Presented at the 32nd International Symposium on Lepton Photon Interactions at High Energies, Madison, Wisconsin, USA, August 25-29, 2025

A search for baryon-number violating decays in $B^+ \to p\Lambda^0$

Matt Bellis and Josephine Swann
Siena University

Brian Shuve
Harvey Mudd College

(On behalf of the BaBar Collaboration)

(Dated: November 28, 2025)

We perform a search for B mesons decaying to final states containing two baryons, specifically a proton and a Λ^0. This decay violates baryon number by ± 2, depending on the charge-conjugation, and would be a signature of physics beyond the standard model. No significant signal is observed and the upper limit is set on the branching fraction at the 90% confidence level of 6.1×10^{-8}.

INTRODUCTION

A search is performed for the decay $B^+ \to p\Lambda^0$. This decay violates baryon number (B) by 2 units (charge conjugation is implied throughout this document, meaning that in addition to searching for $B^+ \to p\Lambda^0$ we also search for $B^- \to \bar{p}\bar{\Lambda}^0$ and all plots are a sum of the searches for these two processes). The Λ^0 candidate is reconstructed in the $p\pi^-$ final state and so the charge conjugation of the proton tags the charge conjugation of the Λ^0. The final state involves all charged tracks such that the B^+ candidate can be completely reconstructed. This decay would be a signature of physics beyond the standard model. The new physics mediating the decay could be directly searched for at the LHC where it would show up as a cascading decay of new high-mass scalar particles, whereas in studying BNV B meson decays at BaBar, an EFT approach would be employed that is insensitive to these details. The flavor/generation dependence of this type of interaction can be constrained by a comprehensive search of multiple processes.

While energy can be converted freely into matter and anti-matter our stable world is made of matter, where we ignore anti-matter produced in β-decay, cosmic ray showers or experimental apparati. The evidence also favors the hypothesis that the bulk of the universe consists of the same charge-conjugation of matter as our solar system [1, 2] The observation of an asymmetric universe requires there to be some amount of baryon number violation [3], beyond what theoretical processes are allowed in the standard model.

Baryon-number violation could occur through processes which mix quarks and leptons. These decays are referred to as $B - L$ processes, where the difference of baryon number and lepton number are conserved, though each is individually violated. In the first SU(5) paper [4], it was shown that proton decay was an implication of this theory. But the experimental bounds on the lifetime of the proton is greater than 10^{32} years. Other processes could violate baryon number by 2 units, for example neutron oscillations in which a neutron oscillates into an anti-neutron. Searches for this interaction have so far produced null results [5, 6].

In this analysis, we perform a search for a meson that decay to two baryons, $B^+ \to p\Lambda^0$. Any result could be incorporated into an effective field theory, however for this analysis we are focused on the experimental result and leave theoretical interpretations to others.

COLLISION DATASET AND MONTE CARLO SAMPLES

The BaBar experiment collected data from 1998 to 2008 from e^+e^- collisions provided by the PEP-II collider and running at the $\Upsilon(4S)$ center-of-mass energy. About 430 fb^{-1} of data were collected, corresponding to approximately 450 million $B\overline{B}$ pairs. Background processes include $q\bar{q}$ continuum $(u\bar{u}, d\bar{d}, s\bar{s}, c\bar{c})$, as well as charged-lepton pair production ($e^+e^-, \mu^+\mu^-, \tau^+\tau^-$). Monte Carlo is produced to simulate these processes, referred to as Simulation Production (SP) throughout this text, at a higher rate than the integrated luminosity of the collision data. The SP is propagated through a simulation of the *BABAR* detector using the GEANT software library so as to understand the detector response to these different physics processes. For this search, we find the primary backgrounds come from the light-quark $q\bar{q}$ processes, referred to as *uds*, and the $c\bar{c}$ continuum production.

To reconstruct the B-meson, we make use of two kinematic variables: the difference between half the CM energy of the colliding beams and the measured CM energy of the B candidate, ΔE; and the energy-substituted mass $m_{\rm ES}$ of the B candidate. When calculating $m_{\rm ES}$, we again make use of half of the colliding beam energy to improve the resolution on the calculated mass of the B candidate: $m_{\rm ES} = \sqrt{[(s/2 + \vec{p_i} \cdot \vec{p_B})/E_i]^2 - |\vec{p_B}|^2}$, where $(E_i, \vec{p_i})$ and $\vec{p}_B$ are respectively the four-momentum of the e^+e^- system and the three-momentum of the B-meson candidate in the lab frame.

Signal Monte Carlo was generated and processed with the same *BABAR* GEANT-based simulation package as the background SP samples. 200,000 events were generated and processed. In these samples, one of the B^+/B^- mesons decays through the decay mode $B^+ \rightarrow p\Lambda^0$, where charge conjugations is implied. The $p\Lambda^0$ final state is generated according to phase-space with a flat angular distribution and the Λ^0 decay is unpolarized. The signal is associated with one and only one of the B-mesons while the companion B-mesons decays according to a standard table of well-established B decays.

The signal peaks in the $m_{\rm ES}$/ ΔE plane in a very small region. We perform a blinded search to minimize human bias and we establish a blinding region in this plane such that roughly 95% of the signal is hidden during the analysis, This is referred to as the *signal region* in this text. Outside of this signal region in this plane is referred to as the *sideband region*.

EVENT SELECTION

The data are initially processed with a minimal set of cuts, designed to identify events with Λ^0 candidates, at the expense of retaining events that do not contain true Λ^0's. We identify a number of selection criteria that will increase the signal purity of our sample by rejecting background events while retaining signal events. To optimize these cuts, we make use of different figures of merit, depending on the type of cuts.

Number of B/Λ^0 candidates. Some small number of events have either multiple Λ^0 candidates or multiple proton's that can lead to multiple B candidates. 99% of the signal SP has only one candidate and so we choose to only analyze events with both one Λ^0 and one B candidate. Roughly 98% of the background SP is retained. While this doesn't significantly reduce the background, it simplifies the analysis with virtually no loss of signal events.

Λ^0 purity - flight length and mass cuts. The Lambda skim is already a loose cut on $\chi^2 > 0.001$ for the geometric vertex constraint for the Λ^0 which provides a fairly clean signal from the skim stage. To further improve the purity, we cut on the *flight length significance*, which is defined as the flight length divided by the uncertainty on that measurement. We define a mass window around in the unconstrained Λ^0-candidate mass distribution (1.115683 ± 0.003 GeV/c^2) and use the sidebands to estimate the number of background events in the window. We use $\frac{S}{\sqrt{S+B}}$ as a figure of merit, where S and B are the number of true and false Λ^0 respectively.. This approach allows us to identify a cut on the flight significance that retains about 82% of the signal while reducing the SP backgrounds by roughly 60% (*uds*) and 35%($c\bar{c}$).

Particle ID (PID) cuts. The *BABAR* analysis software provides six (6) "selectors" for PID that provide increasing levels of purity for each particle hypothesis for charged tracks. We have three (3) charged hadrons in our final state: the proton and pion that come from the Λ^0 decay and the proton that comes from the B decay. For each hadron, we vary the six levels of the selector, which gives us 216 ($6 \times 6 \times 6$) PID combinations. Using the Monte Carlo to estimate the response of background events with false PID hypotheses, we optimize the choice of PID selectors using the Punzi figure-of-merit[7]. This is the most powerful background reduction in our analysis eliminating between 98%-99% of the background SP while retaining 67% of our signal SP.

Anti-baryon cuts. Standard model decays of the B that produce a baryon will also produce an antibaryon. We therefore cut out events which contain an antiproton (charge conjugation is implied). This also removes some signal events when the oppositeside B decays to some baryon-antibaryon final state. We identify the antproton using a PID selector such that retained roughly 90% of the signal events are retained and between 28-44% of background SP is eliminated, depending on the exact physics process.

Summary of cuts. Table. I shows the effects of these cuts, *where the impact of each cut is shown independent of the others*. The last line in the table shows the effect of all the cuts when applied together. Combining these cuts indicates that fewer than 400 events survive in the collision data.

Machine learning. To increase sensitivity to a possible signal, the use of a multivariate discriminator is explored. We make use of the python machine learning library, `scikit-learn`, in particular the `MLPClassifier` neural-net algorithm. We use SP for the background modes and signal modes to train the algorithm.We then make use of the Punzi figure-of-merit to determine the threshold at which to cut on the output variable from the classifier. The PID cut reduces the background so much that we do not have enough events on which to train the algorithm. Therefore, we decide to only apply the following cuts when training: 1) Retain events with only one Λ^0 and one B candidate, 2) apply the Λ^0 cuts on flight length significance and invariant mass. This leaves us with 10's of thousands of events to work with for both signal and background modes. We used six (6) kinematic variables as features for the discriminator.

To determine an optimal cut on the discriminator output, we calculate the Punzi figure-of-merit (FOM) for the signal region using the signal efficiency as calculated with signal SP and the estimated number of background events using the collision data and estimating the background from the sidebands in ΔE. We set $a = 4$ for the FOM.

When all cuts are applied, including cutting on the classifier at the optimal value determined in the previous section, we find that 3 events survive in the collision data in the unblinded region.

EFFICIENCIES AND UNCERTAINTIES

Given some measured yield of signal events, the prescription for turning this into a branching fraction (either observed or an upper limit) is given as follows:

$$\mathcal{B}(B^+ \to p\Lambda^0) = \frac{N_{\text{sig}}}{\epsilon_{\text{sig}} \times \mathcal{B}_{\Lambda^0 \to p\pi^-} \times N_{B\bar{B}} \times 2 \times \mathcal{B}_{\text{chgB}}} \tag{1}$$

Where N_{sig} is the number of signal events extracted by the fit, ϵ_{sig} is the reconstruction efficiency for the full decay path. This is calculated from the signal MC sample and the full GEANT4 detector simulation and is found to be 16.08%±0.38%, $\mathcal{B}_{\Lambda^0 \to p\pi^-}$ is the branching fraction for the decay of the Λ^0 to the final state of interest, $p\pi^-$, $N_{B\bar{B}}$ is the number of $B\bar{B}$ pairs produced in the experiment during the running time over which these data samples were recorded, and $\mathcal{B}_{\text{chgB}}$ is the branching fraction of the $\Upsilon(4S)$ to a charged $B\bar{B}$ pair. The factor of 2 is because each $B\overline{B}$ event has two (2) B mesons which could decay to this final state.

$N_{B\bar{B}}$ is estimated to be $(470.89 \pm 2.84) \times 10^6$. We use the measured branching fractions and associated uncertainties [8]

$$\mathcal{B}(\Upsilon(4S) \to B^+B^-) = (51.4 \pm 0.6) \times 10^{-2}$$

For the branching fractions of the Λ^0, the current branching fraction is [8].

$$\mathcal{B}(\Lambda^0 \to p\pi^-) = (64.1 \pm 0.5) \times 10^{-2}$$

To convert the observed signal yield into a branching fraction, we define a conversion factor f:

$$f = \epsilon_{\text{sig}} \times \mathcal{B}_{\Lambda^0 \to p\pi^-} \times N_{B\bar{B}} \times 2 \times \mathcal{B}_{\text{chgB}}$$

Additional systematics. There are additional systematics which contribute to the reconstruction efficiency.

Tracking. For the tracking systematics, we reference the studies from the BABAR Tracking working group which looked at different topologies. The results of this study can be found in internal BABAR documentation[9]. The systematic uncertainty ($\delta\epsilon$) associated with the identification of charged tracks was estimated to be 0.128% per track. The conservative estimate for an n-topology event is $n \times \delta\epsilon$. Our final state has 3 charged tracks: two protons and pion, and so we assign a systematic uncertainty of 0.384% coming from the tracking.

PID. Previous BABAR analyses have estimated the uncertainties in PID by studying the effects of varying parameters used in the PID algorithms and studying the differences in response between

TABLE I. Summary of cuts on signal SP, background SP, and blinded collision data. *Each cut is performed on the skimmed data and is independent of the cuts before it, save for the last row which enumerates the effect when all cuts are applied.* The table shows the number of events retained and the % retained relative to the initial skim. We do not show numbers for the B^+B^- or $B^0\overline{B^0}$ simulation as all the events are eliminated by the PID cuts.

cut	# signal	% signal	# $u\bar{u}, d\bar{d}, s\bar{s}$	%$u\bar{u}, d\bar{d}, s\bar{s}$	# $c\bar{c}$	% $c\bar{c}$	# data
None	119215	100.0	902421	100.0	320134	100.0	738840
Duplicates	118337	99.3	887707	98.4	311814	97.4	716940
Fit Area	117344	98.4	223095	24.7	80402	25.1	155430
Lambda	98312	82.5	544360	60.3	114237	35.7	371824
PID	79545	66.7	18395	2.0	2362	0.7	9057
Antibaryon	106474	89.3	504989	56.0	231614	72.3	452834
All cuts	59937	50.3	900	0.1	73	0.0	394

collision data and SP. This can be done for each hadron hypothesis. The systematic for proton PID was estimated to be 1% per track Similarly, we estimate a 1.0% systematic uncertainty associated with pion PID. These individual systematics are added in quadrature for the total PID systematic uncertainty. The individual and total systematic uncertainties are summarized in Table II.

TABLE II. Contributions to the statistical and systematic uncertainty on the branching fraction (fractional uncertainties are given in the table).

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**

The conversion factor f is then

$$f = (49.74 \pm 1.18) \times 10^6$$

SIGNAL EXTRACTION

The $m_{\rm ES}/\ \Delta E$ 2D distributions, after unblinding and with all cuts applied, are shown in Fig. 1 and the $m_{\rm ES}$ 1D distributions are shown in Fig. 2. In both of those plots we show the background SP, signal SP, and collision data. We define the *signal region* as those events for which $m_{\rm ES} > 5.27 - 5.28$ GeV/c^2, where 5.28 is the upper value that contains

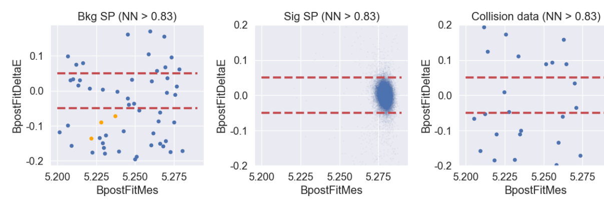

FIG. 1. $m_{\rm ES}/\ \Delta E$ distributions for the background SP modes (left column), signal SP (middle column), and the *unblinded* collision data (right column). For the background SP, the blue dots are from SP-998 (uds) and the orange dots are from SP-1005 ($c\bar{c}$). The top row has all rectangular cuts applied but no cut on the MLP classifier. The bottom row has all rectangular cuts applied and cuts on the MLP classifier to maximize the Punzi figure-of-merit. The red dashed lines demonstrate the region that we will fit to.

95% of the signal SP. The region from $5.2 < m_{\rm ES}5.27$ GeV/c^2 is referred to as the *sideband* or *normalization* region. No significant signal is observed.

Upper-limit calculation. To determine the upper limits, we make use of an approach developed by Rolke et.al [10] which takes into account Poisson fluctuations on both the signal and background, the signal efficiency, and the uncertainty on the efficiency. We use the ROOT implementation of this approach.[11] We define a sideband region where we do not expect signal and use this to estimate the background rate in the signal region. The sideband region does not span the same amount of phase space as our signal region so we define a scaling factor τ which is the relative size of our sideband region to the signal region. We find $\tau = 4.667$. The signal efficiency is 16.08%. with a relative uncertainty of 2.38%. This accounts for total uncertainty on the signal efficiency coming from both systematic

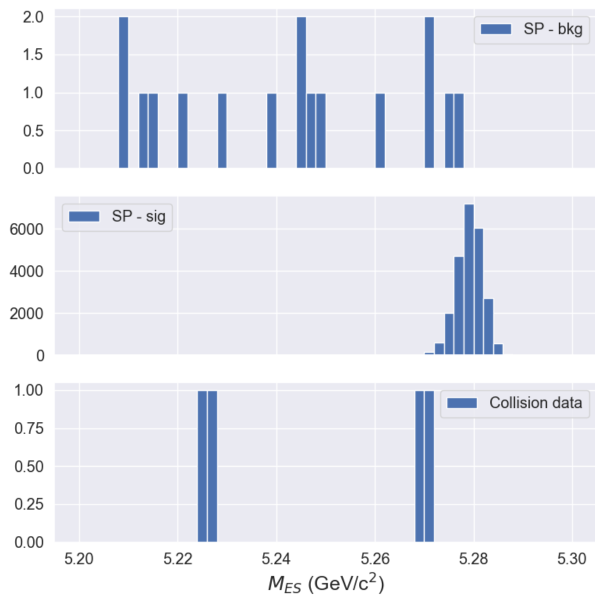

FIG. 2. m_{ES} distributions with the narrower cut on ΔE $(-0.05, 0.05)$ for the background SP modes (top row), signal SP (middle row), and the *unblinded* collision data (bottom row). The left column has all rectangular cuts applied but no cut on the MLP classifier. The right column has all rectangular cuts applied and cuts on the MLP classifier to maximize the Punzi figure-of-merit.

and statistical contributions, as well as the uncertainty associated with our estimate of the number of charged-B mesons produced. This corresponds to an uncertainty on the efficiency of 0.38%.

We find 3 events in the sideband and 1 event in the signal region. The approach from Rolke et al., yields an efficiency-corrected number of signal events at the 90% confidence level of

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

Upper-limit calculation - branching fraction. To convert this to a branching fraction, we use the conversion factor discussed earlier in this article, of $f = (49.74 \pm 2.36) \times 10^6$. We ensure that we do not double-count our efficiency calculations when arriving at the upper-limit on the branching fraction. Our final result then is

$$UL(\mathcal{B}(B^+ \to p\Lambda^0))_{90\%} = 6.1 \times 10^{-8}$$

RESULTS

We have performed the first search for the decay $B^+ \to p\Lambda^0$, a process which violated baryon number (B) by $\Delta B = \pm 2$ and which would constitute new physics. We observe no significant signal and place upper limits on the branching fraction at the 90% confidence level at 6.1×10^{-8}.

[1] P. Coppi, How Do We know Antimater is Absent?, eConf **C040802**, L017 (2004).

[2] G. Steigman, Observational tests of antimatter cosmologies, Ann. Rev. Astron. Astrophys. **14**, 339 (1976).

[3] A. D. Sakharov, Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe, Pisma Zh. Eksp. Teor. Fiz. **5**, 32 (1967).

[4] H. Georgi and S. L. Glashow, Unity of All Elementary Particle Forces, Phys. Rev. Lett. **32**, 438 (1974).

[5] M. Baldo-Ceolin *et al.*, A New experimental limit on neutron - anti-neutron oscillations, Z. Phys. C **63**, 409 (1994).

[6] K. Abe *et al.* (Super-Kamiokande), Neutron-antineutron oscillation search using a 0.37 megaton-years exposure of Super-Kamiokande, Phys. Rev. D **103**, 012008 (2021), arXiv:2012.02607 [hep-ex].

[7] G. Punzi, Sensitivity of searches for new signals and its optimization, eConf **C030908**, MODT002 (2003), arXiv:physics/0308063 [physics].

[8] S. Navas *et al.* (Particle Data Group), Review of particle physics, Phys. Rev. D **110**, 030001 (2024).

[9] `http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/TauEff/R24/TauEff.html`.

[10] W. A. Rolke, A. M. Lopez, and J. Conrad, Limits and confidence intervals in the presence of nuisance parameters, Nucl. Instrum. Meth. A **551**, 493 (2005), arXiv:physics/0403059.

[11] `https://root.cern.ch/doc/master/classTRolke.html`.