A search for baryon-number violation in B-meson decays to two baryons

Matt Bellis [1] and **Josephine Swann** [2] and **Brian Shuve** [3]
mbellis@siena.edu [1] , *jr21swan@siena.edu* [2], *bshuve@g.hmc.edu* [3]

For the BaBar Collaboration
8/27/2025

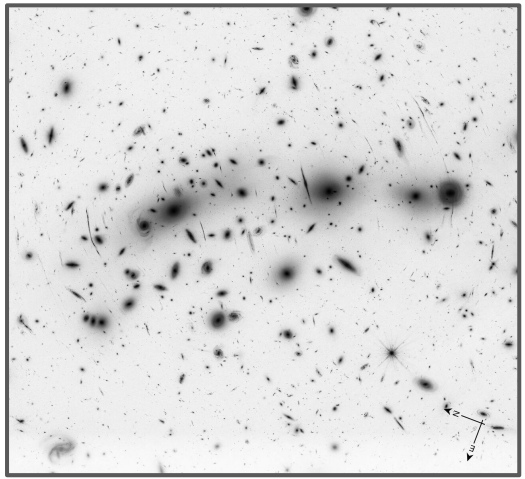
Matter - antimatter asymmetry

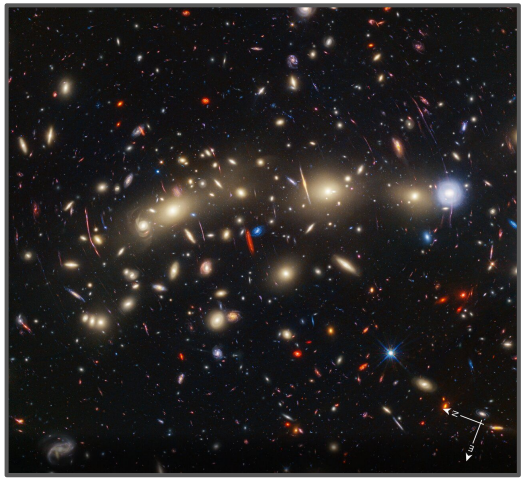
- Universe is asymmetrical wrt matter and antimatter
- Sakharov's 3 conditions:
  - Reactions out of thermal equilibrium ✅
  - CP Violation (quarks) ✅
  - **Baryon Number Violation**
    - Sphalerons (theoretical)

→ Either CP violation in something other than the quark sector…

or there are additional forms of BNV

NASA, ESA, CSA, STScI, J. Diego (Instituto de Física de Cantabria, Spain), J. D'Silva (U. Western Australia), A. Koekemoer (STScI), J. Summers & R. Windhorst (ASU), and H. Yan (U. Missouri)





Early universe: symmetry

time

Less matter, lack of antimatter

"Violation of *CP* invariance, *C* asymmetry, and baryon asymmetry of the universe" A. Sakharov

"On anomalous electroweak baryon-number non-conservation in the early universe" – V.A. Kuzmin, V.A. Rubakov, M.E. Shaposhnikov

What could generate BNV? - *Leptoquarks*

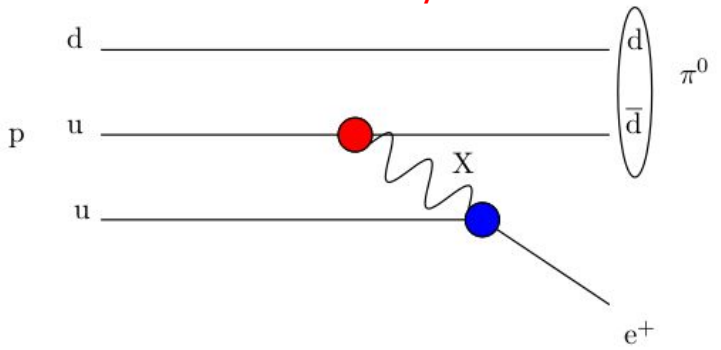
ΔB=1

B-L is the conserved quantity

Lifetime is > 0.96 x 10^{29} years

First generation interactions

Proton decay

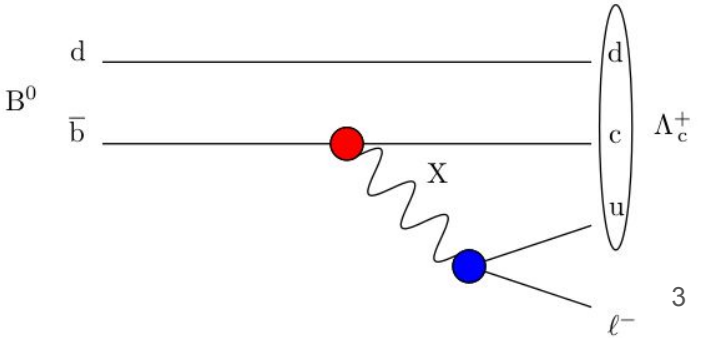


e.g. "Searches for the baryon- and lepton-number violating decays $B0 \to \Lambda + c\ell - B0 \to \Lambda c + \ell -$, $B- \to \Lambda\ell - B- \to \Lambda\ell -$, and $B- \to \bar{\Lambda}\ell - B- \to \bar{\Lambda}\ell -$"

Phys.Rev.D 83 (2011) 091101

Not-first generation interaction

B-meson decay



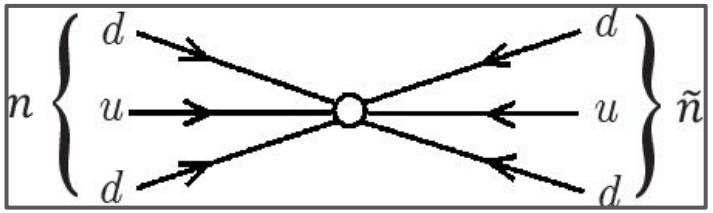
No observation to date

A different type of BNV

Suggestion at March 2024
BaBar Collaboration Meeting

ΔB=2:

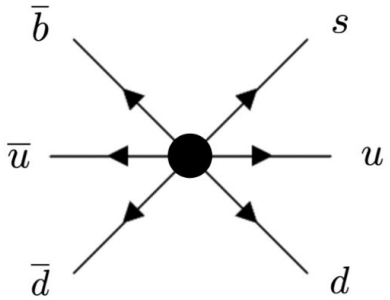
Neutron-antineutron oscillations



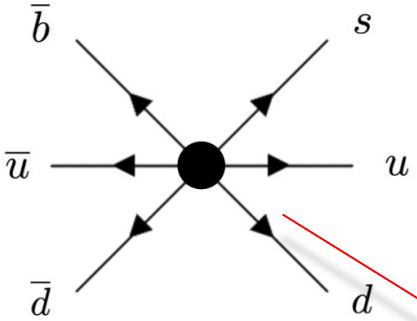
Neutron-antineutron oscillation

This search:

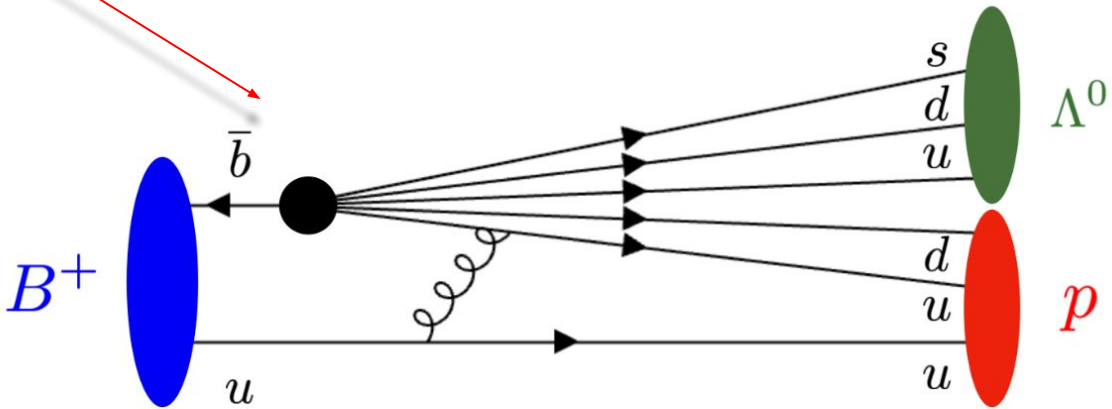
- Proton decay and neutron oscillation places severe constraints on BNV
  - Search in heavy quarks
  - $B^+ \to p^+ \Lambda^0$
- This is a signature in RPV SUSY models



Note: Belle II: $B^- \to \bar{\Xi}_c^0 \bar{\Lambda}_c^-$ $UL_{95\%}(BR) \approx 3 \times 10^{-5}$
Phys. Rev. Lett. 133, (Aug, 2024) 071802

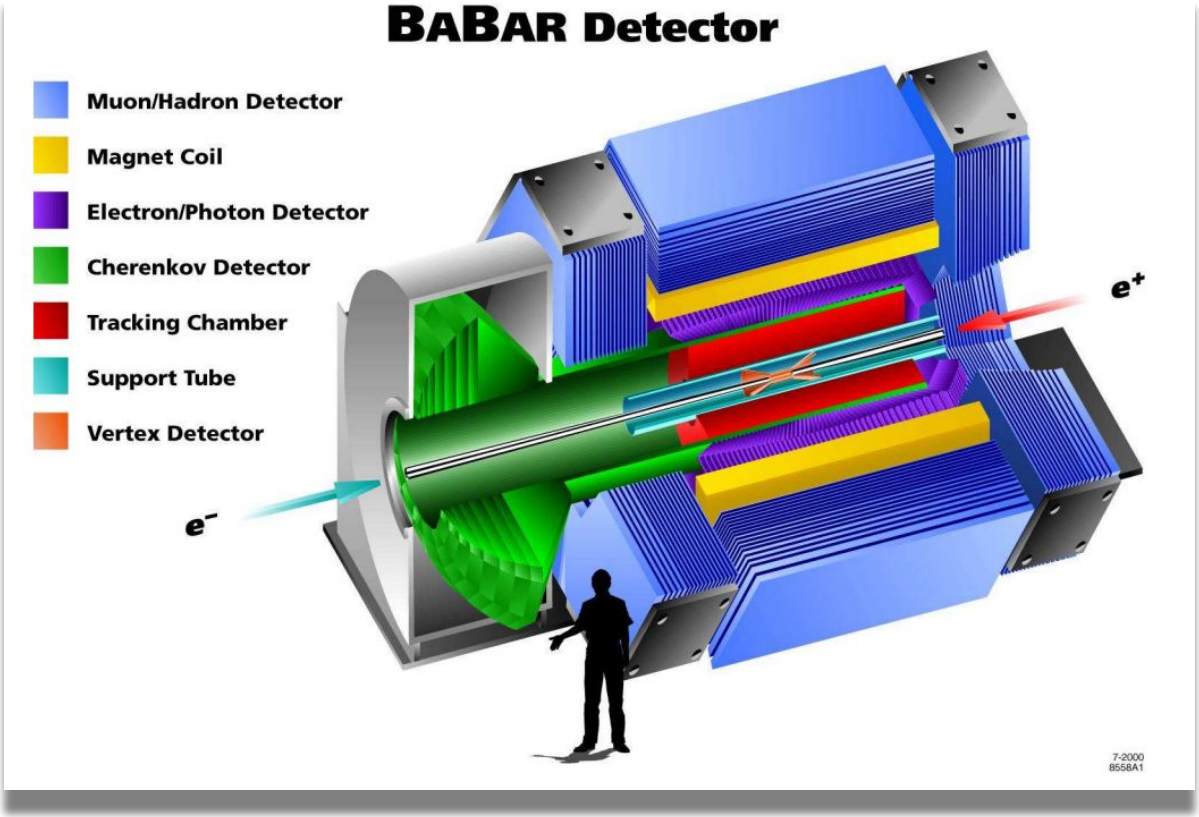


$$B^+ \rightarrow p\,\Lambda^0$$



BaBar

$$e^+ e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$$



Ran at 10.58 GeV CM
(also 10.36 GeV and 10.02 GeV [Y(3S), Y(2S)]

SLAC PEP-II collider - 1999-2008

Sub-detectors

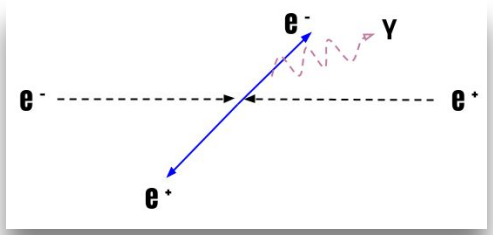
- Vertex detector
  - (position of primary/secondary vertex)
- Tracking
  - (momentum, dEdX)
- Cerenkov
  - (velocity)
- Muon system
  - (hadron/muon discrimination)

Provides input to particle ID hypotheses (e.g. proton/K/pi discrimination)

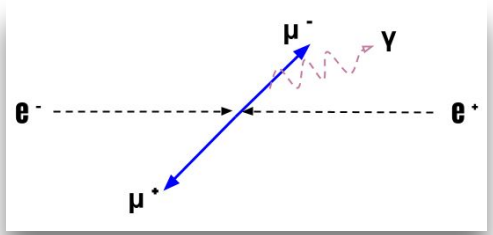
424 fb^{-1} collected at $\Upsilon(4S)$

~230 million $B^0\bar{B}^0$ pairs
~230 million B^+B^- pairs

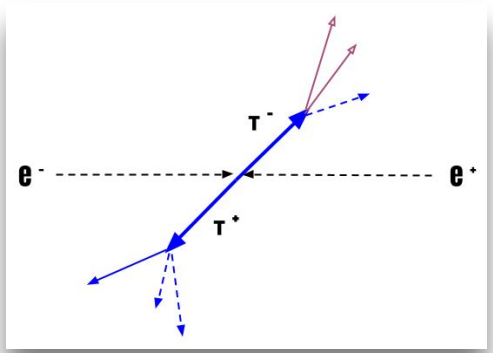
BaBar - e^+e^- collider physics



$e^+e^- \rightarrow e^+e^- \ (\gamma)$

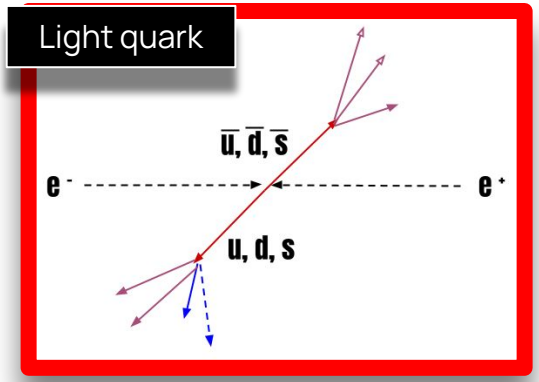


$e^+e^- \rightarrow \mu^+\mu^- \ (\gamma)$



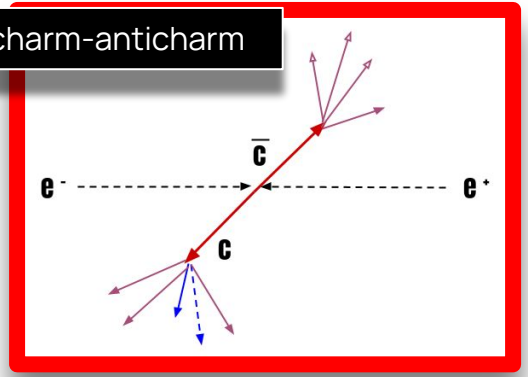
$e^+e^- \rightarrow \tau^+\tau^-$

Light quark

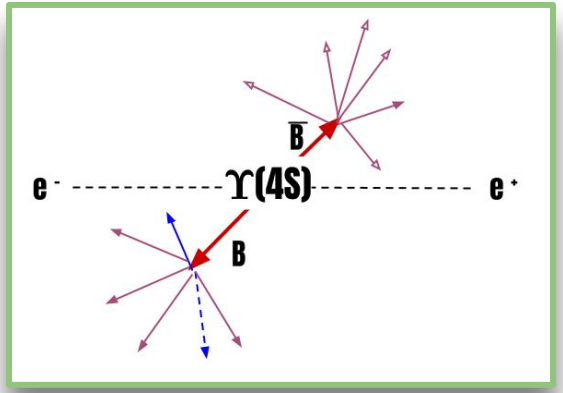


$e^+e^- \rightarrow u\bar{u} \,/\, d\bar{d} \,/\, s\bar{s}$

charm-anticharm



$e^+e^- \rightarrow c\bar{c}$



$e^+e^- \rightarrow \Upsilon\,(4S) \rightarrow B\bar{B}$

Parameter Space

Discriminating variables:

- $M_{ES} = \sqrt{\left(E^*_{\text{beam}}\right)^2 - \left(\vec{p}^*_B\right)^2}$
- **ΔE** = $(\frac{1}{2} E_{\text{beam}} - E_{\text{B-cand}})$

Blinded analysis

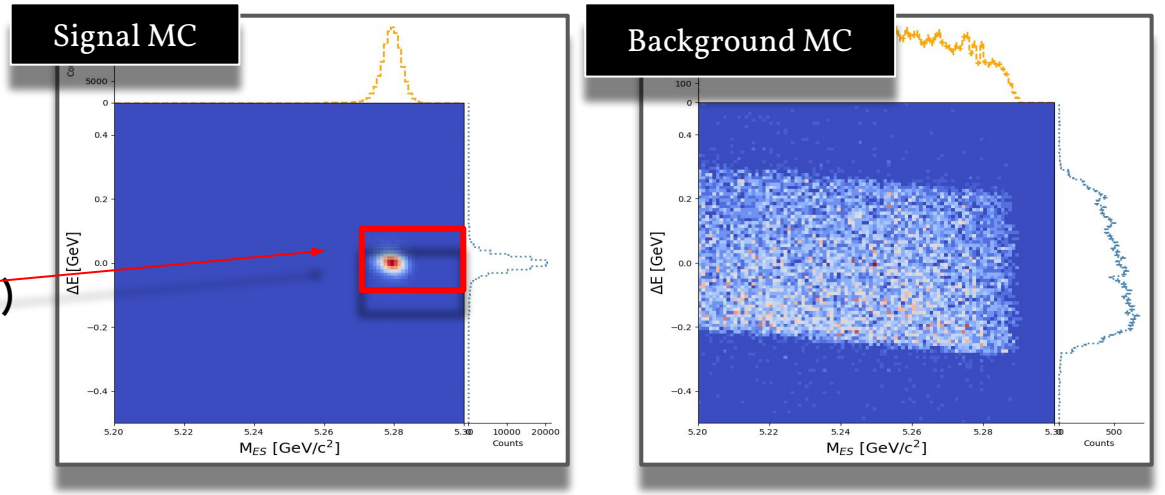
Signal region (blinded for ~~collision~~ data)

- $M_{ES} > 5.27\ \text{GeV}/c^2$
- $-0.07 < \Delta E < 0.07\ \text{GeV}$

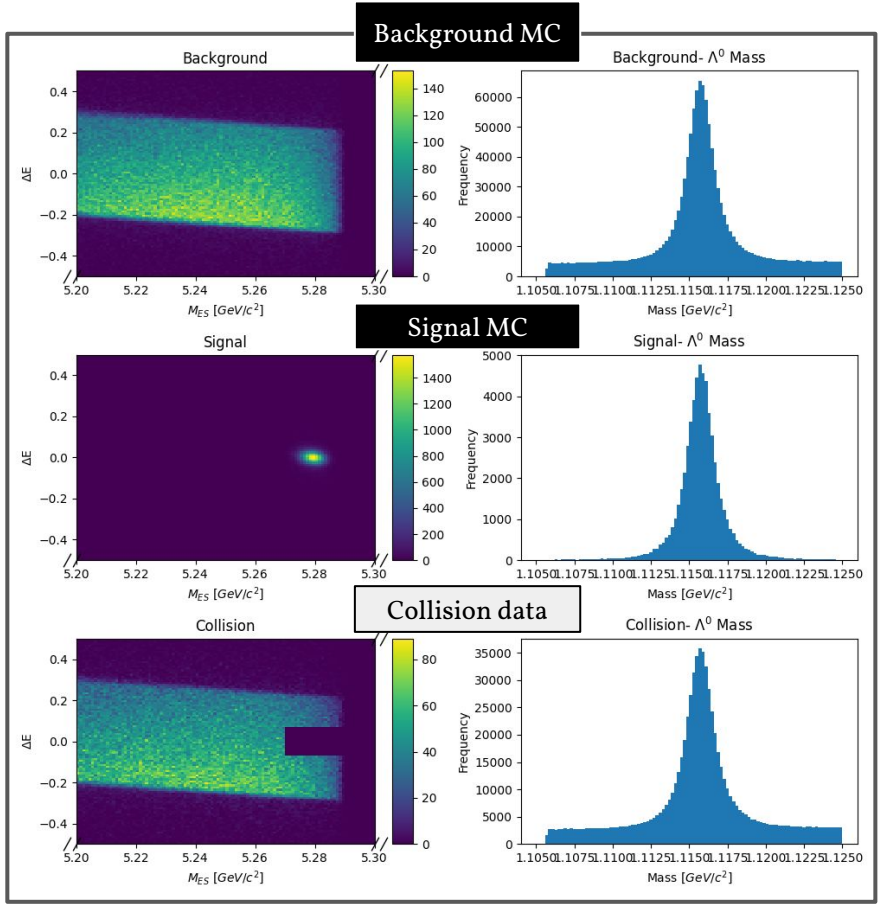
Contains 95% of the signal

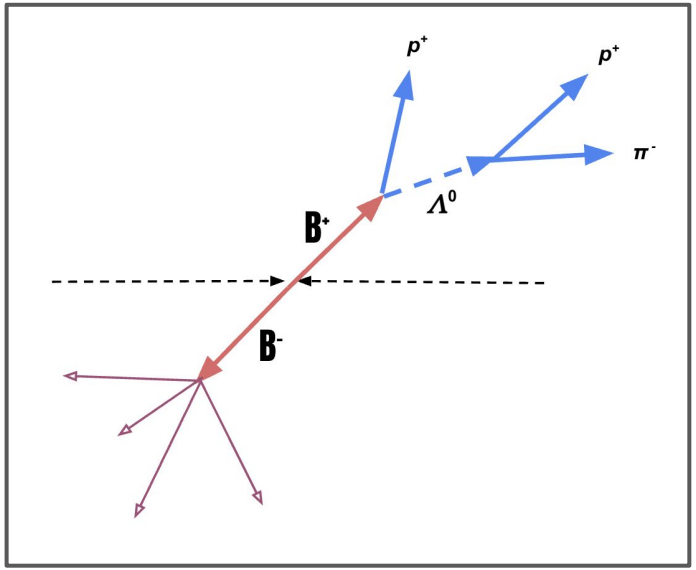
Background is primarily $e^+e^- \to \bar{q}\,q$ ($u\bar{u}/d\bar{d}/s\bar{s}/c\bar{c}$)

Minimal selection criteria applied



Selection criteria





Lambda^0 mass and flight length significance

Λ^0 $c\tau$ = 7.89 cm

Flight significance = $\frac{L}{\sigma_L}$

Vary cut on flight significance and calculate purity of Lambda^0

Mass window retains 93% from signal MC

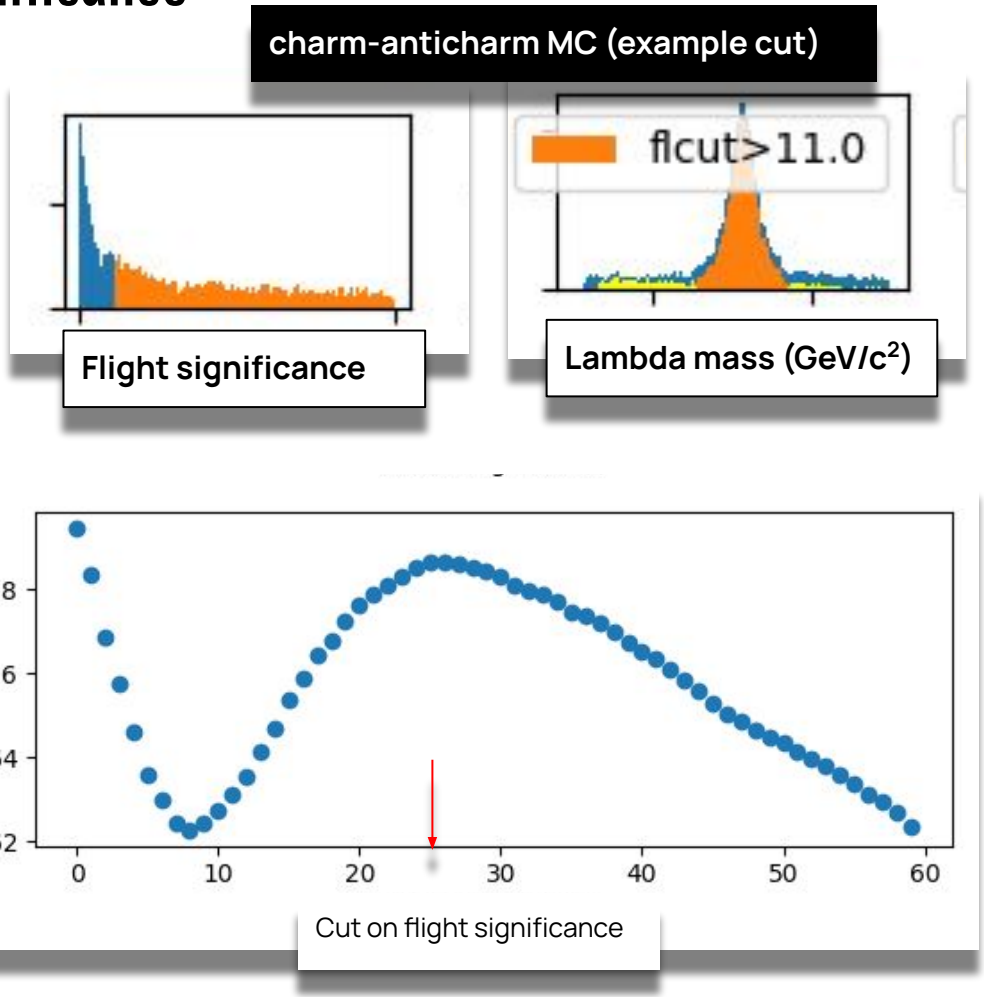
- Nominal Lambda mass +/- 0.003 GeV/c^2

Optimize $\frac{S}{\sqrt{S+B}}$

Cut on flight length significance at 25

Retains ~82% of signal

Reduces background by
- ~65% for charm-anticharm
- ~40% for light quark



Particle ID (PID)

For each PID hypothesis there are 6 levels of purity

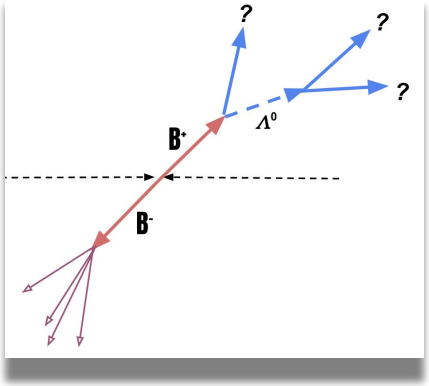
- **`VeryLoose - SuperTight`**

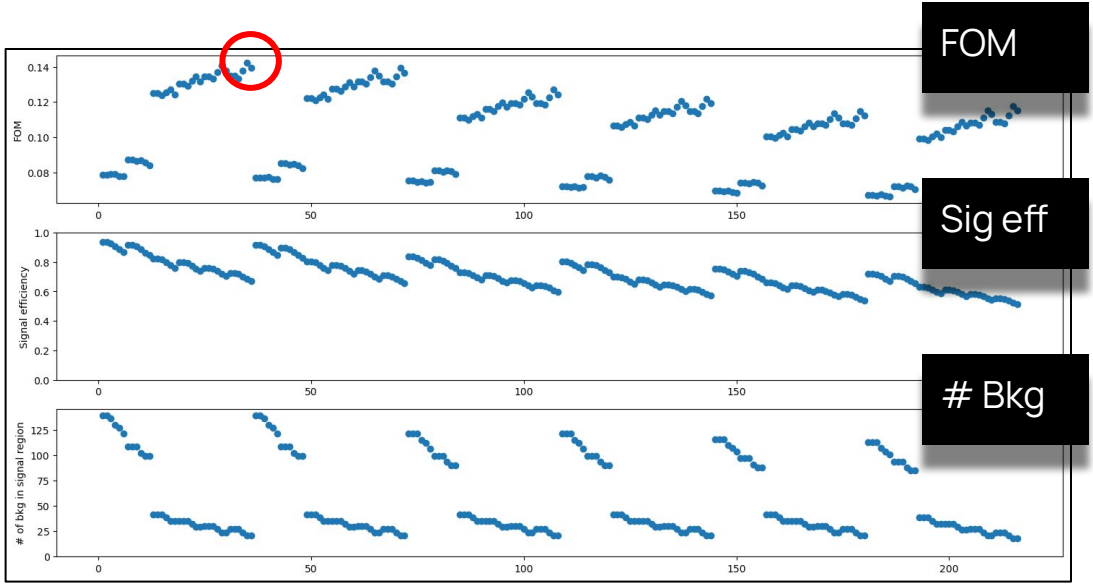
6 (proton) x 6 (proton) x 6 (pion) = 216 combinations

Vary selectors and count number of events in the signal region (M_{ES} and ΔE)

Optimize using Punzi figure-of-merit $\frac{\epsilon(t)}{a/2+\sqrt{B(t)}}$

- ϵ - signal efficiency
- B - # of background events
  - Estimated from collision data sidebands
- a - desired significance of discovery (4)





Retains ~67% of signal

Reduces background by

- 99% for charm-anticharm
- 98% for light quark

Largest impact is from PID of proton from B decay

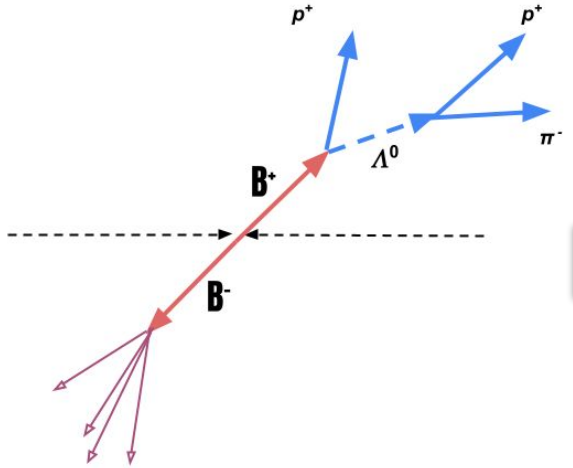
Antiproton veto (charge-conjugation implied)

A **standard model** process will have additional proton candidates in the final state that are *oppositely charged* from the proton candidates in the signal reconstruction

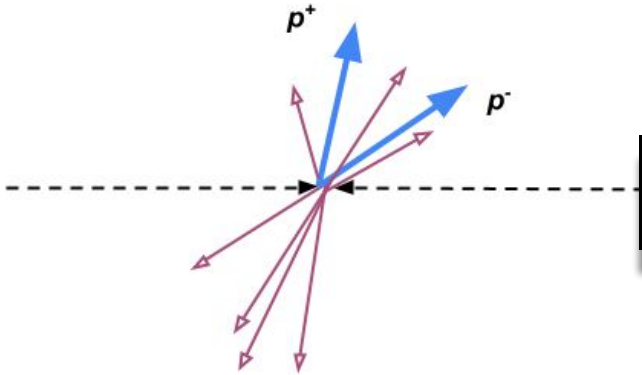
Look for viable anti-protons and cut out those events

Retains 85-90% of signal

Eliminates
- 44% of charm-anticharm
- 28% of light quark



Signal



Background (SM)

Classifier

After previous cuts, ~400 events

Majority of background is *light quark*

Train MLP classifier (scikit-learn)

Relax earlier cuts and use
30k bkg + 30k signal events for training

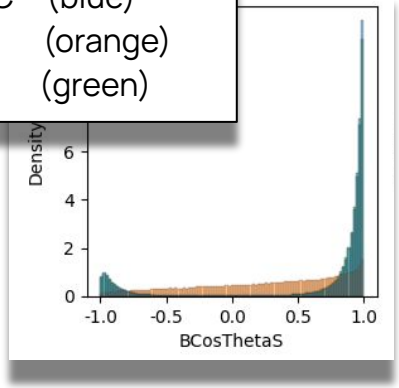
Training variables (6)

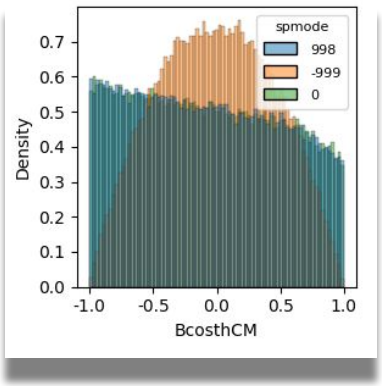
- Two are shown at right
  - B cos(theta) sphericity wrt ROE sphericity
  - B cos (theta) in the center-of-mass

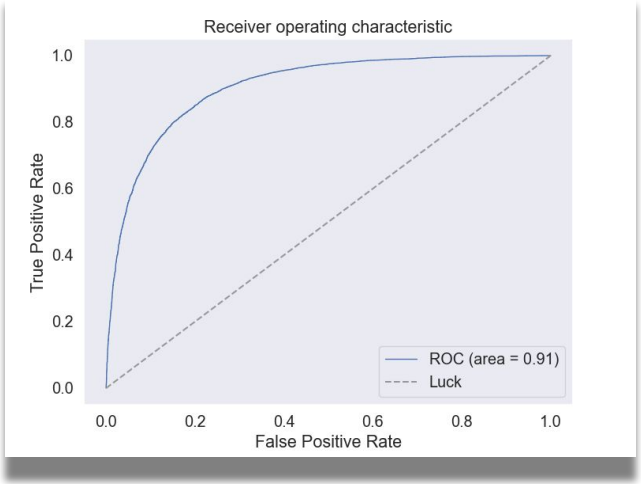
Use Punzi fom to determine selection point

$$\frac{\epsilon(t)}{a/2+\sqrt{B(t)}}$$

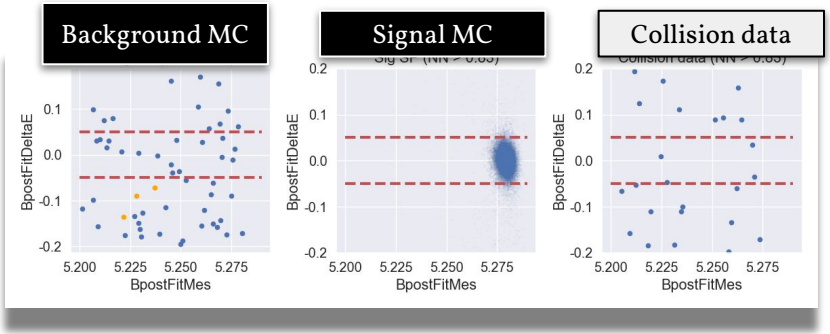
- Background MC (blue)
- Signal MC (orange)
- Collision data (green)







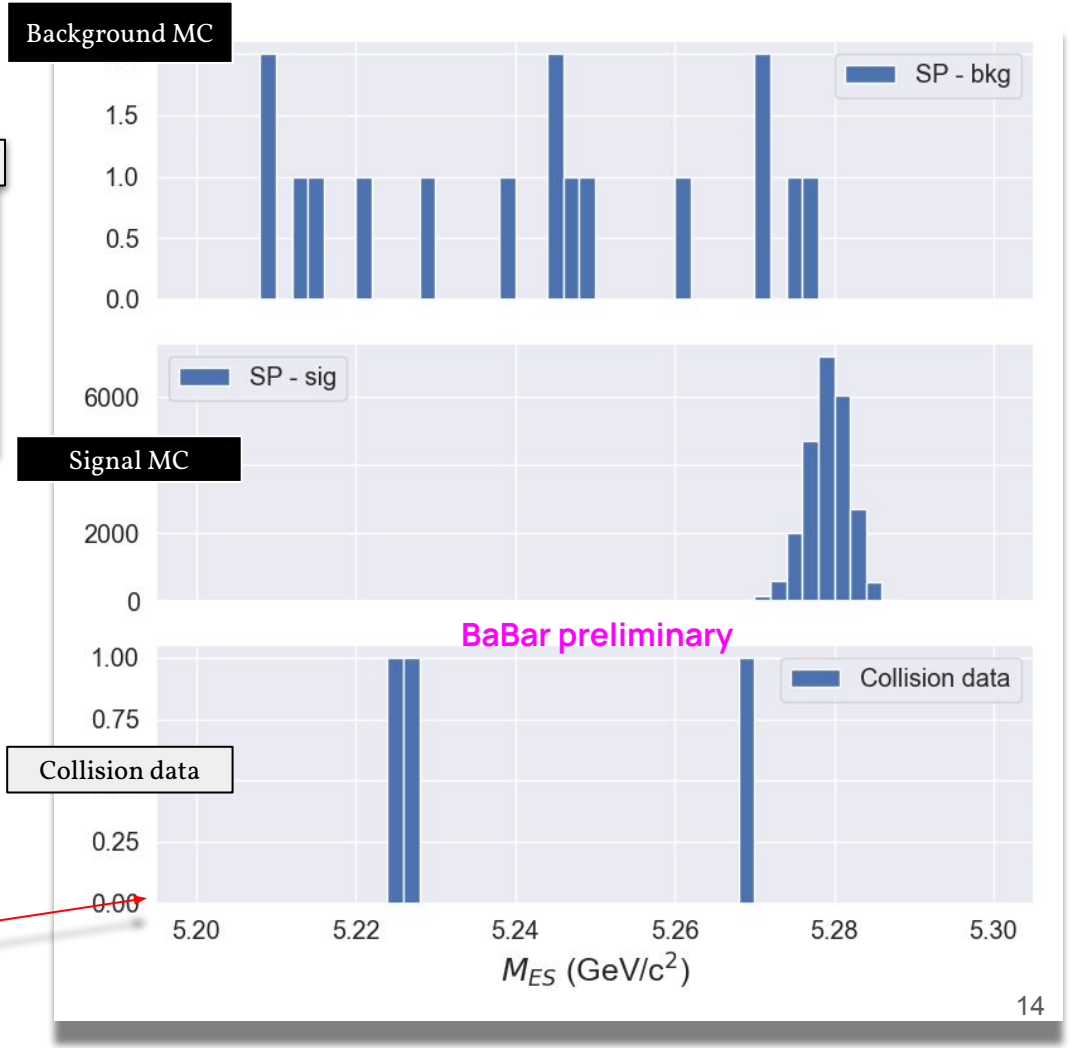
After all cuts



M_{ES} in red bands

(tight cut on ΔE)

3 events survive in data (still blinded)



Systematic and statistical uncertainties

$$\mathcal{B}\left(B^+ \to p\,\Lambda^0\right) = \frac{N_{sig}}{\epsilon_{sig} \times \mathcal{B}_{\Lambda^0 \to p\pi^-} \times N_{B\bar{B}} \times 2 \times \mathcal{B}_{\Upsilon(4S) \to B^+B^-}}$$

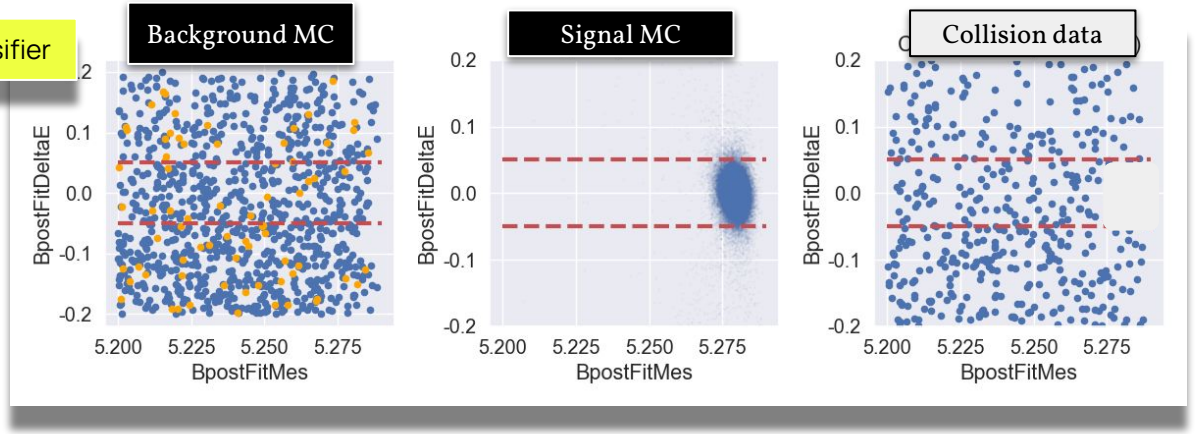
Uncertainties on signal efficiency and conversion from # signal events → branching fraction

As we will see, there are so few events remaining after all selection criteria that these are subdominant contributions to the final resul.

Contribution	
Systematic uncertainties	
B counting (%)	0.60
$\mathcal{B}(\Upsilon(4S) \to B^+B^-)$ (%)	1.17
$\mathcal{B}(\Lambda^0 \to p\pi^-)$ (%)	0.78
Tracking eff. (%)	0.38
PID eff. (%)	1.73
Statistical uncertainties	
Efficiency (MC stat.) (%)	0.34
Total (%)	**2.38**

Blinded

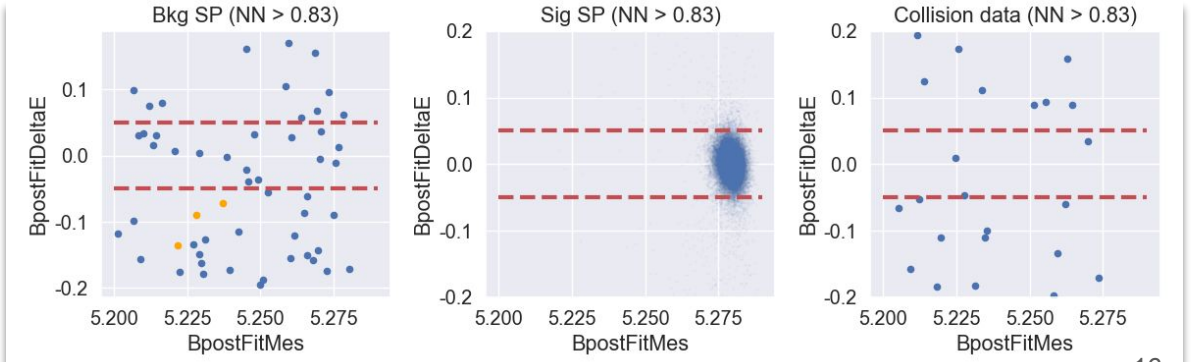
No cut on classifier



Collision data blinded here

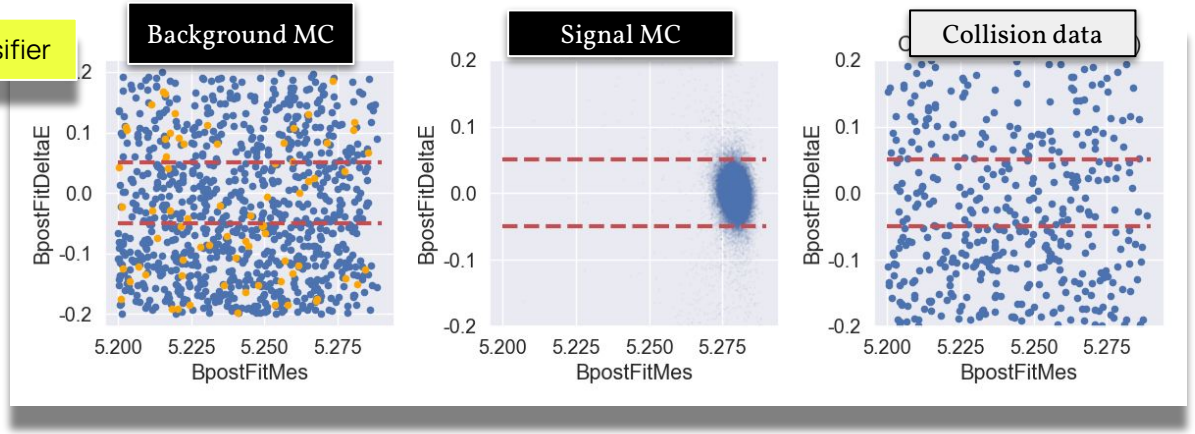
Cut on classifier

BaBar preliminary



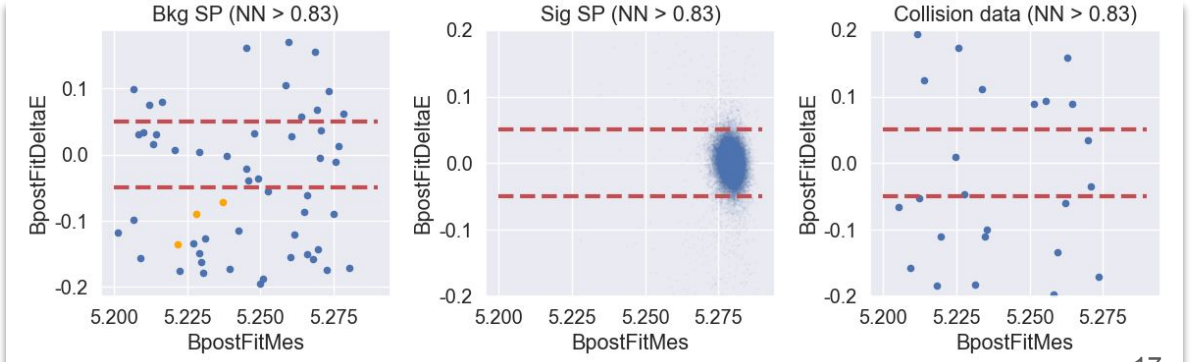
Unblinded

No cut on classifier



Cut on classifier

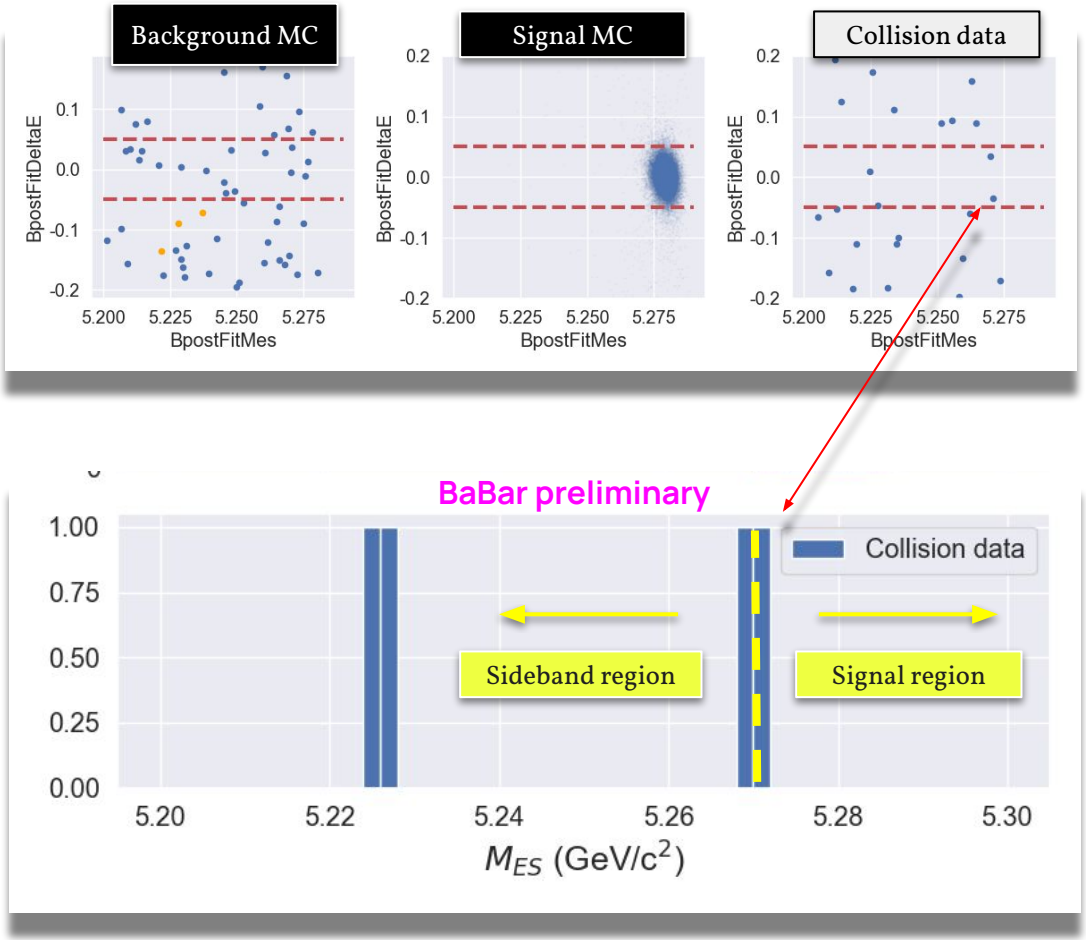
BaBar preliminary



Unblinded

3 events in sideband

1 event in signal region



Upper limit

"Limits and Confidence Intervals in the Presence of Nuisance Parameters" Rolke et al

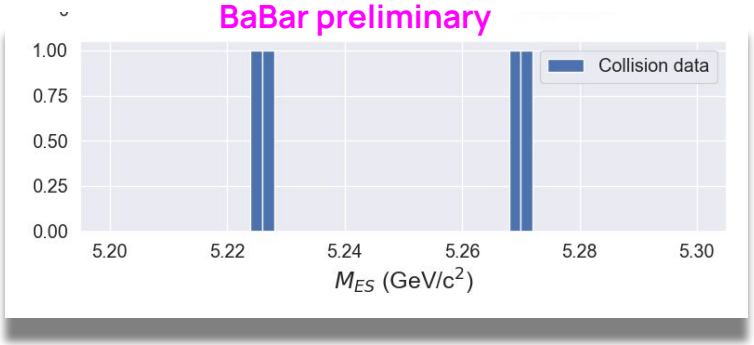
$$\ln \mathcal{L}\left(x, y, z \,||\, \mu, b, \epsilon_s\right) = x \ln\left(\epsilon_s \mu + b\right) - \left(\epsilon_s \mu + b\right) + y \ln\left(\tau b\right) - \left(\tau b\right) - \frac{1}{2}\left(\frac{\epsilon - z}{\sigma_\epsilon}\right)^2$$

Use above function in likelihood ratio test

μ - signal rate
b - background rate
ε - signal efficiency

x - observed events in signal region (1)
y - observed events in sideband region (3)
z - calculated efficiency (16.08%)
$\boldsymbol{\sigma_\varepsilon}$ - uncertainty on signal efficiency (0.38%)

τ - "scaling" to account for relative size of sideband and signal regions



$UL_{90\%}$(eff corr N_{sig}) = 18.96 events

$UL_{90\%}$(BR) = 6.1×10^{-8}

BaBar preliminary

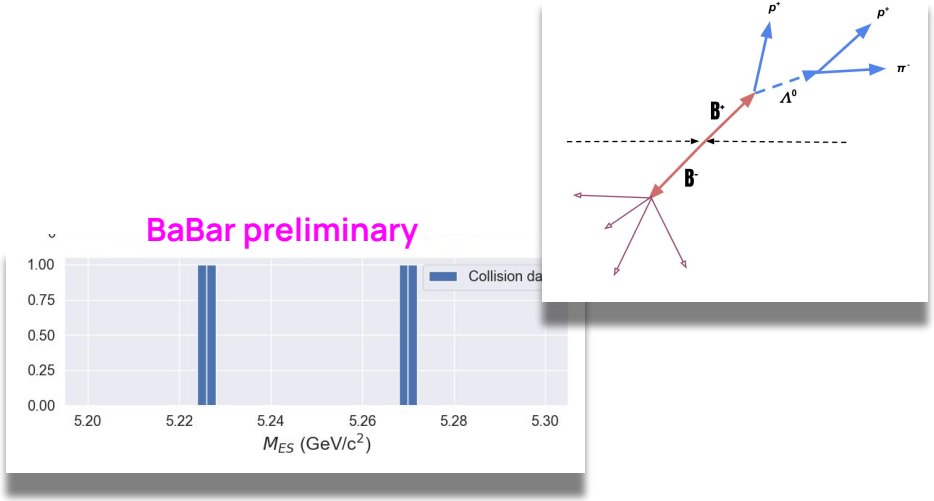
Results

This is the first search for $B^+ \rightarrow p\,\Lambda^0$

Thank you to all the BaBar collaborators who provided support and advice during this analysis!

17 years after shutdown, we are still contributing novel physics measurements

Paper is in progress…



$UL_{90\%}(BR) = 6.1 \times 10^{-8}$

BaBar preliminary

Thank you for your time!

Backup slides

BaBar status

2009 Long-Term Data Storage (LTDA) 1 2 3

Freeze all code (reconstruction, MC, DB, etc.)

Virtual Machines

SLAC (for a while)

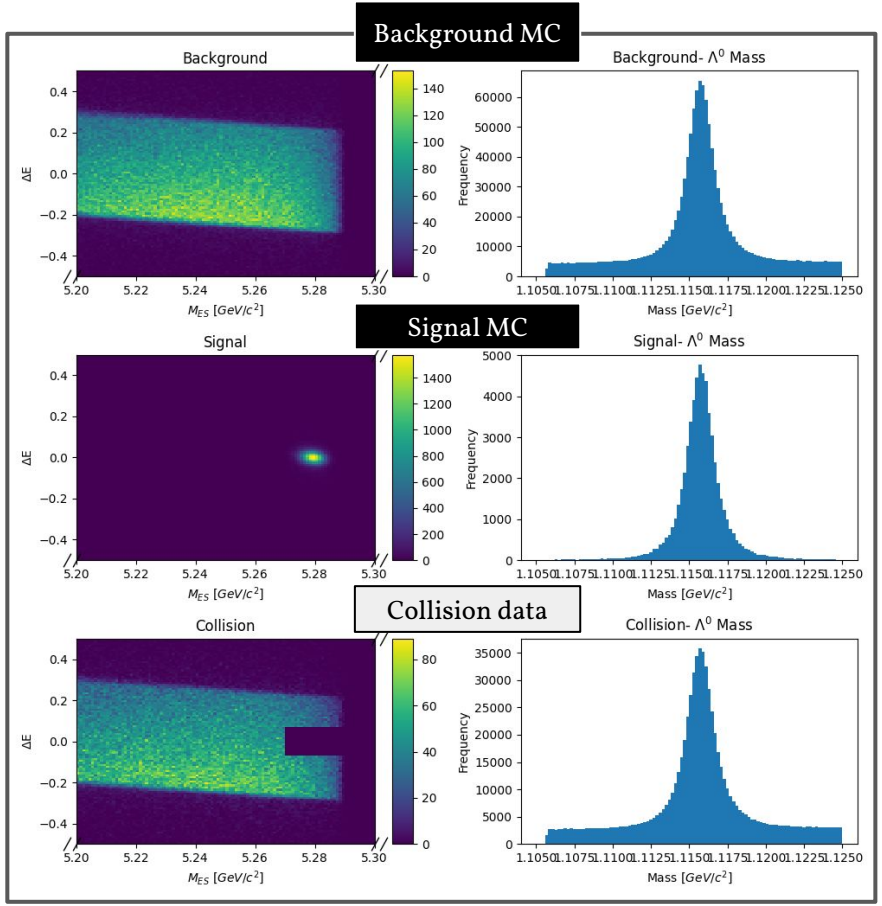
Current status

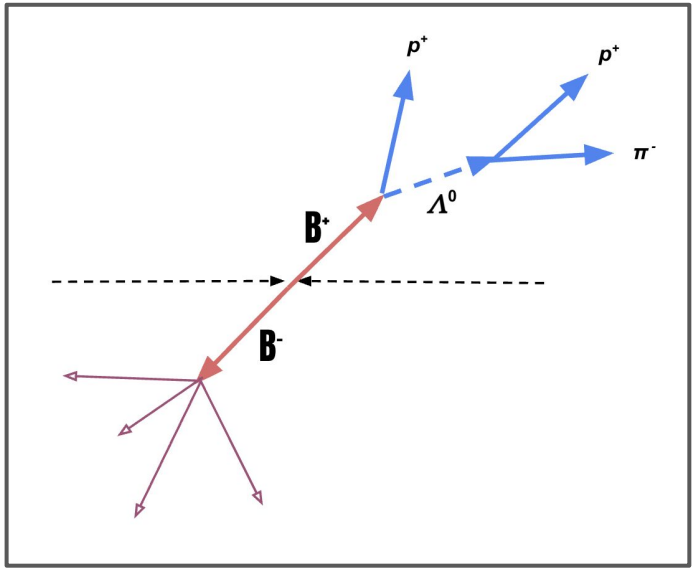
- Code, CPUs, documentation, communications
  - *University of Victoria (BC, Canada)*
  - High Energy Physics Research Computing
  - Google Docs / INSPIRE
  - CERN e-Talk
- Data
  - GridKa (~1.5 PB)
    - "Grid Computing Centre Karlsruhe" (Germany)
- Person-power
  - Marcus Ebert (UVic)

For more details see talk at APS Global Physics Summit 2025

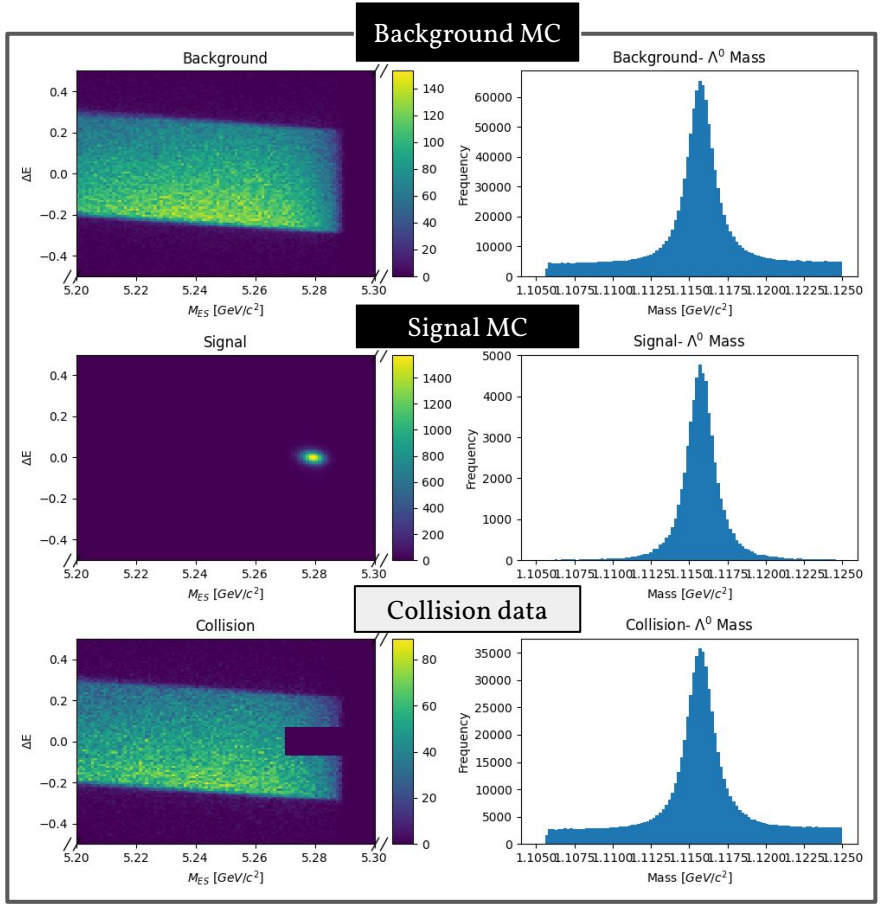
"From tcl to awkward: analyzing old data with new tools"

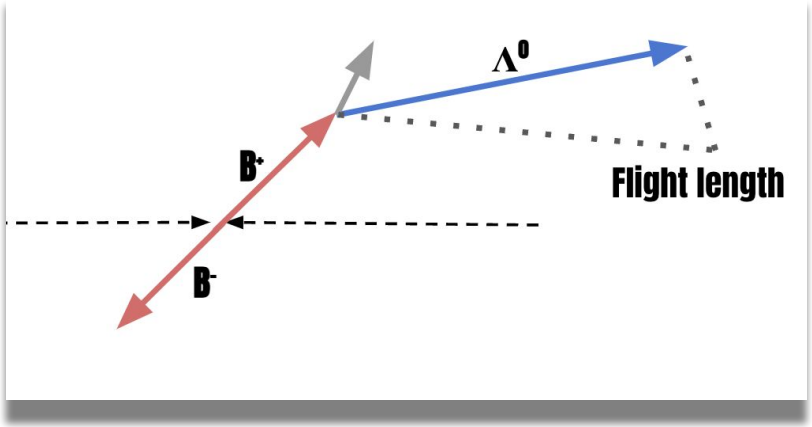
Selection criteria





Selection criteria - Lambda displaced vertex





Λ^0 $c\tau$ = 7.89 cm

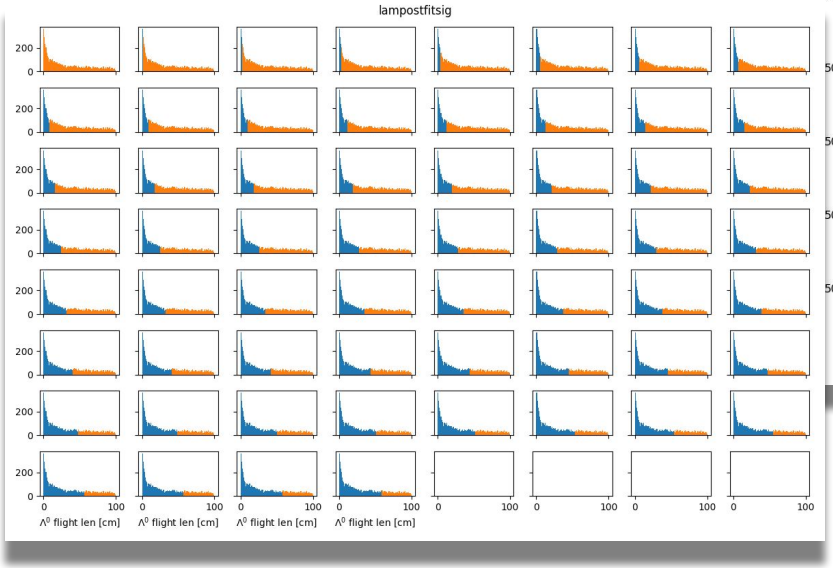
Flight significance = $\frac{L}{\sigma_L}$

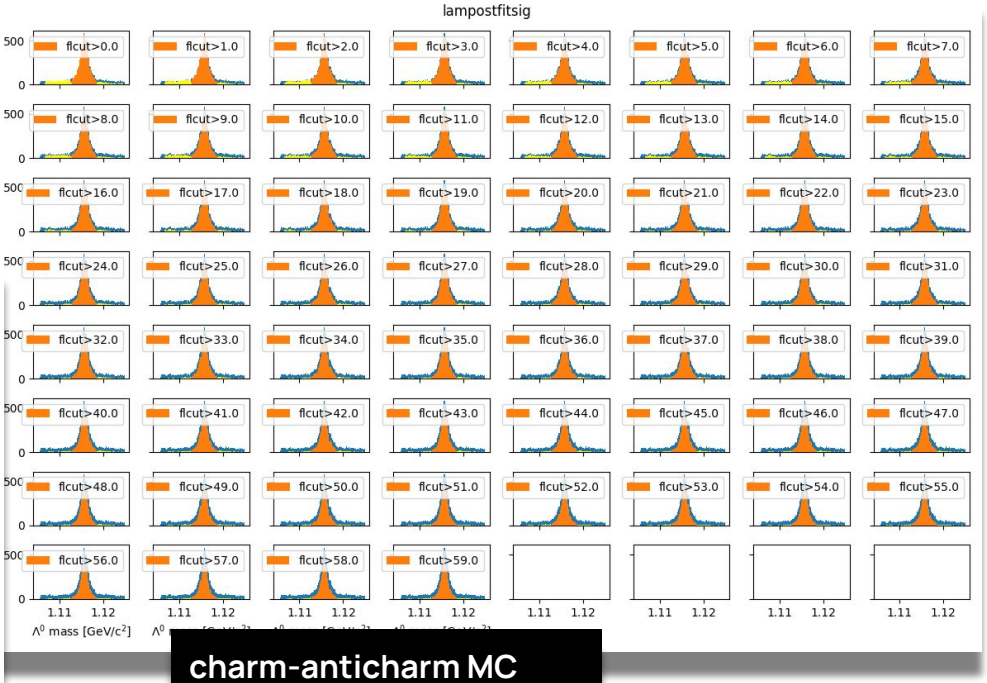
Lambda^0 mass and flight length significance

Make tighter cuts on flight-len and calculate purity of Lambda^0

Mass window retains 93% from signal MC

- Nominal Lambda mass +/- 0.003 GeV/c^2





charm-anticharm MC

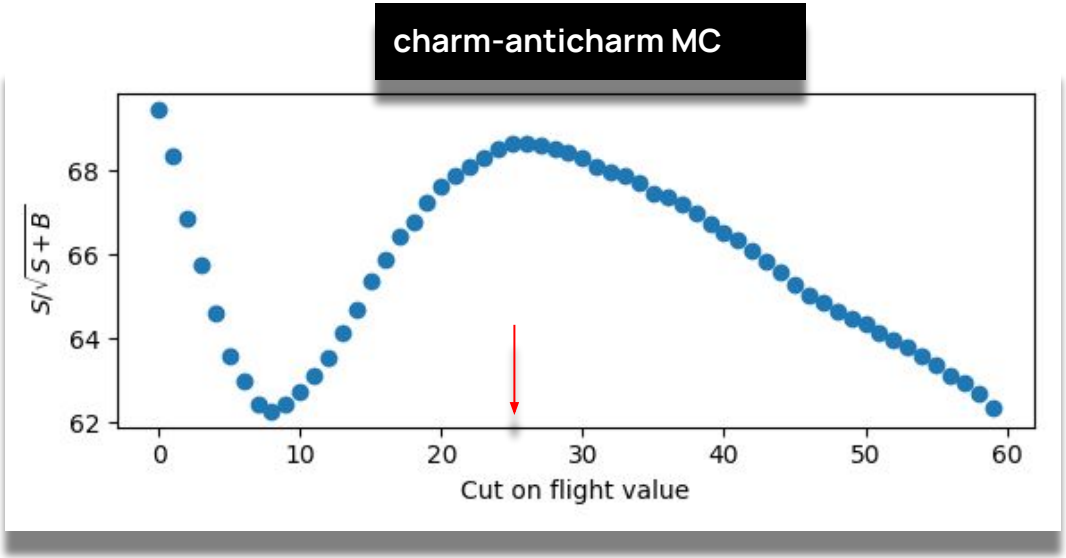
Lambda^0 mass and flight length significance

Optimize $\frac{S}{\sqrt{S+B}}$

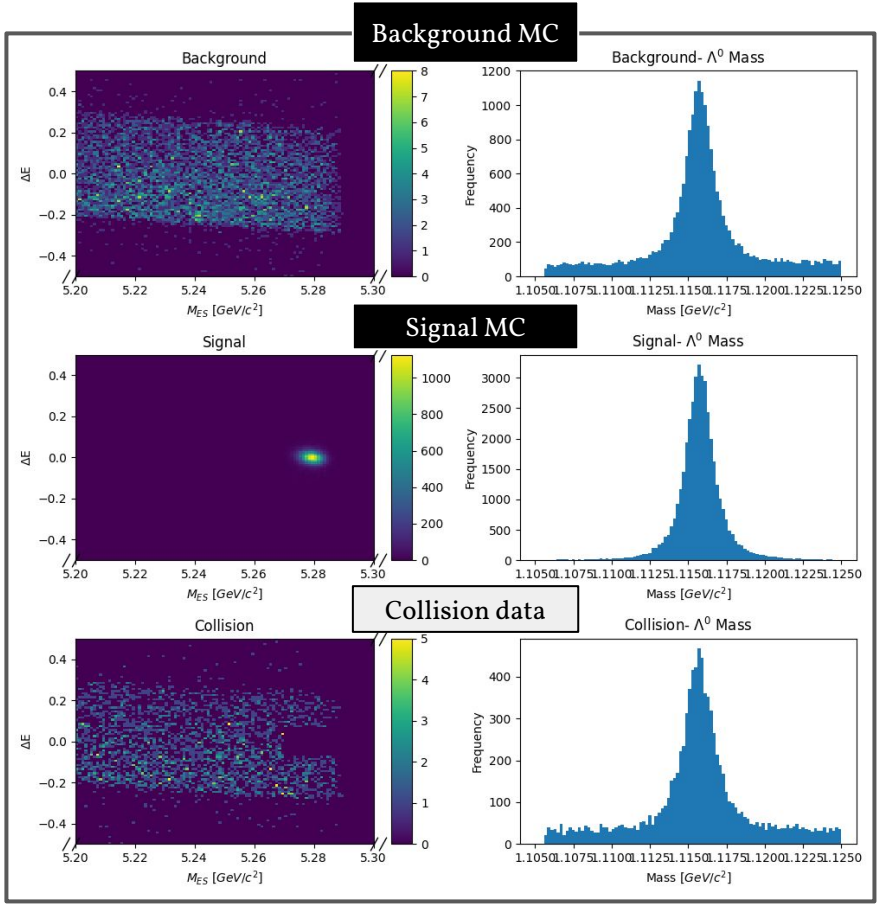
Cut on flight length significance at 25

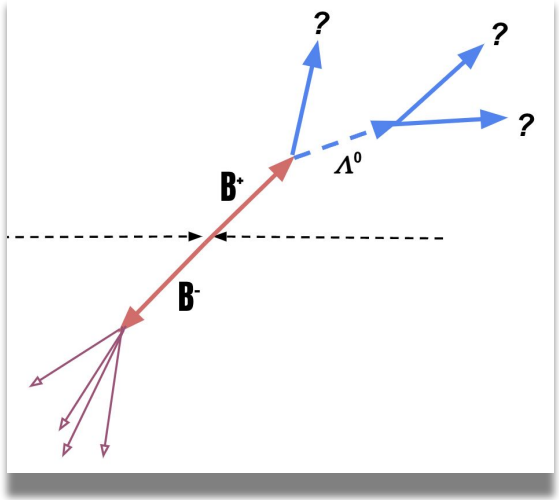
Retains ~82% of signal

Reduces background by
- ~65% for 1005
- ~40% for 998



Particle ID (PID)





Particle ID (PID)

6 proton selectors

- `SuperLooseKMProtonSelection`
- `VeryLooseKMProtonSelection`
- `LooseKMProtonSelection`
- `TightKMProtonSelection`
- `VeryTightKMProtonSelection`
- `SuperTightKMProtonSelection`

6 pion selectors

- `SuperLooseKMPionMicroSelection`
- `VeryLooseKMPionMicroSelection`
- `LooseKMPionMicroSelection`
- `TightKMPionMicroSelection`
- `VeryTightKMPionMicroSelection`
- `SuperTightKMPionMicroSelection`

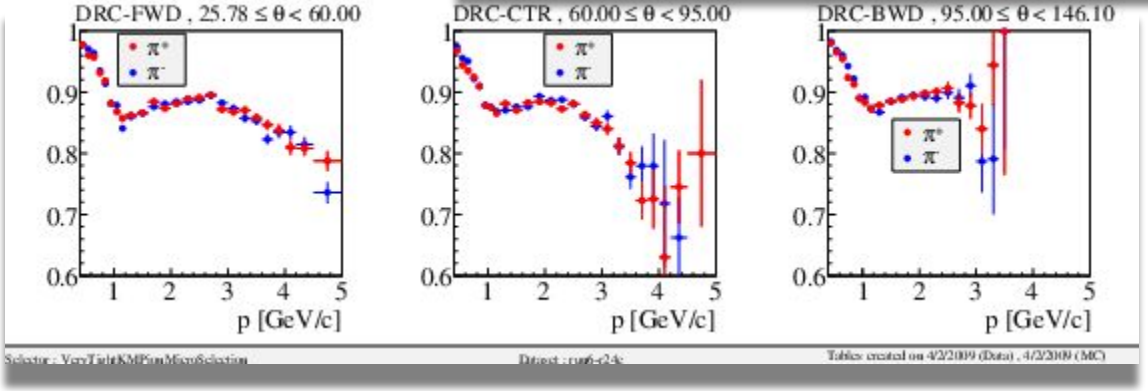
6 x 6 x 6 = 216 combinations

Vary selectors and count number of events in the signal region (M_{ES} and ΔE)

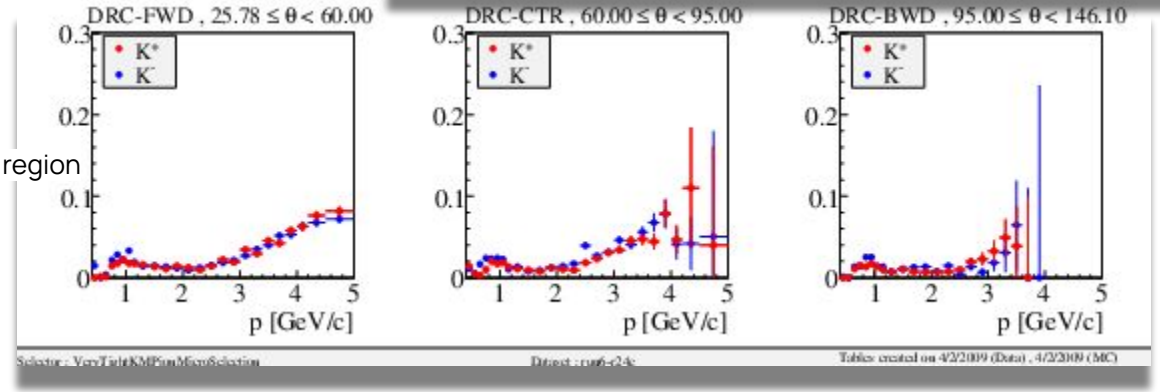
Optimize using Punzi figure-of-merit $\frac{\epsilon(t)}{a/2+\sqrt{B(t)}}$

- ϵ - signal efficiency
- B - # of background events
  - Estimated from collision data sidebands
- a - desired significance of discovery (4)

Pion "VeryTight" selector - true pions



Pion "VeryTight" selector - false pions (kaons)



Particle ID (PID)

6 proton selectors

- `SuperLooseKMProtonSelection`
- `VeryLooseKMProtonSelection`
- `LooseKMProtonSelection`
- `TightKMProtonSelection`
- `VeryTightKMProtonSelection`
- `SuperTightKMProtonSelection`

6 pion selectors

- `SuperLooseKMPionMicroSelection`
- `VeryLooseKMPionMicroSelection`
- `LooseKMPionMicroSelection`
- `TightKMPionMicroSelection`
- `VeryTightKMPionMicroSelection`
- `SuperTightKMPionMicroSelection`

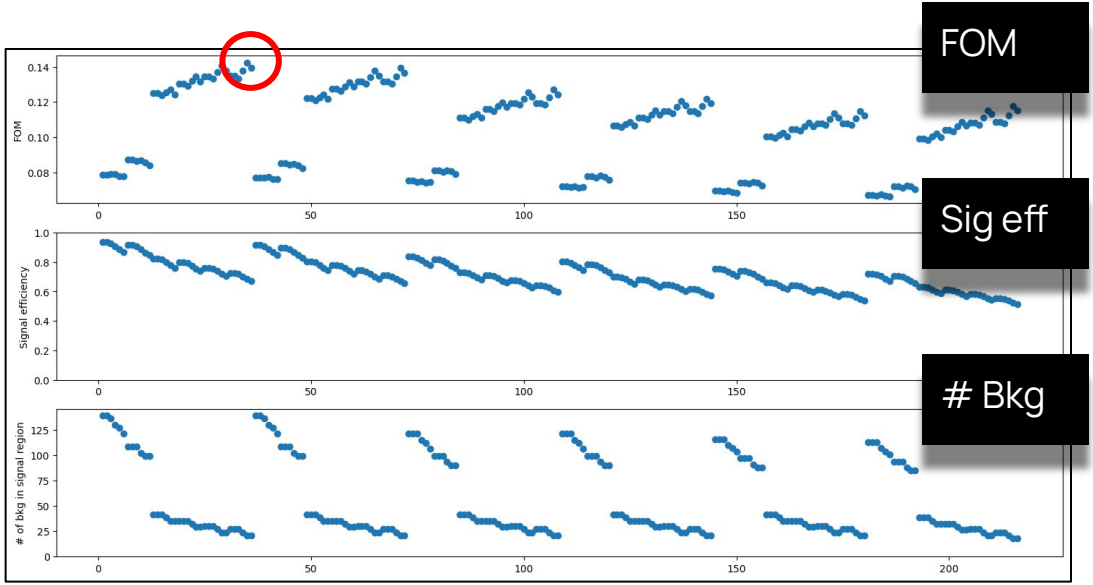
6 x 6 x 6 = 216 combinations

Vary selectors and count number of events in the signal region (M_{ES} and ΔE)

Optimize using Punzi figure-of-merit

$$\frac{\epsilon(t)}{a/2 + \sqrt{B(t)}}$$

- ϵ - signal efficiency
- B - # of background events
  - Estimated from collision data sidebands
- a - desired significance of discovery (4)



Retains ~67% of signal

Reduces background by

- 99% for charm-anticharm
- 98% for light quark

Standard model decays of concern

$B^+ \to K_S\, pi^+$ (BR = 2.4×10^{-5})

$B^+ \to K_S\, K^+$ (BR = 1.3×10^{-6})

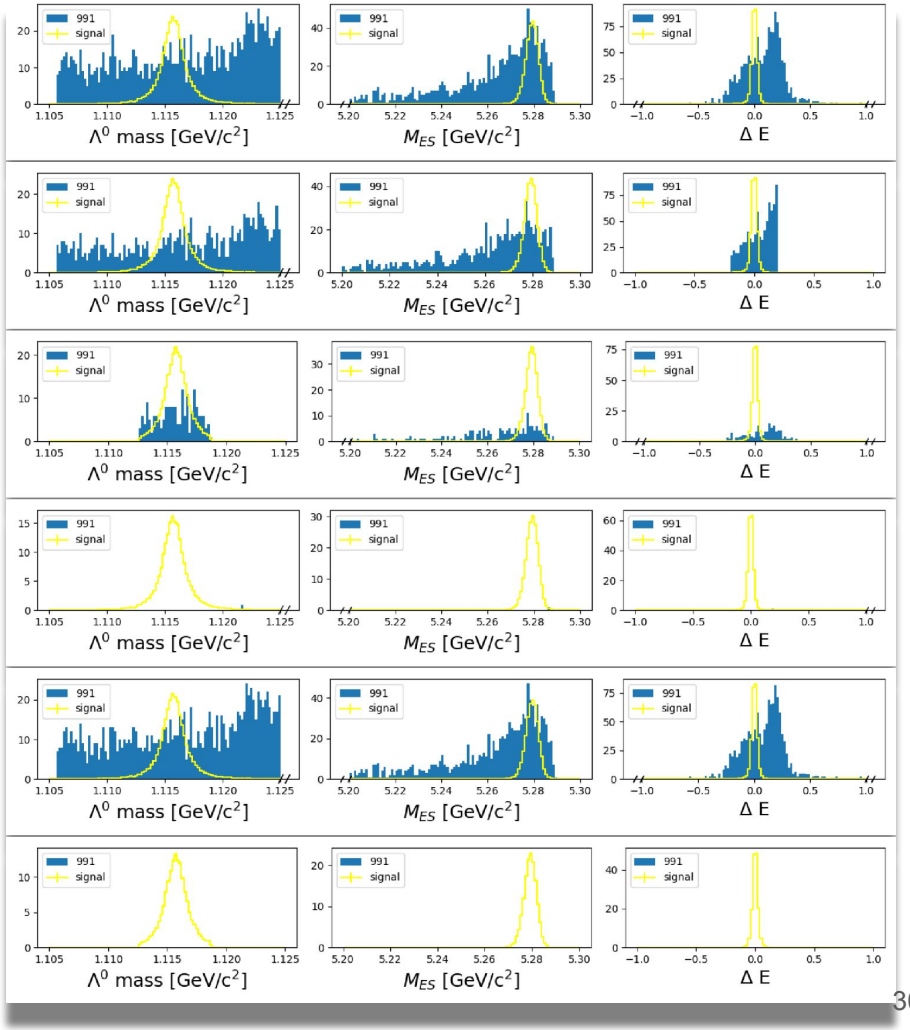
$K_S \to pi^+\, pi^-$

$c\tau$ = 2.7 cm for K_S

Could mis-PID create a peak?

MC contains 70 (1400) times more events than the integrated luminosity

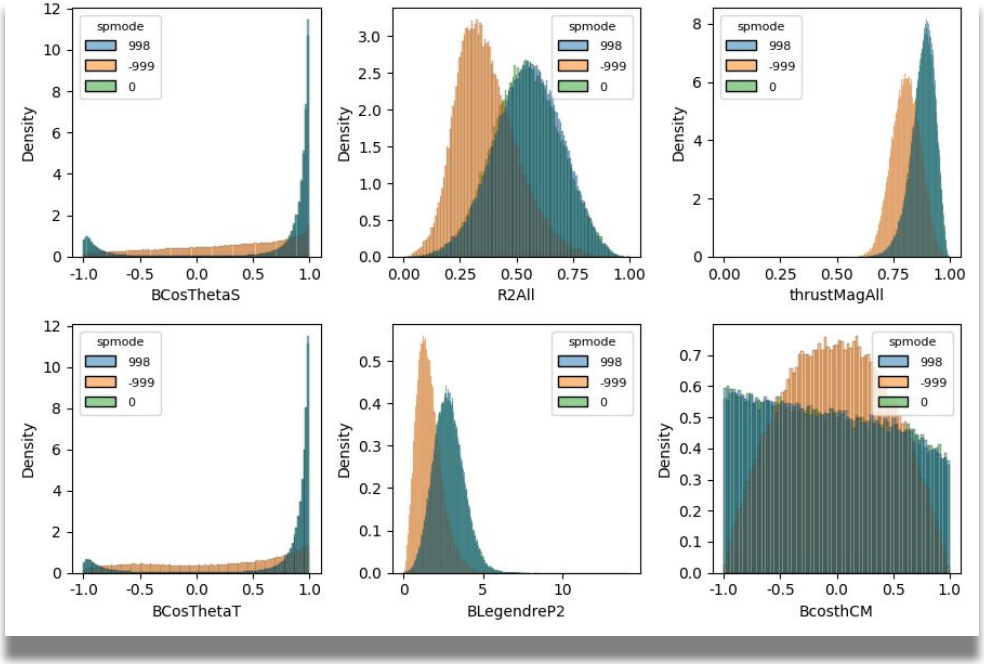
After all “rectangular cuts” (before ML), no events survive from either decay mode



Neural Net

Training variables (6)

- Background MC (blue)
- Signal MC (orange)
- Collision data (green)



- B cos (theta) in the center-of-mass
- B cos(theta) sphericity wrt ROE sphericity
- B cos(theta) thrust wrt ROE thrust
- BLegendreP2
- Thrust for all the charged tracks and neutral depositions of energy.
- R2 for all the charged tracks and neutral depositions of energy.
    - This is the ratio of Fox-Wolfram moments (0 and 2)

Neural Net

After all cuts, ~400 events remain in collision data

Train MLP classifier (scikit-learn)

- Use data upstream of PID cuts
  - Cut events with N>1 B/Lambda candidates
  - Clean up Lambda (flight/mass)
  - Select events in fitting region

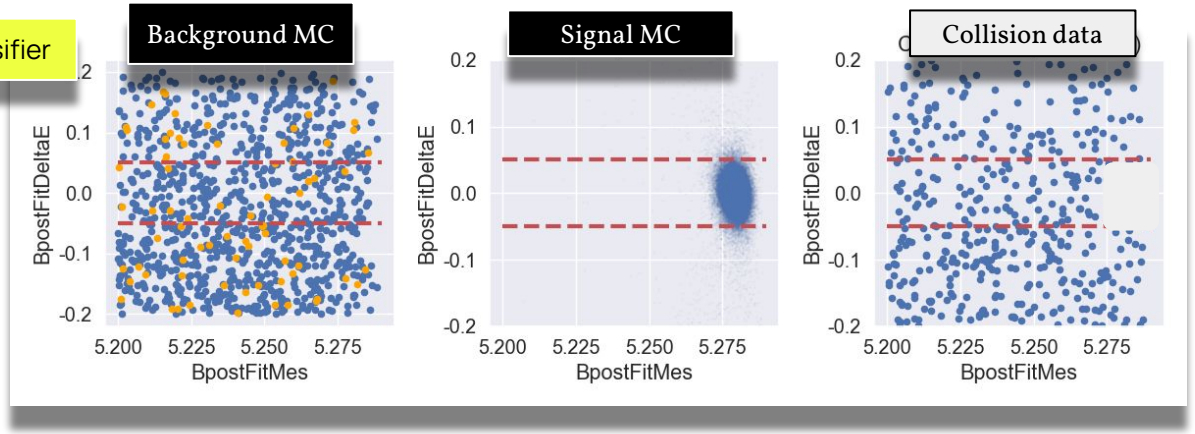
- Majority of surviving background is light-quark
- Train with 30k signal and 30k background

Neural Net

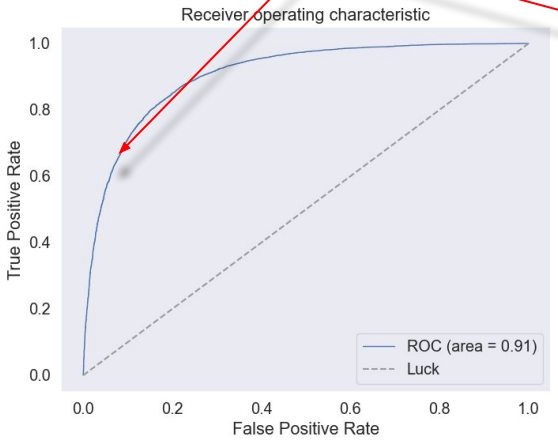
No cut on classifier

Use Punzi fom to
determine selection point

$$\frac{\epsilon(t)}{a/2+\sqrt{B(t)}}$$



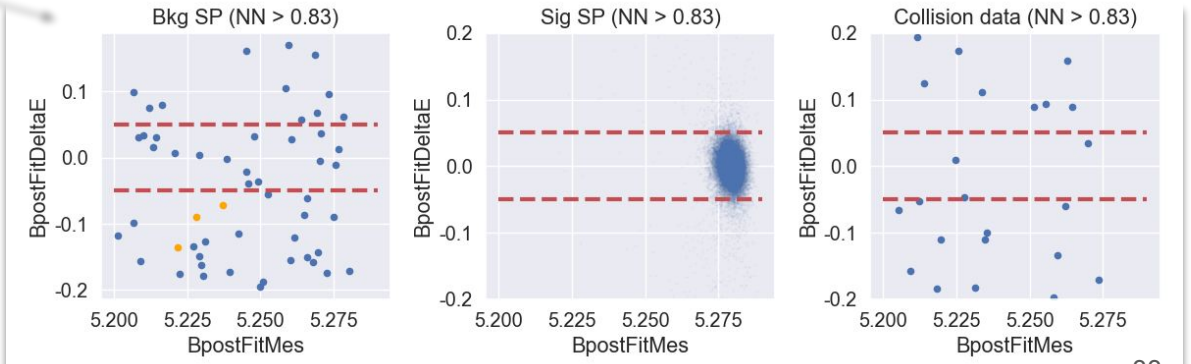
Cut at NN > 0.83



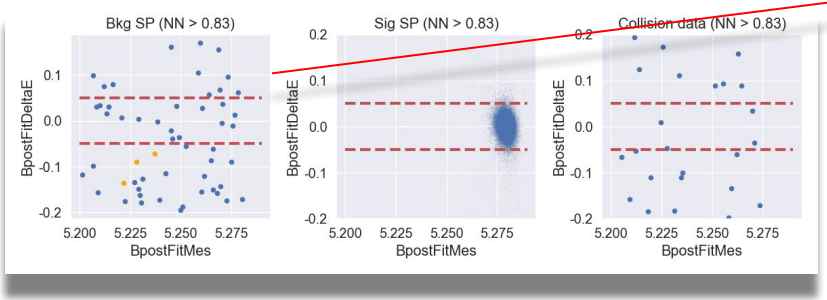
Cut on classifier

Collision data blinded

BaBar preliminary



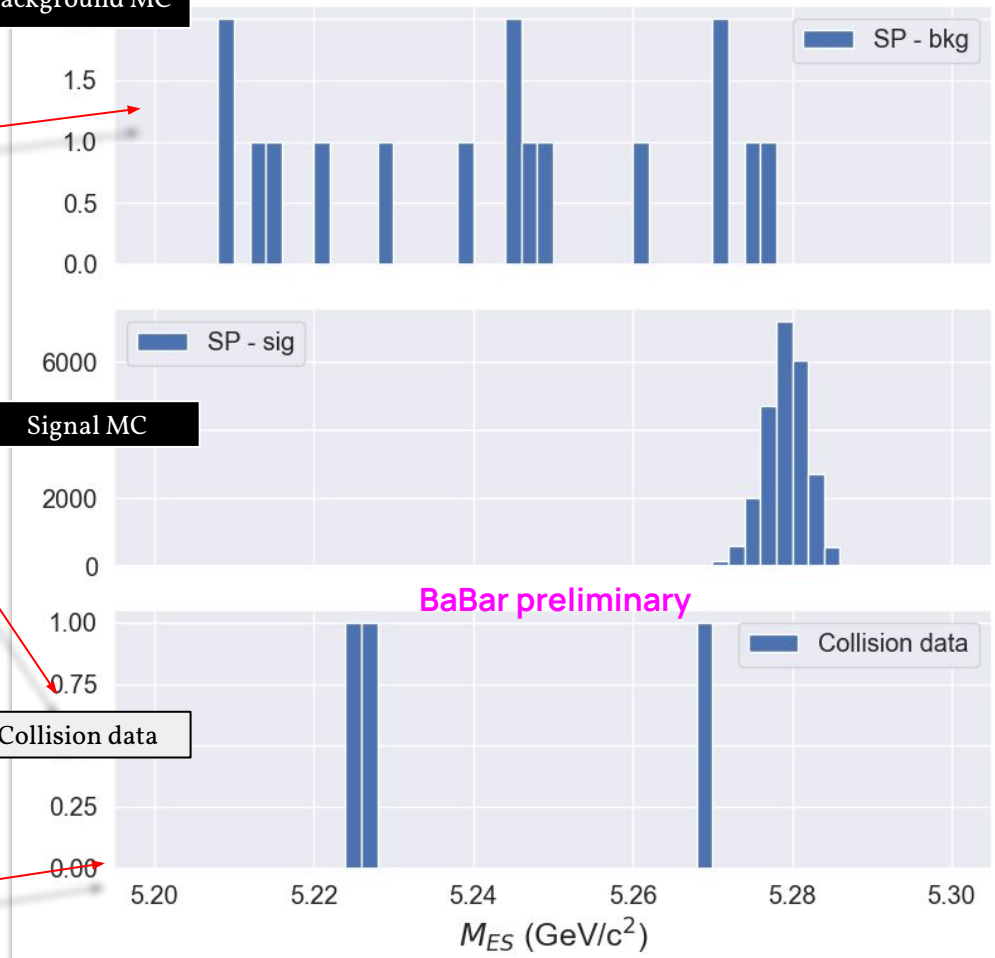
Neural Net



Background MC

Signal MC

Collision data



M_{ES} in red bands

(tight cut on ΔE)

3 events survive in data (still blinded)