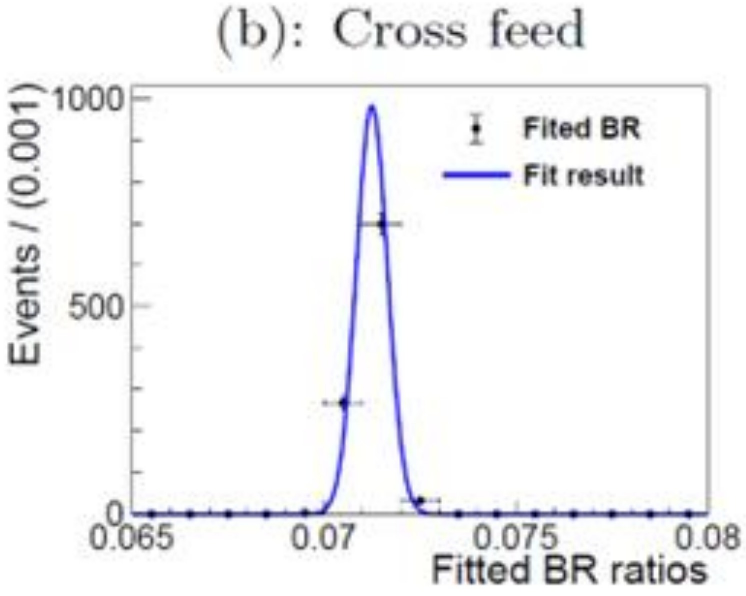


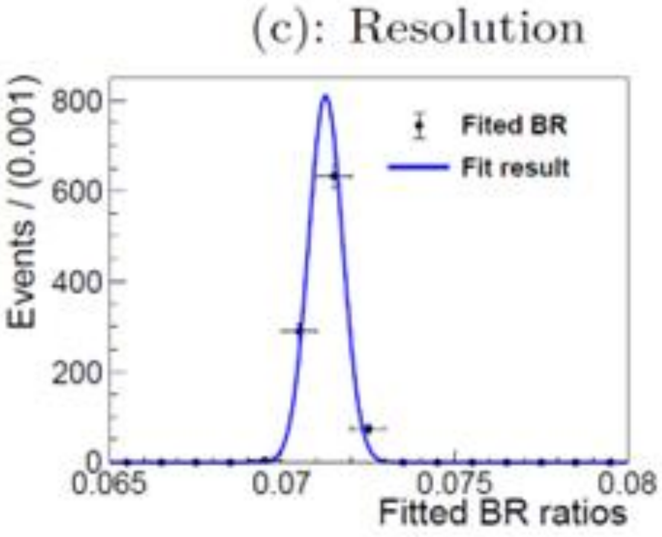
- Convolve the Gaussian(0, $\sigma_{syst.}$)
- significance of $D_{s0}(2317)^+ \to D_s^{*+}\gamma$ changes from 10.13σ to 10.12σ, i.e., we can still claim that the significance of $D_{s0}(2317)^+ \to D_s^{*+}\gamma$ is 10.1σ.

Decay topology around
$M(D_s^{*+}\gamma)\sim 2.26\ \mathrm{GeV}/c^2$

2	$vpho \to c\bar{c}gg\pi^+K^-D^{*-}\boxed{D_s^{*+}a_1^0\gamma^F\gamma^F\gamma^F}, D^{*-} \to \pi^0 D^-, D_s^{*+} \to D_s^+\gamma, a_1^0 \to \pi^0 f_0(600$ $D^- \to \pi^- K_0^0, D_s^+ \to \bar{K}^*K^+, f_0(600) \to \pi^+\pi^-, K_0^0 \to \pi^-K^+,$ $\bar{K}^* \to \pi^+K^-$
3	$vpho \to c\bar{c}gg\rho^0K^+D^{*0}\boxed{D_s^{*-}\gamma^F\gamma^F}, \rho^0 \to \pi^+\pi^-, D^{*0} \to D^0\gamma, D_s^{*-} \to D_s^-\gamma,$ $D^0 \to e^+\nu_e K^-, D_s^- \to K^*K^-, K^* \to \pi^-K^+$
4	$vpho \to c\bar{c}gK^{*+}D^{*0}\boxed{D_s^{*-}\gamma\gamma,} K^{*+} \to \pi^+K^0, D^{*0} \to \pi^0 D^0, D_s^{*-} \to D_s^-\gamma,$ $K^0 \to K_L^0, D^0 \to e^+\nu_e K^-\gamma^F, D_s^- \to K^*K^-, K^{*-} \to \pi^-\bar{K}^0,$ $K^* \to \pi^-K^+, \bar{K}^0 \to K_S^0$
5	$vpho \to c\bar{c}ggg\rho^0K^-\bar{D}^{*0}D_s^{*+}\gamma^F\gamma^F\gamma^F\gamma^F, \rho^0 \to \pi^+\pi^-\gamma^F, \bar{D}^{*0} \to \bar{D}^0\gamma, D_s^{*+} \to D_s^+\gamma,$ $\bar{D}^0 \to \pi^0K^*, D_s^+ \to \bar{K}^*K^+, K^* \to \pi^0K^0, \bar{K}^* \to \pi^+K^-,$ $K^0 \to K_L^0$
6	$vpho \to c\bar{c}g\pi^0\bar{K}^0D^{*-}\boxed{D_s^{*+}\gamma\gamma\gamma,} \bar{K}^0 \to K_L^0, D^{*-} \to \pi^0D^-, D_s^{*+} \to D_s^+\gamma,$ $D^- \to \pi^-\pi^-K^+, D_s^+ \to \bar{K}^*K^+, \bar{K}^* \to \pi^+K^-$

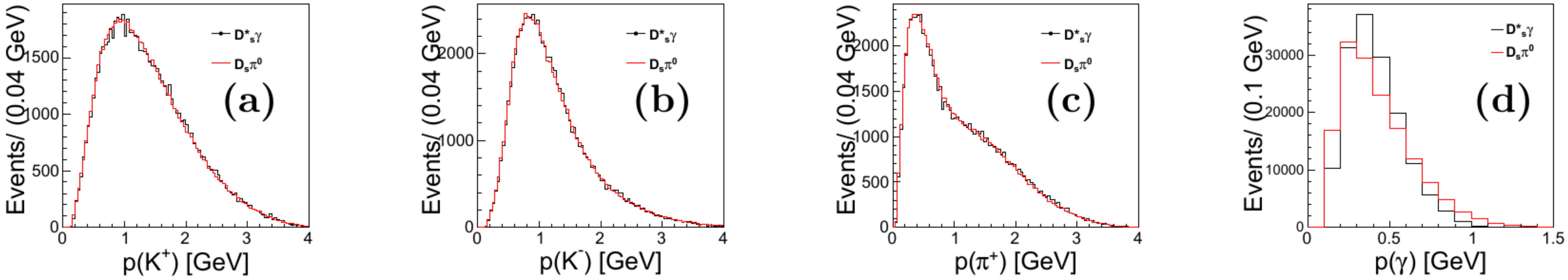






Systematic Uncertainty Study

- The systematic uncertainties due to $\boldsymbol{D_s}$ and **photon** selection, and cut on $\boldsymbol{x_p}$ are cancelled



- **x_p reweighting:**
  - ➢Change the order of the polynomial used to fit x_p distribution
  - ➢Redo the reweighting to get the new detection efficiency ε_{exp}
  - ➢Fit $\mathcal{R}$ again

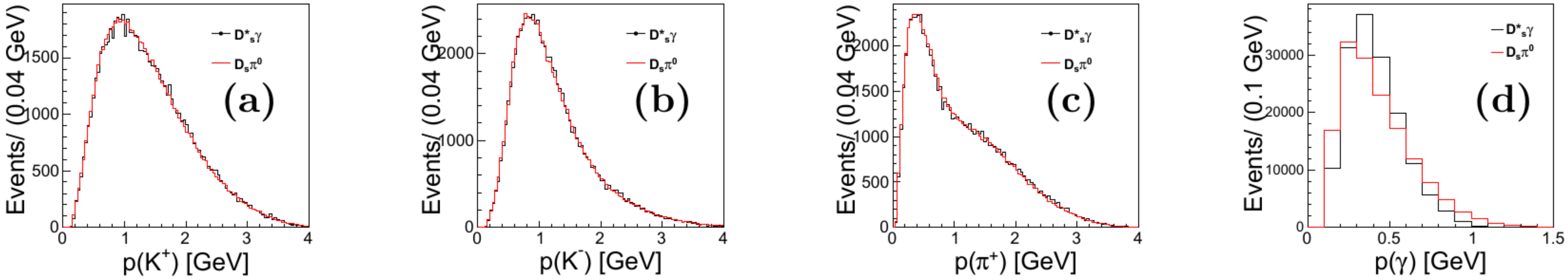
$$N_{\text{exp}}(D_s^{*+}\gamma) = \mathcal{R}\frac{N_{\text{exp}}^{\text{fit}}(D_s^+\pi^0)\varepsilon_{\text{exp}}(D_s^{*+}\gamma)}{\varepsilon_{\text{exp}}(D_s^+\pi^0)}$$

Uncertainties on $\mathcal{R}$:
Difference in Fit $\mathcal{R}$

- **MC statistic:**
  - ➢Vary ε_{exp} by $\pm\sqrt{(1-\varepsilon_{\text{exp}})/(\varepsilon_{\text{exp}}N)}$, N is the number of the used signal MC events.
  - ➢Fit $\mathcal{R}$ again

Systematic Uncertainty Study

- The systematic uncertainties due to $\boldsymbol{D_s}$ and **photon** selection, and cut on $\boldsymbol{x_p}$ are cancelled



- **x_p reweighting:**
  - ➢Change the order of the polynomial used to fit x_p distribution
  - ➢Redo the reweighting to get the new detection efficiency $\varepsilon_{\exp}$
  - ➢Fit $\mathcal{R}$ again

$$N_{\exp}(D_s^{*+}\gamma) = \mathcal{R}\frac{N_{\exp}^{\text{fit}}(D_s^{+}\pi^{0})\varepsilon_{\exp}(D_s^{*+}\gamma)}{\varepsilon_{\exp}(D_s^{+}\pi^{0})}$$

Uncertainties on $\mathcal{R}$:
Difference in Fit $\mathcal{R}$

- **MC statistic:**
  - ➢Vary $\varepsilon_{\exp}$ by $\pm\sqrt{(1-\varepsilon_{\exp})/(\varepsilon_{\exp}N)}$, N is the number of the used signal MC events.
  - ➢Fit $\mathcal{R}$ again

Systematic Uncertainty Study

$$N_{\exp}(D_s^{*+}\gamma) = \mathcal{R}\frac{N_{\exp}^{\mathrm{fit}}(D_s^+\pi^0)\varepsilon_{\exp}(D_s^{*+}\gamma)}{\varepsilon_{\exp}(D_s^+\pi^0)}$$

- **$D_{s0}(2317) \to D_s\pi^0$ mode:**

In the fits to $M(D_s^+\pi^0)$ spectra, By

1. Changing the fit region, order of background polynomial
2. Enlarging the resolution by 1σ (gotten from fits to signal MC)
3. Using the shapes of $M(D_s\pi)$ from $D_s(2460)^+ \to D_s^{*+}\pi^0$ MC directly

➢ Differences in the fitted signal yields: Uncertainties ($\Delta_{\exp}^{N(D_s^+\pi^0)}$) of $N_{\exp}^{\mathrm{fit}}(D_s^+\pi^0)$

Then, in the simultaneous fit to $M(D_s^{*+}\gamma)$ spectra,

By varying the $N_{\exp}^{\mathrm{fit}}(D_s^+\pi^0)$ based on Gaussian($N_{\exp}^{\mathrm{fit}}(D_s^+\pi^0)$, $\Delta_{\exp}^{N(D_s^+\pi^0)}$) in the fits

Widths of the distributions of fitted values of $\mathcal{R}$:
Uncertainties on $\mathcal{R}$:

Systematic Uncertainty Study

- $D_{s0}(2317) \to D_s^*\gamma$ **mode:**

In the simultaneous fit to $M(D_s^{*+}\gamma)$ spectra, by:

1. Changing the fit region, order of background polynomial
2. Enlarging the resolution by 1σ (gotten from fits to signal MC)
3. Changing resolutions of broken signal by 2σ
4. Changing ratios of broken signal to signal yields by 2σ

➢Fit $\mathcal{R}$ again

Uncertainties on $\mathcal{R}$:
Difference in Fit $\mathcal{R}$

Source	$D_s^+\pi^0$	$D_s^{*+}\gamma$
Fit region and background pdf	0.8	1.3
Resolution	0.7	2.5
Cross-feed or broken signal	0.6	0.7
x_p reweighting	0.5	
MC statistic	0.5	
Sum	3.2	

Systematic Uncertainty Study

$$N_{\text{exp}}(D_s^{*+}\gamma) = \mathcal{R}\frac{N_{\text{exp}}^{\text{fit}}(D_s^+\pi^0)\varepsilon_{\text{exp}}(D_s^{*+}\gamma)}{\varepsilon_{\text{exp}}(D_s^+\pi^0)}$$

- **$D_{s0}(2317) \to D_s\pi^0$ mode:**

In the fits to $M(D_s^+\pi^0)$ spectra, By

1. Changing the fit region, order of background polynomial
2. Enlarging the resolution by 1σ (gotten from fits to signal MC)
3. Using the shapes of $M(D_s\pi)$ from $D_s(2460)^+ \to D_s^{*+}\pi^0$ MC directly

➢ Differences in the fitted signal yields: Uncertainties ($\Delta_{\text{exp}}^{N(\mathrm{D}_s^+\pi^0)}$) of $\mathrm{N}_{\text{exp}}^{\text{fit}}(\mathrm{D}_s^+\pi^0)$

Then, in the simultaneous fit to $M(D_s^{*+}\gamma)$ spectra,

By varying the $N_{\text{exp}}^{\text{fit}}(D_s^+\pi^0)$ based on Gaussian($N_{\text{exp}}^{\text{fit}}(D_s^+\pi^0)$, $\Delta_{\text{exp}}^{N(D_s^+\pi^0)}$) in the fits

Widths of the distributions of fitted values of $\mathcal{R}$:
Uncertainties on $\mathcal{R}$:

Systematic Uncertainty Study

- $\boldsymbol{D_{s0}(2317) \to D_s^*\gamma}$ **mode:**

  In the simultaneous fit to $\boldsymbol{M(D_s^{*+}\gamma)}$ spectra, by:

  1. Changing the fit region, order of background polynomial
  2. Enlarging the resolution by 1σ (gotten from fits to signal MC)
  3. Changing resolutions of broken signal by 2σ
  4. Changing ratios of broken signal to signal yields by 2σ

  ➢ Fit $\mathcal{R}$ again

Uncertainties on $\mathcal{R}$:
Difference in Fit $\mathcal{R}$

Source	$D_s^+\pi^0$	$D_s^{*+}\gamma$
Fit region and background pdf	0.8	1.3
Resolution	0.7	2.5
Cross-feed or broken signal	0.6	0.7
x_p reweighting	0.5	
MC statistic	0.5	
Sum	3.2	