Search for the Lepton-Number-Violating B decay at Belle

Oksu Seon
Nagoya University
osseon@hepl.nagoya-u.ac.jp
Motivation
Majorana neutrino

- In Standard Model (SM), neutrino is massless particle.
- The discovery of neutrino oscillation $\rightarrow$ massive neutrino?
- Q: Are the neutrinos Dirac or Majorana type?
- If the neutrinos are Majorana-type, neutrino = anti-neutrino $\rightarrow$ Lepton-number-violation (LV)!
LV in B decays

**B^+ → D^-l^+l^+**: decay through Majorana neutrinos where l is e or μ.

**B^+ → h^-l^+l^+**: CLEO in 2002 [1], 9.2 fb⁻¹ where h = K, K*, π or ρ, set the upper limit of order 10⁻⁶

naive motivation: ex) B^+ → D^-l^+l^+ vs. B^+ → K^-l^+l^+

→ |V_{ud}^*V_{cb}|² ≃ 2500 × |V_{us}^*V_{ub}|²

**Theoretical expectation**

Two theoretical papers appeared in 2010

Heavy neutrinos mass \((mN/m_4)\) in range of 2-5 GeV \([2,3]\)

This is the first measurement of branching fraction of \(B^+ \rightarrow D^- l^+ l^+\) testing the heavy Majorana mass scale.


Analysis
Mt. Tsukuba

KEKB ring (HER+LER)

Belle detector

Linac

KEK Tsukuba site
Mt. Tsukuba

KEK Tsukuba site

KEKB ring (HER+LER)

Lina

Superconducting Solenoid Magnet

1.5T

EM Calorimeter

CsI(Tl) $16\times$

Time-of-Flight Counters

8GeV $e^-$

Extreme Forward-and-Backward Calorimeters

BGO

Silica-aerogel Cherenkov Counters

$n = 1.015 \sim 1.030$

Tracking + $dE/dx$

small cell + He/C$_2$H$_6$

Central Drift Chamber

Silicon Vertex Detector

5 layers Silicon strip sensor

Silicon Vertex Detector

μ / $K^*$ detection

14/15 lyr. RPC+Fe

KL & muon detection
Selection criteria

- Data : 711fb$^{-1}$ (770M $B\bar{B}$)
- MC : signal – 0.6M (using 3-body phase space model) 
  background – $3 \times$ (Data size)

Selection criteria

- eID : eff. : 90% / mis-ID : 0.1%
- $\mu$ID : eff. : 90% / mis-ID : 1%
- $\pi,K$ID : eff. : 80% / mis-ID : 10%
- lepton momentum : $P_e > 0.5$ GeV/c, $P_\mu > 0.8$ GeV/c
- dilepton energy : $E_{ll} > 1.3$ GeV
Definition of signal box

 Beam constraint mass

\[ M_{bc} = \sqrt{E_{beam}^2 - (\sum \vec{p}_i)^2} \]

 Energy difference

\[ \Delta E = \sum E_i - E_{beam} \]
Definition of signal box

 Beam constraint mass

\[ M_{bc} = \sqrt{E_{\text{beam}}^2 - (\sum p_i)^2} \]

 Energy difference

\[ \Delta E = \sum E_i - E_{\text{beam}} \]
Definition of signal box

Beam constraint mass

\[ M_{bc} = \sqrt{E_{\text{beam}}^2 - (\sum \vec{p}_i)^2} \]

Energy difference

\[ \Delta E = \sum E_i - E_{\text{beam}} \]
Background Events

Main sources of background events

- continuum: $e^+e^- \rightarrow q^+q^-$
  (q = u, d, s and c)
- $B\bar{B}$: mis-reconstructed B events

Why?

- Misidentification
- Random combination

4 variables to suppress background:
- KSFW, $\cos\Theta_B^*$, $E_{\text{miss}}$, $\delta z$
Continuum Suppression

continuum events are no resonant, there is a large energy release; jet-like

KSFW

Fox-Wolfram moments:

\[ H_l \equiv \sum_{i,j} |\vec{p}_i||\vec{p}_j| P_l(\cos \phi_{i,j}) \]

Super Fox-Wolfram: Fisher discriminant of Fox-Wolfram, the terms in the summation of \( H_l \) separated by \( B \) candidates and others.
Continuum Suppression

- **continuum events** are no resonant, there is a large energy release; jet-like

**KSFW**

- Fox-Wolfram moments:
  \[ H_l \equiv \sum_{i,j} |\vec{p}_i| |\vec{p}_j| P_l(\cos \phi_{i,j}) \]

- Super Fox-Wolfram: Fisher discriminant of Fox-Wolfram, the terms in the summation of \( H_l \) separated by \( B \) candidates and others.
Continuum Suppression

- **continuum events** are no resonant, there is a large energy release; jet-like

$\cos \Theta_B$

- $\Theta_B$ : angle between B candidate and beam direction in the B rest frame.

- $Y_1^1 (\cos \Theta_B) = \sin \Theta_B$
  
  $= \sqrt{1-\cos^2 \Theta_B}$
Continuum Suppression

- **continuum events** are no resonant, there is a large energy release; jet-like

\[
\cos \Theta_B
\]

- \( \Theta_B \): angle between B candidate and beam direction in the B rest frame.

- \( Y_1^1 (\cos \Theta_B) = \sin \Theta_B = \sqrt{1-\cos^2 \Theta_B} \)
**B̄B Suppression**

**B̄B events** have leptons, which are produced along with neutrinos.

**E_{miss}**

Production of the neutrinos result large missing energy of the background event compare to that of signal.
B̅ B̅ Suppression

- B̅ B̅ events have leptons, which are produced along with neutrinos.

E miss

- Production of the neutrinos result large missing energy of the background event compare to that of signal.

![Graph showing signal and background events with missing energy distribution.](image)
**B̅B Suppression**

- **B̅B events** have leptons, which are produced along with neutrinos.

**δz**

- **δz** : vertex displacement of two leptons in z direction.

Background events have larger distribution than signal because their mothers are usually different.
**$\bar{B} \bar{B}$ Suppression**

- **$\bar{B} \bar{B}$ events** have leptons, which are produced along with neutrinos.

$\delta z$

- $\delta z$ : vertex displacement of two leptons in $z$ direction.

- Background events have larger distribution than signal because their mothers are usually different.
**Optimization**

- **Likelihood:** \( \mathcal{L}_\text{sig(bkg)}(x) = \prod_i \text{PDF}(x)_i^{\text{sig(bkg)}} \)
  
  where \( \text{PDF}(x)_i^{\text{sig(bkg)}} \) is probability density function.

- **Likelihood ratio:** \( \mathcal{R}_L(x) = \frac{\mathcal{L}_\text{sig}(x)}{\mathcal{L}_\text{sig}(x) + \mathcal{L}_\text{bkg}(x)} \)

- **Figure of Merit:** \( \text{FOM} = \frac{N_{\text{sig}}}{\sqrt{N_{\text{bkg}}}} \)

- **Expected number of events**

  \[
  N_{\text{sig}} = \text{Eff} \times N_{\overline{B}B} \times B(B \to D\ell\ell)(= 10^{-6}) \times B(D \to K\pi\pi) \\
  N_{\text{bkg}} = \text{bkg in analysis region} \times \text{scale factor} \left( \frac{\text{signal box}}{\text{analysis region}} \right)
  \]

<table>
<thead>
<tr>
<th></th>
<th>( \text{B}^+ \to \text{D}^+\ell^+\ell^- )</th>
<th>( \text{B}^+ \to \text{D}^+\ell^+\mu^- )</th>
<th>( \text{B}^+ \to \text{D}^+\mu^+\mu^- )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Signal MC</strong></td>
<td># of events in signal box</td>
<td>0.92</td>
<td>1.03</td>
</tr>
<tr>
<td><strong>Bkg MC</strong></td>
<td># of events in wide range</td>
<td>4.3</td>
<td>21.8</td>
</tr>
<tr>
<td></td>
<td># of events in signal box</td>
<td>0.1</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Sideband Data

Data events distribution on $M_{bc} \Delta E$ 2D

<table>
<thead>
<tr>
<th># of background events</th>
<th>$B^+ \rightarrow D^- e^+ e^+$</th>
<th>$B^+ \rightarrow D^- e^+ \mu^+$</th>
<th>$B^+ \rightarrow D^- \mu^+ \mu^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>in wide range</td>
<td>DATA</td>
<td>5</td>
<td>23</td>
</tr>
<tr>
<td>expected by MC</td>
<td>4.3</td>
<td>21.8</td>
<td>38.3</td>
</tr>
<tr>
<td>estimated in signal box</td>
<td>0.2 ±0.1</td>
<td>0.8 ±0.3</td>
<td>1.4 ±0.4</td>
</tr>
</tbody>
</table>
No events observed in signal box.

Using POLE program to calculated U.L. [4] at 90% C.L.

\[
\mathcal{B}(B^+ \rightarrow D^- \ell^+ \ell^+) = \frac{N_{\text{sig}}}{N_{B\bar{B}}} \times \epsilon_{\text{sig}} \times \mathcal{B}(D \rightarrow K\pi\pi)
\]

Table:

<table>
<thead>
<tr>
<th></th>
<th>Dee</th>
<th>Deμ</th>
<th>Dμμ</th>
</tr>
</thead>
<tbody>
<tr>
<td>expected (N_{\text{bkg}})</td>
<td>0.2 ± 0.1</td>
<td>0.8 ± 0.3</td>
<td>1.4 ± 0.4</td>
</tr>
<tr>
<td>(N_{\text{obs}})</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(\epsilon_{\text{sig}})</td>
<td>1.2</td>
<td>1.3</td>
<td>1.8</td>
</tr>
<tr>
<td>(N_{\text{sig}})</td>
<td>2.3</td>
<td>1.7</td>
<td>1.3</td>
</tr>
</tbody>
</table>

\[
\mathcal{B}(B^+ \rightarrow D^- \ell^+ \ell^+) \times 10^{-6} = 2.7 \quad 1.9 \quad 1.1
\]

Conclusion

- First search for $B^+ \rightarrow D^-\ell^+\ell^+$
- No events observed in signal box range
- 90% C.L upper limit of order $10^{-6}$
Conclusion

- First search for $B^+ \rightarrow D^- \ell^+ \ell^+$
- No events observed in signal box range
- 90% C.L upper limit of order $10^{-6}$
Conclusion

- First search for $B^+ \rightarrow D^- l^+ l^+$
- No events observed in signal box range
- 90% C.L upper limit of order $10^{-6}$

We expect Belle II experiment
- will reach the theoretical prediction with much larger data
- will extend this search to other D decay
\[(\beta mc^2 + \sum_{k=1}^{3} \alpha_k P_k) \psi(x, t) = \frac{i}{\hbar} \frac{\partial \psi(x, t)}{\partial t}\]
Thank you!
Back up
FIG. 1: The $t$-type and $s$-type weak amplitudes at the quark level that enter in the process $M^+ \rightarrow M'^- \ell_1^+ \ell_2^+$ (plus the same diagrams with leptons exchanged if they are identical).
<table>
<thead>
<tr>
<th></th>
<th>D⁻e⁺e⁺</th>
<th>D⁻e⁺μ⁺</th>
<th>D⁻μ⁺μ⁺</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Signal MC efficiency</strong></td>
<td>12%</td>
<td>13%</td>
<td>13%</td>
</tr>
<tr>
<td>in signal box</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bkg MC # of events</strong></td>
<td>112</td>
<td>223</td>
<td>114</td>
</tr>
<tr>
<td>in signal box</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bkg MC # of events</strong></td>
<td>4721</td>
<td>9246</td>
<td>4949</td>
</tr>
<tr>
<td>in analysis region</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sources of systematic error

- **Signal MC efficiency**: statistical error of signal MC efficiency
- **Tracking Efficiency**: Track finding uncertainty
- **Kaon(pion) and lepton ID**: Uncertainties of ID efficiency depending on the momentum and requirement.
- **$M_{bc} - \Delta E$ and D mass shape**: Uncertainties due to the difference of shape of MC and data
- **$N_{BB \bar{B}}$**: Luminosity uncertainty

### Systematics Table

<table>
<thead>
<tr>
<th>Source</th>
<th>$\Delta e$</th>
<th>$\Delta \mu$</th>
<th>$\Delta \mu\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC signal efficiency</td>
<td>1.1</td>
<td>1.1</td>
<td>0.9</td>
</tr>
<tr>
<td>Tracking efficiency</td>
<td>5.2</td>
<td>5.2</td>
<td>5.2</td>
</tr>
<tr>
<td>Kaon(pion) ID</td>
<td>1.3</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td>Lepton ID</td>
<td>3.0</td>
<td>3.4</td>
<td>3.5</td>
</tr>
<tr>
<td>$M_{bc}$ and $\Delta E$ shape</td>
<td>1</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>$D^+$ mass shape</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>$N_{BB \bar{B}}$</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td>$B(D^+ \to K^- \pi^+ \pi^+)$</td>
<td>2.7</td>
<td>2.7</td>
<td>2.7</td>
</tr>
<tr>
<td>Cut on $\mathcal{R}$</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Sum</td>
<td>8.6</td>
<td>10.1</td>
<td>8.8</td>
</tr>
</tbody>
</table>

- **Sub decay branching fraction**: Error of the $D^- \to K^+ \pi^- \pi^-$ branching fraction
- **Cut on $\mathcal{R}$**: Uncertainties due to the likelihood requirement of MC and data
Continuum Suppression

**KSFW**

- Continuum events are no resonant, there is a large energy release; jet-like

- Fox-Wolfram moments:
  \[ H_l \equiv \sum_{i,j} |\vec{p}_i| |\vec{p}_j| P_l(\cos \phi_{i,j}) \]

- Super Fox-Wolfram: **Fisher discriminant** of Fox-Wolfram, the terms in the summation of \( H_l \) separated by \( B \) candidates and others.
Continuum Suppression

KSFW

- continuum events are no resonant, there is a large energy release; jet-like

- Fox-Wolfram moments:
  \[ H_l \equiv \sum_{i,j} |\vec{p}_i||\vec{p}_j|P_l(\cos \phi_{i,j}) \]

- Super Fox-Wolfram: Fisher discriminant of Fox-Wolfram, the terms in the summation of \( H_l \) separated by B candidates and others.
Continuum Suppression

KSFW

- Continuum events are non-resonant, thus there is a large energy release.

- Fox-Wolfram moments:

  $$ H_l \equiv \sum_{i,j} |\vec{p}_i| |\vec{p}_j| P_l(\cos \phi_{i,j}) $$

- Super Fox-Wolfram discriminant consists of the terms in $H_l$ separated by $B$ candidates and others.
Continuum Suppression

KSFW

- continuum events are not resonant, there is a large energy release.
- Fox-Wolfram moments:
  \[ H_l \equiv \sum_{i,j} |\vec{p}_i| |\vec{p}_j| P_l(\cos \phi_{ij}) \]
- Super Fox-Wolfram discriminant:
  the terms in \( H_l \) separated by B candidates and others.
Continuum Suppression

**KSFW**

- **continuum events** are no resonant, there is a large energy release; jet-like

- Fox-Wolfram moments:
  \[
  H_l \equiv \sum_{i,j} |\vec{p}_i| |\vec{p}_j| P_l(\cos \phi_{i,j})
  \]

- Super Fox-Wolfram: **Fisher discriminant** of Fox-Wolfram, the terms in the summation of \( H_l \) separated by B candidates and others.
Continuum Suppression

**KSFW**

⚠️ Fox-Wolfram moments:

\[
H_l \equiv \sum_{i,j} |\vec{p}_i||\vec{p}_j| P_l(\cos \phi_{i,j})
\]

⚠️ Super Fox-Wolfram: **Fisher discriminant** of Fox-Wolfram, the terms in the summation of \( H_l \) separated by \( B \) candidates and others.

⚠️ Fisher discriminant: linear combination of discriminating variables. The coefficients are determined to maximize the separation power.

\[
S = \frac{\left( \vec{w} \cdot (\vec{\mu}_{y=1} - \vec{\mu}_{y=0}) \right)^2}{\vec{w}^T\left( \sum_{y=0} + \sum_{y=1} \right) \vec{w}}
\]
Continuum Suppression

\( \cos \Theta_{B^*} \)

\( \Theta_{B^*} : \) angle between B candidate and beam direction in the B rest frame.

\( Y_1^1 (\cos \Theta_{B^*}) = \sin \Theta_{B^*} \)
\( = \sqrt{1 - \cos^2 \Theta_{B^*}} : \) signal \cdot background
Continuum Suppression

\( \cos \Theta_{B^*} \)

\( \Theta_{B^*} \): angle between B candidate and beam direction in the B rest frame.

\( Y_1^1 (\cos \Theta_{B^*}) = \sin \Theta_{B^*} = \sqrt{1-\cos^2 \Theta_{B^*}} \):

![Graph showing signal and background events](image)
Production of the neutrinos result large missing energy of the background event compare to that of signal.

\( E_{\text{miss}} \)

\( \overline{B}B \) events have leptons, which are produced along with neutrinos.

signal \cdot background
$\delta z$ : vertex displacement of two leptons in z direction.

Background events have larger distribution than signal because their mothers are usually different.
**B̅B Suppression**

**δz**

δz : vertex displacement of two leptons in z direction.

Background events have larger distribution than signal because their mothers are usually different.

![Graph showing δz distribution](image)