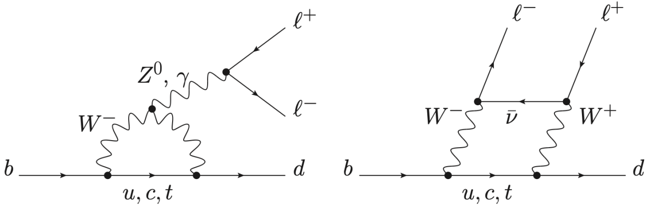
Search for rare $b \to d\ell\ell$ decays at Belle

Seema Choudhury
(Iowa State University, USA)

June 20, 2024

Introduction

- Rare decay $b \to (s, d)\ell\ell$ are suppressed at tree level in the SM
- These FCNC processes propagate through penguin loop and box diagrams



- Previous experimental and theoretical studies mostly focused on $b \to s\ell\ell$ decays
- Signature due to NP maybe uniquely observed in $b \to d\ell\ell$ decays if NP is sensitive to quark flavors
- $b \to d\ell\ell$ transitions are suppressed relative to $b \to s\ell\ell$ processes by the ratio $\frac{|V_{td}|^2}{|V_{ts}|^2} \sim 0.04$
- Typical $\mathcal{B}$ for $b \to d\ell\ell$ decays are $\mathcal{O}(10^{-8})$ or smaller
- $B^0 \to (\pi^0, \eta, \omega, \rho^0)\ell\ell$ and $B^+ \to (\pi^+, \rho^+)\ell\ell$ decays involve $b \to d$ quark-level transition

Experimental Status

Channel	UL or BR	Collaboration	Data size
$B^0 \to \eta ee$	$< 10.8 \times 10^{-8}$	BaBar	428 fb^{-1}
$B^0 \to \eta \mu\mu$	$< 11.2 \times 10^{-8}$	BaBar	428 fb^{-1}
$B^0 \to \omega ee$	–	–	–
$B^0 \to \omega \mu\mu$	–	–	–
$B^0 \to \pi^0 ee$	$< 8.4 \times 10^{-8}$	BaBar	428 fb^{-1}
$B^0 \to \pi^0 \mu\mu$	$< 6.9 \times 10^{-8}$	BaBar	428 fb^{-1}
$B^+ \to \pi^+ ee$	$< 8.0 \times 10^{-8}$	Belle	605 fb^{-1}
$B^+ \to \pi^+ \mu\mu$	$(1.78 \pm 0.22 \pm 0.03) \times 10^{-8}$	LHCb	3 fb^{-1}
$B^0 \to \rho^0 ee$	–	–	–
$B^0 \to \rho^0 \mu\mu$	$(1.98 \pm 0.53) \times 10^{-8}$	LHCb	3 fb^{-1}
$B^+ \to \rho^+ ee$	–	–	–
$B^+ \to \rho^+ \mu\mu$	–	–	–

- $\mathcal{B}(B^0 \to \pi^+\pi^-\mu\mu) = (2.1 \pm 0.5 \pm 0.1) \times 10^{-8}$ using 3 fb^{-1} data sample of LHCb
- LHCb has observed final states with muons and no neutrals
- Electron channels have not been searched yet
- Searches with electrons and muons in the final state allow to probe the lepton-flavor universality in $b \to d\ell\ell$ transitions

Selection Criteria

- 711 fb^{-1} data of Belle corresponds to 772×10^6 $B\overline{B}$ pairs
- Signal MC is generated using BTOSLLBALL decay model

$$M_{\rm bc} = \sqrt{(E^*_{\rm beam}/c^2)^2 - (p^*_B/c)^2}$$

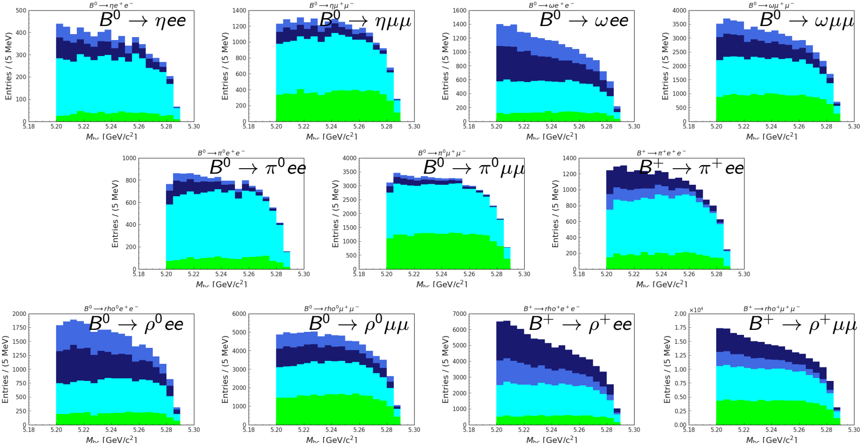
$$\Delta E = E^*_B - E^*_{\rm beam}$$

	Selection	Requirement
Charged tracks	d_0	< 1.0 cm
	d_z	< 4.0 cm
$e^\pm$ candidate	eIDBelle	> 0.9
	$p_{\rm lab}$	> 0.5 GeV/c
	correctBremsBelle	50 mrad
$\mu^\pm$ candidate	muIDBelle	> 0.9
	$p_{\rm lab}$	> 0.8 GeV/c
$\pi^\pm$ candidate	atcPIDBelle(2,3)	> 0.6
π^0 selection	pi0:mdst	$M_{\pi^0} \in [0.124, 0.145]$ GeV
γ	goodBelleGamma	–
ω	$\pi^+\pi^-\pi^0$	$M_\omega \in [0.7685, 0.795]$ GeV
η	$\gamma\gamma$, $\pi^+\pi^-\pi^0$	$M_\eta \in [0.530, 0.559]$ GeV
ρ^0	$\pi^+\pi^-$	$M_{\rho^0} \in [0.650, 0.9]$ GeV
ρ^+	$\pi^+\pi^0$	$M_{\rho^+} \in [0.650, 0.9]$ GeV
Beam energy constrained mass	$M_{\rm bc}$	$M_{\rm bc} > 5.2$ GeV/c^2
Energy difference	ΔE	$-0.15 < \Delta E < 0.1$ GeV
charmonium veto	$J/\psi \to \mu\mu(ee)$	$8.8(8.5) < q^2 < 9.9(10.22)$ GeV$^2/c^4$
	$\psi(2S) \to \mu\mu(ee)$	$13 < q^2 < 14$ GeV$^2/c^4$
Converted photon/π^0 Dalitz decay	$q^2_{e^+e^-}$	> 0.045 GeV$^2/c^4$

- Best candidate is selected using vertex χ^2 probability, BCS efficiency is $70 - 90\%$

- $M_{\rm bc}$ distribution in 6 streams (1 stream corresponds to Belle data size luminosity) of generic MC

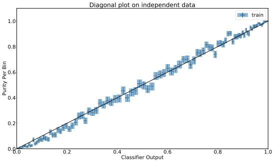
uds, charm, $\rm B^0\overline{B^0}$, $\rm B^+B^-$

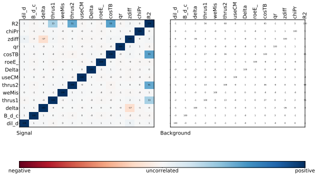


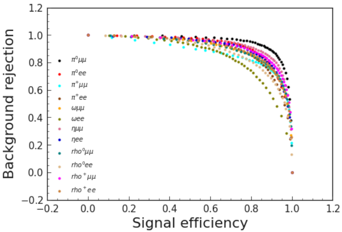
- Mostly continuum background (uds + charm), but $B\overline{B}$ background ($\rm B^0\overline{B^0}$+$\rm B^+B^-$) is non-negligible

Background Suppression

- FastBDT is trained with event shape, vertex quality, and kinematic variables: $|\mathrm{Thrust_B}|$, $|\mathrm{Thrust_O}|$, $\cos^{\mathrm{Thrust}}\theta_B$, R_2, $\cos\theta_B$, $\delta\phi_{\ell\ell}$, $P(\chi^2)$, $\Delta Z_{\ell\ell}$, $Z(X_d - \ell_1)$, E_{vis}^{ROE}, M_{miss}^2/E_{miss}, ΔZ, to suppress background
- Equal amount of signal and background (continuum and generic B) events are used for training
- BDT training is good and there is negligible correlation between input variables

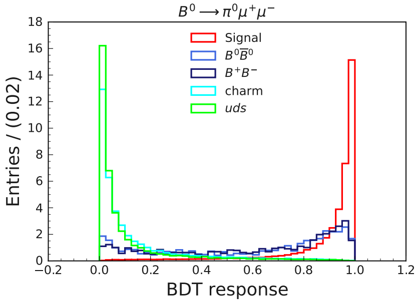


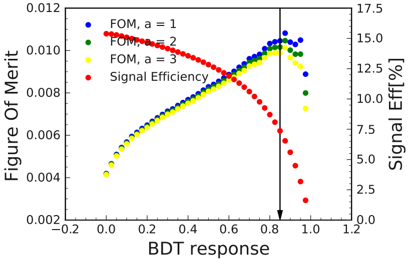




- Each BDT is trained separately for each decay channel, so they expected to behave differently

BDT Optimization





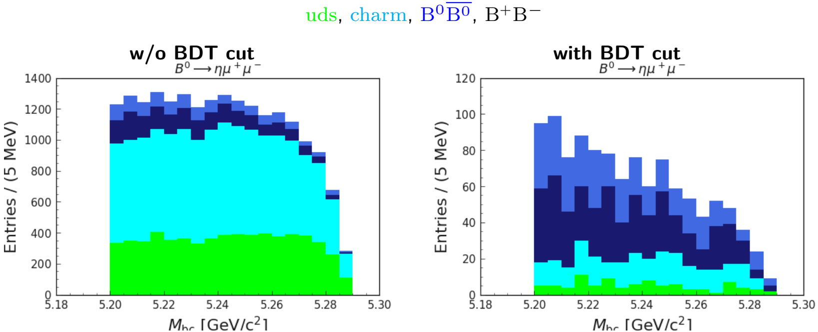
- Continuum background peaks around 0 and signal around 1
- Each decay channel is trained and optimized separately
- FOM is calculated using Punzi's method

$$\text{FOM} = \frac{\varepsilon}{a/2 + \sqrt{\text{B}}},$$

ε= efficiency, B= background in the signal region, & a= significance

channel	BDT cut
$B^0 \to \eta ee$	0.65
$B^0 \to \eta\mu\mu$	0.85
$B^0 \to \omega ee$	0.80
$B^0 \to \omega\mu\mu$	0.90
$B^0 \to \pi^0 ee$	0.75
$B^0 \to \pi^0\mu\mu$	0.90
$B^+ \to \pi^+ ee$	0.80
$B^0 \to \rho^0 ee$	0.80
$B^0 \to \rho^0\mu\mu$	0.85
$B^+ \to \rho^+ ee$	0.90
$B^+ \to \rho^+\mu\mu$	0.95

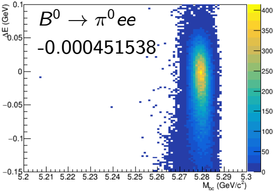
Background Suppression

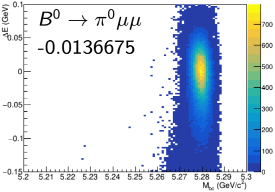


- Optimal BDT requirement reduces the background by $93-98\%$ with $25-55\%$ signal efficiency loss, over the decay channels

Fitting strategy

- Perform 2-dimensional un-binned extended maximum likelihood fit in $M_{\rm bc}$ and ΔE to extract the signal yield.
- Negligible correlation between fitting variables





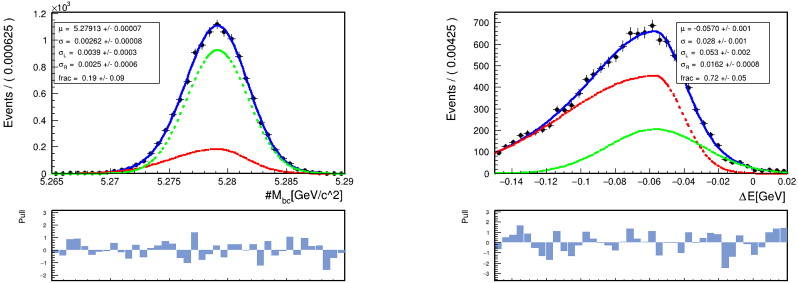
- Fitting PDFs

Variable	Signal	Background
$M_{\rm bc}$	CB or CB+Gaussian	Argus
ΔE	Johnson or Johnson+Gaussian	Chebyshev polynomial

- Signal shape parameters are fixed from signal MC
- Signal shape fudge factor and mean shift are fixed from control sample
- Background shapes are floated

Peaking background for $B^+ \to \pi^+ ee$

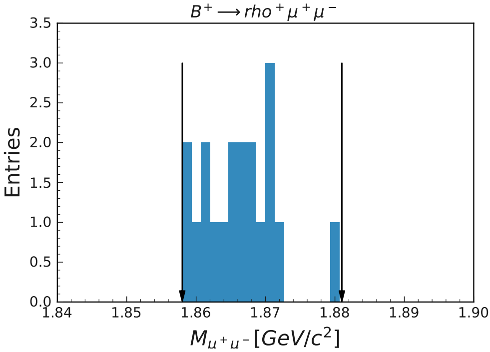
- Background from $B^+ \to K^+ ee$, where kaon misidentified as a pion
- Peaks in $M_{\rm bc}$ signal region, but shifted in ΔE (~ -0.06 GeV)
- Studied using $B^+ \to K^+ ee$ signal MC generated using BTOSLLBALL decay model



- Fit parameters obtained from these fits are fixed in signal yield extraction, with floated peaking yield

Peaking background for $B^+ \to \rho^+ \mu\mu$

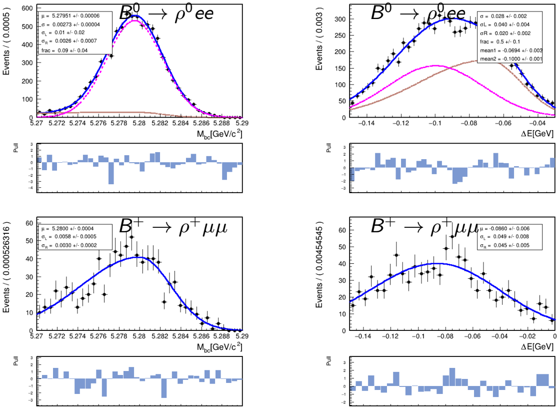
- Background from $B^+ \to \rho^+ \overline{D^0}(K^+\pi^-)$, where K^+ and π^- are misidentified as muon candidates



- Faking particle mass hypothesis is applied to decide on the veto window
- Applied $M_{\mu^+\mu^-} \notin [1.858 - 1.881]$ GeV$/c^2$ to suppress the background

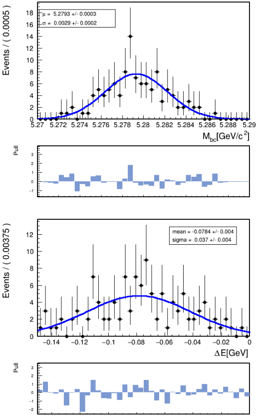
Peaking backgrounds for $B^0 \to \rho^0 ee$ and $B^+ \to \rho^+ \mu\mu$

- Either of the two leptons come from J/ψ, while kaon originated from K^{*0}, K_2^{*0}, K_0^{*0}, etc is misidentified as the other lepton
- Events are shifted in ΔE (~ -0.09 GeV)
- Background is studied using inclusive J/ψ sample which has 100 times more luminosity compared to Belle data sample
- Applied loose LID requirement (> 0.1) on one lepton and standard LID on other

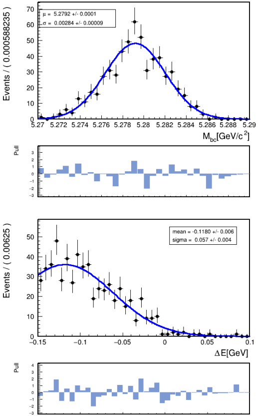


- Fit parameters obtained from these fits are fixed in signal yield extraction, with floated peaking yield

- $B^0 \to \eta K\pi$ background is studied by signal MC using PHSP decay model
- Applied loose lepton ID of > 0.1 for reconstruction
- Peaks in $M_{\rm bc}$ signal region, but shifted in ΔE (~ -0.08 GeV)



- $B^0 \to \pi^0 K\pi$ & $B^0 \to \pi^0 KK$ backgrounds are studied by signal MC using PHSP decay model
- Applied loose lepton ID of > 0.1 for reconstruction
- Peaks in $M_{\rm bc}$ signal region, but shifted in ΔE (~ -0.11 GeV)



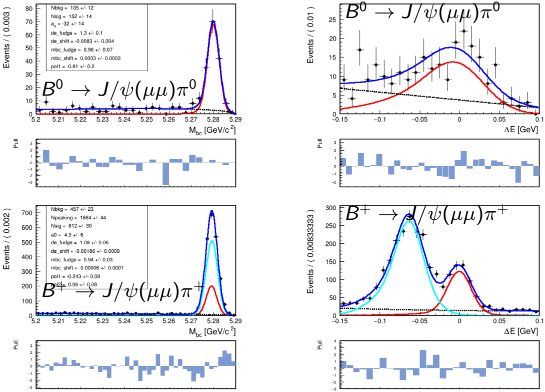
Summary of peaking backgrounds

channel	source
$B^0 \to \eta\mu\mu$	$B^0 \to \eta K\pi$
$B^0 \to \pi^0\mu\mu$	$B^0 \to \pi^0 K\pi$ & $B^0 \to \pi^0 KK$
$B^+ \to \pi^+ ee$	$B^+ \to K^+ ee$
$B^0 \to \rho^0 ee$	$J/\psi \to e$ & $K^* \to K \leftrightarrow e$
$B^+ \to \rho^+\mu\mu$	$J/\psi \to \mu$ & $K^* \to K \leftrightarrow \mu$, $\rho^+\overline{D^0}(K^+\pi^-)$

- Peaking backgrounds are either vetoed or included in the fit by:
  - applying invariant mass vetoes by assigning faking particle mass hypothesis to correct particle
  - studying high statistics inclusive J/ψ sample
  - generating signalMC for peaking contribution

Control sample: $B^{\pm,0} \to J/\psi(\ell\ell)\pi^{\pm,0}$ in data

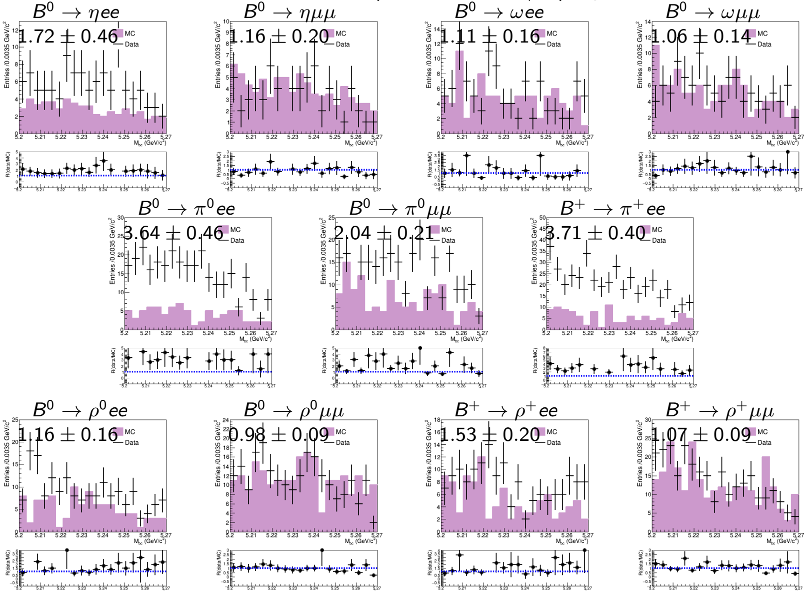
- Performed 2-dimensional fit in $M_{\rm bc}$ and ΔE to extract the signal yield
- Introduced fudge factor and mean shift to the width and mean of signal distributions



channel	$N_{\rm sig}$	ε (%)	$\mathcal{B}$	PDG $\mathcal{B}$
$B^0 \to J/\psi(\mu\mu)\pi^0$	152 ± 14	20.35	$(0.975 \pm 0.090) \times 10^{-6}$	0.990×10^{-6}
$B^0 \to J/\psi(ee)\pi^0$	101 ± 11	12.09	$(1.091 \pm 0.119) \times 10^{-6}$	0.991×10^{-6}
$B^+ \to J/\psi(\mu\mu)\pi^+$	612 ± 30	32.04	$(2.397 \pm 0.118) \times 10^{-6}$	2.337×10^{-6}
$B^+ \to J/\psi(ee)\pi^+$	295 ± 24	17.34	$(2.140 \pm 0.174) \times 10^{-6}$	2.340×10^{-6}

- Calculated $\mathcal{B}$ is consistent with PDG within the uncertainty

- Compared the data-MC in the $M_{\rm bc}$ sideband ($M_{\rm bc}$<5.265 GeV/c^2) region

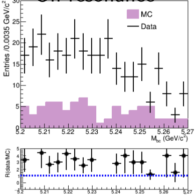


- Factor of 3 − 4 excess data in $M_{\rm bc}$ side-band compared to generic MC for $B^{\pm,0} \to \pi^{\pm,0} ee$ decays

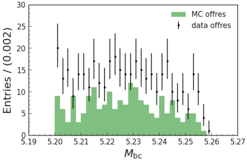
π^0 momentum for $B^0 \to \pi^0 ee$

- data-MC discrepancy for both on- and off- resonance samples

On-resonance

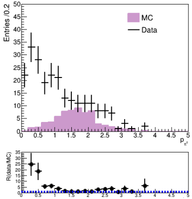


Off-resonance

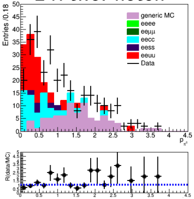


- Most of the difference in data and MC comes from lower pion momentum region
- Contribution from two photon samples

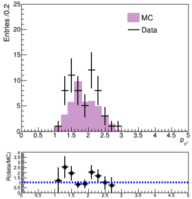
Background distribution



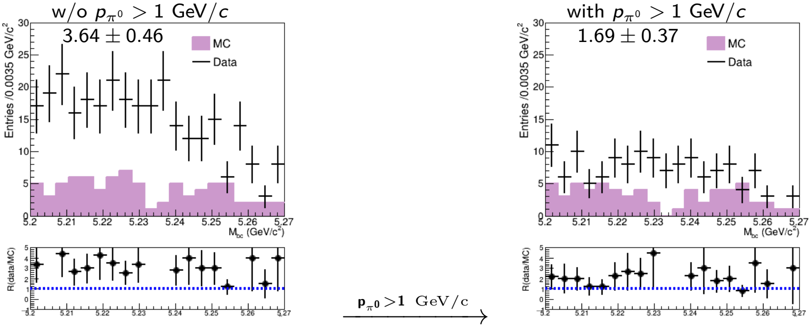
Background distribution
2 x twoPhoton



Signal for control sample



- Signal momentum ranges from $1 - 3$ GeV$/c$
- Applied $p_{\pi^0} > 1$ GeV$/c$ to suppress the two photon background
- Similar background for $B^+ \to \pi^+ ee$ channel, and suppressed using $p_{\pi^+} > 1$ GeV$/c$



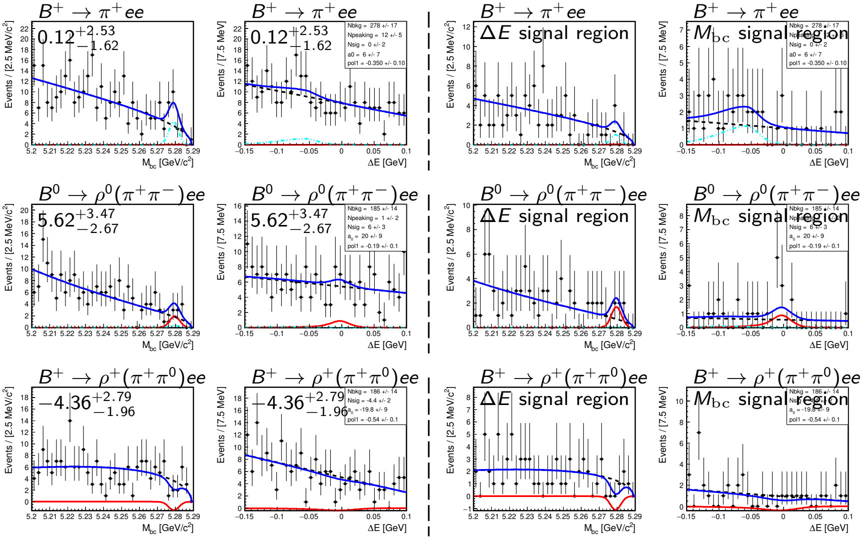
- Background has reduced significantly with pion momentum cut, without significant impact on signal
- Data is now in much better agreement with MC simulation

Data Unblinding

Fit result for $B^+ \to \pi^+ ee$, $B^0 \to \rho^0 ee$ and $B^+ \to \rho^+ ee$ decays

- Signal enhanced projection: $M_{\rm bc} > 5.27$ GeV$/c^2$ for ΔE plot and $|\Delta E| < 0.05$ GeV for $M_{\rm bc}$ plot

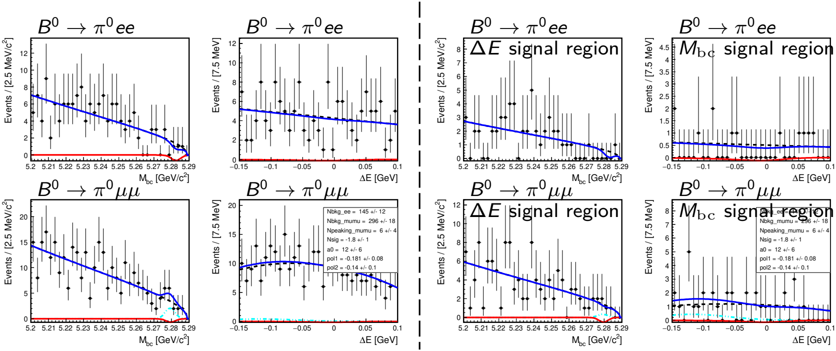
signal, background, peaking background, total



Fit result for $B^0 \to \pi^0 \ell\ell$

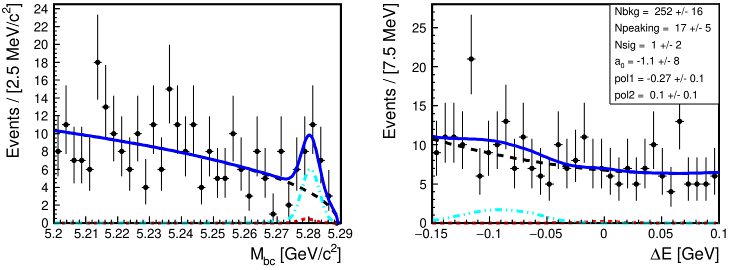
- Signal enhanced projection: $M_{\rm bc} > 5.27$ GeV$/c^2$ for ΔE plot and $|\Delta E| < 0.05$ GeV for $M_{\rm bc}$ plot

signal, **background**, **peaking background**, **total**



- $N_{\rm sig} = 0.48^{+1.04}_{-0.83}$, $1.05^{+1.75}_{-1.28}$, $-1.81^{+1.62}_{-1.14}$, $3.26^{+2.06}_{-1.65}$, and $0.39^{+2.41}_{-1.95}$ for $B^0 \to \eta\ell\ell$, $B^0 \to \omega\ell\ell$, $B^0 \to \pi^0\ell\ell$, $B^0 \to \rho^0\ell\ell$, and $B^+ \to \rho^+\ell\ell$

$B^0 \to \rho^0 \mu\mu$ result



channel	Belle $\mathcal{B}$ (10^{-8})	LHCb $\mathcal{B}$ (10^{-8})
$B^0 \to \rho^0 \mu\mu$	$2.18^{+5.06}_{-3.20}$	1.98 ± 0.53

- Belle quoted uncertainty is statistical only
- Our calculated $\mathcal{B}$ is consistent with LHCb within uncertainty

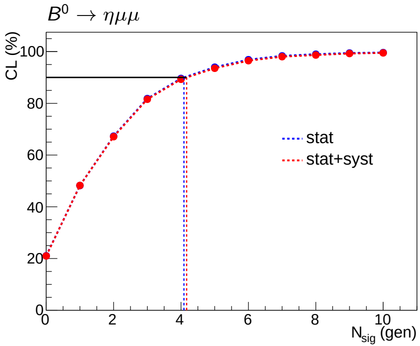
Systematic Uncertainties

source	ηee	$\eta\mu\mu$	ωee	$\omega\mu\mu$	$\pi^0 ee$	$\pi^0\mu\mu$	$\pi^+ ee$	$\rho^0 ee$	$\rho^+ ee$	$\rho^+\mu\mu$
μ	–	0.6	–	0.6	–	0.6	–	–	–	0.6
e	0.8	–	0.8	–	0.8	–	0.8	0.8	0.8	–
π^+	1.0	1.0	1.0	1.0	–	–	0.5	1.0	0.5	0.5
π^0	2.3	2.3	2.3	2.3	2.3	2.3	–	–	2.3	2.3
γ	4.0	4.0	–	–	–	–	–	–	–	–
FastBDT	7.1	6.6	7.1	6.6	7.1	6.6	1.4	1.4	7.1	6.6
MC statistics	0.48	0.37	0.73	0.53	0.34	0.24	0.24	0.53	0.80	0.54
decay model	0.57	0.45	0.75	0.69	0.49	0.76	0.40	0.66	0.81	0.52
mass window	1.05	1.05	1.21	1.21	–	–	–	3.03	3.03	3.03
BCS	0.03	0.11	0.15	0.43	0.21	0.23	0.11	0.02	0.6	0.5
Tracking	0.7-1.4	0.7-1.4	1.4	1.4	0.7	0.7	1.05	1.4	1.05	1.05
PDF shape	0.04	0.04	0.43	0.07	0.10	0.09	0.50	0.20	0.34	0.32
$f^{+-/00}$	2.45	2.45	2.45	2.45	2.45	2.45	2.45	2.45	2.45	2.45
$N_{B\bar{B}}$	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Total	9.35	8.95	8.37	7.91	8.07	7.64	3.56	4.80	8.75	8.29

- Systematic uncertainty varies between $4-9\%$ depending on the decay channel
- Dominate systematic is from FastBDT, because of $B^0 \to J/\psi(\ell\ell)\pi^0$ sample size

Signal yield in the UL ($N_{\rm sig}^{\rm UL}$)

- $N_{\rm sig}^{\rm UL}$ is calculated using frequentist approach
- Generated 10,000 toys for each $N_{\rm sig}({\rm gen}) = 0, 1, 2, ..., 10$ considering data fitting PDF having statistical significance equivalent to data sample
- Performed fit on these sets of events using data fitting PDF
- Added the systematic uncertainty as width of Gaussian to the signal yield
- Calculated the fraction of sets that have a fitted yield greater or equal to the one observed in data fitting for each value of input signal yield
- $N_{\rm sig}^{\rm UL}$ is the value of signal yield with 90% CL



$B^{\pm,0} \to (\eta, \omega, \pi^{\pm,0}, \rho^{\pm,0})\ell\ell$ results

$$\mathcal{B}^{\rm UL} = \frac{N_{\rm sig}^{\rm UL}}{2f^{\pm(00)} N_{B\bar{B}}\varepsilon}$$

channel	$N_{\rm sig}$	$N_{\rm sig}^{\rm UL}$	ε (%)	$\mathcal{B}^{\rm UL}$ (10^{-8})	$\mathcal{B}$ (10^{-8})
$B^0 \to \eta ee$	$0.0^{+1.4}_{-1.0}$	3.1	3.9	< 10.5	$0.0^{+4.8}_{-3.3} \pm 0.1$
$B^0 \to \eta\mu\mu$	$0.8^{+1.5}_{-1.1}$	4.2	5.9	< 9.2	$1.8^{+3.3}_{-2.4} \pm 0.2$
$B^0 \to \eta\ell\ell$	$0.5^{+1.0}_{-0.8}$	1.8	4.9	< 4.7	$1.3^{+2.7}_{-2.2} \pm 0.1$
$B^0 \to \omega ee$	$-0.3^{+3.2}_{-2.5}$	3.7	1.6	< 30.0	$-2.1^{+25.9}_{-20.3} \pm 0.2$
$B^0 \to \omega\mu\mu$	$1.7^{+2.3}_{-1.6}$	5.5	2.9	< 24.3	$7.5^{+10.5}_{-7.3} \pm 0.6$
$B^0 \to \omega\ell\ell$	$1.0^{+1.8}_{-1.3}$	3.6	2.2	< 21.5	$6.3^{+10.5}_{-7.7} \pm 0.5$
$B^0 \to \pi^0 ee$	$-2.9^{+1.8}_{-1.4}$	4.0	6.7	< 7.7	$-5.7^{+3.5}_{-2.7} \pm 0.5$
$B^0 \to \pi^0\mu\mu$	$-0.5^{+3.6}_{-2.7}$	6.1	13.7	< 5.8	$-0.4^{+3.4}_{-2.5} \pm 0.1$
$B^0 \to \pi^0\ell\ell$	$-1.8^{+1.6}_{-1.1}$	2.9	10.2	< 3.7	$-2.3^{+2.1}_{-1.4} \pm 0.2$
$B^+ \to \pi^+ ee$	$0.1^{+2.5}_{-1.6}$	5.0	11.5	< 5.5	$0.1^{+2.8}_{-1.9} \pm 0.1$
$B^0 \to \rho^0 ee$	$5.6^{+3.5}_{-2.7}$	10.8	3.2	< 44.4	$23.0^{+14.2}_{-11.0} \pm 1.1$
$B^+ \to \rho^+ ee$	$-4.4^{+2.3}_{-2.0}$	5.3	1.4	< 46.9	$-38.4^{+24.6}_{-17.3} \pm 3.4$
$B^+ \to \rho^+\mu\mu$	$3.0^{+4.0}_{-3.0}$	8.7	2.9	< 38.2	$13.1^{+17.6}_{-13.3} \pm 1.1$
$B^+ \to \rho^+\ell\ell$	$0.4^{+2.3}_{-1.8}$	3.0	2.0	< 19.0	$2.5^{+14.7}_{-11.8} \pm 0.2$

Summary

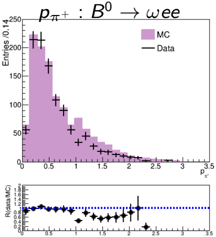
- Searched for rare decays $B^{\pm,0} \to (\eta, \omega, \pi^{\pm,0}, \rho^{\pm,0})ee$ and $B^{\pm,0} \to (\eta, \omega, \pi^0, \rho^{\pm})\mu\mu$ using 711 fb^{-1} data sample of Belle
- No evidence of signal is found and we set UL at 90% CL
- $\mathcal{B}^{\rm UL}$ are $< (3.7 - 47) \times 10^{-8}$, depending on the decay channels
- $B^0 \to \omega\ell\ell$, $B^0 \to \rho^0 ee$, and $B^{\pm} \to \rho^{\pm}\ell\ell$ decay channels are searched for first time
- World's best limits for all decay channels
- Our $\mathcal{B}$ results for ee channels are consistent with $\mu\mu$ channel, and $\mathcal{B}(B^+ \to \pi^+\mu\mu)$, $\mathcal{B}(B^0 \to \rho^0\mu\mu)$ of LHCb: No LFU in $b \to d\ell\ell$ transitions
- Paper is submitted to PRL, and available at arXiv:2404.08133

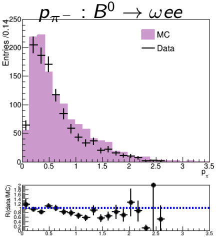
Decay channels of interest

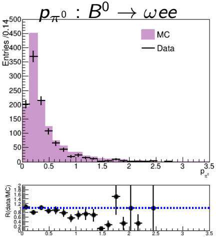
- Decay channels of interest are:
  - $B^0 \to \pi^0 \ell\ell$
  - $B^+ \to \pi^+ ee$
  - $B^0 \to \eta\ell\ell$, $[\eta \to 2\gamma\ (\sim 39\%), \pi^+\pi^-\pi^0\ (\sim 23\%)]$
  - $B^0 \to \omega(782)\ell\ell$, $[\omega(782) \to \pi^+\pi^-\pi^0\ (\sim 89\%)]$
  - $B^0 \to \rho^0(770)ee$, $[\rho^0(770) \to \pi^+\pi^-\ (\sim 100\%)]$
  - $B^+ \to \rho^+\ell\ell$, $[\rho^+ \to \pi^+\pi^0\ (\sim 100\%)]$
- $B^0 \to \omega\ell\ell$, $B^0 \to \rho^0 ee$, and $B^\pm \to \rho^\pm\ell\ell$ will be searched for the first time in Belle

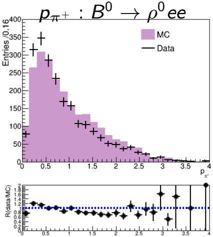
Channel	UL or BR	Collaboration	Data size
$B^0 \to \eta ee$	$< 10.8 \times 10^{-8}$	BaBar	428 fb^{-1}
$B^0 \to \eta\mu\mu$	$< 11.2 \times 10^{-8}$	BaBar	428 fb^{-1}
$B^0 \to \omega ee$	–	–	–
$B^0 \to \omega\mu\mu$	–	–	–
$B^0 \to \pi^0 ee$	$< 8.4 \times 10^{-8}$	BaBar	428 fb^{-1}
$B^0 \to \pi^0\mu\mu$	$< 6.9 \times 10^{-8}$	BaBar	428 fb^{-1}
$B^+ \to \pi^+ ee$	$< 8.0 \times 10^{-8}$	Belle	605 fb^{-1}
$B^+ \to \pi^+\mu\mu$	$(1.78 \pm 0.22 \pm 0.03) \times 10^{-8}$	LHCb	3 fb^{-1}
$B^0 \to \rho^0 ee$	–	–	–
$B^0 \to \rho^0\mu\mu$	$(1.98 \pm 0.53) \times 10^{-8}$	LHCb	3 fb^{-1}
$B^+ \to \rho^+ ee$	–	–	–
$B^+ \to \rho^+\mu\mu$	–	–	–

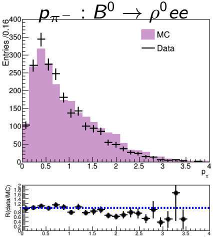
$\pi^{\pm,0}$ momentum for $B^0 \to \omega(\pi^+\pi^-\pi^0)ee$ and $B^0 \to \rho^0(\pi^+\pi^-)ee$ in background





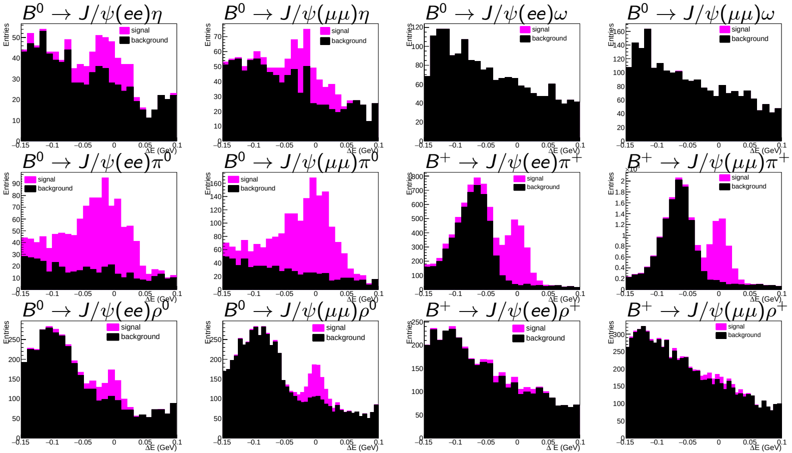






- These modes are not affected by two photon background because of tight constraint on ω, η, ρ masses
- Data-MC are in good agreement, no need of any π momentum cut

- ΔE distribution (stacked histograms) for 6 streams of generic MC
- Applied same selection criteria as the $b \to d\ell\ell$ analysis

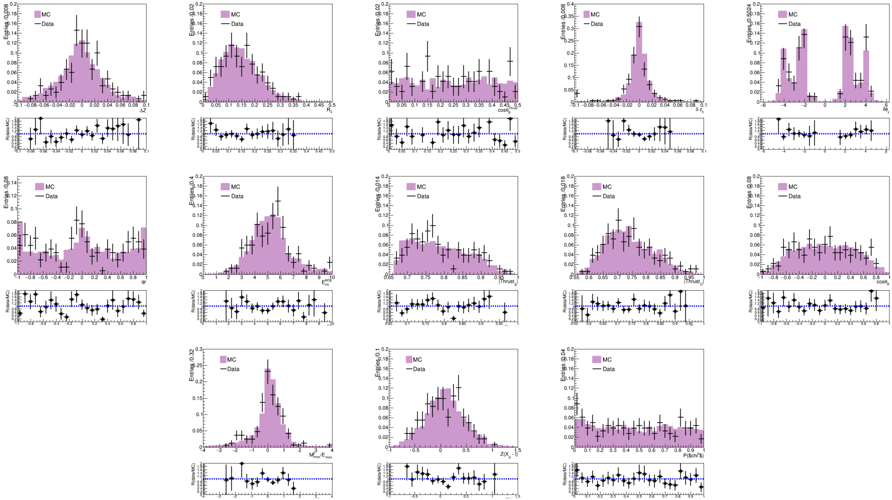


- $B \to J/\psi(\ell\ell)\omega$ is not generated in Belle
- Expect for $B^{\pm,0} \to J/\psi(\ell\ell)\pi^{\pm,0}$, all decays are dominated with background
- Assign separate fudge factors and mean shifts for modes with or without π^0
- $B^+ \to J/\psi(\ell\ell)\pi^+$ and $B^0 \to J/\psi(\ell\ell)\pi^0$ have high statistics containing π^+ and π^0, respectively
- Will use these two samples as control sample for $b \to d\ell\ell$ analysis

Signal yield extraction strategy

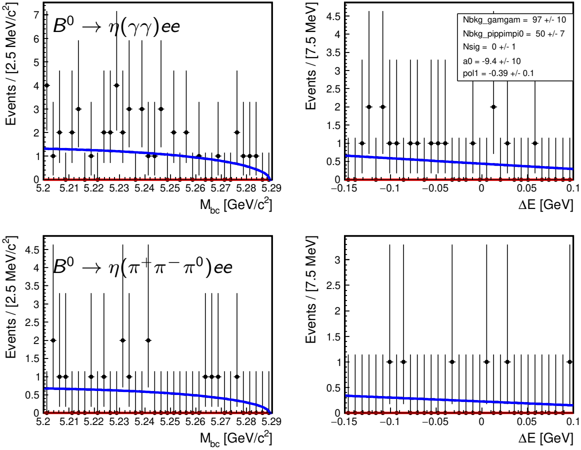
- $B^{\pm,0} \to (\omega, \pi^0, \rho^\pm)\mu\mu$ and $B^{\pm,0} \to (\omega, \pi^{\pm,0}, \rho^{\pm,0})ee$:
  - Two dimensional fit with $M_{\rm bc}$ and ΔE
- $B^{\pm,0} \to (\omega, \pi^{\pm,0}, \rho^{\pm,0})\ell\ell$:
  - Two dimensional simultaneous fit with $M_{\rm bc}$ and ΔE between $B^{\pm,0} \to (\omega, \pi^{\pm,0}, \rho^{\pm,0})\mu\mu$ and $B^{\pm,0} \to (\omega, \pi^{\pm,0}, \rho^{\pm,0})ee$
- $B^0 \to \eta\mu\mu$ and $B^0 \to \eta ee$
  - Two dimensional simultaneous fit with $M_{\rm bc}$ and ΔE between $B^0 \to \eta(\gamma\gamma)\mu\mu$ or $B^0 \to \eta(\gamma\gamma)ee$ and $B^0 \to \eta(\pi^+\pi^-\pi^0)\mu\mu$ or $B^0 \to \eta(\pi^+\pi^-\pi^0)ee$
- $B^0 \to \eta\ell\ell$
  - Two dimensional simultaneous fit with $M_{\rm bc}$ and ΔE between $B^0 \to \eta(\gamma\gamma)\mu\mu$, $B^0 \to \eta(\gamma\gamma)ee$, $B^0 \to \eta(\pi^+\pi^-\pi^0)\mu\mu$, and $B^0 \to \eta(\pi^+\pi^-\pi^0)ee$ samples

- Data and MC are compared in terms of BDT input variables in the $M_{\rm bc}$ signal region



- Shape of the distributions in data is consistent with MC
- Good agreement between data and MC

$B^0 \to \eta(\gamma\gamma, \pi^+\pi^-\pi^0)ee$

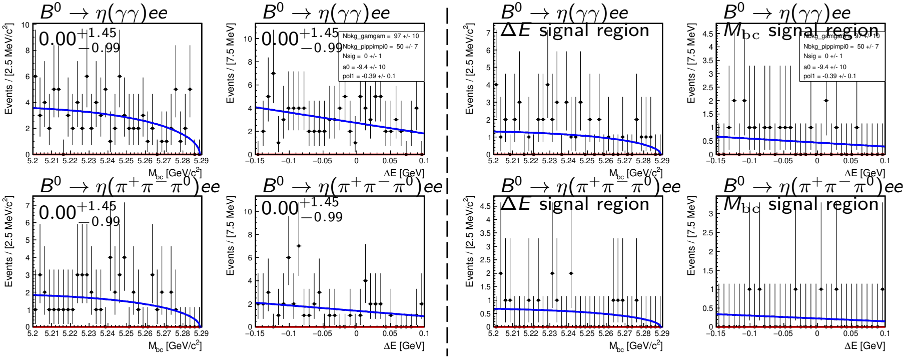


channel	$N_{\rm sig}$	$N_{\rm bkg}(\gamma\gamma)$	$N_{\rm bkg}(\pi^+\pi^-\pi^0)$
$B^0 \to \eta ee$	$0.00^{+1.45}_{-0.99}$	97 ± 10	50 ± 7

Fit result for $B^0 \to \eta ee$

- Signal enhanced projection: $M_{\rm bc} > 5.27$ GeV$/c^2$ for ΔE plot and $|\Delta E| < 0.05$ GeV for $M_{\rm bc}$ plot

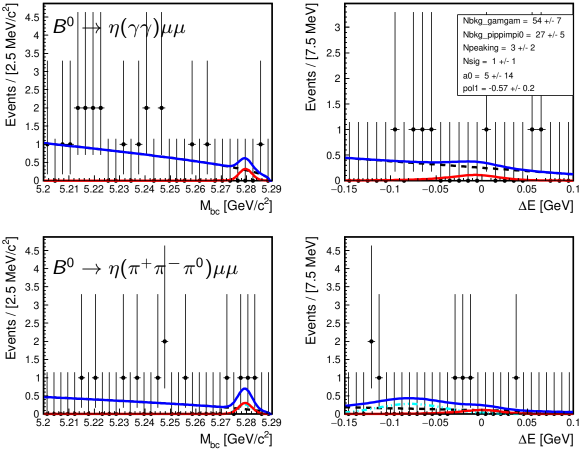
signal, **background**, **peaking background**, **total**



- $N_{\rm sig} = 0.00^{+1.45}_{-0.99}$, $0.84^{+1.51}_{-1.09}$, $-0.26^{+3.23}_{-2.53}$, $1.68^{+2.36}_{-1.63}$, $-2.92^{+1.81}_{-1.40}$, $-0.48^{+3.60}_{-2.68}$, $0.12^{+2.53}_{-1.62}$, $5.62^{+3.47}_{-2.67}$, $-4.36^{+2.79}_{-1.96}$, $2.99^{+4.02}_{-3.05}$
  for $B^0 \to \eta ee$, $B^0 \to \eta\mu\mu$, $B^0 \to \omega ee$, $B^0 \to \omega\mu\mu$, $B^0 \to \pi^0 ee$, $B^0 \to \pi^0\mu\mu$, $B^+ \to \pi^+ ee$, $B^0 \to \rho^0 ee$, $B^+ \to \rho^+ ee$, and $B^+ \to \rho^+\mu\mu$

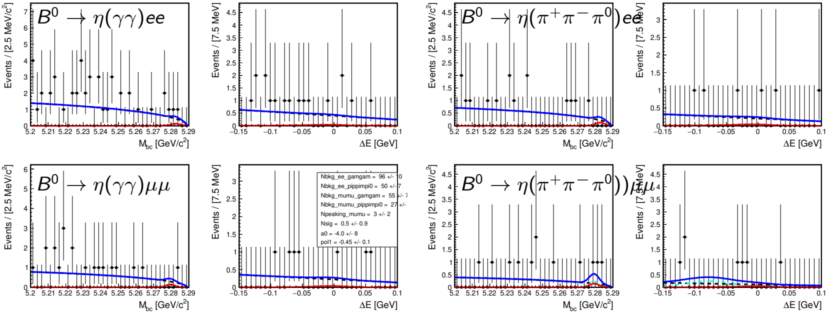
- No significant excess of signal is found for any of the decay channels

$B^0 \to \eta(\gamma\gamma, \pi^+\pi^-\pi^0)\mu\mu$



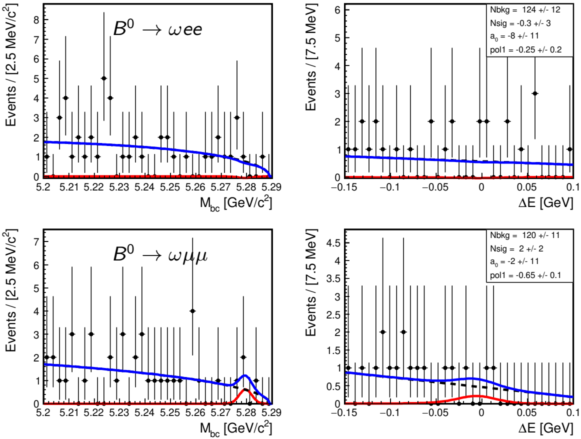
channel	$N_{\rm sig}$	$N_{\rm bkg}(\gamma\gamma)$	$N_{\rm bkg}(\pi^+\pi^-\pi^0)$
$B^0 \to \eta ee$	$0.84^{+1.51}_{-1.09}$	54 ± 7	27 ± 5

$B^0 \to \eta(\pi^+\pi^-)\ell\ell$



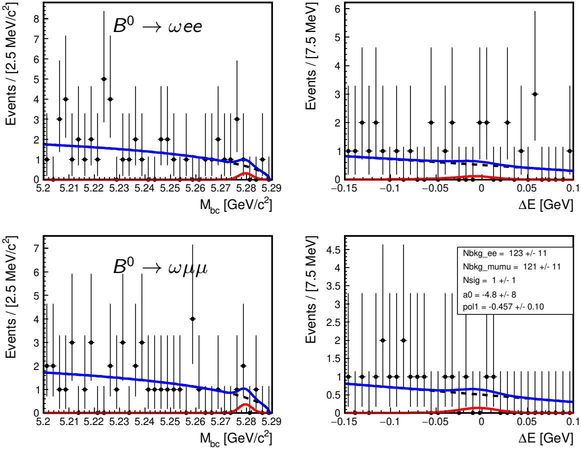
channel	$N_{\rm sig}$	$N_{\rm bkg}(\gamma\gamma)ee$	$N_{\rm bkg}(\gamma\gamma)\mu\mu$	$N_{\rm bkg}(\pi^+\pi^-\pi^0)ee$	$N_{\rm bkg}(\pi^+\pi^-\pi^0)\mu\mu$
$B^0 \to \eta\ell\ell$	$0.48^{+1.04}_{-0.83}$	96 ± 10	55 ± 7	50 ± 7	27 ± 7

$B^0 \to \omega(\pi^+\pi^-\pi^0)ee$ & $B^0 \to \omega(\pi^+\pi^-\pi^0)\mu\mu$



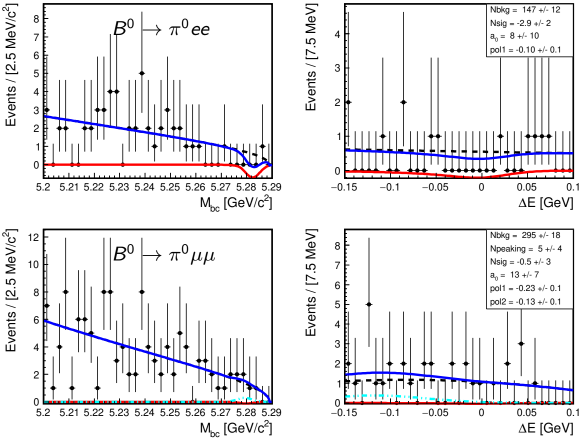
channel	$N_{\rm sig}$	$N_{\rm bkg}$
$B^0 \to \omega ee$	$-0.26^{+3.23}_{-2.53}$	124 ± 12
$B^0 \to \omega\mu\mu$	$1.68^{+2.36}_{-1.63}$	120 ± 11

$B^0 \to \omega(\pi^+\pi^-\pi^0)\ell\ell$

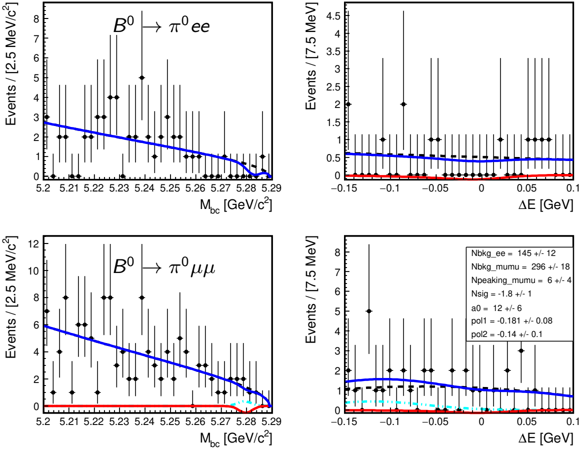


channel	N_{sig}	$N_{\mathrm{bkg}}(ee)$	$N_{\mathrm{bkg}}(\mu\mu)$
$B^0 \to \omega\ell\ell$	$1.05^{+1.75}_{-1.28}$	123 ± 11	121 ± 11

$B^0 \to \pi^0 ee$ & $B^0 \to \pi^0 \mu\mu$

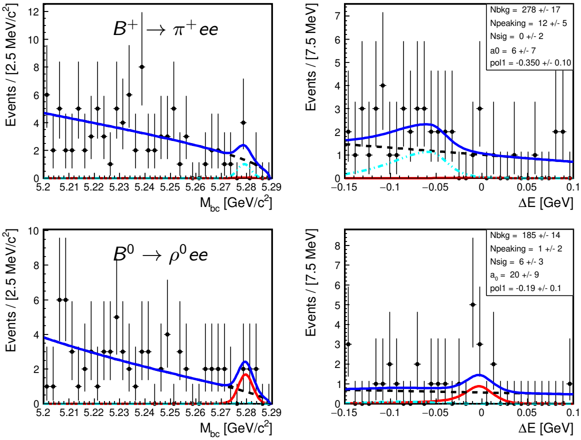


channel	$N_{\rm sig}$	$N_{\rm bkg}$
$B^0 \to \pi^0 ee$	$-2.92^{+1.81}_{-1.40}$	147 ± 12
$B^0 \to \pi^0 \mu\mu$	$-0.48^{+3.60}_{-2.68}$	295 ± 18



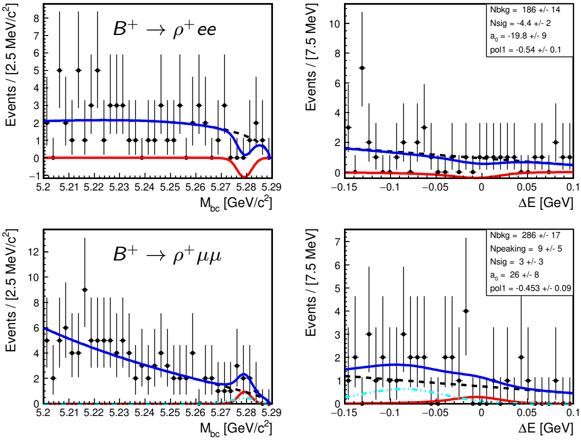
channel	N_{sig}	$N_{\mathrm{bkg}}(ee)$	$N_{\mathrm{bkg}}(\mu\mu)$
$B^0 \to \pi^0 \ell\ell$	$-1.81^{+1.62}_{-1.14}$	145 ± 12	296 ± 18

$B^+ \to \pi^+ ee$ & $B^0 \to \rho^0(\pi^+\pi^-)ee$



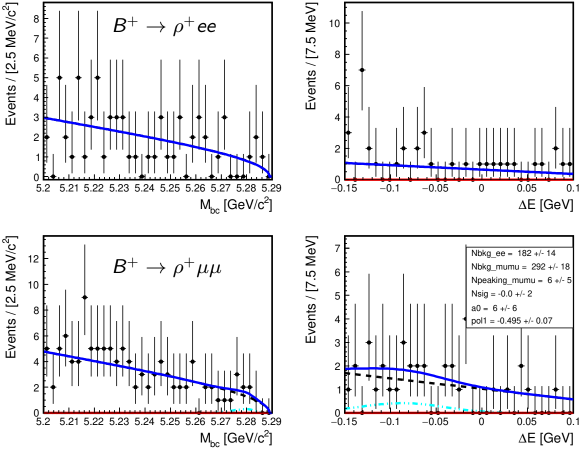
channel	$N_{\rm sig}$	$N_{\rm bkg}$
$B^+ \to \pi^+ ee$	$0.12^{+2.53}_{-1.62}$	278 ± 17
$B^0 \to \rho^0 ee$	$5.62^{+3.47}_{-2.67}$	185 ± 14

$B^+ \to \rho^+(\pi^+\pi^0)ee$ & $B^+ \to \rho^+(\pi^+\pi^0)\mu\mu$



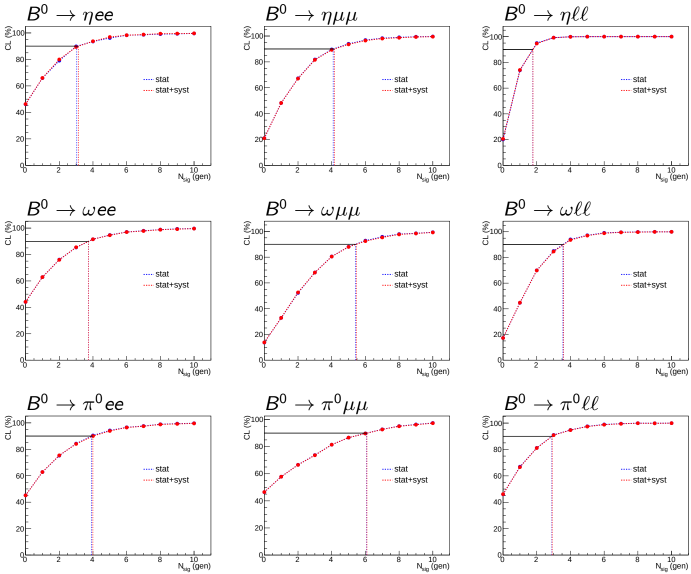
channel	$N_{\rm sig}$	$N_{\rm bkg}$
$B^+ \to \rho^+ ee$	$-4.36^{+2.79}_{-1.97}$	186 ± 14
$B^+ \to \rho^+ \mu\mu$	$2.99^{+4.03}_{-3.05}$	286 ± 17

$B^+ \to \rho^+(\pi^+\pi^0)\ell\ell$



channel	N_{sig}	$N_{\mathrm{bkg}}(ee)$	$N_{\mathrm{bkg}}(\mu\mu)$
$B^0+ \to \rho^+\ell\ell$	$0.39^{+2.41}_{-1.95}$	182 ± 14	292 ± 18

$N_{\rm sig}^{\rm UL}$

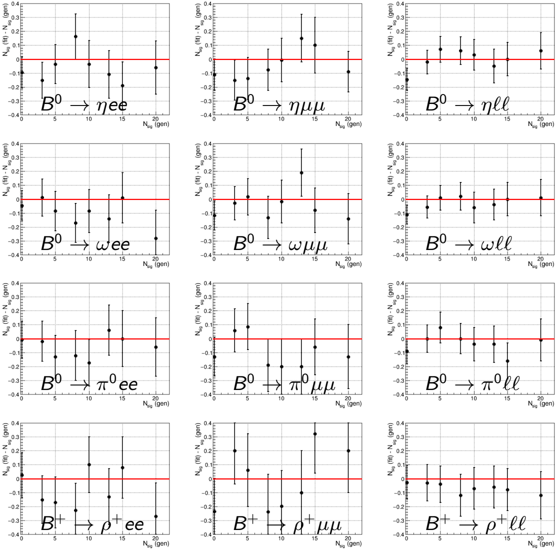


$B^0 \to \eta ee$
$B^0 \to \eta \mu\mu$
$B^0 \to \eta \ell\ell$
$B^0 \to \omega ee$
$B^0 \to \omega \mu\mu$
$B^0 \to \omega \ell\ell$
$B^0 \to \pi^0 ee$
$B^0 \to \pi^0 \mu\mu$
$B^0 \to \pi^0 \ell\ell$
CL (%)
N_{sig} (gen)
stat
stat+syst

Systematic Uncertainties

- lepton identification: 0.3% and 0.4% for muon and electron, calculated using inclusive J/ψ sample
- π^+: 0.5% using $D^{*+} \to D^0(K^-\pi^+)\pi^+$ sample
- π^0: 2.3% using $\tau^- \to \pi^-\pi^0\nu_\tau$
- γ: 2% determined from radiative Bhabha and $B^0 \to K^{*0}\gamma$
- FastBDT: Ratio of efficiency between data and MC from simultaneous fit between the samples having BDT>optimal and BDT<optimal, used $B^{+,0} \to J/\psi\pi^{+,0}$
- MC statistics: Uncertainty in efficiency due to limited MC statistics is $< 1\%$
- decay model: Difference in efficiency by changing BTOSLLBALL decay model to LCSR and QUARK models
- mass window: Difference in signal yield in MC by setting the mass resolution of data to MC for ρ, η, ω resonances is assigned as mass window systematic, $1-3\%$
- BCS: Change in signal yield by changing best candidates selection to random candidate selection
- Tracking: Assigned a systematic of 0.35% per track
- PDF shape: Parameters fixed in the fit are varied by $\pm 1\sigma$ and the difference in signal yield from nominal fit is given as systematic, effect is $< 1\%$
- $f^{+-/00}$: $\mathcal{B}[\Upsilon(4S) \to B^+B^-](\mathcal{B}[\Upsilon(4S) \to B^0\overline{B^0}])$ is 2.4%
- $N_{B\overline{B}}$: Uncertainty in the number of $B\overline{B}$ pair is 1.4%

ToyMC study



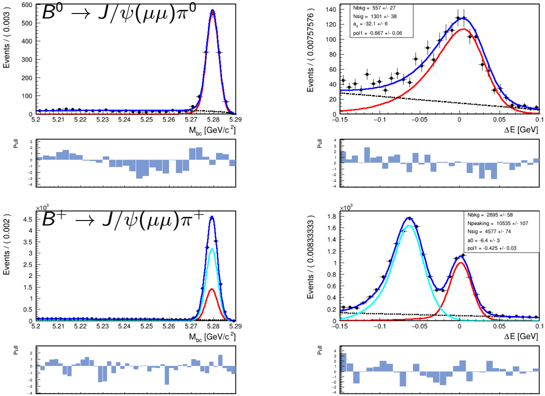
$B^0 \to \eta ee$
$B^0 \to \eta\mu\mu$
$B^0 \to \eta\ell\ell$
$B^0 \to \omega ee$
$B^0 \to \omega\mu\mu$
$B^0 \to \omega\ell\ell$
$B^0 \to \pi^0 ee$
$B^0 \to \pi^0\mu\mu$
$B^0 \to \pi^0\ell\ell$
$B^+ \to \rho^+ ee$
$B^+ \to \rho^+\mu\mu$
$B^+ \to \rho^+\ell\ell$
N_{sig} (fit) - N_{sig} (gen)
N_{sig} (gen)

Charmless peaking backgrounds

- Charmless peaking background is found for the modes with muons in the final state
- Peaking background is studied in rareMC

channel	background	MC ($\mathcal{B}$)	N_{peak} (50×)	N_{peak} (Belle)	status
$B^0 \to \eta\mu\mu$	$B^0 \to \eta K\pi$	1.15×10^{-5}	14	0.28	**included in PDF**
	$B^0 \to \eta\pi\pi$	0.62×10^{-5}	4	0.08	negligible
$B^0 \to \omega\mu\mu$	$B^0 \to \omega K\pi$	0.30×10^{-5}	2	0.04	negligible
	$B^0 \to \omega\pi\pi$	0.10×10^{-5}	0	0	no contribution
$B^+ \to \rho^+\mu\mu$	$B^+ \to \rho^+ K\pi$	0.80×10^{-5}	7	0.14	negligible
$B^0 \to \pi^0\mu\mu$	$B^0 \to \pi^0 K\pi$	0.94×10^{-5}	13	0.26	**included in PDF**
	$B^0 \to \pi^0\pi\pi$	0.95×10^{-5}	3	0.06	negligible
	$B^0 \to \pi^0 KK$	1.89×10^{-5}	24	0.48	**included in PDF**

- Performed 2d fit in $M_{\rm bc}$ and ΔE to extract the signal yield
- Shape parameters of $B^+ \to J/\psi(\ell\ell)K^+$ peaking background for $B^+ \to J/\psi(\ell\ell)\pi^+$ decay is fixed from signal MC



channel	$N_{\rm sig}$	ε (%)	$\mathcal{B}$	decay.dec $\mathcal{B}$
$B^0 \to J/\psi(\mu\mu)\pi^0$	1301 ± 38	21.45	$(1.309 \pm 0.093) \times 10^{-6}$	1.311×10^{-6}
$B^0 \to J/\psi(ee)\pi^0$	773 ± 29	12.20	$(1.368 \pm 0.126) \times 10^{-6}$	1.313×10^{-6}
$B^+ \to J/\psi(\mu\mu)\pi^+$	4590 ± 74	35.18	$(2.817 \pm 0.111) \times 10^{-6}$	2.861×10^{-6}
$B^+ \to J/\psi(ee)\pi^+$	2257 ± 56	17.79	$(2.739 \pm 0.166) \times 10^{-6}$	2.866×10^{-6}

- Fitted yield for each stream is consistent with truth match value within the uncertainty
- Calculated $\mathcal{B}$ is consistent with decay.dec within the uncertainty